Exploring the factors that influence adoption of open government data for commercial service innovation in cities

Giovanni Maccani

PhD 2016

Maynooth University
National University of Ireland Maynooth

Exploring the factors that influence adoption of open government data for commercial service innovation in cities

Giovanni Maccani

Thesis for the Degree of Doctor of Philosophy

Maynooth University, Maynooth, Co. Kildare, Ireland

School of Business

September 2016

Head of School: Professor Peter McNamara

Supervisor of Research: Professor Brian Donnellan

Co-Supervisor of Research: Doctor Markus Helfert

Abstract

City councils produce large amounts of data. As this data becomes available, and as information and communication technology capabilities are in place to manage and exploit huge amount of data, open government data is seen as becoming more and more valuable in several contexts. Information Systems (IS) research on open data has been primarily focused on its contribution to e-government inquiries, government transparency, and open government. Recently, open data has been explored as a catalyst for service innovation as a consequence of claims around the potential of such initiatives in terms of additional value that can be injected into the worldwide economy. Subsequently, there is an active open data services academic conversation among IS researchers. Notwithstanding this, the evidence of impacts of open data adoption is largely scattered and anecdotal. A systematic literature review study highlighted the underexplored users' perspective within this field. On the other hand, previous research demonstrates that existing IS adoption theories should be extended as currently "a large part of the variance in the use of open data technologies is not yet explained" (Zuiderwijk *et al.* 2015, p.437). To address these gaps, an interpretive multiple case study was carried out to explore the factors that influence the adoption of open data for commercial service innovation in cities. In particular, factors influencing this process were inductively derived from the in-depth investigation of three organizations that achieved sustained adoption of open data for the delivery of commercial services across different cities and industries.

This research contributes to the IS open data services debate by proposing a model of factors perceived by open data services businesses as the most relevant in explaining adoption of open government data for commercial service innovation in cities. Adopting an inductive reasoning approach through qualitative methods enabled the generation of thick descriptions that were extracted, clustered, defined, and validated for each component of the model. This approach was critical to capture the complexity of the open data ecosystem perceived by those re-using this data.

With respect to IS adoption theories, by positioning the cross-case findings within those theoretical elements defined in the literature as predicting adoption of IS, a new enriched model has been formulated and is proposed as a contribution to this theoretical debate. Furthermore, this research demonstrates the suitability of interpretive multiple case study research to inductively generate knowledge in this field.

I expect that the findings of this study will be leveraged for stimulating and encouraging successful use of open data, therefore increasing the value of this source as a catalyst for service innovation. Recommendations to city councils are also proposed from the findings of this research.

Acknowledgements

Both my parents, Claudio and Franca, have worked very hard to ensure I have every opportunity in life. This is as much a product of their work as it is mine. This extends to my brother Federico who has never ceased to support and love me.

My heartfelt thanks to Paulina for her unwavering encouragement and fate as well as her patience and understanding.

I would like to express my profound sense of gratitude to my supervisors Professor Brian Donnellan and Doctor Markus Helfert for their constant constructive insights and inspirations throughout the course of this PhD project.

I am grateful to Prof Piero Formica who introduced me to Maynooth University, Intel Corporation, my supervisors of research, and, more generally, to the research world. It was because of Piero that I have started this program.

I acknowledge that this work wouldn't have been possible without the greatly thoughtful help from Prof Anne Huff across all stages of my PhD research program.

A special thanks to Dr David Prendergast, Jim Kenneally, and Prof Martin Curley from Intel for their input and for the great opportunities they have given me during these years.

Thanks to my annual reviewer Prof. Robert Galavan, Dr. Paul Donovan, and Dr. John Cullen for their time in seriously reviewing my work and for providing incredibly valuable feedback.

I especially thank Edward O'Connor for being a great friend during this period. Thanks also to Tadhg Charleton and Reka Petercsak.

I appreciate and acknowledge the contributions to this research from the Business Informatics Group (BIG) at Dublin City University and particulalry I would like to thank Dr Martin Meyer, Dr Plamen Petkov, Phelim Murnion, Dr Karen Carey, Patrick Brandtner, Dr Shuyan Xie, Dr Owen Foley, Dr Mouzhi Ge, and Dr. Lukasz Ostrowski.

Sincere thanks to all participants in this research; to all interviewees, Dublin City Council, and particularly thanks to all the representatives of the cases for their precious, serious, and proactive collaboration.

Last but not least, I would like to acknowledge the generous support and funding received from the Irish Research Council and Intel Corporation for the award which made undertaking my doctoral study possible.

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

This dissertation presents findings from an inductive multiple case study on open government data adoption by private organizations for commercial service innovation in cities. In particular, factors influencing this process were inductively derived from the in-depth investigation of three cases. These represent businesses that achieved sustained adoption of open data for the delivery of commercial services across different cities and industries. Overall, the following research question has been formulated and addressed in this study:

What are the factors that influence Adoption of Open Government Data for Commercial Service Innovation in a City Context?

The findings from this research include the definition of factors, variables and links between factors emerged as relevant in explaining adoption of open government data across the three cases studied. The positioning of these findings within Information Systems (IS) literature on open data services and adoption theories, allowed the formulation of research contributions to both of these academic domains.

Through the analysis of the findings across cases, this research contributes to the IS academic conversation on open data services (Lindman *et al.* 2013a) by addressing two gaps in the existing literature: (1) the underexplored users' perspective within this ecosystem; and (2) the lack of inductive contributions to-date in this field. The factors emerging from this research complement the extensive extant literature focused on the supply-side of this ecosystem, by paying attention to the user perspective which was argued to inhibit open data innovation (Janssen *et al.* 2012).

On the other hand, previous research demonstrates that existing IS adoption theories should be extended as currently "a large part of the variance in the use of open data technologies is not yet explained" (Zuiderwijk *et al.* 2015, p.437). The choice of inductively investigating these phenomena enabled the establishment of additional constructs, specifically defined as predictors of open data adoption. Therefore, by relating the cross-case findings with those theoretical elements defined in the literature as predicting adoption of IS, a new enriched model has been formulated.

Overall, this research proposes an inductively generated model for open data adoption for commercial service innovation in cities. These findings are believed to stimulate and encourage successful use of open data, that has so far largely failed to meet its high expectations (Kaasenbrood *et al.* 2015) (Hjalmarsson *et al.* 2014, 2015) (Susha *et al.* 2015).

This introductory chapter is structured as follows: after describing the structure of this thesis (section 1.1), the main concepts involved in this research are tackled and defined; these include an introduction to open data (section 1.2), the specific conversation on open data services (section 1.3), and a summary of the research contributions' domains (section 1.4).

1.1 Thesis Structure

This thesis is structured to be consistent with the inductive approach advocated by (Huff, 2008), and with the steps proposed for building theory from multiple case studies (Stake, 2006) (Eisenhardt, 1989). Figure 1.1 provides a summary of the structure of this document, and the main outputs of each chapter.

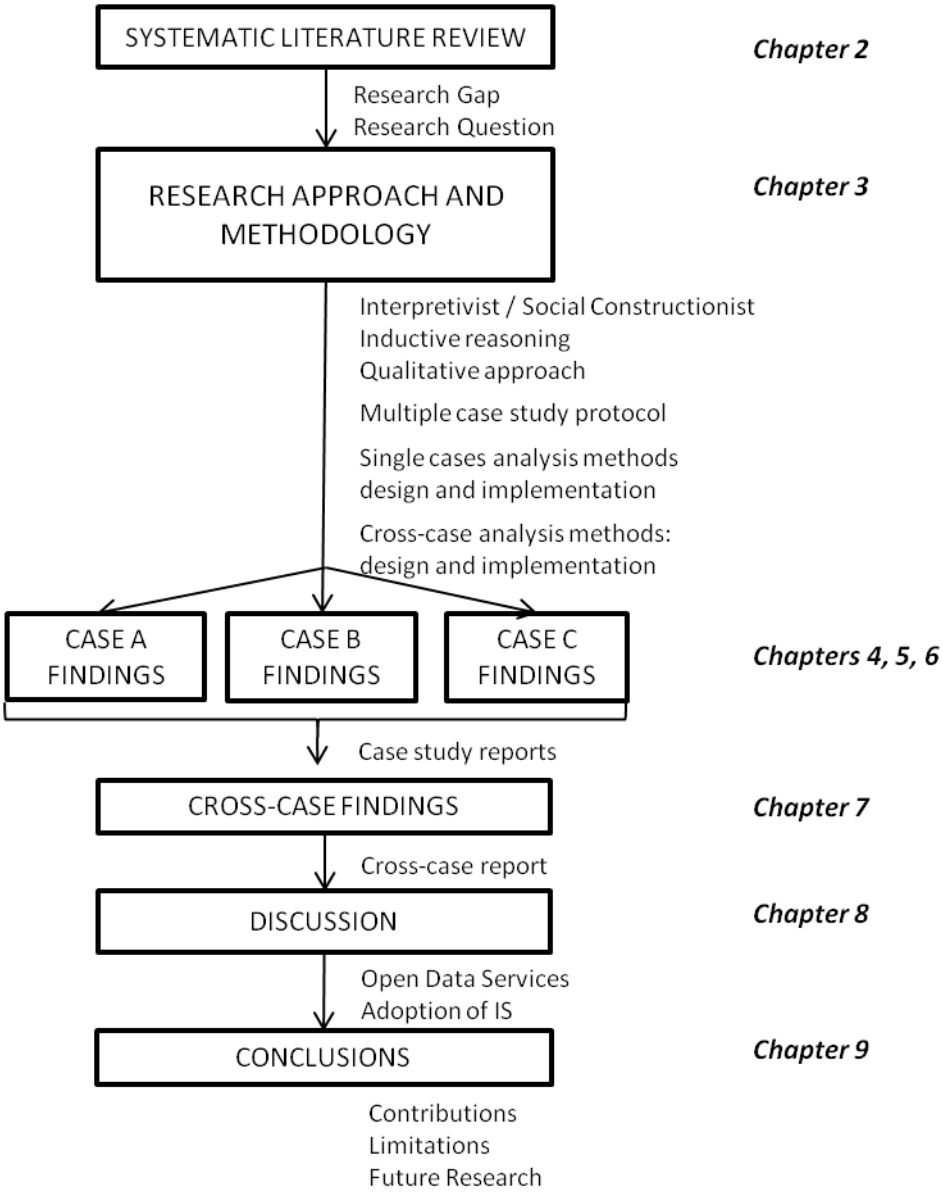

Figure 1.1: Structure of the thesis

As shown in Figure 1.1, Chapter 2 is dedicated to the Systematic Literature Review (SLR) undertaken for this study. Because of the inductive nature of this research, this review focused on identifying research efforts in terms of contributions to the open data services academic conversation. In other words, in order to minimize the risk for inductive researchers of extensively considering the literature at the early stage of the process, the SLR study had the main objective of mapping contributions to the field of interest and identifying opportunities for implementing a research effort. Therefore, Chapter 2 introduces the research gap identified, and the subsequent research question formulated for this study.

Chapter 3 outlines the approach taken and the methodology designed for this research. This chapter is divided into four main segments. The first is dedicated to the philosophical underpinnings of this research, and the key choices in terms of adopting an inductive reasoning through a qualitative approach to data collection and analysis. Subsequently, a comparison and reflection of methodologies that are acknowledged to be suitable with these decisions are provided. This led to the identification of multiple case study as the methodology for investigating the research question. This first section of the chapter is concluded with reflections on the role of IS adoption theories across the stages of this inductive study. The following segment outlines the actual methodology designed and implemented. In particular, it is dedicated to the multiple case study protocol, the cases' selection criteria, the data collection techniques employed, and to aspects about trustworthiness of this research. The third segment of Chapter 3 focuses on the actual analysis undertaken in relation to the single cases. Once all the case studies were carried out, the cross-case analysis was conducted. This phase was complemented by a validation effort undertaken through interactions with relevant actors in the open data services ecosystem outside the cases themselves. Therefore, the fourth and last segment of this chapter focuses on the specific methods and techniques leveraged for this step.

In Chapters 4, 5, and 6, the case study findings are reported. Each chapter includes a description of the case and the techniques leveraged based on the specific situations encountered; subsequently, findings represented by the thick descriptions of factors influencing open data adoption are presented for each individual case.

In Chapter 7 all the findings across cases are structured and presented.

In Chapter 8 the cross-case findings are discussed in relation to the two academic conversations tackled by this study. The inductive findings from this research are positioned within the existing literature on open data services as well as within those relevant theories proposed in relation to adoption of IS. Through this process, specific contributions to these two fields are proposed in Chapter

9. This dissertation ends with concluding remarks as well as with reflections about limitations of this study and future research avenues that are expected to be built on the findings achieved through this research.

1.2 Open Data

The definition of open data was firstly developed by the Open Knowledge Foundation in 2005 as "data that can be freely used, shared and built on by anyone, anywhere, for any purpose" (http://okfn.org/). There are three principles behind this definition:

1. Availability and access (i.e. people can get the data);
2. Re-use and redistribution (i.e. people can re-use and share the data);
3. Universal participation (i.e. anyone can use the data).

Public administrations' data was one of the primary applications of the open data movement, pioneered in the respective projects (http://data.gov) in the US, and (http://data.gov.uk) in the UK. Information on cultural attractions, public equipment and services, businesses available within cities etc., is naturally part of this effort. The general idea is that government agencies produce large amounts of data. As this data becomes digitally available, and as information and communication technology capabilities are in place to manage and exploit huge amount of data, open government data is seen as becoming more and more valuable in several contexts. In their Digital Agenda (http://ec.europa.eu), EU commissioners listed four reasons for promoting open data initiatives:

1. Public data has significant potential for reuse in new products and services. Overall economic gains from opening up this resource could amount to € 40 billion a year in the EU.
2. Addressing societal changes. Having more data openly available will help discovering new and innovative solutions.
3. Achieving efficiency gains through sharing data inside and between public administrations.
4. Fostering participation of citizens in political and social life and increasing transparency of government.

Producing open data from proprietary databases, however, takes significant effort. This is mostly due to the variety of native formats in which the data is produced and stored by the various stakeholders. Some intermediary organizations (e.g. DataPublica in France, Dublinked in Ireland) are being created to help and support the release of open data produced by national and local administrations. They usually proceed by manually establishing catalogues of public administrations' data portals and then semi-automatically feeding them into their warehouses as a collection of datasets. A dataset corresponds to a specific category of information aggregated at a specific level over a specific period

of time (e.g. *premature children born in each department of France in 2005,* http://datapublica.fr). The research presented in (Manyika *et al.* 2013) aimed at quantifying the potential value of re-using open data across seven domains of the global economy: education, transportation, consumer products, electricity, oil & gas, healthcare, and consumer finance. Their findings showed that between three and five trillion US$ can be created as additional annual value across these seven domains (Figure 1.2).

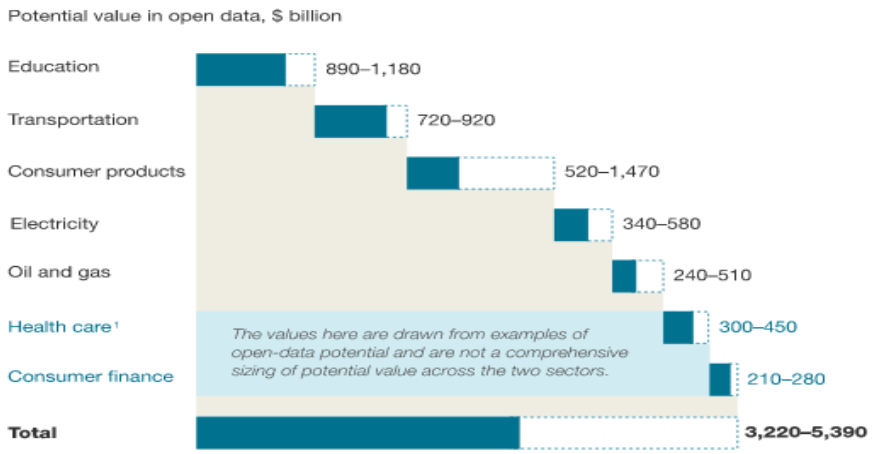

Figure 1.2: Potential value for open data; source: (Manyika et al. 2013)

Academic research on open data (and more generally on public sector information), was initially strongly related to e-government inquiries, addressing aspects of democratic theory, voter participation, democratic deliberation, and open government in a broader context (Amichai *et al.* 2008). Later, the focus included also open data which was seen as vital for the vitality of civic society (Bertot *et al.* 2010). Recent research explores open data as a foundation and catalyst of innovation (Lakomaa and Kallberg, 2013), and particularly service innovation. According to this stream, open data is seen as a source for web-based applications and information services innovations. The importance of open data as an enabler for innovation was presented through five categories (Lakomaa and Kallberg 2013): (1) simulate potential viability to ensure funding: “open data provides data that can serve as a test bed to simulate an application and build evidence for the factual viability of the proposed project”; (2) provide information about potential market: “open data creates an abundance of niche information through all sectors”; (3) reduce development lead time to application market: open data is “often of high quality and well-structured and thus requires less processing before it can be used in applications”; (4) drive innovation beyond applications: the usage of open data can be also indirectly reflected in the final outcome; and (5) enhance existing online services and offerings: open data can be used to create an additional user and customer value within already existing services. All these concepts led IS academic authors to introduce and structure a new research stream named open data services. This novel academic conversation is giving a structure to the investigation of open data

as a foundation for service innovation (see for example “Open Data Services Minitrack” at the Hawaii International Conference on System Sciences, 2013 (Lindman et al 2013a)). This research focuses on open government data released by city authorities, specifically defined in the literature as “all stored data of the public sector which could be made accessible by government in the public interest without any restriction on usage and distribution” (Geiger and von Lucke, 2011). Therefore, the term “open data” in this dissertation refers to this definition. Open data from private institutions, from national level governments, and open data adopted for other reasons than service innovation (e.g. research purposes (Zuiderwijk *et al.* 2015)) are beyond the scope of this study.

1.3 Open Data Services

In the open data services’ domain, a distinction between supply and demand of open data can be made (Lindman *et al.* 2013b). Specifically, the supply perspective (represented in this research by city authorities) aims at releasing data to the public. The demand side builds something useful on top of this source. Access to open data is therefore just the first step within an infrastructure that allows end-users to consume open data services. Hence, for the open data to become valuable there needs to be a chain of steps that take the raw data, make it available to others, and analyse, combine and present data in ways that make it useful for users to interpret as information (Lindman *et al.* 2013b). In (Tammisto and Lindman, 2012), a framework of actors and their roles in the open data services ecosystem is proposed (see Figure 1.3).

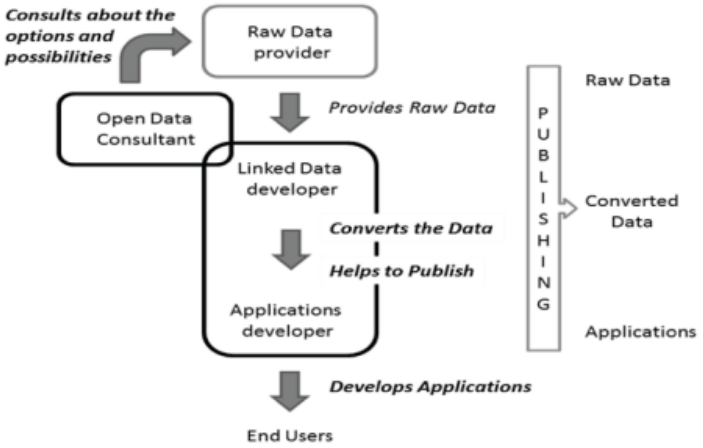

Figure 1.3: Open data services ecosystem; source: (Tammisto and Lindman 2012)

With respect to Figure 1.3, this ecosystem can be summarized as follows: the raw data provider publishes raw data; the linked data producer utilizes the raw data to produce (linked) open data; finally, the application provider utilizes the (linked) open data to produce a valuable application for

the end-user (Latif *et al.* 2009). The open data consultant informs and advices raw data providers about the possibilities of developing and publishing their data.

In (Lindman *et al.* 2013b), the authors adopted Alter's (2010) Work System Framework to describe open data services. According to (Alter, 2002), a work system is "a system in which human participants and/or machines perform work using information, technology, and other resources to produce products and/or services for internal or external customers". Thus, a service system can be seen as a work system that produces services for customers. The Work System Framework is a useful basis for describing and analysing an IT-reliant work system (Alter, 2010). These concepts were linked to open data services, and are represented in Figure 1.4.

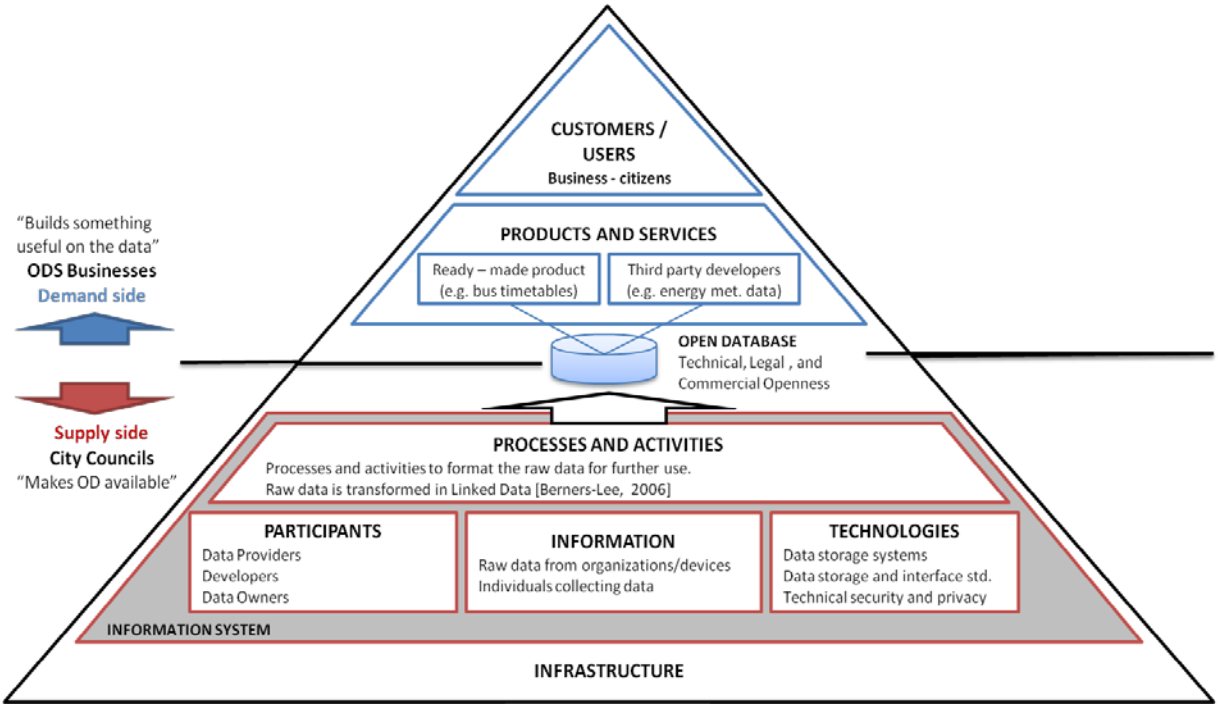

Figure 1.4: Open data services framework; adapted from (Lindman et al. 2013b)

In this framework two main parts can be identified following the distinction with supply and demand side of open data. Hence, there is an information system infrastructure that provides the open data on one side, while on the other side services are built on top of this data for the final end users. Within the supply side, four main factors can be identified within this framework. Firstly, the technological building block includes data storage systems and standards for interfacing these systems. The other three dimensions are: the type of raw data that is collected and transformed (information building block); the processes and activities that are undertaken to reach the development of linked open data (Berners-Lee, 2006); and the actors involved in this work system. Within the demand side of this framework, the data is assumed to be available as linked open data. With data openly available to the

public, designers can apply different models and/or theories to the data, and create new artefacts (Kuk and Davies 2011). Hence, at this stage open data is used to build a service. The range of actors involved in this step can vary from hackers, software consultancies, or even governments themselves. Two main types of services can be distinguished. Firstly, the service developer can take the data as it is and sell it to customers (e.g. apps, such as city bus timetables). Secondly, the open data can be exposed as a service to third party developers (e.g. energy metering data). Ultimately, to make open data services sustainable, there need to be customers. According to Alter (2010), customers are recipients of a work system's services for other purposes than performing work activities within the work system. Thus, customers of the open data services work system are users in terms of people or organizations that either pay for the services or finance them through other means (e.g. advertisement).

As a conclusion of their study, Lindman *et al.* (2013b) identify a research gap for open data services researchers. This was described as follows: "it is entirely unclear how to build a sustainable open data market and establish actors within it" (Lindman *et al.* 2013b, p.1242). However, I learned through the SLR study (Chapter 2) how the conversation around this topic is rapidly evolving and being enriched by several contributions.

In the context of open data services, this research investigates adoption of open data for commercial service innovation in city contexts. In other words, this research analyses new commercial service development (innovation in service products component) from open data technological innovation (innovation in service processes component) in cities' contexts.

1.4 Research Domain and Contribution

In summary, this research investigates adoption of open data for commercial service innovation in cities' contexts. The concepts of open data and open data services (i.e. open data for service innovation) have been clarified in this introductory chapter. Regarding adoption of an information system, this term is defined as a decision to make full use of a technological innovation as the best course of action (Rogers, 2010). Figure 1.5 summarizes the domain of this research, represented by the intersection of these three different fields.

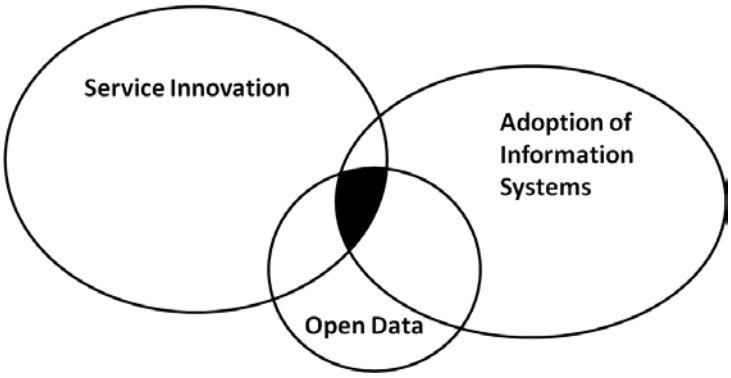

Figure 1.5: Research domain

As anticipated in the beginning of this chapter, contributions to both the open data services and the IS adoption literatures are proposed from this research. It is noted that only commercial open data services are taken into account. The scope of this research reflects a need observed in practice to understand the drivers for enabling local economic growth from the design and delivery of these services. In this way, open data outcomes have so far largely failed to meet those high expectations generated from the claims about its potential (e.g. EU estimates € 40 billion to be injected annually in the European economy from open data). Alternatively, if I had included also those services built as a result of voluntary work (such as the ones studied in (Kuk and Davies 2011), i.e. from one-day hack competition), the complexity of the open data ecosystem would be reduced to a simple interaction between a programmer and an open dataset, limiting, in my opinion, the subsequent findings, contributions, and implications.

2 SYSTEMATIC LITERATURE REVIEW AND RESEARCH GAP

Chapter 2 of this thesis is dedicated to the process that led to the systematic identification of the research gap and the subsequent research question for this study. A Systematic Literature Review (SLR) was implemented with the goal of mapping contributions to-date in the field of open data services (sections 2.1 – 2.4). After giving an overview of the sixty-six papers that were systematically ordered, searched, and selected for this study, a more detailed reflection on those that specifically focused on adoption of open data is provided (section 2.5). From this, a clear research gap is identified (sub-section 2.5.2), and the subsequent research question is formulated (section 2.6).

2.1 Eight-Step Systematic Literature Review

The SLR process undertaken adhered to Okoli and Schabram's (2010) eight-step method. The steps proposed are: (1) Purpose of the Literature Review, (2) Protocol and Training, (3) Searching for the Literature, (4) Practical Screen, (5) Quality Appraisal, (6) Data Extraction, (7) Synthesis of Studies, and (8) Writing the Review. Okoli and Schabram's (2010) method was chosen for its specific focus on IS research. This review was undertaken with the scope of identifying research efforts conducted in the field of open data services and subsequent gaps in the current knowledge base. Based on this objective, a systematic approach for reviewing the literature was found to be suitable. In other words, I acknowledge the presence and value of other forms of literature review (e.g. as theoretical foundation for primary research (Fink, 2013)). In particular, "conventional" (Kitchenham, and Charters, 2007) literature reviews focus on offering a scholarly critique on a specific existing theory (Zorn and Campbell, 2006) as an anchor for the subsequent research (Fink, 2013). However, given the novelty of this topic, this review was not merely focused on grounding the subsequent research, but also on summarizing "existing evidence, identify gaps in current research and provide a framework for positioning research endeavours" (Okoli and Schabram, 2010, p.3). Given these objectives, rigor becomes critical in order to ensure that the review captures all the relevant contributions. In summary, the definition assumed for this SLR is as follows: a systematic and explicit method identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by academic researchers (Okoli and Schabram, 2010).

2.2 Purpose of the Literature Review, Protocol and Training

The first step of the method addresses the fundamental question "why do a literature review?" (Okoli and Schabram, 2010). The purpose of this study was to analyse the progress of academic research in the field of open data services. Specifically, this study focused on reviewing the literature available on adoption and diffusion of open data for the development of services in city contexts. Referring to

Figure 1.5, the literature that was ordered and searched fits not only in the intersection of the three main circles, but also in the intersections “open data – service innovation” and “open data – adoption / diffusion”. These include contributions that are related to (and may overlap with) the focus of this research project, and so that have to be considered when systematically identifying the research gap and the subsequent research opportunity. As suggested in (Okoli and Schabram, 2010), five main sources for searching for the literature were taken into account. To synthesise the data that was collected, I decided to relate the academic studies available in the literature to the open data services work system framework previously introduced (see Figure 1.4). I was the only person conducting this study.

It is noted that the initial SLR study was conducted between September 2012 and January 2013. Given the novelty of the topic, the literature was constantly searched during the overall research program. The literature base supporting this project was updated subsequently.

2.3 Searching for the Literature

In this section, steps 3, 4, and 5, of (Okoli and Schabram, 2010)’s methodology (i.e. “Searching for the Literature”, “Practical Screen”, and “Quality Appraisal” respectively) are presented. Five main sources for searching for the literature were considered. On one side, the first sources considered were journals within the so-called “AIS Basket of eight” (https://aisnet.org/?SeniorScholarBasket). Across these journals, no publications specifically focused on open data were found. On the other hand, according to (Okoli and Schabram, 2010), the “predominant source of literature collection is now internet and particularly: Open Access Databases and Specific Subject Databases” (p.19). Within these two categories, the paper suggests specific websites to be used. Among those proposed, I chose: Google Scholar (http://scholar.google.com/), Directory of Open Access Journals (http://www.doaj.org/), ACM Digital Library (http://dl.acm.org/), and IEEE Xplore (http://ieeexplore.ieee.org/). For this study all of these sources were considered. In particular, the first two represent Open Access Databases, while ACM and IEEE were chosen among those related to specific subjects close to IS. Concerning the definition of the search terms used for searching for the literature across these four sources, the starting point was the research focus, which is re-proposed in Figure 2.1.

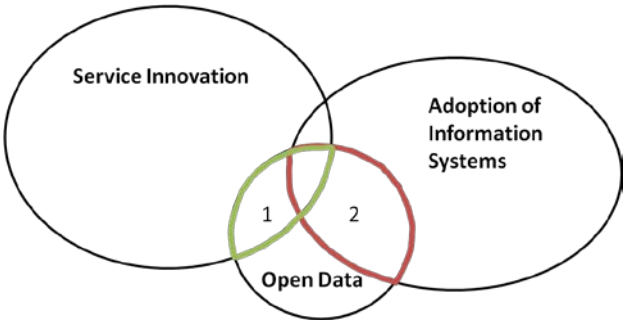

Figure 2.1: SLR, definition of search terms

With respect to Figure 2.1, ensemble 1 (circumscribed in green) and ensemble 2 (circumscribed in red) were objects of this Search for the Literature step. As a result, I decided to use the following search terms: "*Open Data Service(s*)", "*Open Data Innovation*", and "*Linked Data Service(s)*", for intersection 1; and "*Open Data Adoption*" and "*Open Data Diffusion"* for intersection 2. "Linked Data" is also taken into account as a search term, in order to include papers that used the terminology introduced in (Berners-Lee, 2006). Furthermore, given that the objective of the study is to analyse the adoption of open data for service innovation in a city context, I added one more search term to this process: "*Open Data City (Cities)*". This decision was made in order to include those contributions that tackled open data issues and opportunities with a specific focus on cities and their characteristics (e.g. city councils, policies, stakeholders etc.). Because of the commonality of the single words considered (i.e. "open", "data", "service", "innovation" etc.) and the subsequent huge amount of papers initially returned from the sources, I decided to set the search strategy on looking at the presence of these terms either within the title or the keywords of the papers. In other words, I did not consider those papers that include the search terms in the abstract or in the full text only. For example, if running a search on Google Scholar without any constraint and with the keywords "*Open Data Innovation*", more than three million results are returned (search conducted in March 2016). Furthermore, considering that the first definition of open data was developed in 2005, only papers that were published since that year were considered. Finally, only papers published in English were included in this review.

The sequence which I followed was: I first looked at Google Scholar, then at IEEE Xplore, ACM Digital Library, and finally at the Directory of Open Access Journals. As a result of the searching step I found a total of two hundred and sixty-two papers that satisfied the requirements across the four sources. The papers that were found from multiple sources are not duplicated in this count. Due to the large amount of papers, I decided to apply exclusion criteria based on both the titles and the abstracts, before proceeding with the reading of the full texts. This process enabled me to reduce the amount of papers to be read to sixty-eight. These studies were carefully analysed and, as a result, forty-six papers were ordered and searched and considered further. While reading these, a further eighteen

papers were imported into this review from an analysis of the references of those previously selected. This step enabled me to systematically cover the ongoing academic conversation on the domain of analysis. Finally, I had an opportunity to discuss some ideas on how to implement a research effort in this way with experts and researchers in the field of open data services. From this early stage interaction, they suggested to include two more studies in this analysis. A graphical representation of the searching process presented in this section is proposed in Figure 2.2.

Source	Key Terms	Initial Results	1st exclusion criteria	Full paper Exclusion criteria
GOOGLE SCHOLAR	Open_Data Service(s)	85	22	37 Articles ordered and searched
	Open_Data Adoption	7	4	
	Open_Data Diffusion	6	2	
	Open_Data Innovation	25	4	
	Linked_Data Service(s)	38	8	
	Open_Data City	19	8	
DOAJ	Open_Data Service(s)	0	0	0 Articles ordered and searched
	Open_Data Adoption	0	0	
	Open_Data Diffusion	1	1	
	Open_Data Innovation	0	0	
	Linked_Data Service(s)	2	0	
	Open_Data City	0	0	
IEEE Xplore	Open_Data Service(s)	15	5	5 Articles ordered and searched
	Open_Data Adoption	0	0	
	Open_Data Diffusion	0	0	
	Open_Data Innovation	3	2	
	Linked_Data Service(s)	45	4	
	Open_Data City	3	1	
ACM DIGITAL LIBRARY	Open_Data Service(s)	7	4	4 Articles ordered and searched
	Open_Data Adoption	3	1	
	Open_Data Diffusion	0	0	
	Open_Data Innovation	0	0	
	Linked_Data Service(s)	3	2	
	Open_Data City	0	0	
TOT.	--	**262**	**68**	**46**

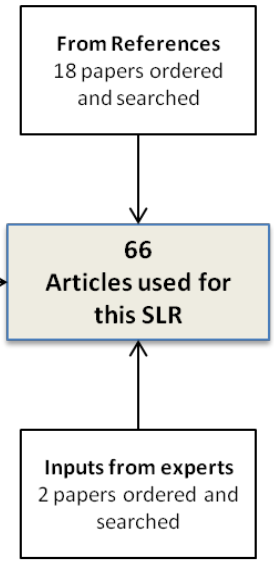

Figure 2.2: Searching for the literature process

Concluding, as a result of this step, sixty-six papers were ordered, searched, and selected for this SLR study. It is noted that step (5) of the method, i.e. "*Quality Appraisal*" (Okoli and Schabram, 2010) was not undertaken in this study. This step is meant to involve a closer examination of the papers in order

to assess their quality. Its main purpose is to reduce the total amount of papers to be considered. However, given the novelty of the topic, and that a manageable amount of papers was obtained as a result of the analysis of the content, there was no need for a further reduction.

2.4 Data Extraction and Data Synthesis

Within the sixth step of the method, all the data needed for the final synthesis were collected. The output of this step was a complete list of concepts from which I could synthesize this review study and identify the research gap/gaps in the field of open data services. In order to keep track of the main concepts that arose from these sixty-six papers, I built an excel table in which the focus of each of paper as well as its main findings were included. Furthermore, I divided the selected papers in three main categories following the classification between supply – and demand – side of open data (see Figure 1.4). In particular, the papers were divided between three groups, defined as:

- Group A: these papers focused merely on the supply side of the open data services framework.
- Group B: papers of this second group cover both the demand and the supply side of the framework. Typically, these studies focus on different solutions for releasing open data and examples of services implemented from these solutions. Also general frameworks describing the open data services ecosystem were included in this group.
- Group C: the studies included here are merely focused on the demand side of the open data services ecosystem.

The number and the percentage of the papers that belong to each group are represented in Figure 2.3 (next page).

Considering the objectives of this SLR, i.e. to systematically review the literature available on adoption/diffusion/usage of open data for service innovation in a city context, Groups B and C represent the key literature for identifying the research gap that this research project aims to fill. In other words, this research focuses on the demand side of the framework. However, for completeness, also the literature that fits within Group A is synthesised.

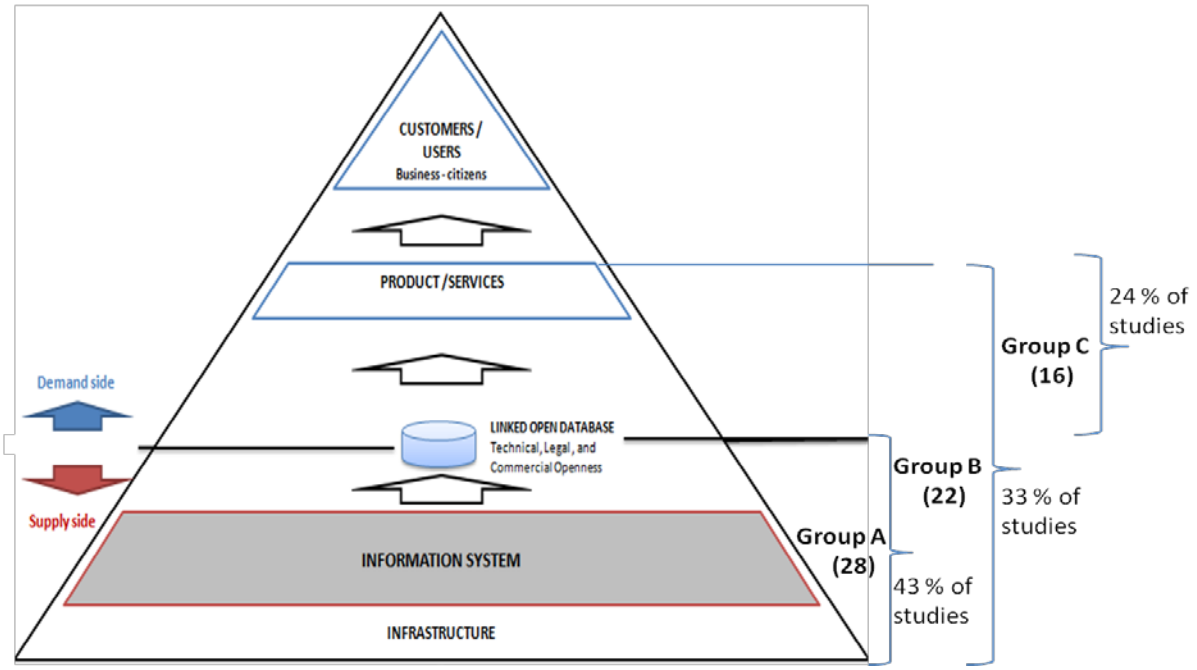

Figure 2.3: Open data services contributions

2.4.1 Group A – Supply Side

As shown in Figure 2.3, the open data services literature to-date is mainly focused on the socio – technical related solutions for making open data available for service innovation (i.e. supply-side of the framework). This section provides an overview of these contributions.

The studies presented in (Cifuentes-Silva *et al.* 2011), (Fox, 2013), and (Cyganiak *et al.* 2010) propose different semantic web solutions to enable and maintain the publication of linked open data. Similarly, in (Aoyama and Kojima, 2013) a new architecture to deliver linked open data is designed. It is based on two distinct layers: resource-oriented architecture and meta-level service brokers. Wilde (2010) contributed to this debate by comparing architectural semantic web principles with systems using the Representational State Transfer (REST) architectural style of the web in the release process of linked open data. Another system design for the release of open data is proposed in (Bellini and Nesi, 2013). The focus is on the ECLAP project in the context of multimedia content on performing arts. Rossi *et al.* (2007) draw scenarios of technological innovation from integrating open data with the concept of open source software. As part of their study, a framework describing adoption of open data standards in the context of open source software is proposed. This specific debate is enriched by Lindman and Nyman (2014) who proposed existing similarities and differences between open data and open source. However, the focus on open data is related (in both studies) to data coming from the private sector, which is beyond the scope of this research. In (Felicetti *et al.* 2012), the authors go beyond the purely technological view. In their paper institutional challenges for open data providers to develop and

deliver linked open data are outlined. Similarly, in (Currie, 2013), a comparative analysis between different open data initiatives is conducted to identify different institutional frameworks (in terms of organizational structures, e.g. number of departments of the municipality involved etc.) for the delivery of open data. Palka *et al.* (2013) focused on the process level when studying open government data release. This study proposes three different classifications schemes at both the strategic and the operational levels. Within these, elements for the design of open data portals are outlined. Several other papers in this group focused on singular projects. For instance, Bedini *et al.* (2014) present the approach followed for the design of the system used in Trento (Italy) for the publication of open data. The authors claim that the system is created for "creating space for innovation for businesses" (p.1). However, the focus of their study is only on the supply side. Isaac and Haslhofer (2011) describe an EU project (Euopeana) for delivering culture-related open data. The same project is presented in (Haslhofer and Isaac, 2011) but from a different angle. In fact, the main contribution of this article is the learning acquired in using URI (Uniform Resource Identifiers) solutions and RDF (Resource Description Framework) storage standards, and the resulting performances in releasing linked open data. Another EU project (i.e. DataBridges) is described in (Herschel and Manolescu, 2012). Their paper aims at providing an overview of the overall program as well as of the main activities that were undertaken for its implementation. In (Frosterus *et al.* 2012), the case studies analysed and described as best practices are Finnish projects on legislation for open data. The purpose of their study is to propose a set of semantic data standards for improving discoverability and interoperability of datasets. Another case study is described in (Oh, 2013) together with a comparative study between different initiatives implemented in New Zealand. This analysis is leveraged to derive implications regarding the format with which open data should be released. In a more comprehensive and holistic way, Stephenson *et al.* (2012) describe an Irish project named Dublinked as a novel open innovation portal to publish and provide access to a wide range of urban data from multiple sources. In (Tinati *et al.* 2013) the contribution of social networks – driven communication of open data initiatives is analysed for the UK open government data movement. From another perspective, in (Lorey, 2013) a novel approach to overcome the situation in which datasets are too often isolated between one and another is outlined. The author proposes "Data as a Service" principles to combine warehousing and query concepts in order to relate different datasets and obtain linked open data. The same concepts are adopted in (Kang and Son, 2008). However, this paper uses these principles to identify and solve data quality management issues. In (Yu *et al.* 2011), the focus is on data sharing in a particular sector (i.e. education), in order to generate useful datasets.

Another clear focus of papers within this group is the investigation of tools for assessing open data suppliers' performances in the release process. Two main topics were identified. Firstly, a framework

for assessing and benchmarking open data initiatives is proposed in (Viscusi *et al.* 2014) from a multiple case study of British, Italian, and Tunisian governments' settings. Secondly, emphasis was placed on the importance of "digital innovation contests" (Juell-Skielse *et al.* 2014, p.247) as a way for data suppliers to create links between open data and "third-party developers". Two papers were found from these authors regarding these contests (Juell-Skielse *et al.* 2014) (Ayele *et al.* 2015). The first proposes recommendations, guidelines and tools for improving their efficiency. The second transforms the identified critical factors and issues into criteria against which these initiatives should be assessed.

Finally, Janssen *et al.* (2012) formulate a fish-bone diagram of benefits and barriers for open data initiatives, and derive a list of "myths" connected to this field. However, the starting point of this study is the "reluctance" of public organization to embrace the idea of open data, and services are not specifically mentioned. Subsequently, Barry and Bannister (2014) referenced these myths, for deriving a categorization of the barriers related to the release of open data. Ultimately, Wang and Lo (2016) imported within this specific debate a deductive study on adoption of open data by governmental agencies examining three specific factors: "Perceived Benefits, Organizational Readiness, and External Pressure" (p.86).

Concluding, two papers were found to be particularly relevant for detailing the research gap tackled in this project, although these investigate the supply-side of the open data services ecosystem. Indeed, both (Janssen *et al.* 2012) and (Barry and Bannister, 2014) represent an attempt of studying adoption and diffusion of open data. Therefore, a more detailed reflection on these studies is provided later in this chapter.

2.4.2 Group B – Supply and Demand Side

As shown in Figure 2.3, twenty-two of the sixty-six papers considered for this SLR tackled both the supply and the demand side of the open data services framework. Many of these papers initially introduce technological – related solutions for the efficient and effective publication of open data. Subsequently, these studies provide examples of usage of open data that was released under the proposed solution. In particular, (Feng *et al.* 2012)'s main contribution is related to technological semantic solution to be used in the release of open data. However, this paper belongs to Group B as a section is dedicated to the description of a service delivered from open data that was released through the proposed solution in the real estate domain. The same structure was identified in: (Guo and Kraines 2010), in which the application proposed facilitates the search process of open data; (Matheus *et al.* 2012) where the application developed is related to government transparency; (Ijima *et al.* 2011) in which examples include a range of mobility services designed and implemented in Japan;

and (Hannemann and Kett, 2010), where the service addresses the integration of open data from the German library service providers.

The second category of papers within Group B is concerned with studies that propose general frameworks. These typically cover all the activities and processes that are undertaken from the raw data to the adoption of open data for different purposes. Hartung *et al.* (2010) present the design of an "open data kit" to act as a support in implementing processes for data collection, storage, query, and application development. The same objectives were pursued in (Tcholtchev *et al.* 2012). However, in this case the authors included architectural principles to demonstrate how ISP (Internet Service Provider) networks contribute to the provision of efficient streaming of "continuous" data. In (Rittenbruch *et al.* 2012) the design of a framework for the development of an "urban integrated open data API" represents the main contribution. A conceptual path from the collection of raw data to the delivery of services is provided. A scenario of services in the urban mobility context is used as an example. A similar focus was addressed in (Gautreau *et al.* 2012) and (Chan, 2013). The key term used across these two papers is "strategy". The first defines a strategy to share environmental-related data at the country level; the second, in a more academic format, concludes that open innovation strategies can be adapted to foster participation and collaboration for e-services design from open data. Furthermore, other general analyses were conducted in (Hausenbals, 2009) and (Bizer *et al.* 2009). These two papers are mentioned together because both aimed at providing a general overview of the concept of open data and linked open data. The first also provides an example of adoption of this source through the description of a BBC project within the context of music. Moreover, two more studies are included in this category: (Lindmann *et al.* 2013b) and (Tammisto and Lindman, 2012). These are considered the seminal contributions upon which the academic conversation around open data services in IS research was drawn. Specifically, the former defined this new research stream through Alter's (2010) Work System Framework (see Figure 1.4) and proposes a research agenda in this field. The latter is the paper in which the business model for open data is defined (see Figure 1.3). Similarly, Latif *et al.* (2009) investigated the value of open data from its "raw data stage" to its usage. Actors involved in this processes helped the authors at identifying potential pitfalls that have to be considered by business engineers in this field. Building on top of these findings two more papers attempted to structure actors and activities in the open data services landscape. Firstly, Lindman *et al.* (2014) proposed an "open data value network" from the investigation of actors and related activities within the complex processes involved in transforming open data into final services. Secondly, Zuiderwijk *et al.* (2014a), take into account the concept of ecosystems (Adner and Kapoor, 2010) to identify "key elements" for creating an open data ecosystem (it is noted that open data was explored beyond service innovation principles). Finally, the last paper in this sub-category is Bonina's

(2013) comprehensive study on open data in terms of: definitions, economic value, business models, and value and ethics in this field.

Moreover, three more papers are included within Group B, and were classified as "isolated" as not related to the two categories mentioned in this sub-section. For example, Janssen (2011) investigated the consequences that the European directive on public sector information had on the principle of freedom of information, and provides insights for policy maker in order to not harm it. Kuk and Davies (2011) analyse "one-day hack competitions" to derive principles on how open data can foster sustainable service innovation. The paper looks at the topic with the lens of Design Science (Hevner *et al.* 2004) to derive five different layers that describe the assemblage of open data complementarities. Finally, in (Foulonneau *et al.* 2014) a state-of-the-art analysis on open data provision and re-usage is proposed. The paper concludes stating that re-users lack interest in re-using this source, that the open data movement is "clearly not a demand-driven initiative but rather an initiative of data providers" (p.8) and claims about the potential of open data are inadequate.

The last and final paper that was grouped within Group B is (Zuiderwijk *et al.* 2014b)'s introduction to the Special Issue on "Innovation through Open Data" of the Journal of Theoretical and Applied Electronic Commerce Research. The paper presents an updated literature review in this field and specifically focuses on theories that were used in research for investigating open data. Given its importance for defining the research gap that this research aims to address, this paper will be tackled in detail later in this chapter.

2.4.3 Group C – Demand Side

This last group of papers includes studies that are related to the demand-side of the open data services framework. As shown in Figure 2.3, sixteen of the sixty-six papers considered for this SLR, were included in this group. Three of them are represented by studies in which innovative services are introduced. These services are either designed on top of open data or integrate open data with other sources of data (e.g. sensor data). In particular, Halb *et al.* (2010) present a prototype for professional online editors to facilitate the production of multimedia content. The service is enabled by linked open data sources. However, open data from the private sector is taken into account and no mention to open government data was found in the paper. In (Savelyev *et al.* 2011), the authors introduce the concept of "volunteered geographic services" as an extension of the idea of volunteered geographic information. Through the combination of open data and semantic technologies, a mobile app is proposed. Once again, however, the focus is beyond open data from the public sector. In (Groen *et al.* 2013) an innovative service for tourists in Amsterdam is proposed. It takes inputs from sensor data

and “crowd-sensed” data, as well as from open data, and provides the tourists with different information about cultural attractions around the city (e.g. queuing time at the Van Gogh Museum).

Several other studies included in Group C represent attempts of investigating service development from open government data. Lakomaa and Kallberg (2013) propose an analysis that was conducted in 2012 aiming at the inclusion of open data as a catalyst for innovation in web-based applications and information services. The authors investigated the perceived importance of open data for IT entrepreneurs. Their findings showed that 82% of the entrepreneurs surveyed claimed that access to open data would support and strengthen their business plan, while 43% see open data as a main requirement for the implementation of their value proposition.

Mainka *et al.* (2015) focus on mobile application development from open data. However, the main goal of this study is to investigate to what extent these applications are developed by the governments themselves and, on the other hand, what is the level of participation of citizens in these processes. In other words, the authors aimed at exploring the level of “crowd participation as a new form of citizens’ engagement in the city’s development” (p.3).

Another sub-group of papers aimed at proposing business models for open data services developments. Two papers, (Ferro and Osella, 2012), and (Ferro and Osella, 2013), presented the progresses achieved from the same research project. The authors used the Business Model Canvas method (Osterwalder and Pigneur, 2013) to propose a set of business models for private entrepreneurs that aim at harnessing and using public datasets for profit-oriented businesses. However, no specific examples of cases that successfully (or unsuccessfully) implemented these business models are provided. In order to address this limitation, Zimmermann and Pucihar (2015), after re-visiting the literature supporting the business model archetypes proposed in (Ferro and Osella, 2013), provide three examples of entrepreneurs who designed open data services with an associated revenue model. Notwithstanding this, data re-used by these cases is not merely open government data but a mixture between data coming from the public and the private sectors.

The last sub-category within this group reflects several attempts of studying adoption, use, acceptance, and diffusion of open data. The remaining eight papers to be synthesised fitted in this sub-group. For example, in (Hjalmarsson *et al.* 2014) and (Hjalmarsson *et al.* 2015) innovation barriers for the market entry of open data services developed at “digital innovation contests” are deductively explored. The two studies are motivated by an observed “limited number of the results developed during these events becoming viable services or products” (Hjalmarsson *et al.* 2014, p.2) The first paper investigates innovation barriers for open data services starting from a list generated from an

extensive literature review on innovation success factors. Surveys were leveraged to quantitatively explore the importance of these barriers for the commercialization of open data services designed at three different Swedish open data innovation contests. The second paper enriches results of the first by extending the one-stage study into a longitudinal one. Results of this research include the listing of twenty-four innovation barriers derived from the literature and the impact of each of these in regards to the market entry of open data services. A theoretical lens to study the adoption of open data is also applied in both (Hielkema and Hongisto, 2013) and (Palka, 2013). In the first article, Porter's model (Porter, 2011) is leveraged to study the forces driving the development of mobile applications from open data. The specific focus is on how "living lab" methodologies can drive forward the urban competition for open data. The second applies Diffusion of Innovations theory (Rogers, 2010) to deductively study the spread of open data. The emphasis is on implementing portals and applications that "foster a stronger democracy". The specific aim does not include service innovation, but citizens' participation and government transparency. Investigating usage of open data in digital services is also the scope of Immonen *et al.*'s study (2014). This explorative study derived a number of opportunities and challenges from eleven interviews of representatives of businesses involved with data-oriented services. However, the focus is clearly on open data from the private sector. The paper concludes with several motivations that should stimulate private businesses to release open data. These are aligned with the principles of open innovation (Chesbrough, 2006), e.g. the emergence of new ideas, new partners, and new business models (Immonen *et al.* 2014).

The last three studies considered for this SLR are complete researches aiming specifically at studying adoption of open data (Susha *et al.* 2015), (Kaasenbrood *et al.* 2015), and (Zuiderwijk *et al.* 2015). Firstly, Susha *et al.* (2015) leveraged Bin's (2013) framework for studying motivations for adopting open data. A survey was distributed and returned by twenty-eight businesses across Netherlands and Sweden. Objectives for open data service development did not include achieving economic value, but exploring and experimenting with open data. The second paper is represented by Kaasenbrood *et al.*'s (2015) study on open government data adoption by private organizations. The paper proposes a framework built from a literature review on open data to investigate the drivers for its adoption. This framework is subsequently leveraged to study open data adoption across five Dutch cases of businesses that re-use open data "as an enhancement of their already existing products" (p.86). The paper concludes by proposing a set of recommendations for the development of open data governmental policies. Finally, the last paper considered for this SLR is the study on adoption of open data proposed in (Zuiderwijk *et al.* 2015). In this paper the authors describe a deductive study testing the Unified Theory of Acceptance and Use of Information Technology (UTAUT) (Venkatesh *et al.* 2003) in predicting the usage intentions of open data for research purposes. Based on a sample of one

hundred and eleven subjects, the constructs proposed in the theory are quantitatively analysed. The final contribution of the paper is framed in relation to the above mentioned theory. In particular, the authors argue that the adapted UTAUT model, and the related predictors of adoption, "account for 45% of the variability of the behavioural intention to use open data" (p.438).

In the last three sub-sections I synthesised the data collected through this SLR. The focus was on highlighting the main contributions of the selected studies. I divided the papers collected in three main groups based on whether they were focused on the supply or demand-side of the open data services framework, or on both. Some reflections on the results achieved through this SLR study are provided in the next section. The research gap that this research project aims to fill will be initially derived and subsequently highlighted.

2.5 Critical Discussion on Existing Open Data Services' Literature

At this stage of the process, I have systematically identified the research efforts that have been implemented in the topic of open data services, with a specific attention to IS research. The SLR was focused on studies that made an attempt to tackle the adoption and diffusion of open data for fostering service innovation in a city context, consistently with the definitions provided in Chapter 1. As previously mentioned, in (Lindman *et al.* 2013b) a literature review was presented, together with a research agenda. Thirty-five research questions were formulated to guide future IS research in this field. The authors concluded that there is a clear research gap, and in particular "it is entirely unclear how to build a sustainable open data market and establish actors within it". Furthermore, "it is of crucial importance to explore from both theoretical and practical perspective, what are sustainable open data network structures and how sufficient revenue can be created" (Lindman *et al.* 2013b, p.1242). However, this SLR study showed how the conversation around this topic is rapidly evolving and being enriched by several academic contributions since the first definition for open data services was formulated.

As documented by this SLR the demand-side of the open data services framework remains significantly underexplored. In other words, through this SLR study I found a plethora of suggestions on how to supply open data to application – or more generally service – developers. A wide range of technical solutions is proposed in the literature to be used by open data suppliers to release their data effectively and efficiently. On the other hand, also a significant amount of papers was found to be focused on defining various types of frameworks aiming at describing actors and activities within the open data services ecosystem. Finally, only a minority of papers was found to be related to the demand-side of this landscape. A number of these studies propose the design of a service (e.g. (Halb *et al.* 2010), (Savelyev *et al.* 2011), (Groen *et al.* 2013)) on top of open data. It can be inferred that the

main purpose of these papers is to demonstrate the potential for open data for service innovation. Furthermore, Ferro and Osella (2012, 2013), and Zimmermann and Pucihar (2015) outlined a set of business models for new potential profit-oriented open data services businesses, partially filling the research gap identified in (Lindman *et al.* 2013b).

From another perspective, some academic researches present empirical studies across different projects or services. In addition to this SLR, I have searched across municipal open data websites of European capital cities and of the top fifty cities in Europe in terms of population. The findings of this search confirmed the statement made in (Kuk and Davies, 2011), i.e. in the open data context the most significant emphasis towards service innovation has been placed to catalyse "civic hacking" taking through weekend-hack days and competitions. High profile competitions, such as "Apps4BCN" in Barcelona and "App4MI" in Milan among many others, have emphasized the potential benefits of releasing data to developers, and enabling them to build services off the back of it. However, as also shown in (Kuk and Davies, 2011), the long-term delivery of these services remains extremely limited. As argued by these authors, of one hundred and thirty distinct projects investigated, only ten remained active one month later. The authors make a very strong statement as a conclusion of their study: "grand claims for the service revolutions that open data may bring about are overstated" (Kuk and Davies, 2011, p.15). In this way, Hjalmarsson *et al.* (2014, 2015) attempted to explore innovation barriers that these services (i.e. open data services developed at hack competitions) face for entering the market, i.e. for becoming sustainable profitable services.

Overall, the literature shows that the potential for open data for service innovation has been demonstrated in several ways (in addition to the claims made by EU and McKinsey mentioned in the introductory section). For example, as documented by Zimmermann and Pucihar's findings (2015), "publicly provided open data are not in conflict with profit-oriented businesses, but they complement each other" and "entrepreneurs and other businesses create added value for the benefit of consumers, other businesses, or public bodies as well and thus for society as a whole" (p.9). Therefore, the usefulness for the private sector of open data for commercial service innovation has been observed (Lakomaa and Kallberg, 2013) and business models solutions have been conceptually proposed (Ferro and Osella, 2012) (Ferro and Osella, 2013) (Zimmermann and Pucihar, 2015). However, the recent literature also acknowledges how "the evidence of impacts of open data adoption is largely scattered and anecdotal" and how "the intersection of open data and innovation is one of the main research gaps in the current open data literature" (Susha *et al.* 2015, p.20).

2.5.1 Discussing papers on Adoption of Open Data

Therefore, an initial research gap is identified in relation to adoption of open government data for service innovation. In this way, the fact that limited knowledge is available on adoption of open data

is well acknowledged in the literature (Susha *et al.* 2015). However, as shown in the last part of the previous sub-section, a few studies attempted to investigate the process of adopting open data. For example, barriers impeding open data –driven innovation (within and beyond service innovation) were proposed in the existing literature. These range from institutional, operational, economic, legal, cultural, to technological barriers (Hjalmarsson *et al.* 2014, 2015) (Janssen *et al.* 2012) (Barry and Bannister, 2013). Therefore, in order to outline a clear research gap, it is needed to look at these studies in detail. In total, ten papers (and nine related studies) were found to be particularly relevant for defining a clear research gap in this research as a result of the SLR. A brief description of each of these nine studies is provided in the following paragraphs. It is noted that other papers initially included in the SLR study focused on adoption and diffusion of open data. These, for example (Immonen *et al.* 2014), were not considered further as the main focus was placed on open data from the private sector, which is not consistent with the scope of this study.

Firstly, in their introduction to the special issue on "Innovation through Open Data", Zuiderwijk *et al.* (2014b) propose a literature review and an emerging agenda for open data research. The focus on re-usage of open data includes, but is not limited to, service innovation. The review included the analysis of one hundred and forty-three papers. Partially confirming the results of this SLR, the authors' findings are stated as follows: "papers on open data revealed that publications were mainly conceptual papers, descriptions of the empirical uses of open data or described the design of technology and systems" (p.3). On the other hand, "the use of theories received considerably less attention" with "institutional and organizational theories [that] were used to investigate policy development, changing systems, and changing organizational cultures and structures" and "democratic theories [that] were used to investigate transparency and trust and participation" (p.4). The paper concludes stating: "this research revealed that little attention is paid to the user perspective, which is likely to inhibit innovation" (p.5), confirming the statement previously introduced about the demand side of the open data services framework being currently under-explored.

Secondly, using Bin's (2013) theoretical framework, Susha *et al.* (2015) investigated the importance of several constructs in driving adoption of open data for service innovation. These include: Perceived Benefits, Perceived Effort, Social Influence, and Personal Characteristics. Open data services are designed and delivered by these companies as "a learning experience, personal interest, reputation and visibility" (p.32). In other words, the adoption and use of open data was studied across existing companies that are "experimenting and exploring the possibilities of open data" (p.33). This study is conducted through surveying entrepreneurs in Sweden and Netherlands. The authors conclude their

paper with recommendations to the public sector that can be summarized as follows: (1) economic value is not the most powerful driver for open data adoption; (2) it is beneficial for the data publishers to liaise with the business community; and (3) public bodies are advised to invest more resources into improving support to businesses (Susha *et al.* 2015).

Thirdly, Kaasenbrood *et al.* (2015) propose a multiple case study of five Dutch companies for investigating open data adoption. For researching this phenomenon, a framework on potential barriers is inductively developed from a literature review on factors driving adoption combined with two initial exploratory interviews. The analysis becomes then deductive tackling multiple case studies. As documented in the paper, "none of the investigating organizations is heavily dependent on the use of open government data. Open government data are used as an enhancement of their already existing products and services. These products, aimed at the end-users are not profitable, nor they are expected to be" (p.86). As a result, the paper proposes a set of recommendations for policy development across four different topics: (1) building a community around the use of datasets; (2) simplifying the process of finding and accessing the datasets; (3) ensure national provenance and highest level of detail; and (4) ensure continuity (Kaasenbrood *et al.* 2015, p.87).

The fourth study selected is Zuiderwijk *et al.*'s (2015) analysis on acceptance and use of open data technologies. This research focuses on open data adopted by students for research purposes. This process is researched leveraging the Unified Theory of Acceptance and Use of Technologies (UTAUT) proposed in (Venkatesh *et al.* 2003). Zuiderwijk *et al.* conclude their study stating: "a large part of the variance in the use of open data technologies is not yet explained. Adoption theories for open data specifically are needed. There is a need for open data specific theories and methodologies that address the idiosyncratic nature of open data, including aspects such as data quality, institutional complexity, legal and economic aspects (…). We recommend that adoption theories specifically for open data are developed" (Zuiderwijk *et al.* 2015, p.437). Moreover, the authors underline how UTAUT should be integrated with "research that has identified causal antecedents of the constructs used within the model in order to provide a greater understanding of how the cognitive phenomena that were the focus of this research are formed" (p.437).

Two more studies, i.e. Hjalmarsson *et al.* (2014, 2015) and Hielkema and Hongisto (2013), aimed at tackling specifically the adoption and diffusion of open data for service innovation from innovation contests and hack competitions. These contests are defined as events in which "third-party developers compete to design and implement the most firm and satisfying service prototype for a specific purpose, based on open data" (Hjalmarsson *et al.* 2014, p.5). The first study was presented through two different papers (Hjalmarsson *et al.* 2014, 2015). This starts from a framework developed from

open innovation principles (Chesbrough, 2011) and a literature review on innovation barriers in IS, defined as constraints or factors that inhibit innovation (Piatier, 1984). The quantitative analysis that follows (as a one stage analysis in the first paper and as a longitudinal study in the second) explores the identified barriers deductively and to which extent these were inhibiting market entry of open data services. Thin descriptions and scores about the importance of each barrier are provided. The authors summarize their contribution as follows: "the results of this study provide contribution to practice as they give to the organizer (of these contests) insights of barriers hampering the long-term success of an inside-out open innovation process" (Hjalmarsson *et al.* 2014, p.13). The authors also acknowledge the need for future research in terms of understanding success factors "more profoundly using qualitative research" (Hjalmarsson *et al.* 2015, p.13). The second study, i.e. (Hielkema and Hongisto 2013), was found to be specifically focused on living labs methodologies as a way to foster "smart city development" (p.190) in Helsinki, Finland. Open data competitions are leveraged as the context to investigate the role of living labs. In particular, two contests are examined using the lens of Porter's (1990) model on competitive advantage of nations. Contributions of this study, however, are mainly framed around the role of living labs for collaborative networking, and the demonstrations of the relevancy of these methodologies in supporting policy goals of smart city developments.

Three more studies were found to be focused on open data adoption and diffusion among governmental agencies. Firstly, Palka (2013) introduces the idea of using Roger's (2010) Diffusion of Innovations theory to study the diffusion of open data portals among public authorities. Contributions from this study were not found in the literature, as the paper included in this SLR was in a research-in-progress format (Palka, 2013) and no further papers were published to-date from this project. However, the focus is clearly made on open data as a way to foster citizens' participation and government transparency. The remaining papers in this last group are (Janssen *et al.* 2012) and (Barry and Bannister, 2013). Both researches aimed at identifying the benefits and the barriers to the diffusion of open data, in terms of obstacles that governments face and need to overcome to release their data and make them "openly" available. Both of these papers used the lens of Institutional Theory (Scott, 1995) for investigating this field. Contributions of these studies include a list of several benefits, adoption barriers and "myths" (Janssen *et al.* 2012) perceived and faced by governments when developing their open data and open government programs.

Following these reflections, of the eight studies on adoption / diffusion of open data found in the current literature base, five are related to the demand side of the open data services ecosystem. These are summarized across five dimensions within the following table. In particular, for each paper, it is outlined: (1) the research approach (i.e. deductive vs. inductive, qualitative vs. quantitative); (2) the

phenomenon of interest; (3) the focus of the study; (4) the theoretical lens (or lenses) applied; and (5) a summary of the nature of each project's contribution.

2.5.2 Research Gap

As summarized in Table 2.1, adoption of open data has been studied across different contexts (e.g. for citizens' participation, for experimenting integration in existing services, etc.). A research gap can now be formulated from the careful analysis of the open data services literature to-date presented in this chapter.

Reference	Research Approach	Phenomenon of interest	Focus	Theory	Contributions
Susha *et al.* 2015	Deductive, quantitative	Open data adoption	Entrepreneurs re-using open data for "learning, visibility, and personal interest"	User Innovation Behaviour Theory (Bin, 2013)	Recommendation for open data suppliers
Kaasenbrood *et al.* 2015	Mixed method approach	Open data adoption	Integration of open data into existing services	Framework developed from literature review and interviews	Recommendation for policy makers
Zuiderwijk *et al. 2015*	Deductive, quantitative	Open data adoption	Adoption by students for research purposes	UTAUT Theory (Venkatesh *et al.* 2003)	Contribution to UTAUT theory and call for inductive studies
Hjalmarsson *et al.* 2014, 2015	Deductive, quantitative	Barriers for market entry of open data services	Open data innovation contests / competitions	Framework from literature review	Thin descriptions of 24 innovation barriers
Hielkema and Hongisto, 2013	Deductive	Mobile application development and living labs	Open data innovation contests / competitions	Porter's (1990) model	Benefits of living labs in supporting smart cities

Table 2.1 Summary of key papers

I conclude that there is a clear research gap concerning the investigation on how open data is adopted for commercial service innovation in city contexts, in order to achieve the benefits that have been claimed in the literature, and surprisingly not yet achieved in practice. Furthermore (this aspect will be extensively addressed in the following chapter), the calls strongly made in (Zuiderwijk *et al.* 2014b), (Zuiderwijk *et al.* 2015), (Hjalmarsson *et al.* 2014, 2015) for moving from deductive and quantitative

towards more qualitative approaches are reflected in the lack of studies that inductively tackled this field so far.

2.6 Research Question

In the introductory chapter of this dissertation, I underlined how open government data programs are being implemented by governments worldwide for different purposes. A major focus for open data initiatives is placed on facilitating and powering the development of new services and/or products that ultimately are expected to enable local economic growth and improved competitiveness of cities. The SLR process revealed the existence of a clear research gap concerning limited knowledge available on adoption of open data by businesses. No studies so far tackled open data adoption for commercial service innovation in city contexts and, more generally, researched adoption of open data in an inductive and qualitative way. Research and practice on open data have been observed to be oriented towards data provision, rather than actual use. Although initiatives are in place for stimulating open data use, such as hack competitions and workshops, the predictors and constructs actually influence organizations' ability and intentions to re-use open data for commercial service innovation remain unexplored. Therefore, the following exploratory research question is formulated for this study:

What are the factors that influence Adoption of Open Government Data for Commercial Service Innovation in a City Context?

It is noted that the concepts of both open government data in city contexts, and service innovation have been defined for the specific purposes of this study in Chapter 1. With respect to the term "adoption" within the research question, detailed definitions will be provided in the following chapter when reflecting upon the usage of theory in this study (see section 3.6).

I believe that investigating these factors is needed to increase adoption, which is observed in the literature as failing to meet high expectations deriving from numerous potential benefits identified (Kaasenbrood *et al.* 2015) (Hjalmarsson *et al.* 2014, 2015) (Susha *et al.* 2015). I expect that understanding the motivations as well as the key drivers for businesses to use open data can provide a representation of the dynamics of open data services development. This is subsequently expected to help interpreting and solving the existing inhibitors that are in place for achieving those outcomes and performances that are stated as part of city councils' objectives for open data initiatives. This research aims at defining what factors are perceived as the drivers in influencing open government data adoption for commercial service innovation by those re-using it. The objective is therefore to generate knowledge for stimulating and encouraging successful use of open data. This is believed to increase the value of open data as a catalyst for innovation.

2.7 Chapter Summary

This chapter presented the SLR study undertaken as part of this research. All the steps followed have been carefully described. An overview of the contributions of the sixty-six papers that were ordered, searched and included in this review, was provided. A more detailed analysis of those specifically focused on adoption and diffusion of open data for service innovation, allowed the formulation of the research gap and subsequent question addressed in this research. The next chapter is dedicated to the key choices made for addressing this research question in terms of: philosophical underpinnings, reasoning, approach, methodology, and analysis methods.

3 RESEARCH APPROACH AND METHODOLOGY

This chapter is dedicated to four main segments related to this research: (1) key research decisions (sections 3.1 – 3.6); (2) research methodology (sections 3.7 – 3.11); (3) the analysis and reporting methods employed for the single cases (sections 3.12 – 3.13); and the method designed and implemented for the cross-case analysis (sections 3.14 – 3.18).

Starting from the research question previously outlined, in the first segment of this third chapter reflections are provided about appropriate and suitable choices in terms of: (1) research philosophy (section 3.1), (2) research reasoning (section 3.2), and (3) research approach (section 3.3). Based on these decisions, the investigation of potential research methodologies is also presented (section 3.4). This is followed by the rationale for choosing multiple case study as the methodological guidance for this research (section 3.5) and by reflections about the role played by adoption theories in this inductive study (section 3.6).

In the second segment, all the key features of the methodology in terms of both design and actual research activities are outlined, with a specific focus on the single cases. Firstly, an overview of the multiple case study protocol designed is provided (section 3.7). Subsequently, the chapter follows with: a description of the cases' selection criteria employed (section 3.8); the single case protocol designed (section 3.9); and the design of the data collection techniques that were leveraged (section 3.10). This segment concludes with some reflections about trustworthiness of this study (section 3.11).

The third segment focuses on the qualitative analysis methods designed and implemented for each single case. The techniques and the process followed for the analysis across cases is presented in Chapter 7, once the findings from each case will be outlined (Chapters 4, 5, and 6). Within this segment, an overview of the single case's analysis method designed is provided (section 3.12). Ultimately, the last section of this segment is dedicated to the techniques leveraged for reporting each of the case's findings (section 3.13).

The fourth and last segment of this chapter is dedicated to the methodology and the actual activities undertaken within the cross-case analysis stage of this research. After reflecting on the consistency of this process with this research's assumptions (section 3.15), the process of merging the cases' findings and the subsequent development of Cross-Case Assertions (CCAs) is outlined (section 3.16). This is followed by section 3.17 which focuses on the validation across cases stage. Finally, the activities implemented for positioning the cross-case findings within the literature are proposed (section 3.18).

3.1 Philosophical Underpinnings

All research is based on some underlying assumptions about what constitutes "valid" research and which research methods are appropriate (Myers, 1997). Hence, in order to conduct research, it is important to know what these assumptions are. According to (Saunders, 2011), philosophical assumptions in research must be established in order to define the "nature of the knowledge" that is aimed to be developed. Across different disciplines, a wide range of research perspectives and paradigms operate concurrently (Burrel and Morgan, 1979) (Astley and Van de Ven, 1983). Given the usually complex and indeterminate nature of research studies, "the existence of a plurality of perspectives enables the exploration of diverse questions and hence adds breadth as well as depth to the knowledge generated" (Orlikowski and Baroudi, 1991). Such decisions are considered to be the basis for the overall research design and implementation (Guba and Lincoln 1994).

In this research, the meaning assumed for the word "paradigm" is "the basic belief system or worldview that guides the investigator not only in choices of methods but in ontologically and epistemologically fundamental ways" (Guba and Lincoln 1994, p.105). Ontology focuses on "what exists", while epistemology considers what human beings can know about what exists (Huff, 2008). According to Burrel and Morgan's (1979) seminal work, by combining different ontological and epistemological positions, a number of different philosophical paradigms can be outlined. In particular, the literature acknowledges the interdependent relationship that exists between epistemology and ontology, and how one informs and depends upon the other. By considering this link, the need to understand the position of the researcher becomes more visible. "If the researcher holds certain ontological positions or assumptions, these may influence the epistemological choices or conclusions drawn" (Flowers, 2009, p.2). It is therefore fundamental to carefully take decisions about these assumptions at this stage of the research. A general idea on how such decisions can lead and influence the overall research process was provided in (Saunders, 2011) and is presented in Figure 3.1 (next page).

Figure 3.1 shows how these perspectives "help ensure that the core of data collection techniques and analysis procedures used in the research undertaken, are both appropriate and coherent" (Saunders, 2011). In general, the results of different combinations of ontological and epistemological choices are classified across three general research paradigms (Guba and Lincoln, 1994) (Orlikowski and Baroudi, 1991): positivist, critical, and interpretive.

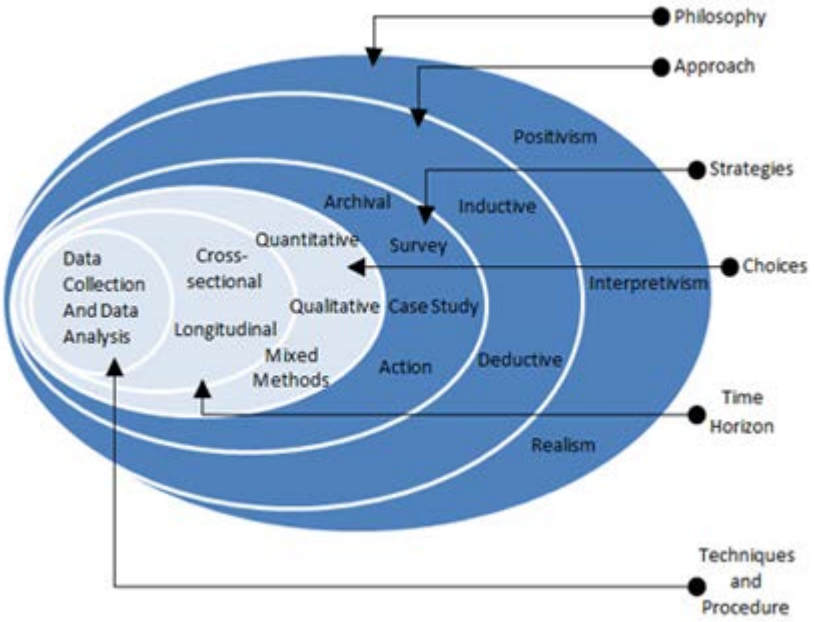

Figure 3.1: Research onion; source (Saunders, 2011)

Positivist research: the positivist perspective assumes that the real world and the social one exist as an objective reality that is unique. This position can be summarized through four main statements (Lincoln and Guba, 1985) (Burrel and Morgan, 1979):

1. The phenomenon of interest is single, visible and tangible, and there is a unique, best description of any chosen aspect of this phenomenon.
2. The researchers and the objects of enquiry are independent entities, and there is a "sharp demarcation between observation reports and theory statements" (Orlikowski and Baroudi, 1991).
3. Nomothetic statements can be formulated. As a consequence, emerging scientific concepts are precise and have a given and fixed meaning.
4. In reality, unidirectional cause-effect relationships exist and can be identified and tested through hypothesis-testing-oriented deductive analysis.

According to Orlikowski and Baroudi (1991), IS research can be classified as positivist if "there is evidence of formal propositions, quantifiable measures of variables, hypothesis testing, and the drawing of inferences about a phenomenon from the sample to a stated population" (p.5). Concluding, positivistic assumptions usually lead to confirmatory type of research (Berg, 2001).

Critical research: critical researchers assume that social reality is historically constituted and that it is composed and reproduced by individuals. "Although people can consciously act to change their social and economic circumstances, critical researchers recognize that their ability to do so is constrained by

various forms of social, cultural and political domination" (Myers, 1997). The main task of critical research is therefore to highlight social critiques, whereby the restrictive and alienating conditions of the status quo are the main focuses. Therefore, a critical position is concerned with criticizing existing social systems and "revealing any contradictions and conflicts that may inhere within their structures" (Orlikowski and Baroudi, 1991).

Interpretive research: interpretive thinking is described by Hatch and Cunliffe (2013) as anti-positivist and by Blaikie (1993) as post-positivist. Indeed, it is contended that there is a fundamental difference between the subject matters of natural and social sciences (Flowers, 2009). A fundamental distinction between the interpretive and positivist world views is the former's primary presumption of social constructionism. According to (Burrel and Morgan, 1979) the world is conceived of as "an emergent social process - as an extension of human consciousness and subjective experience". Thus, the interpretive perspective attempts "to understand the inter-subjective meanings embedded in social life and to explain why people act the way they do" (Gibbons 1987). In other words, the interpretive perspective focuses upon the importance of subjective meanings and social action in the processes through which humans construct and reconstruct their reality (Morgan 1983). In IS, according to (Orlikowski and Baroudi, 1991), "interpretive information systems research assumes that the social world is not given. Rather, the social world is produced and reinforced by humans through their action and interaction". As a consequence, organizations (and more in general social systems) can't be comprehended and/or measured in an objective, universal and deterministic manner. Therefore, interpretive researchers believe that social reality can only be interpreted. From an epistemological point of view, the "social process is not captured in hypothetical deductions, co-variances, and degrees of freedom (i.e. typical positivist approach). Instead, understanding social processes involves getting inside the world of those generating it" (Rosen, 1991).

3.1.1 Choosing a suitable Research Philosophy

The different ways of defining what reality is and how it can be known have been interpreted and adopted differently across several disciplines. However, none of these paradigms is considered to be better than the others. The choice depends merely on the purposes of the study.

The purpose of this research project is exploring the factors that influence the adoption of open government data for commercial service innovation in cities. Through the SLR study I learned that well acknowledged adoption theories have been leveraged for investigating open data adoption in a positivist way (Zuiderwijk *et al.* 2015) (Palka, 2013) (Hjalmarsson *et al.* 2014, 2015) (Kaasenbrood *et al.* 2015) (Susha *et al.* 2015). However, a call for models to be developed for open data specifically has been strongly made across these papers. As a consequence, I believe that a positivist approach can be

excluded from the choice of a potential suitable philosophical position. Indeed, also given the novelty of the topic, opting for a confirmatory type of research from well-established theories is not believed to be a beneficial strategy. Furthermore, several actors, activities, and concepts are involved in the process that takes raw open data and turns it into a valuable experience for end-users. As (weakly) demonstrated in the previous literature, it is the interplay of these actors, activities and concepts within a social system that makes it possible to deliver new open data services. As a result, considering such a reality as something unique and tangible, and that can be known by quantitative analysis is not consistent with the aim of this research.

A critical underpinning can be also excluded as the most suitable perspective from the purposes of this study. In particular, mentioning again its main characteristic in research explained above, it is not the scope of this study to criticize existing social systems and "reveal any contradictions and conflicts that may inhere within their structures" (Orlikowski and Baroudi, 1991).

An interpretivist perspective is considered as the most suitable for this project from both my perceptions and the literature systematically analysed. Thus, from an ontological point of view, as opposed to a critical realist perspective, where the nature of observed reality is partial and immutable (Mir and Watson, 2001), I acknowledge that the reality that is being studied in this research project is a manifestation of individuals' social construction. Studying open data adoption is understood as considering organizations that are made of people that make choices and implement activities in a complex social setting. In this ecosystem, many forces are influencing the process of adoption of open data, and their identification and consideration can't be neglected, rather it is part of the scope of this research. In other words, in the case of open data, the literature acknowledges that the role that is played by the context is fundamental to explain the process that refines open data from a raw dataset to an actual service. Example of contextual elements include hack competitions (Kuk and Davies 2011), the business and consultancy ecosystem (Lindmann *et al.* 2014), the policy context (Janssen *et al.* 2012), the organizational structure of the open data publisher and the people within it (Barry and Bannister, 2013). In this way, an interpretative approach helps to take into account the interplay of the organizations and their environment (Walsham, 1993) and ecosystem. As a consequence, I believe that only through the in-depth study of the experiences of the people that actually went successfully through this process, it is possible to derive significant contribution towards building theory on such an adoption phenomenon. From another perspective, I believe that predefined dependent and independent variables would hide (and oversimplify) the full complexity of this process. From an epistemological point of view, IS interpretive research "aims at producing an understanding of the context of the information system, and the process whereby the information system influences and is

influenced by the context" (Walsham, 1993). As a result, in this exploratory research, I see theory as an act of generation, as opposed to formalization of the underlying reality.

3.1.2 Social Constructionism

The resulting paradigm assumed for this project is social constructionism. Its origins can be traced in interpretivist approach to thinking. Burr (1995) acknowledges the major influence to its development of (Berger and Luckmann 1967). The most important elements of the theory of knowledge of social constructionism are (Leeds-Hurwitz, 2009): (1) the assumption that human beings rationalize their experience by creating a model of the social world and how it functions; and (2) language is the most essential system through which humans construct reality. A major focus of social constructionism is to uncover the ways in which individuals and groups of people participate in the construction of their perceived social reality. In turn it involves investigating how social phenomena are created, institutionalized, known, and enacted by humans. In the IS field, given its nature, the conventional dichotomy between the social and the technical is problematic as technical and social choices are constantly negotiated and socially constructed (Bloomfield and Vurdubakis, 1994).

In the case of this research project, the focus on adoption of open data for commercial service innovation is concerned with the usage of this information system. Users (or adopters) can be seen as social actors as opposed to "traditional" users. Concerning the latter, IS literature studied adoption relying on individualistic cognitive models to carefully examine the criteria that influence the selection of information and communication technologies (e.g. Davis, 1989). A shift towards considering them social actors "sharpens perceptions of how organizational contexts shape ICT-related practices, and at the same time helps researchers more accurately portrait the complex and multiple roles that people fulfil while adopting, adapting, and using information systems" (Lamb and Kling, 2003). Consequently, this IS adoption process is of particular interest since, specifically in the city-level open data domain, it happens in an articulated social context. The characteristic of the IS innovation (open data) is just one of the variables that is expected to explain its adoption. Additionally, the topic under investigation is still emerging, and the idea that information systems emerge from relations among a heterogeneous set of elements is well acknowledged in the literature (Law, 1987) (Fulk, 1993).

Ontologically, the basic assumption is that open data released by cities' authorities diffuses among service developers through an apparently natural process, but that in fact is complex, and contingent on several social actors and activities. From the epistemological point of view, social constructionism has been adopted in IS as a thought concerned with unravelling how these phenomena are constructed (Mitev, 2000). The results of this research will be then social construction of reality, in which the theory developed assumes a correspondence of meaning because a variety of subjects share

a common sense about the adoption process of open data for commercial service innovation (i.e. the reality to be studied). Furthermore, Strang and Meyer (1993) argue that "in a more phenomenological vein, the process underlying adoption can be seen as an inherently sense-making one, where actors jointly construct an understanding of the appropriateness and worth of some practices" (Strang and Meyer, 1993).

However, as for every research paradigm, some criticisms to social constructionism emerge from the literature. Primarily, social constructionism is argued to see individual subjects as "empty vessels filled up through discourse" (King, 2004, p.43). This aspect was addressed in the data collection methods design. However, I believe that it is actually through stimulating the intellectuality of the people that are constructing the reality of adoption of open data, that I can derive theoretical insights. Moreover, in order to overcome this criticism, other forms of data collection were employed in addition to pure discourses with subjects.

In summary, the main characteristics of these ontological and epistemological choices are:

Ontology:

- There are multiple realities (truth is partial).
- Realities exist through different humans' experiences, involving their opinions, knowledge, interpretations and choices.
- In the context of interest, in general terms, adoption of open data happens because users decide to use it for commercial service innovation purposes. These users are the social actors that construct, define, and make sense of this reality.

Epistemology:

- Understanding social process involves getting inside the world of those generating it.
- This research approach operates under the assumption that only through social constructions such as shared meanings and languages, access to reality is possible.
- Reality in this study is investigated starting from the understanding of the meaning given to this world by the open data users that are those social actors that enact its adoption process.

Concluding, a social constructionist approach to research gives me the opportunity to explore how open data is adopted for commercial service innovation purposes. Such a choice had fundamental reflections in the subsequent selection of a suitable research reasoning, approach, and methodology.

3.2 Research Reasoning

In research two general reasoning can be involved: inductive and deductive (also “abductive” (Van de Ven, 2007) was defined as a combination of them). Reasoning in this way is defined in (Goel and Weachter, 2004) as a “the process of evaluating given information and reaching conclusions that are not explicitly stated”. According to Huff (2008), deductive studies start with empirical statements about pertinent phenomena. These statements are translated into hypothesis using the language of well-established theories. Subsequently, observations are made to see whether or not they are true. Philosophically, this underlines a positivistic worldview. On the other hand, Huff (2008) describes inductive studies as those where the meaning is linked to a specific observer in a specific situation. Typically, this approach underlines an interpretivist worldview. The researcher starts with specific observations and measures; subsequently, he or she begins to detect patterns and regularities and formulates some tentative hypotheses to explore; the researcher finally ends up developing some general conclusions or theories.

This study, in accordance with its social constructionist philosophical underpinnings, involves an inductive reasoning. It is noted that this approach is also consistent with the research gap identified through the previously presented SLR study (see Table 2.1). Indeed, such an approach is new to this academic conversation, and is considered as necessary for better understanding open data adoption (Zuiderwijk *et al.* 2015) (Hjalmarsson *et al.* 2015). A further reason why I decided to adopt an inductive approach is its well acknowledged suitability for exploring emerging topics (Zhang and Wildemuth, 2010) such as the focus of this research. Generally, inductive studies can be conceptualized as a set of steps (Glaser and Strauss, 1967) (Martin and Turner, 1986). Huff (2008) summarizes them as follows:

1. Extensively describe an interesting situation without use specialized vocabulary from the existing academic literature.
2. Create a first level of substantive categories by coding these descriptions.
3. Modify and improve codes as additional data are collected, categorized and compared.
4. Expect theoretic insights to emerge as categories stabilize and their relationships become apparent.
5. Conclude empirical observation when new categories are not required to account for further observations.

These steps were taken into account at the time of choosing a suitable research methodology, the subsequent data collection techniques, and for the design of the data analysis process.

3.3 Research Approach

Another choice that is acknowledged as being fundamental in research is between qualitative and quantitative approaches. On one side, quantitative research methods were originally developed in the natural sciences to study natural phenomena. As the name suggests, such studies involve the investigation of phenomena through numbers. This can subsequently entail the usage of mathematical models, statistical ones, or computational techniques (Given, 2008). Examples of quantitative methods include surveys, laboratory experiments, formal methods (e.g. econometrics) and numerical methods such as mathematical modelling (Myers, 1997). On the other hand, qualitative studies were developed in the social sciences to enable researchers to study social and cultural phenomena. Qualitative data sources include participants' observation (fieldwork), interviews and questionnaires, documents and texts, and the researcher's impressions and reactions (Myers, 1997).

A variety of reasons led me to opt for a qualitative choice. Firstly, it is considered consistent with the philosophical paradigm assumed of social constructionism and the choice of adopting an inductive reasoning. Secondly, the emerging concept of open data for service innovation, according with the SLR's findings, must be investigated through an exploratory-oriented study. This is in contrast with a more confirmatory type (from established theories) which would lead to choosing a quantitative approach. Thirdly, as the reality of adoption of open data for commercial service innovation in cities is analyzed from the meaning given by the people that are actively involved in constructing such reality, qualitative research methods are appropriate. Indeed, they are "designed to help researchers understand people and the social and cultural contexts within which they live" (Myers, 1997). Kaplan and Maxwell (1994) argue that the goal of understanding a phenomenon from the point of view of the participants and its particular social and institutional context is largely lost when textual data are quantified. In addition, I believe that hiding the complexity of the context(s) in which open data is adopted and the multiplicity of actors that influence this process, would negatively affect the conclusions. In other words, this study must generate rich data. Unlike most quantitative studies, which are generally limited to thin descriptions of phenomena by way of frequencies, distributions, and statistical patterns of relationships between constructs (Brekhus *et al.* 2005), rich data are deeply implicated in bringing to life the human beings that are the centre of social science research. Thus, rich data is meant as data that enable thick descriptions, thick interpretation, and thick meaning (Ponterotto, 2006). Thick descriptions were described in (Geertz, 1973) as "ways of providing cultural context and meaning that people place on actions, words, things, etc. Thick descriptions provide enough context so that a person outside the culture can make meaning of the behaviour. Thin descriptions by contrast, are stating facts without such meaning or significance". Miles and

Hubermann (1994) have strongly advocated the strengths of qualitative data to generate rich descriptions, and, finally, to investigate the topic "from the inside" (p. 255).

While this choice is usually driven by previous statements made about previously-determined philosophical and reasoning choices, other approaches exist. For example, Patton (2005) argues that qualitative analysis does not implicitly exclude deductive reasoning. In fact, generating concepts or variables from theory or previous studies is also believed as useful for qualitative research, especially at the inception of data analysis (Berg, 2001). Therefore, the word "qualitative" is not a synonym for "interpretive". According to Myers (1997) "qualitative research can be positivist, interpretive, or critical". However, I demonstrated how qualitative approaches are the most suitable for this study. I highlighted how in the context of studying open data adoption for service innovation in cities there is a fundamental need of generate rich data and subsequent thick descriptions to fill the identified research gap.

3.4 Choosing a suitable Research Methodology

At this stage of the research a suitable research methodology had to be identified and selected. This section describes the main methodologies that are consistent with the choices previously made and the objectives for this study. A reflection on the potential suitability of each methodology is presented. The term 'research methodology' is understood as a strategy for inquiry which moves from the philosophical underpinning choices to the design of the actual research process and the subsequent data collection and analysis. What is known so far is that a social constructionist philosophy has been assumed. The approach to theory is inductive, and findings are generated through qualitative data collection and analysis. A brief summary of the implications of these choices can be stated as follows: **this research aims at engaging with the social actors that enact the adoption process of city – level open data for commercial service innovation; through this process, this study aims at qualitatively uncovering the ways through which these social actors construct their perceived reality; the final objective is to generate thick descriptions which are expected to inductively generate theory**.

According to the literature, as there are some philosophical perspectives which can inform qualitative research, so there are a number of qualitative research methodologies. In IS research, Myers (1997) outlined four main qualitative methods: Action Research (Davison *et al.* 2004), Ethnographic Research (Myers, 1999), Grounded Theory (Glaser and Strauss, 1967) (Martin and Turner, 1986), and Case Study Research (Stake, 2006) (Eisenhardt, 1989) (Eisenhardt and Graebner, 2007) (Yin, 2013). However, it could be argued that this list is incomplete. In particular, by looking at the overall management

literature, two more qualitative methodologies should be included; these are: Phenomenology (van Manen, 1990) and Narratives (Polkinghorne, 1995) (Chase, 2005).

The main characteristics of each methodology are provided in the following sub-sections. These are enriched by some reflections upon their potential suitability for investigating the research question formulated for this project.

3.4.1 Action Research

Action research has been defined as an approach that "combines theory and practice (and researchers and practitioners) through change and reflection in an immediate problematic situation within a mutually acceptable ethical framework" (Rapoport, 1970). It "is highly context dependent while attempting to address the specific client's concerns" (Iivari and Venable, 2009). It is noted that a plethora of "versions" in which action research was formulated are available, e.g. (Baskerville and Wood-Harper, 1998) (Baburoglu and Ravn, 1992) (Davison *et al.* 2004). In relation to this research project, action research could be considered as a consistent guidance in terms of its anti-positivistic epistemological nature (Iivari and Venable, 2009). However, there is a main issue which would profoundly weaken this choice. Indeed, unlike this study, action research projects start with a problem formulated by a specific organizational setting (or in a highly participatory approach with the researchers' team) (Nandhakumar *et al.* 2005). Subsequently, the nature of the research inquiry and results are implicitly interlinked with the original setting. It is clear here that the organization is at the core of this research approach. Therefore, action research is excluded from the choice of a suitable research methodology for this study.

3.4.2 Ethnographic Research

According to Creswell (2007), "an ethnographer is interested in examining shared patterns, and the unit of analysis is larger than the twenty or so individuals involved in a grounded theory study" (for details see Grounded Theory below). Hence, the focus of ethnographic research resides upon a cultural group. According to the literature, these groups are typically large. The key for ethnographers is to get people together and let them interact over a period of time. Thus, as an anthropologist (Myers, 1997), the ethnographer spends a long time in the field and "immerses himself in the life of people he studies" (Lewis, 2004). The final goal is to place the phenomena studied in its social and cultural context. Therefore, ethnography is a way of studying a culture-sharing group involving "extended observations of the group, most often through participant observation" (Creswell, 2007). Ethnography was used in several IS studies, e.g. (Harvey, 1997) (Orlikowski, 1991), and it became a widely accepted method among IS researchers (Harvey and Myers, 1995).

In relation to this research, ethnography is found to be suitable in terms of its interpretive philosophical underpinnings and its inductive approach to theory. However, this methodology was rejected. In fact, this research focuses specifically on a domain constituted by different entities (businesses) that are all involved in the process under investigation (i.e. open data adoption for commercial service innovation) but across different contexts (e.g. operating with different datasets). Furthermore, given the early stage of the open data phenomenon, I believe it would be too difficult to find a large number of people from different contexts and join them together for a "long time". Ultimately, from the definitions provided above, it is clear that the most used data collection technique is observation. In the context of people/organizations developing services on top of open data, this technique is believed to have a limited role.

3.4.3 Grounded Theory

In order to develop a new field of knowledge, the literature recognizes that is necessary to develop theory about pertinent phenomena (Glaser and Strauss, 1967). Based on this statement, the grounded theory approach was designed and established as a well acknowledged methodology over the last decades (in IS see for example Special Issue on Grounded Theory of the European Journal of Information Systems (Birks *et al.* 2013)). It is seen as a powerful tool for rigorous theory development, and it is defined as "a systematic methodology involving the discovery of theory through the analysis of data" (Martin and Turner, 1986). This theory is believed to be grounded in the analysis of actual settings and processes (Urquhart *et al.* 2010). In other words, the theory development process is generated starting from the participants who have experienced the phenomenon under investigation. Many of the characteristics related to this study outlined earlier in this chapter are reflected in this method. These include: the fact that it is a theory building-oriented methodology; it represents the perfect example of inductive reasoning; it sees qualitative approaches as being the key to achieve a research contribution; it focuses on the process as experienced and perceived by its participants; and it derives categories (which is related to "factors" in the research question) as its final outcome. Also, from a philosophical point of view, the grounded theory method can be used under constructivist underpinnings (Charmaz, 2006). However, there are some aspects that need further reflections. An important angle to be considered in relation to this project is the role played by existing theories. More specifically, beside the recently introduced academic conversation on open data services (Lindman *et al.* 2013b), this research project aims also at contributing to the broad adoption of IS theories. As a consequence, this conversation needs to play a role in interpreting the qualitative data collected and in giving the right language to explain why and how this project contributes to it. In addition, for the same reasons provided for ethnographic research, it would be practically difficult (if

not impossible) to conduct “fifty or sixty interviews” (Creswell, 2007). Therefore, I believe that grounded theory is not entirely suitable for the objectives of this study.

3.4.4 Phenomenological Research

Phenomenology is primarily concerned with the study of structures of experiences and consciousness (Creswell, 2007). Hence, this method focuses on describing what all participants have in common as they experience a phenomenon. The basic purpose of phenomenology is to analyze individual experiences with a phenomenon and formulate a description of the universal essence (van Manen, 1990). However, the literature acknowledges that its philosophical assumptions rest on some common grounds: the study of the lived experiences of persons, the view that these experiences are conscious ones (van Manen, 1990), and the development of descriptions of the essences of these experiences, as opposed to their explanations or analyses (Moustakas, 1994). While at a first sight this method could be thought as a potential candidate for this study, phenomenology focuses on merely describing phenomena (Moustakas, 1994) (van Manen, 1990) (Creswell, 2007). This research project aims to go beyond the pure description of lived experiences of social actors involved in the adoption of open data. In fact, its main contribution is expected to be reflected in the interpretation and analysis of such experiences. As a result, phenomenology is not believed to be suitable for the purposes of this research.

3.4.5 Narrative Research

Narrative might be the term assigned to any text or discourse, or, it might be text used within the context of a mode of inquiry in qualitative research (Chase, 2005). The specific focus is on the stories told by individuals (Polkinghorne, 1995) (Pinnegar and Daysen, 2007). Like much of qualitative research, narrative inquiry explores life experiences. It describes and analyzes these experiences using the language of “story.” For instance, the field texts or data that narrative researchers gather may be called “stories” of life experience. As every story, these include plots, characters and so forth (Creswell, 2015). More in general, this methodology is defined as a specific type of qualitative design in which “narrative is understood as a spoken or written text giving an account of an event/action or series of events/actions, chronologically connected” (Czarniawska, 2004, p.17). Biographical studies, autobiographies, and life histories are common forms of narrative studies (Creswell, 2007). Analyzing an individual’s narrative is not considered as a suitable strategy for meeting the objectives of this research. Thus, it can be excluded from the choice of a suitable research methodology.

3.4.6 Case Study Research

Case study research is the most common qualitative method used in IS (Orlikowski and Baroudi, 1991). Although there are numerous definitions, Yin (2013) defines the scope of a case study as follows: "a case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin 2013, p.13). Case study research can be positivist, interpretive, or critical. In the literature, while Yin (2013) and Benbasat *et al.* (1987) are advocates of positivist case study research, Walsham (1995), Stake (2006) and (partially) Eisenhardt (1989), are advocates of interpretive, inductive, and in-depth case study research.

Case study research aims at understanding an issue, problem, or phenomenon using the case as a specific illustration (Stake, 2006). Thus, case study research is a qualitative approach in which the investigator explores a bounded system, i.e. a case in a specific setting/context, or multiple bounded systems over time. This investigation is conducted through in-depth data collection involving multiple sources of information. Results include reporting case descriptions and case-based themes. Already at a first sight this methodology seems to consistently fit in the context of this study. In particular, engaging with the social actors of the target phenomenon, uncovering the perceived realities of those involved in the social construction of the process under investigation, qualitatively generating thick descriptions to develop new theory, and having as a result "themes" (factors), are characteristics of both this research and the case study research methodology.

3.5 Rationale behind Adopting Multiple Case Study Research

As a result of the comparative analysis between different methodologies that are consistent with the decisions previously made, case study research is found to be the most suitable for the purposes of this research. In particular, multiple case study was selected. While single-case studies can richly describe the existence of a phenomenon (Siggelkow, 2007), multiple case studies typically provide a stronger base for theory building (Eisenhardt, 1989) (Stake, 2006). Furthermore, this choice comes from the personal belief that in order to understand complex processes such as the adoption of open data, it is useful to look at people, operations, and activities across different contexts. From another perspective, the wide range of open datasets, their characteristics, and all the contextual information are variables that can't be neglected. Thus, considering only one case would, in my opinion, substantially limit the findings of this research. Hence, by selecting a multiple case study approach, the opportunity of studying cases characterized by diversities in their contexts is leveraged in order to strengthen the findings.

In summary, the choice of multiple case study research methodology was motivated by several aspects. These are seen as both strengths of the methodology itself and aspects consistent with the purposes of this research and its philosophical, reasoning, and approach choices. In particular, multiple case study:

- Is a methodology consistent with social constructionism, inductive reasoning, and qualitative approaches (Stake, 2006) (Walsham, 1995) (Eisenhardt, 1989) (Eisenhardt and Graebner, 2007) (Lauckner *et al.* 2012).
- With its diversity across cases, it ensures richness and depth in order to understand the shared phenomenon of interest (Anaf *et al.* 2007) (Flyvbjerg, 2006) (Stake, 2006).
- Enables the exploration of complex situations allowing for the gathering of multiple perspectives, from a range of sources, including contextual information (Stake, 2006) (Flyvbjerg, 2006) (Lauckner *et al.* 2012).
- Is particularly useful when the unit of analysis is a process, which is compatible with the research question of this study (Stake, 2006) (Walsham, 1995) (Lauckner *et al.* 2012).

A further reason why multiple case study was chosen is because of its potential contribution to practice from both a single-and cross-case perspectives. Interpretive case study researchers conduct a form of action research by influencing what is happening in the field of analysis through sharing resulting concepts and interpretations with the people within the case studied (Walsham, 1995). This result is consistent with the objective for this research project in terms of its potential contribution to the world of people, businesses, and other entities involved in stimulating and implementing the creation of new services from open data. Finally, Orlikowski and Baroudi (1991), Benbasat *et al.* (1987) and Myers (1997) argue that case study method is particularly appropriate for the study of IS development, implementation and use within organizations. It is particularly appropriate when theoretical knowledge on the phenomenon under investigation is limited and an understanding is not well developed (Benbasat *et al.* 1987), such as regarding adoption of open data for commercial service innovation in cities.

As every research method, there are critiques and weaknesses in relation to multiple case studies. One of the main limitations of the methodology is argued to be about generalizability (Orlikowski and Baroudi, 1991). However, the fact that “one cannot generalize on the bases of case studies” (Flyvbjerg, 2006, p.221) is argued to be one of the general misunderstandings of these methodologies. In interpretive research, generalization from case studies is seen as a form of “explanations of particular phenomena derived from empirical interpretive research in specific IS settings, which may be valuable in the future in other organizations and contexts” (Walsham, 1995, p.79). From this perspective, I

consider it critical to provide diversity, as suggested by Stake (2006), in the multiple case selection. The issue around "generalization" emerged particularly at the time of implementing cross-case analysis. Further reflections on the suitability of this process in this inductive and interpretive study are provided in section 3.15. Another weakness of multiple case study research is believed to be grounded in the poorly defined data analysis process (Yin, 2013). In order to address this challenge, different sources for conducting data analysis were considered, compared, and chosen prior to the start of the actual data gathering process.

3.6 Theories on Adoption of IS and their role within this study

The central notion of this research is to use a multiple case study to develop theory inductively. "The theory is emergent in the sense that it is situated in and developed by recognizing patterns of relationships among constructs within cases and their underlying logical arguments" (Eisenhardt and Graebner, 2007, p.25). In this way, it is important to reflect upon the use of IS adoption theories and their role along the research process. This section outlines the role of theory in this inductive study (sub-section 3.6.1). Subsequently, reflections about the current literature on adoption of IS and its limitations are proposed and contextualised within this research consistently with the decisions previously taken (sub-section 3.6.2).

3.6.1 Role of theory in inductive multiple case studies

In the context of case study research, Eisenhardt (1989) discusses three main uses of theory. Based on these three positions, Walsham (1995) identifies their relations to the field of IS interpretive case study research:

1. Theory as an initial guide to design and data collection (in IS (Walsham, 1995)): in this case, theory is used in the early stages of the research to take into account previous knowledge. This creates a "theoretical base" (Walsham, 1995, p.76) to inform the topics and approaches of the early empirical work.

2. Theory as a part of an iterative process of data collection and analysis (in IS (Orlikowski, 1993) using grounded theory): in this second case, in a form of grounded theory approach, initial theories are "expanded, revised, or abandoned altogether" (Walsham, 1995, p.76).

3. Theory as a final product of the research (in IS (Orlikowski and Robey, 1991)): in this third case, Eisenhardt (1989) notes that the outputs of case study research can be concepts, a conceptual framework, propositions or mid-range theory.

In order to keep this research consistent with its philosophical assumptions, various degrees of freedom to the field data must be preserved. I believe there is a risk of shifting towards more positivist underpinnings if considering theory in the initial stages of the research (point 1 above). The second position represents a peculiarity of the grounded theory approach. Regarding the third, there is a fundamental need to reflect on the term "mid-range theory" and its philosophical underpinnings. Eisenhardt (1989) sees results of the case study approach as something that can be tested formally through a positivist study. In relation to this research both the first and the third positions are considered partially appropriate. Therefore, some further considerations are needed.

In order to justify the need of theoretical building process, at the time of formulating the research question, it needs to be proven that existing theory either "does not address the research question at all, or does so in a way that is inadequate or likely to be untrue" (Eisenhardt and Graebner 2007, p.26). This aspect has been highlighted as part of the SLR study previously presented (e.g. (Zuiderwijk *et al.* 2015)). However, the same concept can be looked from another angle. In particular, two different natures of research questions are described as consistent in the inductive theory building process through multiple case study research (Eisenhardt and Graebner 2007) (Lee *et al.* 1999). These are:

1. Theory-driven research questions: these questions are believed to lead to outcomes that extend existing theory. In this case, a researcher "has to frame the research within the context of this theory and then show how inductive theory building is necessary. Typically, the research question is scoped within the context of an existing theory, and the justification rests heavily on the ability of qualitative data to offer insight into complex social processes that quantitative data cannot easily reveal" (Eisenhardt and Graebner 2007, p.26).

2. Phenomenon-driven research: the second type of research question focuses more on the importance of the phenomenon and emphasizes the lack of explanations given by existing theory. Here the research question is broadly scoped to give the researcher more flexibility. "The justification rests on the phenomenon's importance and the lack of viable theory and empirical evidence" (Eisenhard and Graebner, 2007, p.26).

Once again both of these aspects on how to justify the theory building process in multiple case study research are partially related to this study. The concepts of both the theory ("adoption") and the importance of the phenomenon of interest ("open data for commercial service innovation") are highlighted in the research question.

As a result of these reflections, theories concerning adoption of IS innovations were considered in two different stages of this research:

1. **Prior to the data collection and analysis**: within the open data services ecosystem, this project investigates the adoption process. As a consequence, the main high level concepts to bind the research within this constraint need to be considered. The term "high level concepts" is here written in order to emphasize the need of avoiding the risk of constraining too much the research within the existing theory. This potential pitfall was identified by Glaser and Strauss (1967) when they state that covering "all the literature before commencing research increases the probability of brutally destroying one's potentialities as a theorist" (p.253). This approach enabled me to preserve certain degrees of freedom when collecting and analysing data for capturing the factors that influence the adoption of open data. Similarly, Eisenhardt (1989) argues that in the context of theory-building research, studies should begin "as close as possible to the ideal of no theory under consideration and no hypothesis to test (...) because preordained theoretical perspectives or propositions may bias or limit the findings" (p.536).

2. **Within the interpretation of the data collected**: a more extensive consideration of the theories on adoption of IS innovations was undertaken once all the data have been collected and partially analysed. This process is found to be consistent with the inductive reasoning assumed in this research (see section 3.2). In particular, referring to (Huff, 2008), existing theory plays a key role at the fourth step of the general inductive process, that is: "expect theoretic insights to emerge as categories stabilize and their relationships become apparent". According with (Miles and Huberman, 1994) (Quaddus and Xu, 2005) (Berg, 2001), once the adoption factors emerged from the case studies, these factors were compared with those from the literature. Through this process, theoretical contributions to the adoption of IS academic debate emerged (see sub-section 9.1.2).

3.6.2 Reflections on current adoption theories and positioning within this study

Prior to data collection and analysis

IS literature on adoption of innovations is plentiful (Nakicenovic and Grubler, 2013) and several IS studies have focused on identifying factors that influence the adoption of particular innovations. Most of these studies used the models proposed by Ajzen and Fishbein (1980), Davis (1989), and Venkatesh *et al.* (2003). According to these models, a user first has to make a decision on whether to adopt an innovation. To this purpose, information is collected about the innovation and it leads to the formation of perceptions about the innovation. In line with these perceptions, a decision to adopt or reject the innovation is made (Rogers, 2010). Furthermore, adoption of an IS innovation sees the need for the user to infuse, routinize and implement the innovation itself (Roger, 2010) (Saga and Zmund, 1993) (Agarwal and Prasad, 1997). In general terms, these authors suggest that some external factors affect

the perceptions about an innovation, which in turn impact the decision of adopting (or rejecting) the innovation. This model is generic in nature and is likely to be applicable in most IS innovation adoption processes (Quaddus and Xu, 2005). As a consequence, the term adoption in the definition of the focus for this study is structured as follows:

"External Factors" →"Perceptions" → "Adoption".

This simple scheme enabled me to bind the data collection and analysis within the adoption process without limiting both the flexibility of the study and the quality of the findings. It is noted that, according to the previous literature (Quaddus and Xu, 2005), the definition of the term "external factors" was deliberately kept general, i.e. those factors that are not (but influence) perceptions, i.e. the cognitive phenomena influencing an adoption process (Venkatesh *et al.* 2003). Both the terms external factors and perceptions were rigorously defined using theoretical language at the time of positioning the findings of this research within the existing IS adoption literature (see Chapter 8), i.e. "within the interpretation of the data collected" (see point 2 above).

In addition, further reflections were conducted on the actual adoption factor. Sustained adoption and the related concept of diffusion of innovations is described in the literature as a multifaceted phenomenon that takes place in a variety of ways over time (Wolcott *et al.* 2001). Rogers (2010) suggests a five-stage innovation diffusion model: (1) agenda setting; (2) matching; (3) redefining / restructuring; (4) clarifying; and (5) routinizing. Others proposed different stages for adoption and implementation. An example is Kwon and Zmud's study (1987) which sees the sustained adoption process as a number of steps including: initiation, adoption, adaptation, acceptance, use, incorporation. Xu and Quaddus (2005) see it instead as: initiation, adoption, pilot implementation, organic growth, organizational implementation, and diffusion (sustained use).

Considering these reflections, the research question originally formulated for this study was consistently enriched with a number of sub-questions:

RQ: **What are the factors that influence Adoption of Open Government Data for commercial Service Innovation in a city context?**

a. **What are the external factors and perceptions that influence adoption of open government data for commercial service innovation in a city context?**
b. **What are the links between these factors?**
c. **What are the adoption stages from the initial generation of the idea until achieving sustained adoption of open data for commercial service innovation?**

It is noted that addressing RQa and RQb led to the emergence of a model describing adoption of open data across the cases studied, whereas from RQc the definition of a number of stages across cases was achieved.

Within the interpretation and analysis phase of the data collected

The objective at this stage of the research was to position the findings from this multiple case study (i.e. factors, variables and links influencing adoption of open data for commercial service innovation in cities) within those constructs that were found to be relevant for studying adoption of IS across several theories. As a result of this process, contributions to this literature are outlined. It is also noted that both the terms external factors and perceptions assume a precise meaning and are defined with theoretical language through this exercise. These concepts, according to the inductive nature of this study, were deliberately kept general in order to enable me to preserve certain degrees of freedom when collecting and analysing data (see previous paragraph).

Research in the area of adoption, diffusion and acceptance of technologies is argued to be a mature field in IS literature (Hu *et al.* 1999). Several theoretical frameworks have been formulated across disciplines to study these phenomena. In their MISQ paper, Venkatesh *et al.* (2003) propose a "Unified Theory of Acceptance and Use of Technology" (UTAUT) integrating eight widely used theories: (1) Theory of Reasoned Action (Fishbein and Ajzen, 1975), (2) Technology Acceptance Model (TAM) (Davis, 1989) and its extension (Venkatesh and Davis, 2000), (3) Motivational Model (Davis *et al.* 1992), (4) Theory of Planned Behaviour (TPB) (Ajzen, 1991), (5) Combined TAM and TPB model (Taylor and Todd, 1995), (6) Model of PC Utilization (Thompson *et al.* 1991), (7) Diffusion of Innovations Theory (Rogers, 2010), and (8) Social Cognitive Theory (Compeau and Higgins, 1995). This theory is developed addressing a fundamental limitation of these models identified as being focused on "individual-oriented information technologies as opposed to more complex and sophisticated" (Venkatesh *et al.* 2003, p.427) ones. Clearly, open data falls consistently within the scope of the UTAUT. Also, this research focuses on voluntary contexts of adoption, as opposed to mandatory settings, and both contexts are included in the UTAUT model. Furthermore, the UTAUT theory has been already leveraged for deductively investigating open government data adoption (Zuiderwijk *et al.* 2015) (although it is noted that the focus of this study was to deductively investigate open data adoption for research purposes, as opposed to this research that focuses on commercial service innovation). Notwithstanding this, Zuiderwijk *et al.* (2015) conclude their study stating: "a large part of the variance in the use of open data is not yet explained. Adoption theories for open data specifically are needed. There is a need for open data specific theories and methodologies that address the idiosyncratic nature of open data, including aspects such as data quality, institutional complexity, legal and

economic aspect. We recommend that adoption theories specifically for open data are developed" (p.9). Another reason why this theory was chosen among others is because of the call for future research stated within its seminal article (Venkatesh *et al.* 2003); in this way, the authors underline how UTAUT should be integrated with "research that has identified causal antecedents of the constructs used within the model in order to provide a greater understanding of how the cognitive phenomena that were the focus of this research are formed" (Venkatesh *et al.* 2003, p.470). This aspect was also anticipated in relation to previous adoption theories, argued to retain a parsimonious structure (Venkatesh, 2000) (Venkatesh and Davis, 2000).

The first segment of this chapter presented four key research decisions made for this study. Firstly, reflections about different philosophical underpinnings led to the choice of adopting an interpretivist position. The resulting paradigm assumed is social constructionism. Secondly, an inductive approach to theory has been chosen from reflecting on both existing studies and the exploratory nature of the research question. Thirdly, qualitative data collection and analysis have been selected for enabling the generation of rich data that enable thick descriptions of factors influencing adoption of open data. Fourthly, multiple case study has been selected as the suitable methodology for addressing the research question. Table 3.1 provides a summary of these choices and the main reasoning associated.

Key research decision's levels	Choices	Main reasoning
1. Philosophy	Interpretivist / Social Constructionist	Open data adoption is a highly complex and contextual process. Need of capturing the interplay of open data adopters and their contexts, i.e. the overall ecosystem.
2. Reasoning	Inductive	Exploratory nature of the research question. Emerging field of research. Lack of viable theory. Need of generating theory.
3. Approach	Qualitative	Need of generating rich data, i.e. enabling thick descriptions.
4. Methodology	Multiple case study	Ensures richness and depth in order to understand the phenomenon of interest. Enables exploration of multiple contexts. Suitable for studying IS development, implementation and use.

Table 3.1: Summary of key research decisions

It is noted that, in addition to the reasoning provided in Table 3.1, each decision was taken considering its suitability with those previously made. Indeed, multiple case study research is suitable for qualitative approaches, which in turn are suitable for inductively generating theory.

Ultimately, at this stage, the (sometimes) ambiguous role played by existing theory in interpretive inductive multiple case study has been clarified in relation to this research. The next section focuses on the actual methodology designed and implemented for carrying out these case studies.

3.7 Multiple Case Study Research Protocol

In every proposed case study approach the first step is always about the detailed definition of the phenomenon of interest (Eisenhardt, 1989) (Eisenhardt and Graebner 2007) (Stake, 2006) (Yin, 2013). Stake (2006) coined the term *quintain* to generally describe the target to be studied. This study involves the inductive investigation of the factors influencing adoption of open government data for commercial service innovation in city contexts.

The second step involves the selection of the cases which are found to be suitable for studying the *quintain* (Stake, 2006). Subsequently, Stake (2006) outlines the need of structuring the staff of people that is working at the research. In this study there was no need to undertake this activity as I played the role of the director, data gatherer, and analyst (limitations deriving from the presence of a single researcher are acknowledged in the last chapter of this dissertation).

Once the cases were properly selected, their investigations were carried out separately and independently. Stake (2006) stresses the importance of concentrating "on each single case almost as if it is the only one" and that multiple case study research should involve the investigation of "one case at a time" (p.1). At this point, I entered the field with multiple data collection techniques previously designed. The first objective was to understand the case. Reflecting on what each case study is and what it is not, is important for epistemological reasons in relation to qualitative studies. There is an inside and an outside of the case (Stake, 2006). In other words, the individual cases were studied to learn about their self-centring, complexity, and situational uniqueness. Indeed, a multiple case study of a phenomenon "is not so much a study of the *quintain* as it is a study of cases of what they tell us about the *quintain*" (Stake, 2006). In this way, two different aspects of each single case were considered, that are the contextual data and those related to the case's operations. Contextual information refers to the "situationality of the *quintain*" (Stake, 2006) and so to all the peripheral characteristics of the case and its activities. In this research it includes, for example, details about the company, number and skills of people within it, open dataset used, city authority that released it, type of services offered and so on. On the other hand, data about the case's operations are meant as the

ensemble of information in relation to activities that are involved in the adoption process of open data. Once all the relevant data had been gathered and systematically stored, the first round of qualitative analysis and interpretation with respect to the specific single case was conducted. From this step, a number of tentative assertions about adoption factors, related variables, and links specifically related to the individual case were outlined. With these preliminary findings, Stake (2006) suggests conducting a validation effort. This activity involves a step back to the case in order to assure that the right information and interpretations have been obtained. The final step (if no needs for further data collection initiatives arise from the validation stage) for the single case study was the writing of the case report.

Within each case study report, a number of factors, variables, and relationships between factors were available. At this stage, according to Stake (2006), the final report of each case study is often presented as intact (see Chapters 4, 5, and 6), accompanying a cross-case analysis with some more emphasis on the binding concepts or ideas (see Chapter 7).

Before reaching the conclusions, however, another stage was planned. In particular, Stake (2006) suggests the opportunity for the case study researcher to gather further data outside the cases themselves. In this way, I chose to integrate the findings from the multiple case study with additional interviews to domain expert and participants in the open data services ecosystem. To this final stage the name “validation across cases” is assigned in this study. A further exercise was finally undertaken through enfolding the literature. As part of this step, the inductive findings across cases were positioned within both the open data services literature and those theories developed for IS adoption. By comparing the findings from this research with these two bodies of literature, contributions to these academic conversations emerged.

The process followed for this research is mainly designed upon the work presented in (Stake, 2006). However, other multiple case study methodologists’ approaches have been reviewed and taken into account as well. Particularly relevant for the scope of this project is the similar approach proposed by Eisenhardt (1989). The author proposes an eight-step process for inductively build theory from multiple case study. These steps are: (1) getting started; (2) selecting cases; (3) crafting instruments and protocols (i.e. data collections techniques); (4) entering the field; (5) analyse the data: the data collected need to be analysed from both the single and the multiple case study perspectives; (6) shaping hypotheses; (7) enfolding literature; and (8) reaching closure. A summary of the multiple case study protocol including the main steps followed is presented in Figure 3.2.

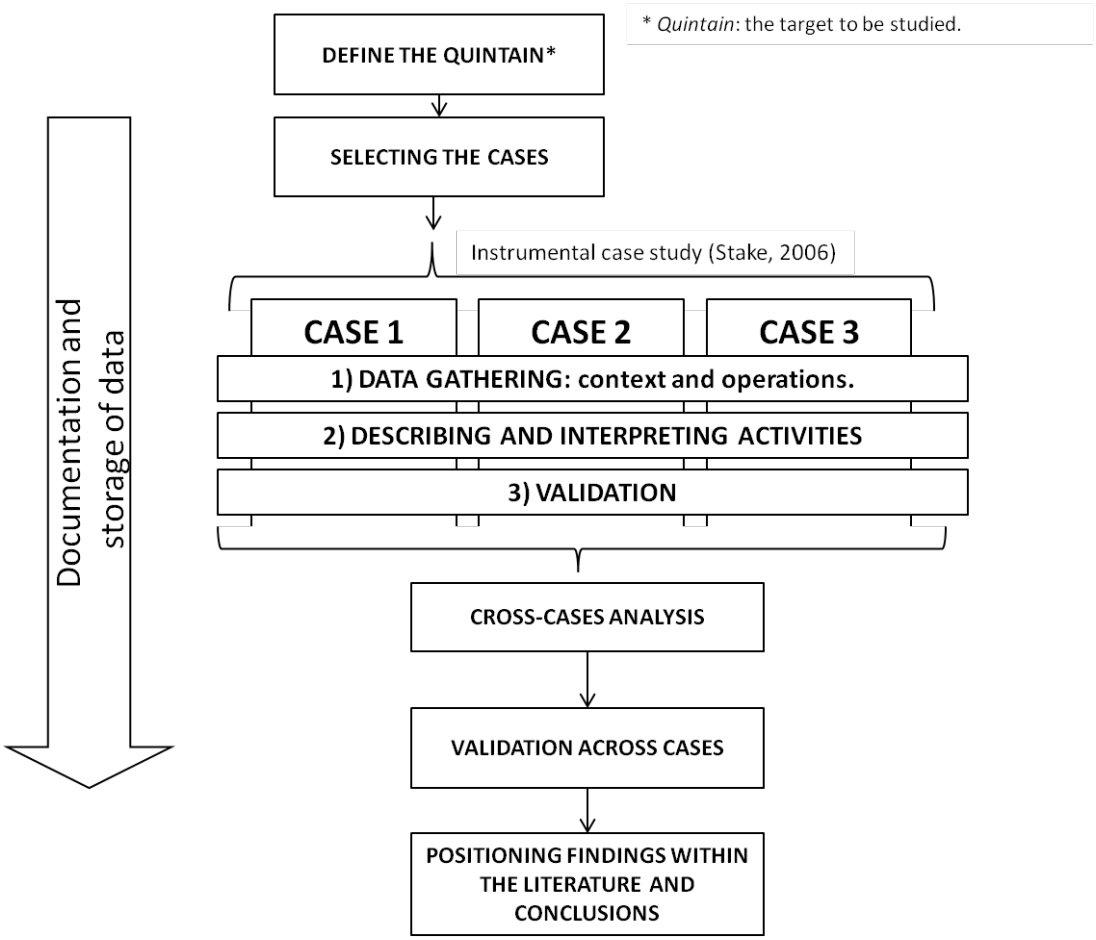

Figure 3.2: Multiple case study protocol

Note: Figure 3.2 anticipates that for this research three cases were selected. A dedicated section about the selection of cases is available below.

3.8 Case Study Selection

As a result of reviewing different approaches proposed in the literature, three cases that are sustainably adopting open data for the delivery of commercial services were selected. This section is provided to describe this process.

The proper selection of cases is considered crucial by all the case study methodologists (Stake, 2006) (Eisenhardt, 1989) (Yin, 2013) (Walsham, 1995). Eisenhardt (1989) argues that "while the cases may be chosen randomly, random selection is neither necessary nor even preferable" (p.537). As a consequence, criteria need to be clearly established to select the most suitable cases.

The first basic requirement for a case to be targeted was its involvement in the design and implementation of commercial open data services on top of one or more open datasets released by one or more city authorities. Eisenhardt and Graebner (2007) argue that cases can be historical accounts of the phenomenon of interest, but are more likely to be contemporary descriptions of "recent events" (p.25). Pettigrew (1988) suggested, since the usual number of cases that can be

studied is limited, to take into account extreme situations in which the phenomenon of interest is transparently observable. Therefore, the potential sample of cases was initially reduced to organizations where services are developed from open government data and related revenue models are in place, and where open data is a main requirement for the nature of these services. In other words, those organizations in which open data plays a marginal role within the actual services delivered were excluded.

Another fundamental aspect is stressed to be the diversity that cases need to have across contexts (Miles and Huberman, 1994) (Stake, 2006). Placing each case within their contexts and studying their relationships and influence to the *quintain* is critical (Stake, 2006). Encompassing all these features, Stake (2006) proposes three questions to be asked in relation to the suitability of the cases. These are: (1) is the case relevant to the *quintain*? (2) Do the cases provide diversity across contexts? (3) Do the cases provide good opportunities to learn about complexities and contexts?

In relation to the first question, the *quintain* "needs to be prominent in these cases" (Stake, 2006). Information-rich cases are recommended for achieving in-depth understanding of the central phenomenon (Patton 1990). Differently, the second question was addressed by selecting cases that are re-using open datasets from different domains. Finally, cases were selected in relation to how much time I could spend with them, and how much accessible they were (Stake, 2006). Hence, a final criterion used to select the cases was related to their accessibility. By this, I mean the availability of the subjects within the targeted organization to: reveal relevant information about the company, its activities, its history, and its future strategies; dedicate an appropriate amount of time for the data collection process; ensure the availability of subjects from all levels of the company to engage with me as the researcher; and show interest in the research and see a potential value for them by contributing to it.

After defining the criteria that have to be employed to select the cases, another decision to be taken is about the number of cases to be studied. Stake (2006) proposes to consider between four and ten cases. However, from a social constructionist perspective, one of the risks related to multiple case study research is to reduce complex cases to a few comparable variables, resulting in the loss of the idiosyncrasies of individual cases (Lauckner *et al.* 2012). In order to mitigate this risk, Creswell (2007) suggests that no more than four cases should be examined to allow individual cases to be adequately explored. According to this perspective, this study examined three cases in which open data is extracted, elaborated, and translated into a valuable experience for end users. A summary of the demographics of each case is provided in Table 3.2.

	Case A	Case B	Case C
Context	Urban planning	Culture and entertainment	Parking
Year of Foundation	2010	2011	2012
Size	6 staff	1 staff, 2 volunteers	3 staff, 2 “interns”
Location	Ireland and US	Ireland and Germany	Ireland , Spain, (Australia)
Open Data Source	Local City Councils (CC)	Local CC & Intermediary Org.	Local CC & Intermediary Org.
Revenue Models	Freemium and premium	Freemium and premium	Freemium, premium, subscription

Table 3.2: Cases demographics

As shown in Table 3.2, all the cases satisfied the selection criteria. In particular, all the cases represent a situation in which sustained adoption of open data for the delivery of one or more commercial services is in place. Furthermore, open data represent a key requirement for the delivery of these services. Also, given the different domains and locations in which services are delivered across cases, diversity across contexts is ensured. Finally, all the cases showed interest in the study and agreed to share information and documents, and to allow an appropriate amount of time for this research.

3.9 Single Case Study Protocol

This section aims at showing the general stages undertaken for each of the cases studied. It is noted that each step outlined in this section is explained in detail later in this chapter.

Clearly, each case study started with a **first contact (stage 1)** with one of its representatives. This contact happened through emails for all of the cases. The first step was about a brief conversation with the subject in order to introduce myself and the purposes of my research. After receiving a positive feedback, a first face-to-face meeting was proposed and scheduled. As argued by Stake (2006), “researchers should anticipate what perspectives on meaning and activities perceived by different people might be”. To do this, the first stage was enriched by a careful **review and analysis of the documents publicly available (stage 2)** regarding the case. This included the company’s website, and other material available on the internet. After collecting this range of general information about the company, I was able to conduct a **first reflection about the suitability of the case (stage 3)**. In particular, by looking at this data, I was able to verify if the company is actually delivering commercial open data service(s), the detailed descriptions about the services offered, and other relevant information such as the dataset(s) used, the city (or cities) the company is taking the data from etc. When the meeting of the requirements for being selected as a case for this study was ensured, a **first face-to-face meeting (stage 4)** was scheduled and conducted. In all the cases, the CEO of the company was the main interlocutor until this stage. At the first encounter the main focus was about collecting

internal documents and a further range of contextual data and general information about the company and more detailed characteristics of the open data service(s) offered. Getting information about the context was critical in order to make clear the "situationality of the *quintain*" (Stake, 2006). After these four stages, the characteristics of each case became evident as well as the activities in which the specific case is involved. Moreover, the main planned data collection methods were described. Technical aspects about the interviews were agreed (i.e. time, recording, type of interview, instances of questions that were asked). Other ethical aspects were mentioned in accordance with the University's regulations. In relation to this aspect, in order to minimize the probability of lack of trust in my regard, it was mentioned that the data collected is for research purposes only, and the interviewee and the case appear anonymously in this academic dissertation and potential related academic publications. This encounter was leveraged in the next stages in terms of specific effort that were made in order to minimize social dissonances with the subject(s). Based on the output of the first meeting a **second reflection upon the suitability of the case (stage 5)** was undertaken. This involved the perceived (and actual) availability of the subjects to reveal documents and relevant information, and the positive feedback received in relation to the detailed purposes of this research. At this stage I was able to **structure the general contextual information (stage 6)** about the company.

Once I was familiar with the case and its activities, I was able to answer in detail the fundamental question about **"what can be known"** from the case (Stake, 2006) **(stage 7)**. With this information I did "**enter the field**" (Eisenhardt, 1989) **(stage 8)**. In this stage I spent an appropriate amount of time in order to place each case's activities in its own situation (Stake, 2006). In multiple case study research, one of the most important aspects is to show how the phenomenon appears in different contexts. Thus, I was involved with the people within each case to study in-depth what the case does in terms of its internal activities and its functioning. This understanding was enriched from observing people conducting their activities and by having informal conversations with them. Field notes were taken to store the relevant information collected. After entering the field of the case, I conducted the **semi-structured interviews (stage 9)**. The original protocol (see sub-section 3.10.4) was modified in relation to the data collected and structured so far in the process. In this way, from the "what can be known" stage, changes to the original protocol were made in terms of both content and language. Once the first interview cycle was conducted, the need for deepening on some aspects emerged for all of the cases. Therefore, based on the analysis of the first interviews, a new protocol was designed specifically to investigate aspects that remained unclear or under-explored. The data collected constituted the input for the **qualitative analysis of the data (stage 10)**. At this stage, tentative assertions about each case were formulated. In other words, a tentative model describing the adoption of open data for each case emerged. These preliminary findings were the main object in

question during the final **validation stage (stage 11)**. This is meant as a set of activities within the case for ensuring that correct interpretation of the data has been obtained (Stake, 2006). After reviewing, (sometimes) correcting, and enriching the conclusions about each case, the case study reports were written and the adoption factors model for the specific case was outlined. The stages described above are summarized in Figure 3.3. In practice, each case study lasted between three (Case B) and four months (Case C) from the first contact until the last adoption model for the company had been established.

Ultimately, I reported to the CEOs of the companies about my involvement with the local council and the smart city officer within it, who showed interest in seeing the final findings of this research. I underlined several times that there is not any contract or agreement for which a follow up with the council will actually happen. However, I specified how the city council officer reported to be interested in engaging with the companies as a step to push the local open data ecosystem forward. This aspect was well received by the CEOs of the cases, and, in my opinion, stimulated them to seriously and proactively conducting this collaboration.

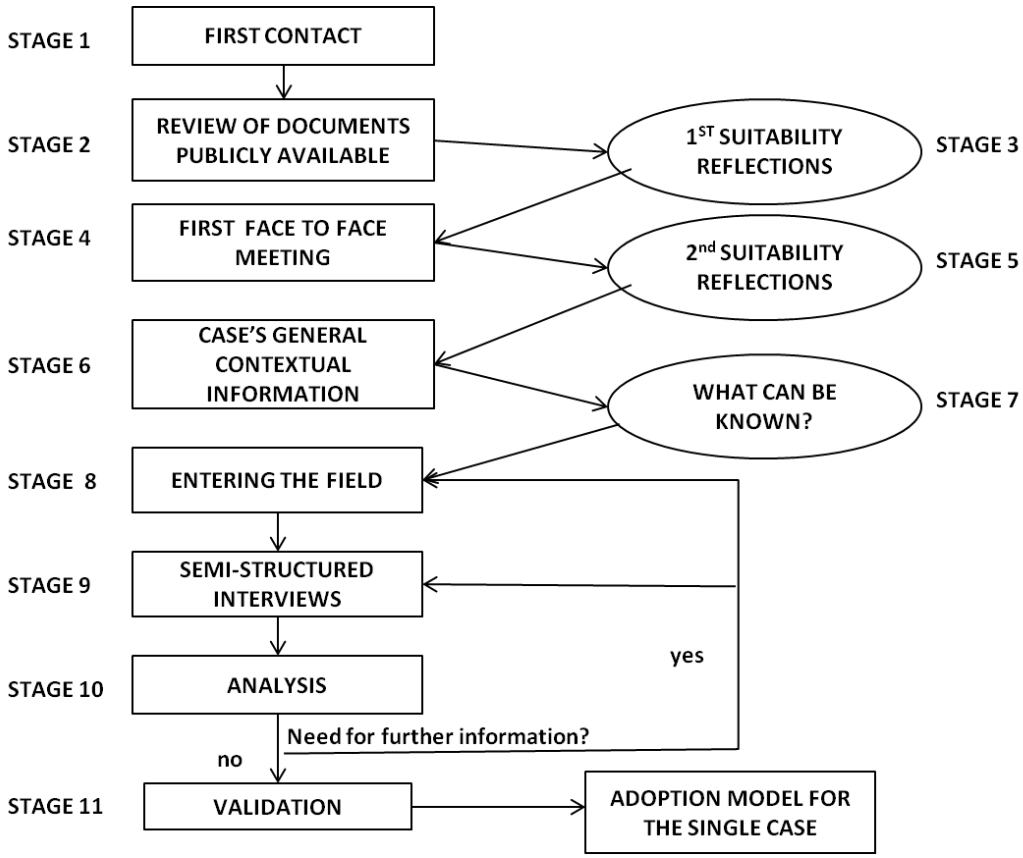

Figure 3.3: Single case study protocol

3.10 Data Collection Methods

In the context of a case study, Yin (2013) and Stake (2006) emphasize the importance of establishing a specific framework to structure data collection. Starting without a plan is acknowledged to be a road to failure. However, it is noted the having a plan for collecting the data is different from having preconceived ideas. While the more the study is quantitative the more similar should be the methods across cases, inductive qualitative case study researchers usually combine multiple data collection methods (Eisenhardt, 1989) and keep the design of the process flexible. Yin (2013) argues that evidence for case studies may come from six different sources: documents, archival records, direct observation, participant observation, interviews and physical artefacts.

While observation was a natural source for collecting data during the time I personally spent within the cases, a few more reflections are needed regarding the other sources of evidence proposed. In particular, in relation to the nature of the cases to be studied, archival records and physical artefacts were not found to be a valuable source for gathering qualitative significant data. In relation to the former, given the early stage of the open data service phenomenon (the first open government dataset was released in US in 2009), the companies that were studied were start-ups (Case B) or recently founded ones (Case A and Case C). As a consequence, archival records were not even available. Furthermore, the journey from taking open data from a raw dataset to the design of one or more valuable services for end-users was not found to involve specific physical artefacts.

Hence, informal meetings (i.e. observation), document analysis and interviews were the main sources for the data collection process across the three cases. Also, it is noted that the final validation stage was also part of the data collection process.

3.10.1 Observation – Informal Meetings

Part of the data collected across the cases was in the form of field notes taken during the time I personally spent on site. This source of data played a key role especially in the first stages. When collecting general information about the organizations, their services, and the people within them, several questions were asked to the cases' people. Another important phase of these meetings took place as an actual observation step. In particular, cases' people agreed to show me (from their computers) technical features of the services including: the general platform and the coding behind it, the datasets and databases, as well as how general database' management activities are implemented.

3.10.2 Documents Analysis

Yin (2013) argues that useful insights can be gained through documents analysis. The author highlights five main documents that are often used in case study research: (1) organizational website; (2) newspaper articles; (3) company reports and regulations; (4) policies; and (5) e-mail correspondences. In this research, I aimed at initially collecting all the documents available for each case. Firstly, contextual information about each case was collected and structured from the company's website. Secondly, I found on the internet a plethora of articles/interviews related to the cases. These were searched, ordered, and those found to be the most relevant, reliable, and complete were selected. A positive feedback on this selection from the cases' CEOs was also achieved. Thirdly, documents such as executive summaries, financial reports, regulation and statute of the companies were believed to be crucial in understanding how the organizations function, operate, and relate themselves with the overall ecosystem. The websites and the related open data portals of those city councils from which the data is re-used were also taken into account. Furthermore, the written communications between me and the people within the companies (relevant for Case B only) constituted another source of qualitative data. Finally, given the availability of all services as web applications, I personally experienced them to get further familiarity with the cases' activities and outcomes.

All documents played a key role in achieving a general understanding of the company and its context and in shaping "what can be known" from the case. Subsequently, this information was leveraged for customizing the interview protocol, as well as for constituting evidence for the final findings.

3.10.3 Qualitative Interview

In interpretive IS case studies, Walsham (1995) argues that interviews are the primary data source, "since it is through this method that the researcher can best access the interpretations that participants have regarding the actions and events which have or are taking place, and the views and aspirations of themselves and other participants" (Walsham, 1995, p.78). In addition, "interviews are highly efficient ways to gather rich, empirical data" (Eisenhardt and Graebner, 2007, p.28). A research interview is an exchange of views between two people – one of whom is in the role of researcher – who are talking about a topic of common interest (Kvale, 2008). The topic of interest is typically the interviewee's experience of his or her life world, as it is lived and constituted in awareness (Polkinghorne, 2005). The qualitative interview method is used in qualitative research of all kinds, whether positivist, interpretive, or critical (Myers and Newman, 2007) (Hesse-Biber and Levy, 2006) (Klein and Myers, 1999) (Myers, 1997). This method was chosen because it distinguishes itself from other research approaches by engaging participants directly in a conversation with the researcher in order to generate deeply contextual, nuanced and authentic accounts of participants' outer and inner

worlds, that is, their experiences and how they interpret them (Schultze and Avital, 2011). Here reality is not discovered but interpreted and constructed. There are various types of interviews. Some of these (the most used) are (Fontana and Frey, 2000):

- Structured interviews: within this type there is a complete script that is prepared and structured before the interview takes place. There is no space for improvisation.
- Semi – Structured Interviews (SSI): in this type the script is incomplete. Some questions might have been prepared before the actual interview, but there is room and need for improvisation.
- Group interviews: in this type two or more people are interviewed at once. These can be structured, semi-structured, or unstructured.

For this research project I adopted the SSI type. The purpose of qualitative interview-based research is to describe and clarify people's experiential life "as it is lived, felt, undergone, made sense of, and accomplished by human beings" (Schwandt, 2001, p.84). As experiences are usually not observable, the interview needs to help research participants reach beyond the superficial layers of their experience in order to generate informative and novel accounts of the phenomena of interest. As a result, this enabled me to be actively engaged in the production of data to uncover people's lived world and experiences prior to scientific knowledge (Kvale, 2008). Indeed, one of the researcher's objectives is to activate and stimulate the interviewee's interpretive capabilities (Holstein and Gubrium, 1995). In other words, SSIs give room to improvisation, which in turn stimulates the interpretive aspect of the conversation. Thus, in the SSI process everyone is an interpreter and "in interpretive research the aim is to understand the complexity of the human sense making process, and the processes by which inter-subjectivity is obtained as the situation is constantly changing" (Nicholson and Sahay, 2004).

In summary, as outlined in (Schultze and Avital, 2011), three main characteristics are associated with the SSI method. These are as follows:

1. Grounding the interview in the participants' own experiences: the method relies on the participants' descriptions of their lived experience. This helps in keeping the interview grounded in actual events and settings. The recounting of lived events allows the researcher to visualize the participants' social world and its meaning (Alvesson, 2003).
2. Acknowledging and valuing participants' narrative (re)construction of their experience: this implies that the data that is generated in an interview needs to be seen as narratives that are produced in the moment rather than facts or established stable meanings.

3. Providing an explicit framework for guiding the participants to articulate and interpret their experiences: a framework needs to be defined to structure the conversation in such a way that guides the interview through his or her introspective journey while honouring his or her freedom of thought and expression. Thus, a SSI protocol was developed prior to the beginning of the data collection.

3.10.4 The Semi-Structured Interview Protocol

As a first step in defining the SSI protocol, the main areas of information needed had to be outlined. The targets were conceptualized mainly from the objectives for this research. According to the role of adoption theories in this research (see section 3.6), theoretical language and concepts were excluded at this stage. Thus, the areas of information needed for this research project were the starting points from which the SSI protocol was defined. These areas are:

- Individual perception and understanding of open data.
- The adoption process of open data for commercial service innovation.
- The motivations for adopting open data.
- The factors that influence the sustainable establishment of the open data in the company's processes and the relationships between those factors.
- The barriers to put in place an open data service-oriented business.
- Factors that would encourage people to use open data for commercial service innovation.
- Factors that would prevent people from using open data for commercial service innovation.
- Required resources and facilitating factors for implementing open data services.

Overall, the SSI protocol designed was divided in six stages. As the first stage (Table 3.3), an estimated time of five minutes was allocated to get personal information about the interviewee in the so called "situating the actor" step (Myers and Newman, 2007).

STEP 1: SITUATING THE ACTOR	
#	Questions
1	*What is your name/age/gender/nationality?*
2	*What is your background?*
3	*What are your main experiences?*
4	*What role are you playing within the organization?*
Estimated time allocated: 5 minutes	

Table 3.3: SSI protocol step 1, situating the actor

The second stage (Table 3.4) was meant to explore the individual perceptions around the open data ecosystem. This step is believed to be fundamental in order to reflect on the language used by the interviewee and to subsequently mirroring it when formulating the next stages' questions.

Additionally, it was important to ensure that there was not divergence regarding the understating of the main concepts involved (i.e. open data, open data services).

STEP 2: INDIVIDUAL PERCEPTION ON OPEN DATA	
#	Questions
1	*How would you define open data?*
2	*Do you believe open data has a great potential for new services development in cities?*
3	*How did you get involved with the open data ecosystem?*
Estimated time allocated: 5 minutes	

Table 3.4: SSI protocol step 2, individual perceptions on open data

The third step (Table 3.5) of this process was about investigating the actual motivations for adopting open data. These also include the process through which awareness of open data was achieved, as well as the reasons why it was selected for implementing a service.

STEP 3: MOTIVATION OF ADOPTING OPEN DATA AND DATASET ATTRIBUTES	
#	Questions
1	*Why did you choose open data as the foundation of your business, instead of, for instance, looking for other sources of data?*
2	*In the choice of the specific dataset(s), which variables did you consider?*
4	*Did you talk and discussed with other people/organizations in this field before initiating your business? If yes, how and to which extent did they influence your choice?*
Estimated time allocated: 15 – 20 minutes	

Table 3.5: SSI protocol step 3, motivations of adopting open data and dataset attributes

After motivations to adopt open data have been inquired, the next step of the SSI was initially designed for investigating the processes for designing the actual open data services and for achieving sustained adoption of this source. These types of information were constructed using as a starting point data about the lived experience of the interviewee (step 4), as a representative of an organization that adopted open data for commercial service innovation.

STEP 4: LIVED EXPERIENCE INFLUENTIAL FACTORS	
#	Questions
1	*Which were the skills that were needed to transform the dataset into an actual service?*
2	*Which were the resources needed to design and deliver the open data service?*
3	*Which were the main factors that influenced the establishment of open data in your organization's processes and practices?*
4	*Which were the obstacles that you have encountered to put in place an open data service-oriented business?*
5	*Was any external action critical in the design and delivery of your service?*
Estimated time allocated: 20 minutes	

Table 3.6: SSI protocol step 4, lived experience influential factors

The following step (Table 3.7) was defined to leverage the opportunity of engaging with individuals with great expertise in the field under analysis. In other words, in this stage I aimed at building on the

subject's generative capacity in an "appreciative" form of inquiry (Cooperrider and Srivastva, 2005). As stated in (Schultze and Avital, 2011), "the appreciative interviewing process is designed as a retrospective inquiry that catalyses a prospective act". In this way the interviewee is taken through a journey in which he or she has the opportunity to relate their most outstanding personal experiences (step 4) to generate hopeful aspirations and desired futures about the diffusion (or not) of open data for commercial service development in the broader ecosystem.

Finally, the last stage of the SSI is the exit step. Here, also potential further interactions were briefly discussed (a question in the form of "would you recommend another person, within or outside your organization, to be interviewed?" was asked).

STEP 5: OPINION FACTORS AS DOMAIN EXPERT AND PRACTITIONER	
#	Questions
1	*In your opinion, which are the factors that would encourage people to use open data for commercial service development?*
2	*In your opinion, which are the factors that would prevent people to use open data for commercial service development?*
3	*Which other external factors, in your opinion, would increase the diffusion of open data for commercial service development?*
Estimated time allocated: 20 minutes	

Table 3.7: SSI protocol step 5, opinion factors as domain expert

From a review of the IS literature on SSI, several academic recommendations were collected and taken into account when designing and implementing such processes. These are: each interview should last for about 1 hour (Quaddus and Xu, 2005); the interview needs to be transcribed the same or the following day in order to be able to reflect on the body language and other cues fresh from memory (Quaddus and Xu, 2005); additional aspects include dress appropriately, show empathy, being relaxed when listening, and respond appropriately to answers (by nodding, smiling etc.) (Myers and Newman, 2007); ensure the absence of pauses during the performance (Myers and Newman, 2007); speaking and dressing likewise the interviewee is recommended; and use mirroring questions and answers: mirroring means taking the words and phrases that the interviewee uses in constructing a subsequent question or comment (Shim *et al.* 2002).

At the stage of starting the SSI processes what has been defined in (Stake, 2006) as the "situationality of the *quintain*" (i.e. the specific case-related context around the phenomenon to be studied), was defined for all the cases. From the information gathered through this process, additional elements to be considered and to be tackled more in depth through the SSI process emerged. In other words, the original SSI protocol was enriched with new aspects. These were not considered at the time of structuring the protocol and were believed to upgrade the quality and completeness of the data

gathering process. For example, one aspect that was found to be common across cases was their international experience. Case A delivers its services in Ireland and US, Case B in Ireland and Germany, and Case C operates in Ireland and Spain (see Table 3.2). Hence, it became relevant to explore the differences that are in place in the process of re-using open data across different countries. Also, further questions were added to step 4 of the protocol to investigate the presence of variables that are considered when targeting new cities. An additional question was added in relation to which city is considered the best in their experiences and which one the worst and the reasons why the interviewee says so. Ultimately, additional questions were added to the protocol in relation to individual cases based on the specific situation encountered. These are detailed in the specific case-related chapters (Chapters 4, 5, and 6, for cases A, B, and C respectively).

It is noted that interviewees from cases A and C agreed to be audio recorded. Regarding Case B, the CEO explicitly requested not to be recorded.

3.10.5 Addressing Potential Pitfalls of the Method

In IS literature a range of eight potential pitfalls are acknowledged to be concerned with the SSI method (Myers and Newman, 2007) (Fontana and Frey, 2000) (Hermanns, 2004) (Miles and Huberman, 1994). These are: (1) artificiality of the interview, (2) lack of trust, (3) lack of time, (4) level of entry, (5) Hawthorne effects, (6) constructing knowledge, (7) ambiguity of language, (8) and (8) elite bias. These potential pitfalls were addressed in a similar way for all the cases. Table 3.8 provides details about each potential pitfall as well as actions I have personally undertaken to address them. Furthermore, the literature acknowledges that interviews can go wrong in terms of the possibility for the interviewer to offend or unintentionally insult the interviewee (Hermanns, 2004). In relation to this research, fortunately none of the interviewees felt as being insulted or offended.

In order to provide a description of the interviews across the cases, (Myers and Newman, 2007)'s framework is adopted. This was designed to describe SSIs as a whole. It was derived from (Goffmann, 1959)'s theory on face to face interactions which was drawn from the dramaturgical model. Qualitative SSIs are seen as a drama as they have: (1) *the stage*: this variable includes the location where the interview takes place but also dresses, ranks, sex and age of the people involved (i.e. actors) and other things that might be used such as pens, notes or recording equipment. (2) *The actors*: these are typically the interviewer and interviewee. (3) *The audience*: both of the actors can be seen as audience. More broadly, also the academic community can be seen as an audience. (4) *The script*: the interviewer has a framework of questions (SSI protocol). (5) *The entry*: impression management is very important; the goal is to minimize social dissonances and to make the interviewee feel comfortable. (6) *The exit*: exiting the stage might involve preparing for next performance. All the interviews with all

the cases are described following this model. Tables for cases A, B, and C are presented in Appendices 2.1, 2.2, and 2.3 respectively.

	Potential pitfalls	Description	Actions undertaken for all the cases
1	Artificiality of the interview (Myers and Newman, 2007)	The interviewee is a stranger, and needs to create opinions under time pressure.	The purposes of the research had been previously detailed. I also asked to allocate two hours for each SSI in order to avoid time pressure concerns.
2	Lack of trust (Myers and Newman, 2007)	This potential pitfall might result in the interviewee's choice to not divulge sensitive information.	The clarification of ethical aspects of this research ensured the subjects that no speculative use (or that can negatively impact the case) of the data collected would be undertaken.
3	Lack of time (Myers and Newman, 2007)	It might result in incomplete data gathering.	No time constraints were fixed, and the time allocated was almost twice the time estimated.
4	Level of entry (Myers and Newman, 2007)	It might be difficult to reach the managerial level.	This risk did not exist as I approached the CEO and Founder of each case.
5	Hawthorne effects (Myers and Newman, 2007)	The researcher may intrude upon the social setting and potentially interfere with people's behaviours.	The investigation was focused on the interviewees' experiences and ideas. By not stating personal opinions, I ensured not to intrude upon the cases' social settings.
6	Constructing knowledge (Fontana and Frey, 2000)	After gathering the data, the researcher has to construct the knowledge and the interviewee might be in a situation in which he or she has to reflect on aspects that he or she had never considered so explicitly before.	A specific meeting was scheduled for re-visiting again interpretations from the data collected (validation meeting) and ensure that correct interpretation of the data has been achieved.
7	Ambiguity of language (Myers and Newman, 2007)	It is not always obvious that interviewees fully understand the questions.	No misunderstandings happened during the processes as language was kept simple and mirrored the companies and the subjects' own languages.
8	Elite bias (Miles and Huberman, 1994)	If only high status people are interviewed, a broader understanding of the phenomena might be missed.	Data coming from all levels of the organizations complemented the data collected from the interviews. In addition this data contributed in refining the interview protocol.

Table 3.8: Qualitative interviews' potential pitfalls

3.10.6 Single Cases – Validation

The last step for each case study was the validation stage. In interpretive case studies no two observers construct knowledge in the same way (Stake, 2006) and, as a consequence, complete confirmation is not possible. To reduce the likelihood of misinterpretation, two of the most common strategies are redundancy of data gathering and procedural challenges to explanations (Denzin, 1989). The former has been addressed within the data collection stages. As part of the latter, a validation effort was implemented. Stake (2006) emphasizes its importance in terms of a "process of gaining assurances" (p.33) by critically reviewing what is being said and interpreted. Thus, validation is meant here as efforts to assure that the right information and interpretations have been obtained. According to Stake (2006), two results are possible: (1) confirmation that the observation means what we think it means; and (2) ideas about the observation are interpreted differently from different people. To this, I perceive that a further possibility should be added. It refers to the potential identification of the need of gathering more data if one or more statements are found as incomplete and/or unclear.

According to the methodology designed for this study, this validation stage happened after all data gathered was qualitatively analysed and preliminary findings have been structured for each case. At this stage, a last meeting was scheduled at the cases' sites with the main goal of ensuring that the data collected were interpreted correctly. The purpose of these meetings was to review the overall analysis process, resulted for each case in the definition of a number of tentative assertions about factors, variables, and links between factors influencing adoption of open data. When validation about these tentative assertions was achieved, these were translated into final assertions about each case. On the other hand, in the case of tentative assertions found to be unclear, inconsistent, or incomplete, a decision on whether dropping these from the findings, or informing the need for more data collection was made in accordance with the participants from the cases. In addition, an extensive review was made across all concepts retained during the analysis as no sufficient evidence was found to motivate their inclusion in the preliminary findings. It is noted that this stage of the research aimed also at addressing potential limitations in relation to single researchers undertaking inductive multiple case study investigation (Stake, 2006).

This last meeting with each case ensured that correct interpretation of the data collected was obtained. A general positive feedback was received from the analysis of the data collected. However, a few changes were proposed by the cases' subjects and addressed subsequently in the final findings about each case. During these processes questions such as "would you like to add something in relation to this aspect?" were continuously asked to the subjects. From the generally negative

responses received it could be inferred that no more data collection was required and final findings about each case could be outlined.

In this section all the data collection techniques leveraged in this research were outlined. It is noted that as part of the cross-case analysis stage, further data collection has been undertaken during the validation across cases stage. This phase of this research is described in section 3.17. Table 3.9 provides a summary of the data collection across the three cases studied.

	Case A	**Case B**	**Case C**
People encountered	4	2	3
Positions	CEO, developer, lead developer, business developer.	CEO / Founder, volunteer	CEO, Product Manager, Employee
Number of meetings	7	5	6
Observation	8 hours	2 hours	3 hours
Number of SSIs	2	2	4
Interview participants' positions	CEO	CEO	CEO, Product Manager
Total interviews' duration (time recorded)	3,5 hours	4 hours*	5 hours
Number of internal documents	3	4	3
Number of external documents	7	7	8
Validation meeting	2 hours	2 hours	3 hours

Table 3.9: Single cases data collection summary

*it is noted that for Case B interviews were not recorded.

3.11 Trustworthiness of Multiple Case Study

As in all research, consideration must be given to construct validity, internal validity, external validity, and reliability (Yin, 2013) (Stake, 2006) (Eisenhardt and Graebner, 2007). All these aspects were achieved through a range of research activities/approaches that are described throughout this document and are briefly re-captured in this section.

Construct validity has traditionally been defined as the experimental demonstration that a test is measuring the construct it claims to be measuring (Brown, 2000). In this way, in accordance with (Yin, 2013) and (Stake, 2006), validation efforts both within cases and at a cross-case analysis level were undertaken. Furthermore, keeping records of the data collected and building chain of evidences were considered as cornerstones throughout the duration of this multiple case study project. Another final suggestion to improve construct validity is to use multiple sources to provide evidences (Eisenhardt, 1989) (Miles and Huberman, 1994). The range of data sources leveraged in this project is reflected in this principle.

Internal validity is concerned with the degree of the researcher inferences in the determination of causal relationships (Yin, 2013). Hence, the coding procedures employed to analyse the data were designed and undertaken for "creating and assigning categories, and exploring connections between them" (Dey, 1999, p.146). This specific approach is argued in IS research to assist "interpretive researchers in establishing the causal relationships that ultimately produce theory" (Andrade, 2009, p.49). In addition, I followed Eisenhardt's (1989) suggestion of "enfolding literature" to positioning the final inductive findings across cases. As explained within section 3.18, pattern matching logics were employed and subsequently strengthened the results (Eisenhardt, 1989) (Yin, 2013). Partially linked to this concept is external validity which represents the generalization of the outcomes from the multiple cases (Yin, 2013). The role of generalization in interpretive studies is controversial (Walsham, 1995). A particular reflection on these issues is provided in section 3.15. However, again by applying pattern matching logic to look at the interpretations achieved from the cases, external validity is expected to increase (Yin, 2013).

Ultimately, in relation to case study reliability, Yin (2013) and Miles and Huberman (1994) recommend to carry out two research activities: develop a case study protocol and a case study database. A protocol was developed in this project in relation to the overall multiple cases methodology, to the single case study, and to the main data collection methods employed. Furthermore, all original data and information about the cases were securely stored and continuously re-captured at the time of constructing the findings.

In this second segment of Chapter 3, the research methodology employed in this study has been described in detail with a specific focus on the single cases. Initially, the multiple case study research protocol has been introduced. The focus of the chapter shifted to the single cases (it is noted that sections 3.14-3.18 are entirely dedicated to the cross-case analysis method). After presenting the case selection criteria, the single case's protocol has been formulated, including the definition of the different stages involved in investigating each case. At a further level of granularity, the data collection techniques employed have been described. Details about the implementations of these techniques across the three cases are provided within the chapters dedicated to each case study. Ultimately, aspects about trustworthiness of this multiple case study research have been outlined and addressed. The next section focuses on the data analysis method and the respective processes conducted for each case.

3.12 Single Cases: Qualitative Data Analysis

With respect to multiple case studies it is suggested to carry out two stages of analysis (Stake, 2006) (Eisenhardt, 1989). The first deals with single case data, while stage two deals with the cross-case

perspective. This section focuses on the analysis conducted for the single cases. The methods and techniques employed in the cross-case analysis are presented sections 3.14-3.18.

Data analysis means "a search for patterns in data" (Neuman, 2005, p.426). Once a pattern is identified, the qualitative researcher moves from the description of a historical event or social setting to a more general interpretation of its meaning. This is consistent with the ultimate goal of inductive case studies, which is "to uncover patterns, determine meanings, construct conclusions and build theory" (Patton and Appelbaum, 2003, p.67). Analysing data is at the heart of inductive case studies, but is "both the most difficult and least codified part of the process" (Eisenhardt, 1989, p.539). The reason why it is found to be difficult is that it is not fundamentally a mechanical or technical exercise, but it is a dynamic, intuitive and creative process of inductive reasoning, thinking and theorizing (Basit, 2003). Miles and Huberman (1994) suggest to choose suitable techniques and to develop a subsequent analysis strategy prior to data collection. These tools must be selected based on the objective of the research. Patton (1990) argues that in analysing the data from interviews and documents (i.e. the main data sources for this study) the objective is on identifying, coding and categorize the patterns encapsulated in this data.

The common threat is that all qualitative modes of analysis are concerned primarily with textual analysis (Myers, 1997). These texts are central to the generation of insights (Pettigrew, 1988). In this research, interviews were transcribed, documents and the cases' websites were available in textual form, and data from observation, informal meetings, and discussions on site was collected and stored in field notes. Therefore, qualitative data emerging from the cases was textual, non-numerical, and unstructured.

This research aims at identifying factors and associated variables as well as the links between these factors that influence the adoption of open data for commercial service innovation. This implies that from a wide range of data collected from each case, the analysis process had the goal of encapsulating it within meaningful categories and links. In this way, the amount of data taken into account is expected to decrease as the analysis proceeds. This process does not imply that the body of data becomes smaller and manageable because there is less to deal with, rather it is "the result of interpretation and organization of this data" (Kaplan and Maxwell, 2005). To this process the term "coding" is assigned in the literature (also the terms "data condensation" and "data distillation" were used to describe this process (Tesch, 1990)). It involves subdividing the data as well as assigning categories (Dey, 1993). Darke *et al.* (1998) provide a useful definition for the word "coding" which particularly meets the objectives for this study; they define coding as "the assignment of themes and concepts to a selected unit such as a sentence taken from the interview transcripts [as well as

documents and field notes]. The concepts are combined into related categories, links between categories are identified and verified against the data, and selective coding attempts to integrate the categories into a theory which accounts for the phenomenon being investigated" (Darke *et al.* 1998, p.284). It is noted that coding and analysis are not synonymous, though coding is a crucial aspect of analysis (Kaplan and Maxwell, 2005). Coding is one of the significant steps taken during analysis to organize and make sense of textual data (Basit, 2003). Codes are argued to be links between locations in the data and sets of concepts or ideas, and they are in that sense heuristic devices, which enable the researcher to go beyond the data (Coffey and Atkinson, 1996). All in all, coding is found suitable in relation to the interpretive nature of this study. In fact, "coding facilitates the organization, retrieval, and interpretation of data and leads to conclusions on the basis of that interpretation" (Lockyer, 2004).

One further decision to take is whether manual or electronic (i.e. facilitated by the use of specific software) coding is appropriate. I see manual coding as the most appropriate way of analysis for this study. This choice was informed by looking at some examples of application of both ways, and by personally trying one of the software available (NVivo). While in quantitative researches the use of software is highly recommended due to the usual huge amount of data to be analysed, for qualitative research an optimal decision does not exist. Some variables that might affect this choice are acknowledged to be time and funds availability, and size of the project (Basit, 2003). I chose manual open coding primarily because my interpretation and control of the outcomes are, in my opinion, enhanced as opposed to a situation in which software is employed.

In general, six main steps are involved in this process (Stake, 2006) (Miles and Huberman, 1994) (Walsham, 1995) (Darke *et al.* 1998) (Berg, 2001). These are:

1. Manually review the transcripts and other documents, line-by-line and sentence-by-sentence, to uncover key patterns/themes and produce key words/phrases (inductive process).
2. Produce labels/categories of these key words/phrases. Identify high level factors and corresponding variables.
3. Look for relationships among the factors from each source of data.
4. Develop raw tables of factors variables and their links for each case.
5. Validation phase and discussion of retained concepts.
6. Develop the final case-related adoption model.

These steps have been adapted to this study and are shown in Figure 3.4. This process should be inserted in the single case study protocol design in the box "analysis" (see Figure 3.2).

The first step, with all data available as text, was about reviewing these pages line-by-line. Dey (1993) argues that the first stages of any initial categorization of the data are bound to be rather slow and tentative. The Semi-Structured Interviews' (SSI) transcripts were initially considered. The choice of starting with this source of data was made as a review of the documents collected and of the field notes taken was already implemented and it informed the SSI process.

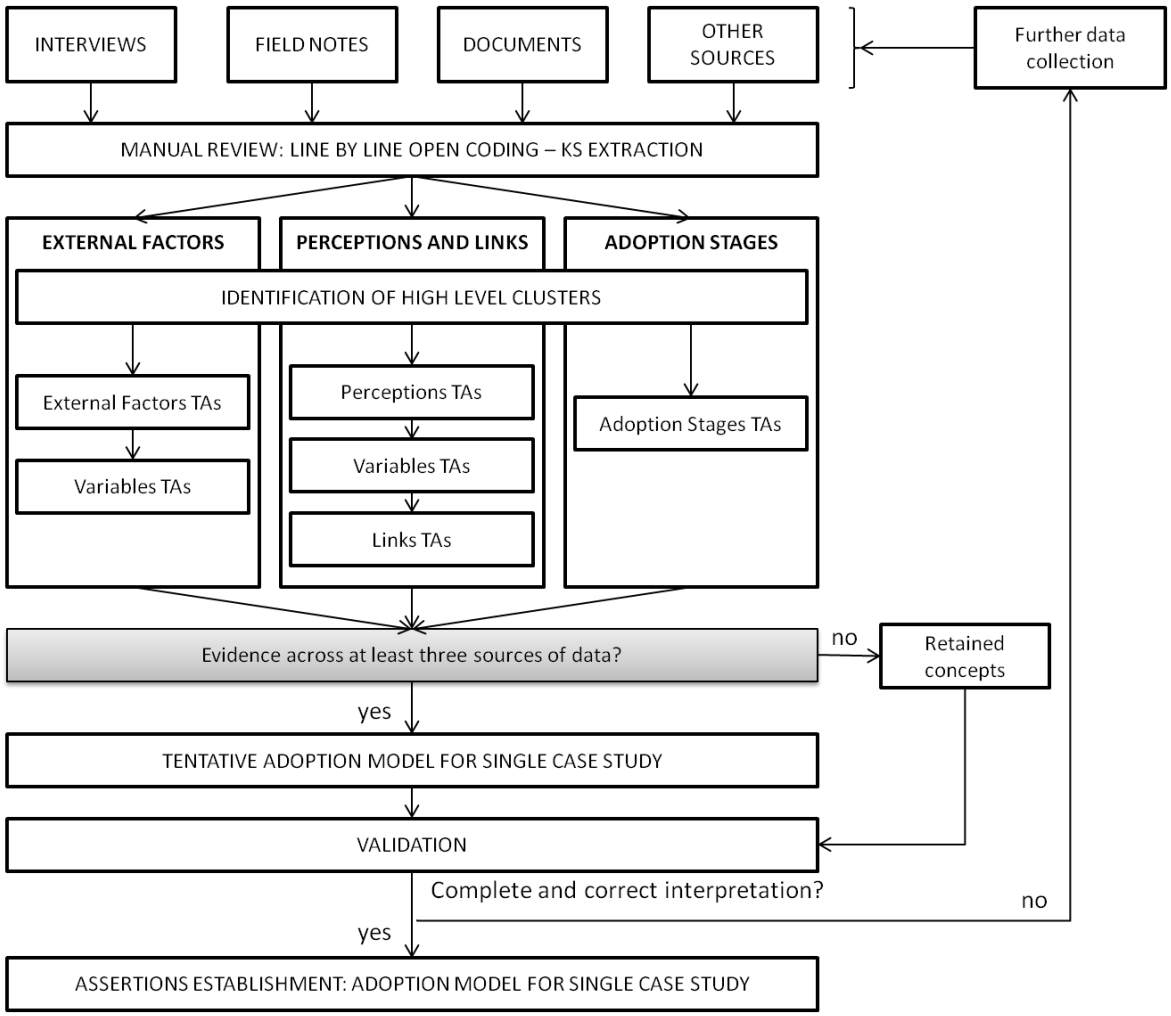

Figure 3.4: Single case qualitative analysis

Glaser and Laudel (2004) contend that the core and central part of the process is what they call "extraction" (p.194). Through this process, the initial material was reduced and a new basis of information separate from the original text came into existence. These concepts are indicated as Key Statements (KS) in this document. In order to keep track of all evidence collected, a code has been assigned to each KS in the form of: "#of question, #of statement" (e.g. the code "23,3" indicates that this specific statement is the third derived from question number 23 of the SSI). This labelling process of the KS became crucial when recursively going back to the original data collected at both the single - and the cross – case analysis stages. These KS (i.e. codes) are seen as tags or labels for allocating units of meaning to the descriptive or inferential information compiled during this study (Glaser and Laudel,

2004). Codes usually are attached to pieces of varying-sized words, phrases, sentences or whole paragraphs, connected or unconnected to a specific setting (Miles and Huberman, 1994). Mayring (2000) describes the summary procedure as attempts to reduce the material in such a way as to preserve the essential content and by abstraction to create a manageable corpus which still reflects the original material.

Therefore, the objective was to extract all the key themes and subsequently to categorize them in relation to the three main concepts of IS adoption (see section 3.6), i.e. KS related to external factors, perceptions, and to the adoption stages. Some of the KS emerged from the line-by-line review were found to be related to one or more of these categories (i.e. external factors, perceptions, adoption stages). These were considered as critical because they ingrain information about potential links between those categories. KS clustered within both external factors and perceptions were labelled as "E->P". An example of KS that fell into this category is (from Case A) "with right standard people will adopt it". This statement encompasses information about external factors (i.e. in first approximation that a standard for the data needs to be developed) that leads to perceptions (i.e. the perceived usefulness of open data which will make "people" adopting it for service innovation purposes).

The analysis proceeded separately between external factors, perceptions, and adoption sub-factors. At first, for all categories, the initial KS were grouped based on their similarities. This exercise was conducted with the objective of defining clusters under each category to have as a result more manageable units of analysis. By tackling these clusters individually, a number of concepts emerged from grouping similar KS, information from the documents collected, and the field notes.

Within each cluster defined, the next step involved explaining, clarifying, and annotating the identified factors. From this exercise, Tentative Assertions (indicated as TAs in this document) about each case were formulated for those resulting factor or factors (some clusters identified a single tentative factor, whereas some others were further divided into additional categories), variables, and links, that were supported by strong evidence (i.e. by at least three sources of data). It is noted that, consistently with the role of theory in this research (see section 3.6), KS about links were tackled as part of the analysis of perceptions. On the other hand, some of the TAs were found not to have strong confirmation across the different sources of evidence. These were highlighted and retained at this stage of the analysis. Their inclusion or exclusion was discussed at the final validation meetings. In other words, for some of the variables more data was needed to be collected to explore their relevancy in the specific case's adoption process. Thus, a dedicated session in the final validation stage was allocated to investigate these variables in-depth.

Through the aggregation of all TAs defined in relation to factors, variables, and links, a tentative adoption model for the single case in question was formulated. These preliminary findings, together with the ensemble of retained concepts, were brought into the validation meeting (see sub-section 3.10.6). The preliminary tentative model was reviewed and changed accordingly. If some tentative aspects were found to be unclear or incomplete and/or new concepts were emerging from these interactions, the possibility of undertaking additional data collection processes was considered. On the other hand, when correct and complete interpretation of the data collected was ensured, those TAs that were validated were translated into actual assertions about the case.

After the validation meetings, thick descriptions about factors, variables and links between factors influencing adoption of open government data for commercial service innovation were available for each case and represented the input for the next stage of this research, i.e. the cross-case analysis. The writing of the case study report represented the last step undertaken for each case study.

3.13 Single Case Study Report

Stake (2006) defines a case study report as "a summary of what has been done to try to get answers [to the research question], what assertions can be made with some confidence, and what more need to be studied" (p.14). The issue on how to report field work is believed to be important in all research, but is argued to be "particularly critical in interpretive case studies" (Walsham, 1995, p.78). In this context, the subject reported does not simply involve facts, but also the researcher's interpretations about some people's interpretations.

An acknowledged challenge for writing qualitative analyses' results is to keep the balance between description and interpretation. According to Patton (2005), these reports should provide "sufficient description to allow the reader to understand the basis for an interpretation, and sufficient interpretation to allow the reader to understand the description" (p.503). To which extent these aspects should be considered is the main challenge (Zhang and Wildemuth, 2010). Furthermore, according to Walsham (1995), an interpretive case study report should encompass reporting on both the collection of field data and the data analysis. The author argues that a case study report should include: the research site chosen; the reasons for this choice; the number of people who were interviewed; what hierarchical or professional positions they occupied; what other data sources were used; over what period the research was conducted; how the field interviews and other data were recorded; how they were analysed; and how the iterative process between field data and theory took place (p.79). Based on these considerations, a case study report was developed for each case. These documents included five main sections:

1. Introduction: this section included reasons why the case was found to be suitable as well as an overview of all the interactions that took place between me and the case and of all the data collection methods employed.
2. Contextual data: here all the general information about the company and the context in which it operates were reported. It included history of the company, organizational structure, information about its people, their experiences, and their roles within the company, and a description of the case's open data service(s).
3. Data analysis and interpretation: the overall coding process was reported.
4. Validation meeting(s).
5. Case study assertions: in this last section the final version of the adoption model for the specific case was outlined. The emerging thick descriptions of each component were also provided.

Once completed, these case study reports were sent to the respective CEOs.

3.14 Cross-Case Analysis

At this stage of the research, the single case study reports have been finalised (see Chapters 4, 5, and 6). The next step was about conducting analyses across the cases studied aiming at the development of Cross-Case Assertions (CCAs). These constitute the main contribution of this research to the academic conversation on open data services. Results achieved from the cross-case analysis were also the starting point for positioning this research's findings within the literature on IS adoption.

The cross-case analysis for this research has been designed (mainly) based on Stake's (2006) approach. Among the options proposed in his book, "track 2" has been chosen because of its focus on "cross-case assertions" as opposed to emphasizing "the situationality of the *quintain*" (track 1). Indeed, track 1 is interpreted to be not aligned with the philosophical stances adopted in this research. It is described as "the track emphasizing the various situations and findings of the cases" (Stake, 2006, p.6). However, one of the main activities in this track is described as rating cases' utility and findings in relation to their importance for understanding the *quintain* (Stake, 2006). I believe that all the cases are equally important for the understanding of the phenomenon of interest and activities involving "rating" cases' contributions towards this goal would be a relevant approach if critical positions were assumed. Notwithstanding this, general concerns emerge from the literature in relation to undertaking cross-case analysis using an interpretivist perspective. Therefore, before embarking into the cross-case analysis process, some considerations are needed about the consistency of this process with the philosophical underpinnings assumed for this research.

3.15 Cross-Case Analysis Philosophical Consistency

Although cross-case analysis is acknowledged to be suitable for inductive studies (Eisenhardt, 1989) and Stake's (2006) methodology "falls within the interpretive/constructivist paradigm" (Lauckner *et al.* 2012, p.5), there is an ongoing debate in the literature about the consistency of comparing cases under interpretivist philosophical underpinnings. This debate is based on the difference between deeply contextualized and particularistic case knowledge, and multiple case study research (Foreman, 1948) (Allport, 1962) (Molenaar, 2004). In other words, "counter- arguments stem from an epistemological conviction that case knowledge emerges from a dense descriptive study of the particularities of a case" (Khan and Van Wynsberghe, 2008). Therefore, comparison is believed to obscure case knowledge including knowledge not pertinent to the comparison (Peattie, 2001).

The main issue emerges around the concept of generalizability (Khan and Van Wynsberghe, 2008). However, in contrast to the positivist understanding of generalizability, new concepts have emerged to extend and amplify the impact of a single case beyond the case itself (Yin, 2013) (Becker, 1990) (Smaling, 2003) (Flyvbjerg, 2006). For instance, while findings from case studies can't be generalized in a probabilistic sense, these may be still relevant to other contexts (Goetz and Lecompte, 1984). In the context of interpretive research, Khan and Van Wynsberghe (2008) list a number of references, including Stake's (2006) approach, "that can be invoked to rationalize cross-case analysis" (p. 7). In this way, the paper highlights the importance on practical concerns for case study researchers to be considered before embarking in cross-case analysis processes. These include three aspects:

1. Preserving the essence of the cases: Tesch (1990) argues that the complexity of meaning resulting from the peculiarities of each case is in danger to get lost when content is reduced to enable comparison across cases. Stake (2006) counter-argues that it is possible to learn from both the uniqueness and the commonality of a case. In particular, "by providing ample contextualized details of the cases [see Chapters 4, 5, and 6] and findings of cross-case analysis [see Chapter 7], a researcher can conceivably preserve the uniqueness of a case and convey the value of their engagement with a cross-case analysis" (Khan and Van Wynsberghe, 2008, p.7).

2. Reducing or stripping the case of context: in cross-case analysis, the contextualized origins of each case are in danger of being lost as cases are compared. However, according to Ayres *et al.* (2003), the reduction of some contextual detail is consistent with the goal of implementing a cross-case analysis, i.e. to identify themes across cases. For this research it becomes useful to focus on Tesch's (1990) conceptual understanding of cross-case comparison. He describes it as a "de-contextualization and re-contextualization" of cases. Therefore, the proposed process is: (1) separate case study data into units of meaning (de-contextualized because they are separated

from the individual cases) and (2) re-contextualize them as they are later integrated and clustered into themes.

3. Selecting appropriate cases to compare: the cases selection criteria have been defined in section 3.8 of this dissertation. In relation to cross-case analysis, the literature suggests that a few cases should be studied in cases-oriented approaches (as opposed to variable-oriented approaches that are not consistent with the purposes of this research), and that comparison of cases is possible if these are diverse across contexts.

These three points were addressed in this research. Particular attention within the cross-case analysis process is placed in relation to points (1) and (2) above. In order to be consistent with the scope of this study, the uniqueness of each case and the related context were kept when merging the cases' findings and subsequently developing CCAs. These CCAs resulted enriched by thick descriptions. These descriptions were formulated from the specific cases that contributed to the assertions' establishment. All factors and variables that weren't included in the cross-case findings (as lacking evidence from across the three cases studied) were retained as contributions to the specific context where these emerged.

Based on these reflections and the methods proposed in (Stake, 2006) (Eisenhardt, 1989) (Eisenhardt and Graebner, 2007) (Lauckner *et al.* 2012) (Darke *et al.* 1998), this study employed five main steps for the cross-case analysis and finalization of the findings. These steps are:

1. Revisit each case's specific transcripts with the developed factors, variables and links.
2. Merging the case findings. Use the concept of "merging" (Stake, 2006) in integrating the variables, factors and links, and develop Cross-case Tentative Assertions (CCTAs).
3. Validation across cases.
4. Develop the final Cross-Case Assertions (CCA) and the related combined model for open data adoption.
5. Positioning these assertions within those from the literature and expect contributions to emerge (i.e. enfolding literature (Eisenhardt, 1989)).

These steps are described in the following sections.

3.16 Merging the Cases Findings and CCTA Development

After thoroughly reviewing each case study report, the scope of the second step (see list above) was to cluster the findings from the single cases according to their similarities. "For merging findings into clusters [the researcher] combines, one by one, the finding strips according to their similarities.

Findings similar in topic get placed close together; those dissimilar are placed farther apart. Even if findings are contradictory, any two that are on the same topic are placed in the same cluster" (Stake, 2006, p.60). In this sentence, "finding strips" refer to the assertions developed for each of the cases. The analysis was conducted separately for external factors, perceptions, and adoption sub-factors.

The first step that was undertaken was an extensive review of all the single case reports with the objective of outlining all the factors emerged and looking for immediate correspondence across cases. This refers to the first "de-contextualization" process (Tesch, 1990) outlined above. For each factor three possible situations were found: (1) the factor appears in all the three cases studied; (2) the factor appears in two of the three cases studied; (3) the factor emerged for one case only. These situations were translated into different levels of cross-case analysis. For each level, different strategies were adopted:

- **Level 1**: for the first level, the analysis focused on merging similar variables within those factors that emerged across the three cases. These were expected to constitute final assertions across cases.
- **Level 2**: regarding the second (i.e. of factors emerged from two of the three cases), further confirmation was sought across variables from other factors from the "missing case" (i.e. the case for which the specific factor did not emerge). If further evidence was found, the resulting factor was established and a CCTA was formulated. Otherwise, reflections in that way were postponed to the validation across cases stage.
- **Level 3**: in relation to the third level of cross-case analysis (i.e. of factors emerged from one case only), all the variables within the remaining clusters were taken as the unit of analysis aiming at the definition of additional factors. In other words, all the variables emerged were listed and clustered together based on their similarities. From the clusters that arose, CCTAs were formulated for those concepts that were supported by considerable evidence across the three cases. According to Stake (2006), those for which merging was not achievable were retained.

This process is summarized in Figure 3.5 (next page). As shown in this figure, for each level of analysis, two possible outcomes were in place: (1) if sufficient evidence across cases was found for a specific concept (i.e. factor, variable, or link), this was translated into a CCTA; (2) if sufficient evidence across cases was not found for a specific concept, this was retained at this stage of the analysis. It is noted that concepts retained at Level 1 were added as additional units of analysis for Level 2, and the same was done between Level 2 and Level 3. This activity was implemented because these retained concepts could be seen in first approximation as specifically related to the contexts from which they emerged. However, before deriving any conclusion in this way, the relevancy of these concepts was

sought among other factors. For example, given that the single cases' models were developed independently, a situation in which a specific variable emerged within different factors across the cases was possible.

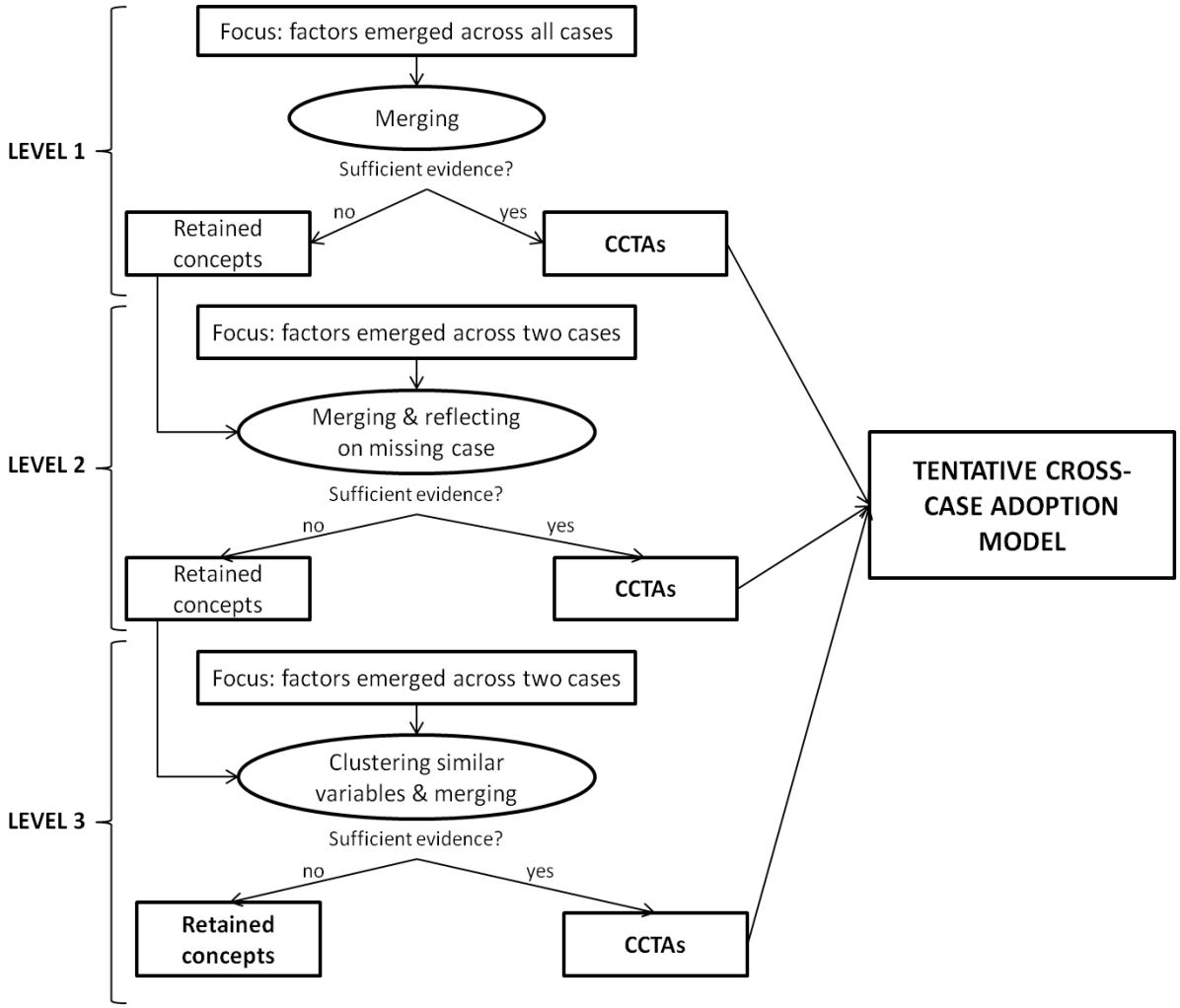

Figure 3.5: Levels of cross-case analysis; development of CCTAs

As a result of these steps, CCTAs were developed. These assertions encompassed one or more of the merged findings. According to Stake (2006), each CCTA was formulated considering: "a single focus, an orientation for understanding the *quintain*, and evidence to support it" (p.62). Through the aggregation of these CCTAs, a preliminary model describing adoption of open data across cases was developed. This, together with the ensemble of concepts retained after level 3, represented the input for the validation across cases stage.

3.17 Validation across Cases

According to the protocol designed for this research (see section 3.7), a further stage of the cross-case analysis was undertaken with the goal of achieving confirmation of the conclusions about the *quintain*

formulated from the merging of the findings of the single cases studied. This process and its purposes can be compared to those implemented for validating the findings within single cases (see sub-section 3.10.6). Similarly, here the objective was to assure that the picture achieved is as clear and meaningful as possible. "The picture is the picture of the *quintain*, which was not studied directly, and its cases, which were. It is a sort of validation step and follows a classical strategy – seeing if new views are consistent with what is already well known about the case and the quintain" (Stake, 2006, p.77). To do this, Stake (2006) suggests going further "afield" (p.77). Therefore, consistently with the multiple case study protocol (see section 3.7), additional field work was conducted.

The starting point for this stage was the preliminary findings achieved so far in the cross-case analysis process. These were represented by CCTAs about factors, variables and links between factors that influence the adoption of open data for commercial service innovation across cases. Also, those variables that were retained at the previous steps were re-considered as part of this stage.

Thus, for this validation across cases three main objectives were defined:

1. Extensive review of the preliminary findings: this stage of the research is meant to ensure that correct interpretation was achieved (Stake, 2006). Stake (2006) argues that at this stage "we need to oblige them [i.e. the participants in this process] to say what the conclusions mean to them" (p.77).
2. Further investigation of the retained variables and factors and discussion on whether these should be included or considered as context-dependent.
3. Leverage the expertise of the participants for collecting additional data about the adoption of open data for commercial service innovation.

3.17.1 Validation Participants

What becomes critical in this process is selecting the right people. The main requirement that participants in this effort needed to meet was being representatives of the various stakeholders within the overall open data ecosystem. A certain level of expertise from these individuals was required in order to meet the objectives of this validation stage. Thus, subjects involved in this effort needed to have knowledge about open data and open data services, and a certain level of experience in this field, in order to be able to: (1) interpret the preliminary findings of this research and the supporting evidence; (2) provide experience-based opinions motivating inclusion or exclusion of those factors and variables that were retained at the previous steps; and (3) provide additional comments, critiques, and details about the adoption of open data for commercial service innovation. Therefore, subjects involved in this effort needed to be "information-rich participants" (Miles and Huberman, 1994).

Based on the objectives of this stage of the research, purposive sampling has been chosen as a suitable method for selecting participants. Purposive sampling is defined as an ensemble of strategies "designed to enhance understandings of selected individuals or groups' experience(s) or for developing theories and concepts. Researchers seek to accomplish this goal by selecting information rich cases, that is individuals that provide the greatest insight into the research question" (Devers and Frankel 2000, p.264). Besides being widely used in qualitative research when looking for information-rich participants (Miles and Huberman, 1994) (Patton, 2005) (Devers and Frankel, 2000) (Sarantakos, 2012), purposive sampling enables "the researcher to select suitable respondents who have the knowledge of the research topic so that it would be of most benefit to the study exercise" (Baccarini *et al.* 2004, p.289). However, the "list of purposive strategies is virtually endless" (Palys, 2008, p.697), and the particular strategy should be selected based on the objectives of the actual study. Several attempts have been made to provide categorizations of different purposive sampling strategies (Sandelowski, 1995) (Coyne, 1997) (Kuzel, 1999) (Patton, 2005). Among those proposed, a combination of two strategies has been selected for selecting participants for this validation effort. These are: (1) maximum variation sampling (Patton, 1990) (Sandelowski, 1995) (Coyne, 1997); and (2) stakeholder sampling (Palys, 2008).

The basic principle behind maximum variation sampling is to gain greater insights into a phenomenon by looking at it from all angles (Patton, 1990). The literature acknowledges that purposeful sampling strategies provide direction for sampling but often depend on some knowledge of the setting being studied. In this research, this knowledge has been developed through the inductive investigation of multiple cases. Palys (2008) argues that maximum variation sampling involves "searching for individuals who cover the spectrum of positions and perspectives in relation to the phenomenon one is studying" (p.697). In this way, maximum variation sampling is a technique that is used to develop a wider picture of the phenomenon. Clearly, this is suitable for the aim of this stage of the research. In summary, a validation of the findings of this research is achieved through interactions with individuals involved at all levels of the open data ecosystem.

On the other hand, stakeholders sampling "involves identifying who the major stakeholders [of the phenomenon of interest] are" (Palys, 2008, p.697). Through this study, a range of critical entities and stakeholders was identified from reflecting on the findings emerged from both the single cases and the cross-case preliminary findings. In particular, the actors / entities that emerged as playing a key role in this ecosystem are: city councils; intermediary organizations; people involved in community-oriented initiatives; researchers; and other actual or potential open data re-users.

Following these techniques, eight subjects were selected and approached. The overall group was heterogeneous across the different entities and stakeholders of this ecosystem. In particular, the following subjects were selected: two open data services business representatives; one manager responsible for the local intermediary organization; two open data researchers with different backgrounds (i.e. sociology and information systems); one software developer interested in discovering open data; and two people from different city authorities. In relation to the latter, these two people were representatives of different situations: (1) city council manager involved in the related open data strategy; and (2) representative from a city council that has not embraced open data yet, but that is researching this field. Therefore, maximum variation of stakeholders across the different entities identified within the open data services ecosystem was achieved consistently with the sampling methods adopted.

A summary of these participants is presented in Table 3.10. It includes: their backgrounds, level of involvement in the open data ecosystem, and the duration of the validation meeting (excluding the time employed for introductions to me and my research). It is noted that all conversations were audio recorded. All interactions happened through face to face meetings and similar conditions to those created for the SSIs for the single cases were in place (see sub-section 3.10.5). Only in one case (i.e. for S8, see Table 3.10 below), this interaction happened through a virtual call.

In terms of the actual process followed, all subjects were initially contacted. In case of a positive feedback on their availability for this study, the objectives of this research were presented. Furthermore, prior to the actual validation meeting, where possible, a document including the summary of the research objectives, the cases, and the preliminary findings was given to the subjects. This enabled: (1) significant time saving at the time when the actual validation process was undertaken; and (2) giving the time for the subjects to thinker with my study, the cases, and the cross-case interpretations and findings. For four of these eight subjects this document was shared prior to the actual meeting. Regarding the remaining subjects, these two activities were undertaken during the same day. The specific activities implemented are presented in the following sub-sections.

#	Background	Role	Description	Duration
S1	Business	Re-user / demand side	Founder of a business that re-uses open data for delivering a service to help users in locating retailers.	45 minutes
S2	Business	Re-user / demand side	Founder of a business that integrates open data in the development of an API for cameras in cities.	50 minutes
S3*	Intermediary organization	Supply side	Manager of an intermediary organization that publishes some of the city's open datasets on its own portal. Involved in engagement initiatives.	80 minutes
S4*	Researcher	Academic research	Research activities conducted in relation to open data as a civic engagement tool. The subject is also involved with the council's initiative.	110 minutes
S5*	City council	Supply side	City council representative involved in the related open data strategy.	95 minutes
S6	City council	Supply side	City council representative of a city that did not embrace open data yet (but is apparently placed during this process).	25 minutes
S7	Software developer	Potential re-user	This person was attending an open data meeting organized by the local authority. The main goal for him was to explore open data as a potential valuable source for new software development opportunities.	40 minutes
S8*	Researcher (sociology)	Academic research	Sociologist with several years of experience in open data research and practice for both Irish and Canadian cities.	120 minutes

Table 3.10: Participants in validation across cases

*summary document was given prior to the meeting

3.17.2 Review of the Preliminary Cross-Case Findings and additional Data Collection

The process that was followed was the same for all of the participants. The first step was about clarifying ethical aspects of these interactions. Subsequently, once the subjects achieved familiarity with this research and its objectives, the cases' history, services and the findings, the recorded period started. The reason why particular attention has been given to the actual cases was motivated by the fact that when discussing the cross-case preliminary findings, continuous referrals were made to the original evidences supporting these statements (i.e. the cases). Therefore, the first step was about presenting all the preliminary findings, factor by factor, variable by variable, and link by link. When discussing their relevancy (enriched by questions such as "would you like to add something on this aspect based on your experiences?") the subjects were relating these concepts to their personal experiences. In this way, the thick descriptions achieved from the integration of the three cases' findings were enriched with those emerging at this stage. This stage of the process involved on average around 80% of each interaction.

Overall, strong confirmation was found among the preliminary factors, variables, and links presented. When presenting each CCTA, a significant amount of comments was made by the subjects approached. In reporting the cross-case findings (see Chapter 7) the aspects that emerged at this stage are highlighted. These led to an augmented understanding of the factors, variables, and links previously derived. On the other hand, the importance of a number of factors, variables, and links was simply confirmed by these subjects. Once validation of factors, variables, and links was achieved, these were translated from CCTA into final Cross-Case Assertions (CCA).

3.17.3 Review of Retained Variables

When familiarity was achieved with the single cases, the cross-case preliminary factors, variables, and links, all the previously retained variables were discussed. For each of these forty-one variables and three adoption sub-factors, two different situations were possible: (1) the variable was integrated in the final findings as a consequence of additional data collected from these subjects; (2) the subjects agreed that the variable had to be excluded as its relevancy is limited to the actual context in which it emerged.

As a result of the analysis of this step, five variables and one adoption sub-factor were included in the final cross-case findings. The remaining thirty-six variables and two adoption sub-factors were finally excluded from the cross-case findings. Their relevancy was established as limited to the specific context in which these emerged. A full list of these finally retained concepts, and the related contexts, is provided in Appendix 7.

In summary, these eight interactions led to significant results for this research:

- The final adoption model, the language used within it, and the interpretations ingrained from the single cases' findings, were strongly validated by all the subjects approached.
- Two tentative external factors were merged into a single one.
- Five of the previously retained variables were integrated into the final findings.
- An additional link was established within the model.
- Confirmation about the retention of thirty-six variables and two adoption sub-factors was achieved (see Appendix 7).
- The subjects contributed in enriching the understanding of some of the factors, variables, and links defined. These contributions are highlighted (together with the reference of the specific contributor) within Chapter 7, i.e. when presenting the cross-case findings.

Concluding, for those factors, variables, and links for which validation was achieved, CCAs were formulated. Subsequently, the final inductively developed model for adoption of open data for commercial service innovation across cases could be outlined.

3.18 Positioning Cross-Case Assertions within the Literature

The last and final stage of the cross-case analysis (and of this project) was about positioning the cross-case findings within the literature. According to both the methodology designed for this study (Eisenhardt, 1989) (Eisenhardt and Graebner, 2007) (Miles and Huberman, 1994) (Quaddus and Xu, 2005) (Berg, 2001) and its inductive nature (Huff, 2008), all the CCAs formulated were positioned within the relevant literature. This process was also labelled as "enfolding literature" (Eisenhardt, 1989). Through this process, specific contributions to the existing knowledge base were rigorously identified.

This exercise was conducted in relation to the two main bodies of literature this research aims to contribute to, i.e. open data services and IS adoption. An extensive review of theories related to IS adoption and of the open data services literature informed this process. The inductively generated findings from this multiple case study were compared with the current understanding of open data services and of how IS innovations are adopted.

With respect to the open data services literature, all aspects emerged from this study were compared with those mentioned across the papers considered for the SLR process (see Chapter 2). In this way, more specific contributions could be outlined.

In relation to adoption theories in IS, the method selected for applying this "theoretical lens" to the findings was pattern matching logic (Yin, 2013). Yin (2013) recommends the usage of this technique to relate the identified patterns (observed patterns, i.e. the findings across cases) among concepts identified in the literature or theory (theoretical patterns). This technique to link data with the existing literature is well established and acknowledged to be appropriate for inductive and interpretive studies (Cambpell, 1975). It is defined as involving "the comparison of predicted patterns and/or effects [i.e. from IS adoption theories] with the ones that have been empirically observed [i.e. the inductive findings of this multiple case study], and the identification of any variances or gap" (Baskarada, 2014, p.16). Through this process theoretical contributions to the IS adoption academic conversation emerged.

The activities and specific methods employed for these exercises are presented in Chapter 8, i.e. when discussing the cross-case findings.

3.19 Chapter Summary

This chapter described the research approach and methodology employed in this study. In summary, an interpretivist philosophical position was assumed. Subsequently, the choice of adopting an inductive reasoning through a qualitative approach was discussed and established. Based on these decisions, different potentially suitable methodologies have been presented and compared against the objectives of this research. These reflections led to the selection of multiple case study as the suitable methodological guidance for addressing the research question.

With respect to the methodology that was selected and employed, the overall multiple case study protocol has been initially presented. The chapter followed with the description of how cases were selected and the design of the single case protocol. Subsequently, all the data collection techniques leveraged for the single cases were outlined. The description of the methodology continued with a specific focus on the methods and techniques adopted for qualitatively analysing the data about each case. Aspects related to reporting the cases studied were tackled in the last section of this segment.

Ultimately, the last segment of this chapter presented the methodology designed and implemented for analysing the findings across the cases studied. Initially, a discussion on the suitability of conducting cross-case analysis in this research was provided. Subsequently, the specific steps of this process have been described from both the conceptual and practical perspectives. Furthermore, the activities implemented to validate these interpretations were presented. In summary, according to the methodology designed for this research, further field work was conducted. This involved the selection and subsequent interaction with eight subjects. These activities led to gain assurance about the final findings derived in this research. Finally, the methods employed for positioning these findings within the existing literature were outlined.

Chapters 4, 5, and 6 focus on reporting the investigation and the findings from cases A, B, and C respectively. It is noted that the overall methodology designed for the single cases presented in this chapter is contextualised to the specific situations encountered in dedicated sections within the following chapters.

4 CASE STUDY A REPORT

This chapter is dedicated to the description of the activities and the findings related to Case A.

The first section provides an introduction to the case, its history, the people working within it, and the detailed characteristics of the open data services offered. The specific interactions happened as part of this case study are described in section 4.2. It is noted that the methodology employed followed the single case protocol (see section 3.9). Therefore, the main focus of section 4.2 is on describing what techniques and questions were added to the original protocol based on the specific situation encountered. Subsequently, also the analysis process is contextualised to the specific context of this case study (section 4.3).

The core of this chapter is represented by section 4.4 in which the findings achieved from the investigation of Case A are presented. All assertions formulated are listed and explained. Thick descriptions generated for each factor, variable, and link between factors are provided. The chapter concludes with a summary of the findings achieved from this case. In this last section the adoption model generated for Case A is outlined. This is presented together with the adoption stages that explain the process followed by Case A, from the initial generation of the idea, to the achievement of sustained use of open data for commercial service innovation.

4.1 Introduction to Case A

Case A represents the first case study that was carried out as part of this research. The company proposes a set of open data services in the context of urban planning in cities. These services analyse building permit data for the home remodelling and construction market.

Case A's starting point was the identification of a business opportunity which arose from the experience of working with governments, and particularly on urban planning, of the CEO and founder. "I was frequently asked to search for planning data and found this very labour intensive, as each council has its own site and its own system. We thought that it would be a lot easier if all the boundaries were removed and the projects were located on the map with an easy search facility which incorporates keywords to filter out the applications you are looking for, so that's what we did" – the CEO explained. The company was founded in 2010. Due to internationalization of activities, the company was re-named in 2011. After its launch in Ireland, it moved to United States. Currently, the services are delivered in Dublin and nine major cities in US. The CEO also highlighted that Case A "is hoping to have the planning data for another sixty US cities in the next year or so".

A team of six people is working within the company. According to the CEO, “I am CEO, while another person heads up our business development in the US as he is based in Palo Alto (California, US). We also have three developers with also graphic design skills”. Ultimately, another person is responsible for front-end activities and marketing initiatives. I had the opportunity to meet four of these people and to investigate their position within the case, as well as their background and previous experiences. This information is presented in the following table.

Subject	Position	Background
1	CEO and founder	Degree in civil engineering; experience in the construction sector, as an engineer and project manager.
2	Business development (US)	Entrepreneurial background
3	Lead developer	Software developer; graduated in electronic systems.
4	Developer	Developer / IT engineer; marketing; experience in a major IT company.

Table 4.1: Case A's people

All in all, the services provided aim at making urban planning information easier to find. Case A visualizes planning and building permit data to enable users to know what permits are being requested near or on their property, in real-time. The service charges retailers, property owners and construction companies monthly fees to get leads and understand what's happening in the building environment in their cities. This enables the display of what is being planned in their area, and at what stage in the planning process the application is at. The service has additional features which allow users to: be notified of new applications in their area; identify what works require planning permission; access a directory of professionals; and receive news feeds of planning-related information from local area plans to strategic infrastructure. The service is available to end-users through a web application. However, “we have built the app in such a way that it can be easily tweaked for both Android and IPhone” (developer). In particular, the company offers four different features within its open data service. For each of them the main characteristics and the revenue models (if in place) are explained:

- Planning map (free service): through an interactive map of the selected cities, all information about all planning applications in the area and their status are provided. This information includes: planning reference number; planning address; planning application description; and link to the drawings. According to one of the developers, “each application is identified with a small site notice icon on the map with an information window providing details on the proposed project, and a link at the bottom providing direct access to the planning drawings and documents associated with that file”.

- Get notified (freemium model with upgrade subscription): by signing up and specifying a zip code or an area through the interactive map, the user can get alerts when planning applications are received within that zone. Users can upgrade to wider radii as a paid service. The upgraded service is suitable mainly to business users.
- Planning list features (premium, subscription service): in a premium account version, weekly planning lists for the local authority are provided. An advanced search tool is available including variables such as: local authority, planning status, reference number, street address, and keywords (e.g. "extension"). "This process saves users valuable time reading through planning list looking for the project leads that suit them" – the business developer explained.
- 3D house (free service): the last feature of the service integrates all information about planning regulations related to different parts of a house. The information is provided through an interactive view of a common house. By selecting each item (e.g. garden, roof) the system provides answers to relevant questions about specific planning requirements and regulations.

Beside the delivery of these services for the end-users, the case also applies a layer of data analytics upon the usage of the services above described. In other words, new data is generated from analysing the users' interactions with the case's systems. According to the CEO, specific observations can be undertaken on "how many people look for information on a specific area, in what times of the year, who are these people etc". This information is leveraged, aggregated, and provided back to the related city council. Through this process, the case contributes towards improving the management of the overall urban planning context within the city. Some examples of these contributions are the identification of trends in the context, location of inefficiencies, and assistance in managing the context by enabling more informed decisions.

Case A's services are enabled by several open datasets. These include:

- Planning application datasets: these multiple datasets contain information about urban planning applications in the city. For each application, this information encompasses: (1) reference number; (2) date application made; (3) short description of the proposed development; (4) date application is registered to authority; (5) application address; (6) "apptype decision" (i.e. grant, refuse, split decision, further information, clarification of further information); (7) long description of proposed development; (8) date of final grant of planning permission; (9) date of planning authority decision; (10) extension of time; and (11) stage of the application. According to one of the councils' website, "the planning application data is updated every night".
- Local area development plans: this dataset provides policies and objectives to guide how and where development will take place in the city over the next six years. This includes spatial

information on the local area plans for areas subject to large-scale development within the lifetime of the overall city plan.

- Urban planning and construction regulations: this dataset is critical for the delivery of the "3D house" feature of the service. This data is taken from the department of environmental planning and urban development (in Ireland).

Figure 4.1 depicts a summary of the open data services delivered by the case.

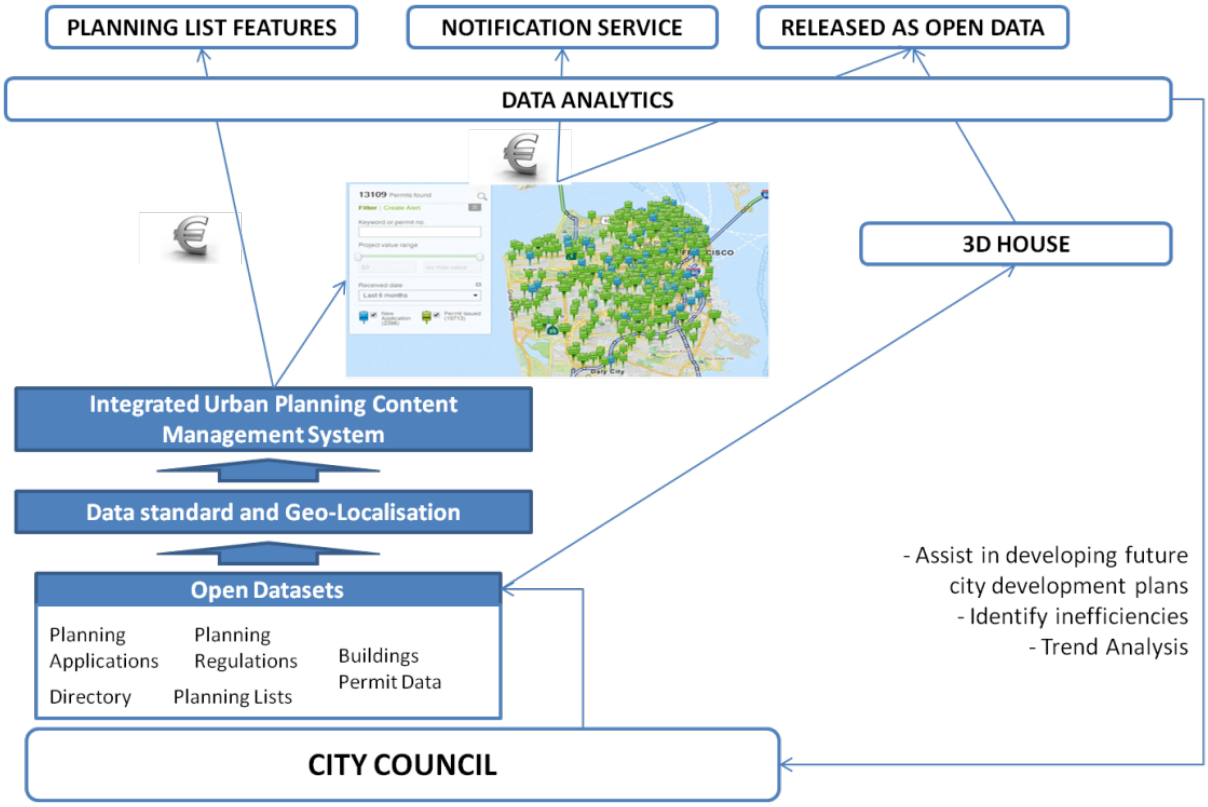

Figure 4.1: Case A's open data services

It is noted that where the "€" sign appears a form of revenue model is in place.

4.2 Contextualization of Methods

According to the methodology designed (see Chapter 3), three main sources of data were leveraged for this case study. Both internal and external documents were collected and analyzed. In particular, three internal documents were received by the CEO and seven external ones were selected among those found on the internet. A description of each document, its content, the year in which it was created as well as its length is provided in Appendix 1.1.

Observation was undertaken through several informal meetings at the case's site. Data about these interactions was collected and stored in field notes. Case A represented the case where this data collection technique has been most exploited. I had the opportunity to meet four of the six people

working within it, who I found very cooperative and willing to share information and experiences. Several discussions were undertaken during the time I spent on site. Notwithstanding this, the CEO was found as the only person with an articulated understanding of open data, and, therefore the only potential interviewee. The developers contributed substantially to this research. However, these were found not to have knowledge on open data specifically. “For us it is just a dataset” – one developer stated.

From the data initially collected and structured some important aspects emerged. Thus, for each of the steps of the SSI protocol (excluding step 1, i.e. “situating the actor”), additional questions were included. In other words, the SSI protocol was customized to the specific context of Case A.

A recurring theme that was observed is the process of expansion of the market of the company. It is clear that there is a specific approach to “get councils on board” (or “brought it to the steps of city hall, bringing that stale political jargon to life”). As a consequence, a set of questions was added to the protocol for this case regarding these types of activities. The purpose was to know more about this process and to explore different scenarios that happened as part of the case’s experiences (e.g. investigating situations where councils were rejecting the idea, and explore the reasons for such decisions). Furthermore, it emerged that a role was played by those so called hack competitions (e.g. “we won an enterprise award from the city council for the use of their data and another council has us up on their website”, CEO). Thus, investigating the specific role of such initiatives towards the adoption of open data was another variable considered when gathering the data. Concerning step 3, a clarification was needed regarding question 2 (*“In the choice of the specific dataset(s), which variable did you consider?*”). This question was targeted for those cases in which the idea followed the analysis of the dataset. Here, the idea preceded the availability of the data. Hence, the same question was asked about the interviewee’s opinions on this stage, and not referring to his personal lived experience within the company. In addition, motivated by controversy between the licenses assigned across the open datasets re-used, the following question was added to the protocol: “*did you feel any threat of lawsuit or violations?”* As part of step 4, it emerged from initial interactions the importance of establishing relationships with the local municipalities. Thus, understanding how these interactions happen, who is involved within them, and how people within the case approach those councils, were an additional goal for this step.

The first interview lasted for approximately one hour and thirty minutes. The interview was transcribed the same day in which it was conducted. At the time of writing the transcripts, the need of clarification of some aspects emerged. In particular, the concept of “contracts” and the strongly stressed importance of the “data standard” informed the need of conducting an additional interview

with the subject. This further interaction gave me the opportunity of revisiting the first SSI with the original interviewee and with two of the case's developers. Regarding each of the questions asked, I listed the major concepts emerged and interpreted. Feedback was sought to achieve confirmation of these interpretations. As a result, further concepts emerged from this stage and enriched those already extracted. Overall, this meeting lasted for almost two hours.

Through the qualitative analysis of the data collected, preliminary findings were outlined and constituted the main input for the final stage of this case study, i.e. the validation meeting. This interaction happened in the case's offices and lasted for slightly more than two hours. The preliminary findings were shown, reviewed, and discussed with the CEO and two software developers.

4.3 Case A Analysis

The analysis process undertaken for Case A strictly followed the method previously designed (see section 3.12). The line-by-line open review exercise conducted on the SSI transcripts led to the extraction of two hundred and eighty-one Key Statements (KS) (see Appendix 3.1). From the first coding step, one hundred and eighty-one statements were grouped under the external factors' category, while for perception and adoption one hundred and nine and sixty-three KS arose respectively (see second column in Appendix 3.1). Finally, seventy-three KS were extracted in relation to links between factors. The analysis proceeded separately between external factors, perceptions, and adoption factors.

4.3.1 External Factors Analysis

As a result of the first round of analysis, I isolated one hundred and eighty-one KS related to external factors. After several attempts, I classified them in five different clusters (see third column in Appendix 3.1) that are:

- ***Cluster A****:* external factors related to open data and open dataset characteristics.
- ***Cluster B****:* external factors related to the city. The city is generally meant as the entity that owns and publishes open data.
- ***Cluster C:*** external factors specifically related to the company.
- ***Cluster D****:* external factors related to the interaction between city councils and the company.
- ***Cluster E****:* other external factors.

Cluster E was added to the list to ensure that these categories were collectively exhaustive, and so that all of the KS fell into one of these five clusters. Some of these codes were interpreted to be related to more than one category. For instance, KS 15,1 (see Appendix 3.1), i.e. "looking at other companies

showed the need for cooperatively creating a standard for the open dataset", contained insights about: external factors related to the need of a standard for the open dataset (Cluster A), that has to be developed by the company (Cluster C), in collaboration with the data supplier (Cluster D). In turn, these clusters became more manageable units of analysis and were tackled individually. For each of these, similar KS were again grouped together. Subsequently, for each concept emerged as supported by several KS, additional evidence was sought across documents and field notes. Those for which strong evidence across sources was found were translated into Tentative Assertions (TA) about factors (i.e. in first approximation the cluster) and variables (i.e. the emerging concepts within each cluster) about the case. Alternatively, those for which strong evidence was not found were highlighted and retained at this stage. The following table gives an overview of the clusters emerged, the number of tentative factors and variables, and the number of retained variables for each cluster.

	Cluster	Number of KS	Tentative Factors	Tentative Variables	Retained Variables
External Factors	A: Open Data Characteristics	38	1	6	0
	B: City Related Factors	99	1	7	3
	C: Organizational Factors	60	1	6	1
	D: Business-Council Relationship	58	1	3	0
	E: Other External Factors	57	3	10	0

Table 4.2: Case A, external factors analysis

As a result, seven tentative external factors emerged as influencing the adoption of open data for commercial service innovation for Case A. In total, thirty-six variables were defined as describing these factors and insights on possible links with subsequent perceptions generated were outlined as well. In addition, four variables were highlighted and retained as further reflection or data collection was needed. For these TAs no sufficient evidence was found. Thus, a decision about their inclusion (or exclusion) was postponed to the final stage of this case study, i.e. the validation stage.

4.3.2 Perceptions Analysis

From the first round of open coding analysis, one hundred and nine KS were classified as perceptions (Appendix 3.1). As done for external factors, other sources of data were included and constituted evidence for the emerging concepts. The first step was to provide a useful taxonomy comprised of different categories to focus the analysis of those KS related to perceptions. In particular, four main classes of perceptions emerged. These are:

- ***Cluster F:*** perceived usefulness of open data.
- ***Cluster G:*** perceived open data services business opportunity.
- ***Cluster H:*** perceived city opportunity.

- ***Cluster I:*** other perceptions.

The approach followed for the clusters defined in relation to the perceptions generated was similar to the one used for external factors. TAs were formulated when evidence was found across at least three sources of data. These were included in the preliminary findings. The remaining concepts were retained at this stage, highlighted, and re-discussed during the final validation meeting at the case's site. As done for the external factors, a summary table (also considering links) is here provided.

	Cluster	Number of KS	Tentative Factors	Tentative Variables	Retained Variables	Tentative Links
Perceptions	F: Perceived Usefulness of Open Data	61	1	3	1	2
	G: Perceived Business Opportunity	35	1	5	0	4
	H: Perceived City Opportunity	57	1	3	0	5
	I: Other Perceptions	7	1	4	0	2

Table 4.3: Case A, perceptions analysis

As shown in Table 4.3, four tentative factors described by a total of fifteen tentative variables were defined. These represent the tentative perceptions that are generated from the external factors previously emerged and upon which decisions about adopting (or rejecting) open data are taken. One variable was retained and its inclusion was discussed during the final validation meeting. Finally, a total of thirteen tentative links emerged in relation to these perceptions.

4.3.3 Adoption Process Analysis

The last and final category to be analysed was related to the actual adoption stages Case A went through from the initial generation of the idea to achieving sustained use of open data for its services' delivery. Evidence for this factor came mainly from one meeting that took place at the case's site. In addition, sixty-three KS arose from the SSI processes were also classified as describing the adoption process. Finally, further evidence was sought and found in all the documents collected and analysed. The analysis resulted in the tentative definition of six adoption sub-factors. These stages and the related number of supporting KS are summarized in the following table.

	Adoption Stages	Number of Key Statements
1	Idea Generation	4
2	Adoption	16
3	Pilot Implementation	21
4	Commercialization	3
5	Growth	29
6	Sustained Use	21

Table 4.4: Case A, adoption stages analysis

As a result, eleven TAs were formulated in relation to eleven factors that influence the adoption of open data for commercial service innovation for Case A. Furthermore, each of these factors was tentatively defined at a further level of detail. In total, fifty-one variables were identified. Regarding the actual adoption process, six stages emerged and were outlined in relation to the different milestones the case went through. Ultimately, those statements that were found as not strongly supported by evidence were highlighted and some time was allocated for reflecting on them at the final validation stage of the analysis. Therefore, a preliminary model for open data adoption for Case A was formulated and was brought into the final validation meeting.

4.3.4 Validation Meeting

In this last stage of the case study, the tentative findings were presented to the CEO and the other members of the company. Initially, the preliminary adoption model for Case A was shown to the CEO. All the eleven factors defined were found as realistically describing Case A. The variables describing these factors were found exhaustive and appropriate. Only in one case the CEO wanted to add a further characteristic and as a result a further variable was added to the preliminary findings.

Once the consistency of all factors and variables was ensured, the focus shifted to those five variables that were extracted and interpreted from the data collected but for which strong evidence was not found. The result of this discussion led to the integration of three of these five variables in the final findings. The remaining two concepts were excluded from the analysis. Before moving to the final stage, a discussion started from a question asked by the CEO around the lack of a “volunteering component” in the preliminary model. Accordingly, an additional factor was added as a result.

The final stage of the meeting was about reviewing the proposed links between the factors defined. This step was taken as the last stage of this meeting because it was important to ensure that the subjects were familiar with all factors and variables within them. The starting point was a list of insights gained from the data collected, structured and interpreted. Each of these links were explained, reviewed and changed accordingly with the feedback received. Overall, a significant approval was gained regarding the links proposed. Five more links were proposed during the meeting and one was excluded as it was found to be not relevant.

This last meeting with the case ensured that correct interpretation of the data collected was achieved. Given the generally positive feedback received, no further data collection was recommended. Some factors and variables were slightly changed (mainly their terminology as opposed to their nature) and five more links between factors were added. After this meeting, the final findings about the case could

be formulated. These were structured, written within the case report, and shared back with the CEO of the case.

4.4 Case A Findings

This section focuses on the description of the findings achieved from Case A. Three main sub-sections are dedicated to the three main levels of analysis, i.e. in relation to: external factors, perceptions, adoption sub-factors. In reporting these findings, a list of the assertions formulated for each case is provided. Each assertion is presented starting with an introduction to the factor and its supporting evidence. It is noted that, for space reasons, only a few examples of tables of evidences are provided. The remaining tables are presented in Appendix 3. The reporting continues with the description of each variable defined within the specific assertion (i.e. the specific factor). The thick descriptions generated for each variable are provided. The outcome of the final validation meeting is also integrated with particular attention to those variables and links that were initially retained and subsequently re-discussed. Those variables for which more reflections were needed (i.e. initially retained as not enough evidence was found in the data collected) are underlined in the text.

Within the perceptions' sub-section also links between factors are presented. When reporting each assertion about perceptions two types of links are highlighted: (1) the enabling external factors that emerged as drivers for the generation of these perceptions; and (2) the link between each perception with either the adoption factor or another class of perceptions (i.e. the factor in question is not directly linked with adoption but mediated by another set of perceptions).

It is noted, that references to the documents (i.e. D1, D2, etc.) reflect the documents listed in Appendix 1.1. KS are referenced in Appendix 3.1. Where quotation marks appear ("") a quote from one of the interviews undertaken is made.

Overall, the findings demonstrate that adoption of open data for commercial service innovation for Case A is influenced by a total of seven external factors and five perceptions. These emerged as being described by a total of sixty variables. Seventeen links were established between factors.

4.4.1 External Factors

From the analysis of Case A, seven external factors were extracted and validated. A total of thirty-four variables were established across these factors. Two variables were finally excluded from the analysis according to the feedback received from the case at the validation meeting.

- ***Assertion 1: Open Data Characteristics influence the adoption of open data for commercial service innovation***

From the first step of Case A's analysis, thirty-eight KS were classified within Cluster A, i.e. in relation to the characteristics of open data and open datasets. Within this cluster, six variables emerged from the data collected as drivers of open data characteristics that influence its adoption. These, together with the supporting evidence, are listed in the following table and subsequently described.

ASSERTION 1					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
1.1	*Free Online Availability*	2,2; 2,4.	D1; D3.	D6; D7; D8; D9.	Yes
1.2	*Machine Readability*	2,3.	D1; D3.	D6; D9.	Yes
1.3	*Open License*	2,5; 8,8; 8,9; 10,1; 10,2; 10,3.	D4.	D5; D9.	Yes
1.4	*Standard*	8,7; 8,8; 11,3; 11,8; 13,5; 14,5; 14,6; 15,1.	D1; D4.	D5; D6; D9.	Yes
1.5	*Real-Time Data*	5,9; 8,1; 8,2; 8,3; 8,4; 8,5; 9,4; 32,4; 32,6; 32,8; 35,2.	D1; D4.	D6; D9.	Yes
1.6	*Data Accuracy*	11,2; 11,4; 11,6; 11,7; 5,8.	D1.	D6; D9.	Yes

Table 4.5: Case A Assertion 1, open data and dataset characteristics

***1.1 Free Online Availability*:** the sources of open data for Case A are the online portals that city authorities develop as the tool to enable people to access the datasets. Open data must be re-usable free of charge. Although this requirement is the essential part of the actual definition of open data (http://okfn.org/), it was found to be a fundamental characteristic.

***1.2 Machine Readability*:** this aspect was found to be critical mainly at the time of observing the activities within the case. All the people encountered agreed on its importance. Having data available with formats that cannot be processed by technologies would constitute a "huge barrier" for its re-usage. Indeed, the data re-used would have to be loaded into the company's systems manually, resulting in a time-consuming process.

1.3 Open License: another recurring variable across all the data collected is about the license that has to be assigned to each open dataset. Research about legal aspects (i.e. about "what can be used and what can't be taken") is one of the first activities undertaken in the process of turning a raw dataset into a service. The case complies with the European Data Protection Law.

***1.4 Standard*:** as the CEO affirmed, "the dataset is so important because of the data standard". This means "normalize the data and give it a common structure so it can comply with the company's systems and operations". The standard includes names, addresses, district codes, etc. (i.e. all information related to each planning activity). As a consequence of data that is stored manually by the city authority, different formats are found within and across datasets. The amount of work

resulting from unstructured data is believed to prevent people from re-using it. Cities and users need to cooperate to define and develop such standards. “No one else is going to do it”.

1.5 Real-Time Data: this variable encompasses the need of having current data that is frequently updated (“near real-time data”). Since the scope of the case is to “provide stakeholders information (and notify them) of what is happening” (D4) in terms of urban planning activities in the city, a stagnant dataset would not be a useful source of information. According to the CEO, “near real-time data is the only way this open data ecosystem can work”.

1.6 Data Accuracy**:** open data released by the public authority needs to be real and precise. Fixing bad data is one of the first activities that Case A undertakes when tackling an open dataset. In some cases, “many inconsistencies” are found in the data, which lead to “endless data cleaning processes” (developer).

- ***Assertion 2: City and City Authorities’ Characteristics influence the adoption of open data for commercial service innovation***

Cluster B represents the ensemble of those KS that are related to the city where open data is leveraged and its related authority (i.e. the open data supplier). Ninety-nine KS were imported for the analysis from the first step. Of the ten variables initially emerged, three (i.e. those underlined below) needed to be re-discussed with the case during the final validation meeting. As a result, one of these variables was finally excluded (variable 2.9 below). Appendix 3.2 provides a summary table of those variables defined, retained and excluded using the same format of Table 4.5 above.

2.1 Open Data Strategy and Commitment**:** across all the data sources, one of the key themes that emerged is that cities need to “come on board”. Open data needs to exist and be released and therefore be available to the public. Hence, cities need to develop and maintain an open data strategy from the collection, through the storage, to the release of datasets in near real-time. As part of this strategy, cities need to promote open data as valuable data and mostly as “real data”. They need to spread awareness of what open data is, i.e. a mean for understanding and structuring knowledge about the huge amount of information that is collected and stored within the city authority (D1). The commitment of municipalities in relation to open data positively influenced the case to adopt it for commercial service creation.

2.2 Internal Law Development: transparency laws are currently hot topics among policy makers in both US and Europe. In other words, “cities should define their own laws to have public sector information published by default”. However, this concept was briefly mentioned during the SSI and was not found across other sources. Only D6 partially confirmed this aspect, with, however, no specific details. Thus, the inclusion of this variable needed to be re-discussed at the final validation meeting.

Here, the CEO confirmed its importance. Also, he provided some examples of cities that are embracing open data as an internal mandatory process ("mainly in the US").

***2.3 Organizational Structure*:** starting from the high level, an IT department should be well established within councils and IT-skilled people should work within it. The presence of an established IT team is believed to be a key to "put things in place". A lack of an IT department within the council leads to difficulties in getting the right skills to standardize the data. Moreover, organizational positions for people within the authority devoted exclusively to open data have to be established. Ultimately, the presence of a city CIO was also found to be critical by the CEO because this figure fosters internal open data initiatives (idea mainly generated from his previous experience in the public sector), and also serves as the person to be approached when targeting a new city.

***2.4 Awareness of Open Data*:** one of the main challenges faced by Case A so far was about "educating cities about what open data is, how to use it, legal information, the standards, and how to release the data". Moreover, cities should be aware of the benefits that the specific open data service can bring to both the city and the local authority. Indeed, most of the times cities do not understand how engaging with users and service developers from their data can be a "win-win" project. Also, capabilities about the releasing process should be in place within city councils.

***2.5 Budget and Resources*:** this variable was established even though its evidence came "only" from two sources of data (field notes and the SSI). However, also considering what emerged in the previous points, it is inferred that as budget and resources allocated to open data programs increase, the effectiveness of the impact of open data among service developers is expected to increase subsequently. This variable was interpreted from the fact that many of the cities that refused to release open data to Case A (and so to embrace an open data strategy) motivated their choice with lack of budget and resources reasons.

***2.6 Engagement Initiatives*:** according to the CEO "just releasing open data and wait that something happens is not enough". Cities need to promote their datasets by engaging with potential users and demonstrating them what can actually be done with open data. Examples of initiatives include "showcases", "workshops" and other forms of events opened to the public where creation of new services on top of their open datasets is stimulated and facilitated. These schemes were also found (in the case's experience) to give cities the support needed for releasing the data. "They need to do it by engaging new people". In addition, by being open for engagement and collaboration, cities learn what the users' needs are and can react accordingly.

2.7 Size of the city: Case A works with open data from several cities. Cities that are targeted are those for which a market for the service exists. In this way, small cities are unlikely to be a good place for

businesses that want to turn open data into commercial services. Therefore, bigger cities are more likely to be places where open data services are developed.

***2.8 Support Creation and Sustainability of Open Data Services*:** while "co-funding open data services projects is not doable", support by cities to those that successfully re-used their open datasets should be given by promoting these services (e.g. through posting them on the website). These types of promotion do not only enable the city to get more users involved (by demonstrating that there is actual value in its open datasets), but also allows open data services companies to approach other cities by saying "if one council is doing it, why shouldn't you do that?" Providing cash prizes at competitions is another example of early-stage support.

***2.9 Cities should use Open Source Software*:** this tentative variable emerged from only one KS. No details were recorded about why open software should be used, and which actual programs are suggested by the case. Hence, this aspect was brought back to the case because further data needed to be collected in these regards. At the final validation meeting, this aspect was unclear to all the participants. Thus, this variable was excluded from the analysis.

***2.10 Cities shouldn't outsource Open Data Service creation**:* one of the main barriers perceived by the case is about cities that "rely on big companies to do all the work". Many cities to-date outsourced "to big multinationals" the creation of the datasets in first place and of services on top of them subsequently. According to the CEO, building their own apps internally is not found as beneficial as "they should stimulate and facilitate new business creation". This aspect was briefly tackled during the interviews, and no more evidence was found to support this concept. When re-discussing it, the CEO argued that, based on his experiences, "too often the city councils rely on big companies to build apps for them". These activities are implemented to "demonstrate to the public that they are doing open data". These situations preclude the rise of new start ups or SMEs that design and develop new commercial services on top of open data.

- ***Assertion 3: Organizational Skills Factors influence the adoption of open data for commercial service innovation***

The definition given for Cluster C was intended to encompass all the KS that are related to the company's people and their skills. Sixty KS that emerged from the interviews were included in this cluster. Seven variables were initially defined from the analysis of these KS. One of these (variable 3.7, underlined in the following list) was found as not sufficiently supported by evidence and was finally excluded after the validation meeting. The remaining six variables were established and validated. These variables are related to sets of skills that were found as critical for Case A to successfully design and deliver its open data services. A table with all the variables supported by evidences from all the

sources of data was developed (see Appendix 3.3). It is noted that in Case A no physical resources were needed along the process apart from normal “offices' equipment” (e.g. desks, computers).

3.1 Geographic Information Systems Skills: strong evidence was found in Case A for the need of internal Geographic Information Systems (GIS) skills. The company is involved in taking the open datasets and mapping the various planning information. In addition, “in Europe it is very difficult to geo-locate information”, referring to the scarcity of GIS skills among city authorities.

3.2 Programming Skills**:** in order to develop the open data standard (the CEO called these abilities also “data normalization skills”) and the web and mobile applications, programming skills are needed. Half of the people within Case A are involved in developing software solutions. According to the CEO, “this is our core skill”.

3.3 Knowledge of Legal and Policy Context**:** this range of knowledge needed to be acquired by the case in order to be able to comply with legal requirements when re-using open data. It includes being able to understand the licenses assigned to the datasets and knowing the policies that are regulating public sector information re-usage. Specific knowledge about data privacy in the urban planning domain should be also part of the company and was critically leveraged by Case A.

3.4 Data Analytics Skills: data analytics skills emerged as critical across several sources of data. These skills are needed when extracting, transforming, and managing the datasets. As one developer stated, “we need data analytics to fix bad data”. This set of skills was observed to be useful for inspecting, transforming, cleaning, and modelling the data. Also, data analytics processes are undertaken to extract information from the usage of the open data service. These processes give the case the ability to detect knowledge about the urban planning domain such as “how people look for this information”, when and where they do it and so on. Among other outcomes, the company is able to provide the city council with information that allows augmenting the understanding of the urban planning domain (e.g. trends, identification of inefficiencies).

3.5 Graphic / Web Design and Data Visualization Skills**:** given that the open data services are delivered by Case A through a web application, it is inferred that web and “front end” graphic design skills become essential. Open data is visualized on the map within the web application, and data visualization skills are needed to do it.

3.6 Communication Skills**:** communication skills for the open data service business emerged to be crucial in order to: approach new cities; “sell” (i.e. convince cities about the actual value provided by the service) them its services; convince them to embrace an open data program and to allocate budget and resources; establish contracts to ensure long-term commitment; teach them what open data is and how it can be used. All in all, the case needs to convince cities to “come on board” (D2).

***3.7 Companies have to cover the entire Process*:** Case A is involved in all the activities from the extraction of the data, through the analysis and mash-up of the information, to the final delivery of a valuable service to a range of end-users. The fact that the case covers all these sets of activities emerged as important from the SSI process. However, no further evidence was found to support this idea during the final validation meeting. Therefore, this variable was excluded.

- ***Assertion 4: Cities-Businesses Collaboration Factors influence the adoption of open data for commercial service innovation***

A separate cluster related to the various interactions between city councils and Case A was defined given the importance that was assigned to these situations. Collaboration, partnership, engagement, and contracts arose as key words for the successful delivery of the open data services by Case A. The need of working together with city councils and to establish mechanisms for optimizing these relationships was emphasized by all people within the company. Furthermore, the ability of Case A to create, establish, and to maintain these relationships is considered as the cornerstone for perceiving open data as useful in the long-term. The nature of this collaboration is described through three variables that were derived from the data collected. Supporting evidence for this assertion is provided in Table 4.6.

ASSERTION 4						
Variables		**Key Statements**	**Int. Doc.**	**Ext. Doc.**	**Field Notes**	**Valid**
4.1	*Partnership*	3,1; 4,3; 4,4; 7,5; 7,8; 8,1; 8,6; 9,3; 13,12; 14,5; 15,3; 16,8; 20,1; 20,2; 20,3; 20,4; 20,5; 20,6; 20,7; 20,9; 20,10; 20,12; 20,13; 21,1; 21,6; 21,8; 24,2; 24,3; 24,4; 26,1; 27,1; 28,6; 29,4; 31,6; 33,4; 34,2;	D1; D2; D3;	D7; D8;	Case A operates as a partner of the city. It establishes "win-win relationships".	Yes
4.2	*Commitment*	14,11; 34,6; 34,10; 34,11; 34,12;	--	D5; D8;	Concept confirmed also by the business developer.	Yes
4.3	*Contracts*	21,2; 32,1; 32,2; 32,3; 32,5; 32,7; 33,1; 33,2; 34,2; 35,4;	--	--	A further meeting was scheduled for investigating in-depth this notion.	Yes

Table 4.6: Case A Assertion 4, city-business collaboration factors

***4.1 Partnership*:** one of the key factors that emerged throughout the case study is that Case A works very closely with the city authority that owns the open data. The case's ability to engage with cities and to get them involved in the process emerged to be a crucial aspect. The key point is the "win-win" type of these relationships. Indeed, "a general belief that engaging in partnerships increases the internal workload" for the councils is in place. However, regarding Case A, its services improve

substantially internal urban planning processes and reduce significantly the amount of work the city is in charge of. Cooperating with cities is found as critical for several reasons: develop the standard, educate the cities, and ensure long-term commitment.

***4.2 Commitment*:** this variable is meant to emphasise the effort that is requested to the city councils towards ensuring a long-term provision of accurate and standardized "near real-time" open data. The focus has to be about showing what the company does and demonstrating that "there would be a win-win relationship". In addition, long-term commitment's impact goes beyond the single business-city interaction. Indeed, commitment from one city is leveraged for targeting additional city councils (e.g. "if another one is doing it, why shouldn't you do it?").

***4.3 Contracts*:** commercial contracts are sometimes stipulated between Case A and the city authority it is working with ("mainly in the US") as a further assurance of the long-term commitment of the council in the data provision. The city council in exchange for the service that the case is providing to the city, commits to release open data on a daily basis and with a certain degree of quality and accuracy levels. The stipulation of contracts enables overcoming many of the barriers that exist to put in place an open data services-oriented business – according to the CEO.

- ***Cluster E: Other External Factors***

The last external factors' cluster encompassed all the remaining KS. Differently from the previous clusters, in this case there was not a common underlying aspect across these KS. After several attempts of grouping together these fifty-seven KS, three factors emerged in relation to: (1) *External Inspiring Factors*, (2) *Individual Factors*, and (3) *Community Factors*. These are tackled in the following paragraphs. The evidences supporting each of these three factors are represented in Table 4.7 with the same format used so far.

Cluster E1, External Inspiring Factors: one of the aspects that led Case A to leverage the use of open data was the nature of the context of urban planning. According to the CEO, "in far too many cities, if you want to know about construction plans and other real estate development data, you'll need to prepare for a dusty slog through paper documents or a battle with arcane city-records tech". Hence, maturity of digital solutions in this context, and opportunities to gain efficiency and efficacy are aspects that significantly influenced the development of the open data services business, and so the adoption of open data for commercial service innovation purposes. In summary, the name *External Inspiring Factors* was assigned to these conditions of the context and includes three main variables; these are: ***Maturity of Digital Solutions (5.1); Opportunity to Gain Efficiency (5.2);*** and ***Opportunity to Gain Efficacy (5.3).***

OTHER EXTERNAL FACTORS					
Factors	**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Field Notes**	**Validated**
External Inspiring Factors	1,1; 1,5; 1,7; 7,14;	D1; D2; D3;	D5; D7; D8;	The idea that the ODS offered substantially improves the current practices was common among all people.	Yes
Individual Factors	6,1; 6,2; 6,3;	D1; D4;	D7; D9;	When having informal conversations on site, it emerged several times that the CEO/founder of Case A and his skills and experiences were the reasons why "it exists and it works".	Yes
Community Factors	13,13; 14,8; 14,9; 14,12; 14,13; 15,1; 15,3; 16,8; 21,3; 21,5; 21,6; 21,7; 22,9; 24,1; 24,3; 24,4; 28,3; 29,1; 31,5; 34,1; 34,9; 34,10; 34,11; 34,13; 34,14; 34,15; 34,16; 35,8; 36,1; 36,2; 36,3; 36,5; 36,7; 36,8;	D1; D3; D4;	D8;	The importance of champions has been underlined by most of the people within at the case at the question "what are in your opinion other factors that would facilitate new service development from Open Datasets.	yes

Table 4.7: Case A Other External Factors

For example, if the process of visualizing data and having information about urban planning activities in cities had been lean, fast, efficient, and effective, probably Case A wouldn't have been created. An important source supporting the definition of variables 5.2 and 5.3 is given by D1: Case A provides "an excellent tool [that] is rapidly expanding due to the added efficiency it brings to local authorities by reducing counter-time for their staff" (D1). In relation to variable 5.3, the opportunity of augmenting efficacy in the urban planning context was leveraged by the case to establish partnerships (and sometimes contracts) with the city councils. In this way, Case A is able to "guide planning decisions" and "proper trend analysis" to, for example, "see what effect a new supermarket will have on the requirement for other new construction in a given area" (D1). Therefore, the following assertion was formulated:

Assertion 5: External Inspiring Factors influence the adoption of open data for commercial service innovation.

Cluster E2, Individual Factors: in Case A's experience, the availability of open data followed the generation of the business idea. The CEO founded the company after working in the field of urban planning within governments for several years. "Then open data happened", he stated. A deep

reflection on this, led me to consider individual factors as a potential contributor for the adoption of open data. Thus, a tentative assertion was formulated around these concepts and later confirmed at the validation stage:

Assertion 6: Individual Factors influence the adoption of open data for commercial service innovation.

Within this factor, a number of variables were identified and confirmed. These are: ***Attitude to Work with Governments (6.1); Attitude toward Open Data (6.2); Experiences in the Field (6.3);*** and ***Entrepreneurial Spirit (6.4).*** Gaining specific knowledge about the context was the starting point from which the CEO generated the initial business idea. Knowing specifically how urban planning information is collected, stored, requested, analysed, and used was acknowledged by all people within Case A as one of the main cornerstones for its success. This knowledge is leveraged in all stages of interactions with the cities. Firstly, cities are approached and the "right language is needed to get them on board". Secondly, having an *Attitude to Work with Governments "*is essential if you want to educate them about open data" (e.g. about which data to release and how to do it, as well as about how to tackle internal resistances to change). Thirdly, having *Experiences in the Field* enabled the CEO to know exactly "what was working well and what wasn't" and which were the needs of the market and the opportunities that could be exploited. Finally, the experiences and commitments of the CEO outside the case to foster the open data ecosystem and his *Entrepreneurial Spirit* (emphasized in D9) were also found as additional variables that contributed to the successful and sustainable adoption of open data by Case A.

Cluster E3, Community-Oriented Factors: as the last class of variables, thirty-three KS led to consider a range of other entities and initiatives, that are external to both the company and the city council, that contribute to the adoption of open data for designing new commercial services. These include: ***Hack Competitions (7.1); Champions (7.2);*** and ***Community-Oriented initiatives (7.3).*** In summary, variable 7.3 includes initiatives such as "hackathons, coder dojo, and others" with the objective of engaging with the public and spreading awareness. Other goals are stimulating re-usage of open data and demonstrating the value that can be achieved from raw open datasets. In the CEO's opinion, these initiatives could create an ecosystem within which standards and demonstrations can be implemented. Evidence was also found concerning the role that *Champions* (variable 7.2) can play in the open data services ecosystem. Spreading awareness and demonstrations across both city authorities and the general public are believed to be the main goals for these subjects. According to the CEO, champions have the power of getting cities involved and data subsequently released. In relation to *Hack Competitions* (7.1), this experience played a key role for developing the overall

business. Although a general idea was previously generated, it was at the time of participating in the local hack competition that open data was adopted as a source for developing the solution. By winning this competition, Case A could also leverage the prize for the initial launch and developments. In the CEO's opinion, these programs shouldn't involve just potential re-users and governments, but also people working in the domains where the open datasets belong to in order to "help re-users to enter their contexts".

Therefore, the following assertion was formulated encompassing these three variables:

Assertion 7: Community-Oriented Factors influence the adoption of open data for commercial service innovation.

In total, seven assertions were formulated in relation to seven external factors that emerged as influencing Case A in adopting open data for commercial service innovation. A total of thirty-four variables was extracted, defined, validated, and established across these factors.

4.4.2 Perceptions

With respect to the analysis about perceptions, starting from the four initial clusters emerged (see sub-section 4.3.2), four tentative factors described by a total of fifteen tentative variables were defined and subsequently validated. Furthermore, one additional factor was added to the final findings from the validation stage interaction. A total of thirteen links emerged in relation to these perceptions.

- ***Assertion 8: Perceived Usefulness of Open Data influences the adoption of open data for commercial service innovation***

The first perceptions' cluster encompasses sixty-one KS related to perceptions about the usefulness of open data for commercial service innovation. Three variables supported by strong evidence emerged in relation to the perceived value of open data for: providing the content; enabling short time development of a new service in the context; and being a source of reliable near real-time information. In addition, a further variable arose. However, since strong evidence was not initially found, this needed further reflections and discussions with the case's people before being finally integrated in the case's findings (variable 8.4). Table 4.8 provides an overview of these variables and the related supporting evidence.

ASSERTION 8					
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Validated**
8.1	*Provision of Content*	2,2; 2,4; 2,6; 5,5; 5,8; 5,12; 10,6; 11,1; 11,3; 11,5; 13,4; 13,5; 14,1; 14,9; 22,11; 22,14; 34,10; 36,11.	D1; D2; D3; D4.	D5; D6; D7.	Yes
8.2	*Time Saving / Short Time development*	1,3; 1,4; 11,2; 11,3; 14,4; 14,6; 16,3.	D1.	D9.	Yes
8.3	*Enabling Dynamic Reliable Information*	2,3; 2,5; 5,9; 8,1; 8,2; 8,3; 8,4; 8,5; 8,7; 8,8; 8,9; 8,11; 9,4; 10,1; 10,2; 10,5; 11,1; 11,2; 11,4; 11,6; 11,7; 11,8; 14,1; 14,4; 14,6; 16,3; 20,2; 20,3; 20,4; 20,5; 20,6; 21,2; 32,1; 32,2; 32,3; 32,4; 32,8; 33,1; 34,3.	D1.	D5; D6; D7; D9.	Yes
8.4	*"Opens the Council's doors"*	8,12; 13,3; 16,8.	--	--	Yes

Table 4.8: Case A Assertion 8, perceived usefulness of open data

8.1 Provision of Content: almost all documents and several KS highlight the fact that open data is perceived useful mainly because it provides content. The service aims at visualizing structured urban planning information and open data is actually this information, i.e. the content of the service.

8.2 Short Time Development / Time Saving*:* open data (if consistent with Assertion 1) is found to enable significant time saving. This includes avoiding the implementation of the entire range of activities that would be needed to manually gather and store this information. Thus, having structured datasets made available makes it "easier for citizens to understand it and for companies and developers to create interesting things" (D5).

8.3 Enabling Dynamic Reliable Information: Case A implements efforts and activities towards developing a standard because this data is perceived "reliable and a dynamic source of valuable information" for delivering the final service. Trust in the data being delivered in the future is also fundamental in order for this source to be sustainably re-used. Case A ensures long-term availability of dynamic (i.e. "near real-time") open data with contracts in some cities, while others are selected in relation to the level of maturity of their open data programs.

8.4 "Opens the Council's Doors": the last variable that emerged was initially supported by only three KS and no evidence was found across other sources. In particular, the CEO of the case pointed the attention on how Case A is leveraging the fact of actually working with open data to facilitate the establishment of relationships with the various city councils. Given that cities want to showcase the value of their open data programs, being a case of open data re-usage is "extremely well-welcomed" by local governments. When re-discussing this aspect at the final validation meeting, the CEO stressed the importance of mentioning that the company was working on open data "for getting the councils' attentions". This is because "everybody [i.e. the councils] wants to say we are doing open data" and

at this “early stage of development” all the councils welcome new re-users. This enables councils to promote their open data programs with actual successful examples of SMEs being created from there. This is believed (by councils) to improve the city’s image as a “transparent and smart city”.

Assertion 8 links

With respect to the links from and to this set of perceptions, the data collected across all sources initially suggested that only *Open Data Characteristics* were directly linked with this factor. Although this link was confirmed at the validation meeting, three more causal links from other external factors emerged from this discussion. In particular, both *Individual Factors* and *Community-Oriented Factors* were found to be connected to *Perceived Usefulness of Open Data*. These links are defined based on personal awareness of open data and its potential (the former) and collective awareness achieved among the public involved in the community-oriented initiatives, trainings, and hack competitions (the latter). In other words, deep awareness of open data is believed to be needed in order to perceive it as useful for commercial service innovation purposes. A further link emerged with *Organizational Skills*. In this way, the developers acknowledged that significant technical skills need to be in place in order to access, understand, and, more generally, re-use open data. This new link has been subsequently established. On the other hand, all people agreed on the direct link between this factor and the adoption of open data. Figure 4.2 provides a summary of these links.

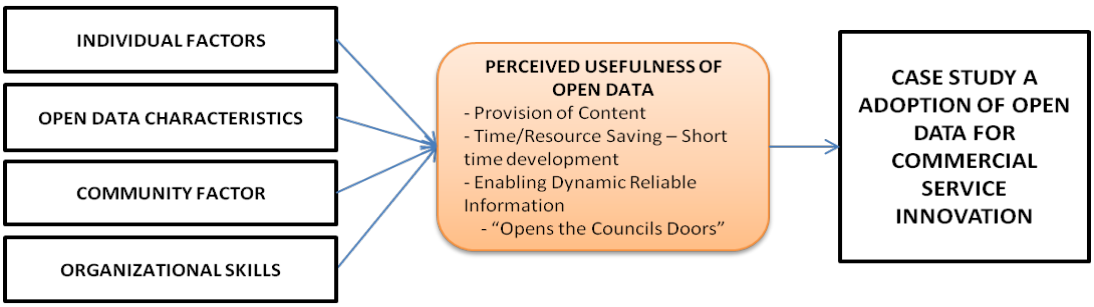

Figure 4.2: Case A Assertion 8 links

- ***Assertion 9: Perceived Business Opportunity influences the adoption of open data for commercial service innovation***

Regarding Cluster G, a number of variables emerged from the thirty-five KS categorised within it. Strong evidence was found for five of those. In general, Cluster G included concepts about perceived business opportunities that significantly influenced the case when the decision of adopting open data was made. The rationale behind this finding is clear: open data is adopted for commercial service innovation if business opportunities and business models can be thought and implemented from this

data. The five variables that are included in this set of perceptions and the supporting evidences are presented in Appendix 3.4.

***9.1 Benefits for End-Users*:** the perceived business opportunity is mainly generated from the development of a valuable service. This is understood as a service that provides benefits for the end users. Benefits include users empowered with more informed decisions about urban planning. With Case A's open data service: the time to get urban planning information is almost eliminated; the quality of this information is significantly augmented; and all the intermediaries that worked at planning activities across different departments of the city council are skipped, and resources are saved subsequently.

***9.2 Partnership Strength*:** the level of partnership that is established with the public authority is found as another fundamental variable that can increase the perceived business opportunity (especially where commercial contracts are in place). Also, the more the company can be engaged with the public authority, the more the chances of establishing a sustainable open data services' business increase.

9.3 Level of Engagement: as level of engagement increases, so does the perceived business opportunity. As opposed to the previous variable (which focuses on particular relationships with the city authority), here engagement is related to the overall open data services ecosystem. Thus, the CEO believes that participating in initiatives such as hack competitions, trainings, showcases, and other community-oriented initiatives increases the perceived opportunity of creating a new business based on open data. According to the CEO, this process is made of two fundamental steps: (1) gain awareness of what open data is, and (2) learn from the experiences of the people and companies that have already undertaken the journey to providing a profitable service based on this source.

9.4 Scalability: Case A's strategy for enabling business growth is to expand the service across cities. According to the CEO, "once the service is designed and delivered in the first city, its nature and the standard are shown to others to convince them to come on board". The key for enabling a scalable solution relies both on the data and on the company's systems and operations. The latter are designed as an "IT architecture" upon which the data is loaded and visualized and the key for achieving scalability is to have data that is compatible with these systems and processes. Thus, scalability is found critical in perceiving opportunities to expand the service across cities, i.e. to enable growth, i.e. to increase perceptions about business opportunities from open data re-usage.

***9.5 Demonstrability*:** the last variable associated with perceived business opportunity is related to the ability of demonstrating the validity and the value of the proposed open data service. Case A, at the time of approaching a new city, starts with expressions like "here it is what other cities are doing", or "here is the prize we won in the hack competition". In this way, according to the CEO, demonstrations are critical to get new "cities on board".

Assertion 9 links

As stated in the descriptions of variables 9.2 and 9.3, *Individual* and *Community-Oriented Factors*, and the presence of a strong collaboration and partnership were derived from the data collected as contributing to the creation of this set of perceptions. In addition, *External Inspiring Factors* are connected with these perceptions. Concerning the latter, if the level of maturity of solutions in the urban planning context is low, and opportunities to gain efficiency and efficacy in these processes are in place, perceptions about business opportunities are generated. Furthermore, the only link proposed that was found to be not consistent during the validation meeting was the one between *Perceived Business Opportunity* and the adoption factor. This choice was motivated by the fact that a perceived business opportunity in the context of urban planning data does not imply that open data is adopted for service innovation. Solutions might be not directly related to the open data ecosystem – one of the developers observed. Alternatively, once the business opportunity is perceived, perceptions about the usefulness of open data to exploit this opportunity are formed.

A further link emerged during the validation stage. In particular, *City-Related Factors* were found to be linked with this set of perceptions. The initial activities made by the case when targeting a new city are to look at its size, the subsequent size of its market, and the level of maturity of open data within the related council. Thus, characteristics of the cities were fundamental parameters that led to the consciousness that a sustainable business based on open data could be developed in that specific urban environment. The following figure summarizes the links established for this set of perceptions.

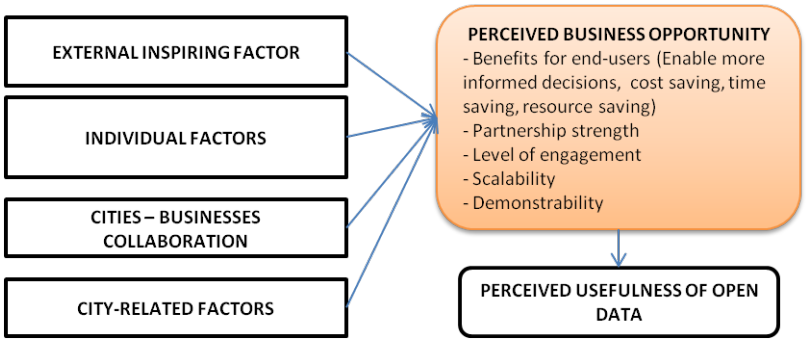

Figure 4.3: Case A Assertion 9 links

- ***Assertion 10: Perceived City Opportunity influences the adoption of open data for commercial service innovation***

The definition of Cluster H was supported by fifty-seven KS that emerged from the data collected. These were related to potential city opportunities. In total, three variables were found to be supported by strong evidence across all other sources of data collected and were subsequently validated (details are provided in Table 4.9). Perceiving opportunity for the city was found to be a key for Case A to

perceive open data as useful for commercial service innovation. According to the data collected, if an opportunity for the city is believed to be in place, perceptions about the commitment of the city council in sustaining the provision of open data are generated. In other words, the data collected suggests that in a hypothetical situation in which cities did not perceive these opportunities, they would not be stimulated in implementing efforts for releasing open data with the characteristics outlined above and on a sustainable basis.

ASSERTION 10						
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Field Notes**	**Validated**
10.1	*Improve City Image*	5,13; 7,19; 9,1; 20,12; 20,13; 21,7; 23,3; 24,3; 24,4; 28,1; 30,2; 30,6; 33,3;	--	D8;	Thought shared by the CEO and the lead Business Developer.	Yes
10.2	*Economic Growth*	5,13; 5,14; 7,15; 7,19; 22,8; 22,9; 22,11; 28,1; 30,6; 33,3;	--	D8;	Thought shared by the CEO and the lead Business Developer.	Yes
10.3	*Internal Efficiency and Efficacy Gain*	7,15; 7,16;7,17; 13,9; 13,10; 13,11; 22,5; 30,6; 33,2;	D1; D2; D3;	D5; D7; D8;	All people underlined the benefits that the service brings to the city council's operations.	Yes

Table 4.9: Case A Assertion 10, perceived city opportunity

***10.1 Improve City Image*:** as the CEO underlined several times, "it is all about transparency" (D8). Thus, opportunities for cities to embrace the open data services ecosystem include improving their public image by promoting transparency of governmental activities and citizens' engagement.

10.2 Economic Growth: another opportunity that might emerge for the cities that embrace open data and undertake an open data program is the rise of start-ups and SMEs (such as Case A). These are expected to actually exploit open data and create value on top of services designed from it. A further confirmation of the importance of this variable came from the CEO who stated: "if you (i.e. the city council) outsourced the creation of [open data] services to big companies (...) why would they do this?" This statement was related to the potential missed opportunity (by governments) of fostering local economy growth.

***10.3 Internal Efficiency and Efficacy Gain*:** one of the most visible opportunities that the city had from the involvement with Case A was a gain of efficiency and efficacy in internal context-related operations. In particular, the solution that the open data services introduced regarding urban planning information clearly led the cities to save time, resources, and costs and increase the overall efficiency

in this domain. In addition, the outcomes achieved from the data analytics stages (e.g. trend analysis) "enable local authorities to increase efficacy" (D1).

Assertion 10 links

In total, three causal links were identified for this factor. The main enabler for these perceptions is found to be the presence of a relationship between the open data services business and the city authority. Indeed, establishing a "win-win relationship" with the council entails delivering a service that is beneficial for internal efficiency and efficacy improvements in the context. Furthermore, establishing relationships with the business was found as a way for the councils involved to promote their open data programs, and therefore (following the interpretation of the CEO) to achieve a better image as a "transparent and smart city". The second link is related to *Community-Oriented Factors* enabling this set of perceptions. This connection was established and validated based on the rationale behind these initiatives. Indeed, these programs are seen as valuable for the possibility that actual and potential re-users have for "educating the council". In addition, the characteristics of the open data supplier also influence the perceived opportunity by the same entity. In other words, even though evidence did not directly emerge from data collected, it is inevitable to think that city councils' variables (such as budget, resources and positions allocated to open data etc., see Assertion 2) influence this set of perceptions. This link was proposed at the final validation meeting and, after a short discussion, its inclusion in the final findings was agreed by the case's people.

On the other hand, the fact the cities perceive an opportunity leads the case to perceive open data as a useful source for service innovation. Following the reasoning provided by the CEO, the fact that cities perceive these opportunities lead the case to believe that an effort from the council will be in place to sustainably release open data in the future. This, in turn, was found as a fundamental condition to perceive open data as useful for commercial service development. As done for the previous clusters, a summary picture of this factor, its variables and the links is here provided.

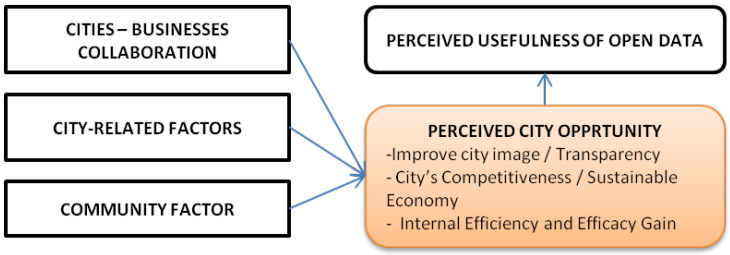

Figure 4.4: Case A Assertion 10 links

- ***Cluster I: Other Perceptions***

Unlike its respective cluster for external factors, Cluster I included only seven KS that arose from the SSI processes. These are: 1,3; 1,4; 1,6; 7,4; 7,9; 7,10; and 36,4 (see Appendix 3.1). Apart from the latter that underlines the importance of voluntary use of open data, the other six KS focus on technical characteristics of the service. In particular, three main concepts could be interpreted from these statements. The first revolves around the lean nature of the services offered. The second stresses how the open data service enables the users to get the information needed. Thirdly, the service is perceived as "easy-to-use". All these perceptions about users of the service found considerable evidence across the different sources of data. Details are shown in the following table.

ASSERTION 11					
Variables		**KS**	**Int. Docs.**	**Ext. Docs.**	**Validated**
11.1	*Lean Service*	1,3; 1,4; 7,9.	D1; D4.	D5; D6; D7.	Yes
11.2	*Being Able to Get Information*	1,3; 1,4; 1,6; 7,4; 7,9.	D1; D2; D3; D4.	D5; D6; D7.	Yes
11.3	*Easy to Use*	1,3; 1,4; 7,10.	D1.	--	Yes

Table 4.10: Case A, Cluster I

As shown in Table 4.7, three new variables emerged from the data: ***Lean Service (11.1); Being Able to Get Information (11.2);*** and ***Easy to Use (11.3).*** The common denominator is interpreted to be a set of perceptions about the user-friendly feature of the open data services offered perceived by the service's users. These three key features of the service emerged as variables that positively influence the sustainable adoption of open data for commercial service innovation. As a consequence, sufficient evidence was found to formulate the following assertion:

Assertion 11: Perceived User-Friendly of the open data service influences the adoption of open data for commercial service innovation.

Assertion 11 links

Regarding the links to this factor, *Organizational Skills* are found to be the natural cause of achieving a perceived user-friendly open data service. On the other hand, if the chances of designing a service enabling these perceptions are in place, open data is considered as useful for new commercial service development.

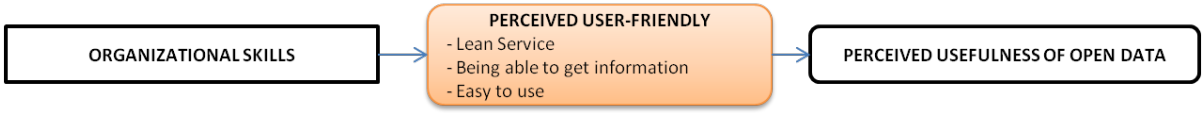

Figure 4.5: Case A Assertion 11 links

- ***Additional Factor: Perceived Voluntariness***

At the time of the validation meeting, once all TAs were reviewed, another discussion started from a question asked by the CEO. The absence of the "volunteering component" in the overall model was highlighted. Particular importance was given by the CEO to aspects about voluntary use of open data. He argued that beside all these factors, someone can just take the initiative and develop a commercial open data service voluntarily from an open dataset available online without any contract or partnership. From this reflection a further assertion was formulated. In relation to this assertion, and in accordance with the case's people, a further factor named *Perceived Voluntariness* was added to the final findings.

Assertion 12: Voluntary Use of Open data influences the adoption of open data for commercial service innovation.

Concluding, as shown in Figure 4.6 below, no specific links emerged between external factors and this set of perceptions. Following the reasoning provided, this was found to be directly linked to the adoption factor.

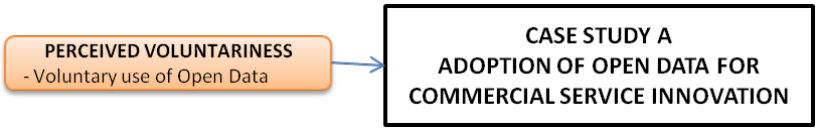

Figure 4.6: Case A Assertion 12 links

Assertion 12 represented the last perception factor identified for Case A. In this sub-section, a total of five perceptions factors, sixteen variables, and seventeen links that emerged from the analysis were presented.

4.4.3 Adoption Stages

The last and final category analysed was about the actual adoption stages the company went through from the initial generation of the idea to achieving sustained use of open data for its services' delivery. As reported in the analysis section (sub-section 4.3.3), evidence for this cluster came mainly from one of the face-to-face meetings that took place at the case's site. In addition, sixty-three KS that arose from the SSI processes were classified as describing the adoption process. Finally, further evidence was sought and found in all the documents collected and analysed. As a result of the analysis, six sub-factors emerged and were established as the different steps Case A went through. Appendix 3.8 provides supporting evidence for each of them.

ADOPTION STAGES					
Stage		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Field Notes**
1	*Idea Generation*	6,4; 7,7; 7,13; 7,18;	D1;	D9;	All stages were derived and discussed proactively with the case's CEO.
2	*Adoption*	2,6; 6,4; 7,1; 7,2; 7,3; 7,7; 7,11; 7,13; 10,1; 11,2; 12,1; 13,1; 14,5; 15,4; 16,4; 28,2;	D3;	--	
3	*Pilot Implementation*	4,2; 7,2; 7,12; 8,10; 11,1; 11,9; 11,10; 12,1; 13,1; 13,6; 13,7; 16,4; 16,5; 16,6; 16,7; 20,9; 20,11; 22,9; 23,4; 31,6; 34,4;	--	--	
4	*Commercialization*	7,18; 14,2; 16,8;	D1;	D6;	
5	*Growth*	1,2; 3,4; 4,2; 8,10; 9,2; 13,8; 14,5; 14,7; 15,3; 16,8; 16,9; 20,1; 20,2; 20,3; 20,4; 20,5; 20,6; 20,7; 20,8; 20,11; 22,5; 23,5; 28,2; 28,3; 28,4; 28,5; 28,6; 31,6; 34,5;	D1; D2;	D5; D6;	
6	*Sustained Use*	1,2; 3,3; 3,4; 3,9; 13,2; 14,7; 14,10; 15,3; 16,10; 16,11; 16,12; 20,1; 20,2; 20,3; 20,4; 20,5; 20,6; 22,14; 25,6; 34,7; 34,8;	D1; D4;	D5;	

Table 4.11: Case A, Adoption Stages

Stage 1, Idea Generation: as the first stage, the initial idea of designing a solution was generated for the urban planning context in the city where the current CEO was living and working. His experience in this domain and his previous works within the local authority, led the subject to identify several inefficiencies in how the process of getting urban planning information was conducted. Hence, the first impression was about conceiving a business opportunity related to the provision of a solution that was able to provide value to both the council as well as to potential end-users (i.e. citizens and construction companies). "Open data did not exist" at this stage. Therefore, the main challenge for the subject was about getting the data to develop such an application. What was found was a situation in which different sections of the council where responsible for different activities in the process of getting urban planning information. For example, one department was responsible for the management of planning applications (i.e. "submission, check documentation, review of the city masterplan, and review of the application", D1), another for "issue planning clarifications" (D1), and a further one for "issue planning permit" (D1). After identifying these various departments, the CEO asked for this data and "educated the council on the data that was needed to design the service".

Stage 2, Adoption*:* in reporting how awareness of open data was actually achieved and how it was subsequently adopted, the CEO repeated a few times "well, open data happened". The greater attention that the council was giving towards this initiative stimulated the CEO to push the authority to release integrated data in the urban planning domain. The first step followed was to explore legal constraints in re-using this data. Secondly, the company wanted to ensure that "near real-time data would have been made available in the long-term". During this period the CEO took part in the local

hack competition (2012). At that time, he entered the competition knowing which data he was going to re-use. The result was the first basic version of the application developed. Furthermore, winning the competition was an additional push to develop the service further.

Stage 3, Pilot Implementation: with a basic application developed at the hack competition, and with the prize won, the CEO officially launched the company. The open data service was expanded in order to cover the overall city. Working together with the council enabled Case A to have as a result data publicly available and consistent with the standard proposed. The resulting service was delivered for the first city. Different tests and subsequent improvements were undertaken.

Stage 4, Commercialization: after the pilot was implemented and tested, the company started to conceptualize its current revenue models. In particular, the solution, once sustainable provision of data was ensured, "was ready for the market". The first step was about identifying the potential customers. According to the documents collected, the case's "customers are banks, retailers, telco's and those involved in the construction industry" (D6). As a result, the "freemium" and the "premium" revenue models were established for the service's features *Get Notified* and *Premium Listing* respectively (see section 4.1).

Stage 5, Growth: the case's growth process consists on expanding the service across additional cities. According to the business developer, "once done it for the first city, the service and the data standard are shown to others to convince them to come on board". In other words, new cities are targeted and approached and the service is proposed. Assertion 2 covers the specific aspects that cities should have to be considered as a potential target by Case A. Once targeted, these cities are approached and "educated" in terms of the data and standard needed. The first topic covered is about convincing the council that an effort in publishing the data would be paid back in terms of significant improvement of internal processes in the context. Then, the original system is scaled to the new urban environment. "Today the service is delivered in ten major cities worldwide" (D4).

Stage 6, Sustained Use: ultimately, Case A achieved a stage where it is sustainably using open data for delivering a commercial service across different cities. In this way it is critical to ensure the council's long-term provision of "near real-time data". In some cities this process is supported by contracts (especially in US). In Europe, by leveraging the "win-win relationship" concept (see Assertion 4), the service became integral part of the councils' operations. Continuous improvements of the general usability and design of the service are being undertaken by the case.

4.5 Case Study A Conclusions: Adoption Factors Variables and Links

Concluding, from Case Study A, twelve assertions were formulated in relation to twelve different factors influencing adoption of open data for commercial service innovation in cities. Each of these

factors is described by a set of variables (sixty in total). Seventeen links between factors were derived. The resulting adoption model for Case A is depicted in Figure 4.8 (next page).

In addition, six main stages emerged to be critical in describing the process through which Case A went from the initial generation of the idea for the open data service, to achieving sustained use of open data for the delivery of the services. These stages are represented in the following figure.

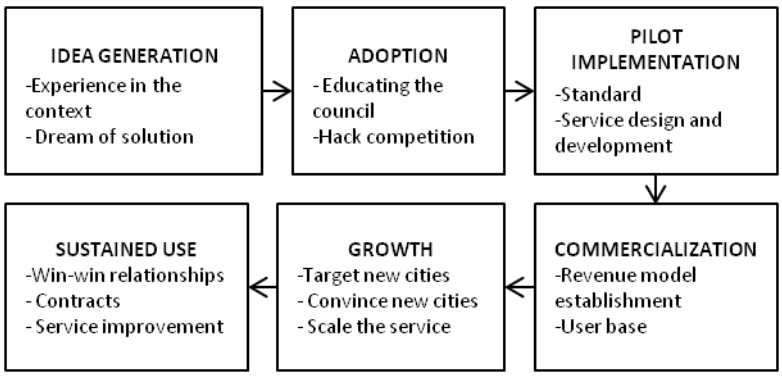

Figure 4.7: Case A adoption stages

These results concluded the first case study. Ultimately, it is noted that the evidence that supported the definition of each factor/variable/link was stored and track has been kept in the "data reduction" process. By doing so, potential issues about the "situationality of the *quintain*" (Stake, 2006) were addressed when incorporating results from the other cases.

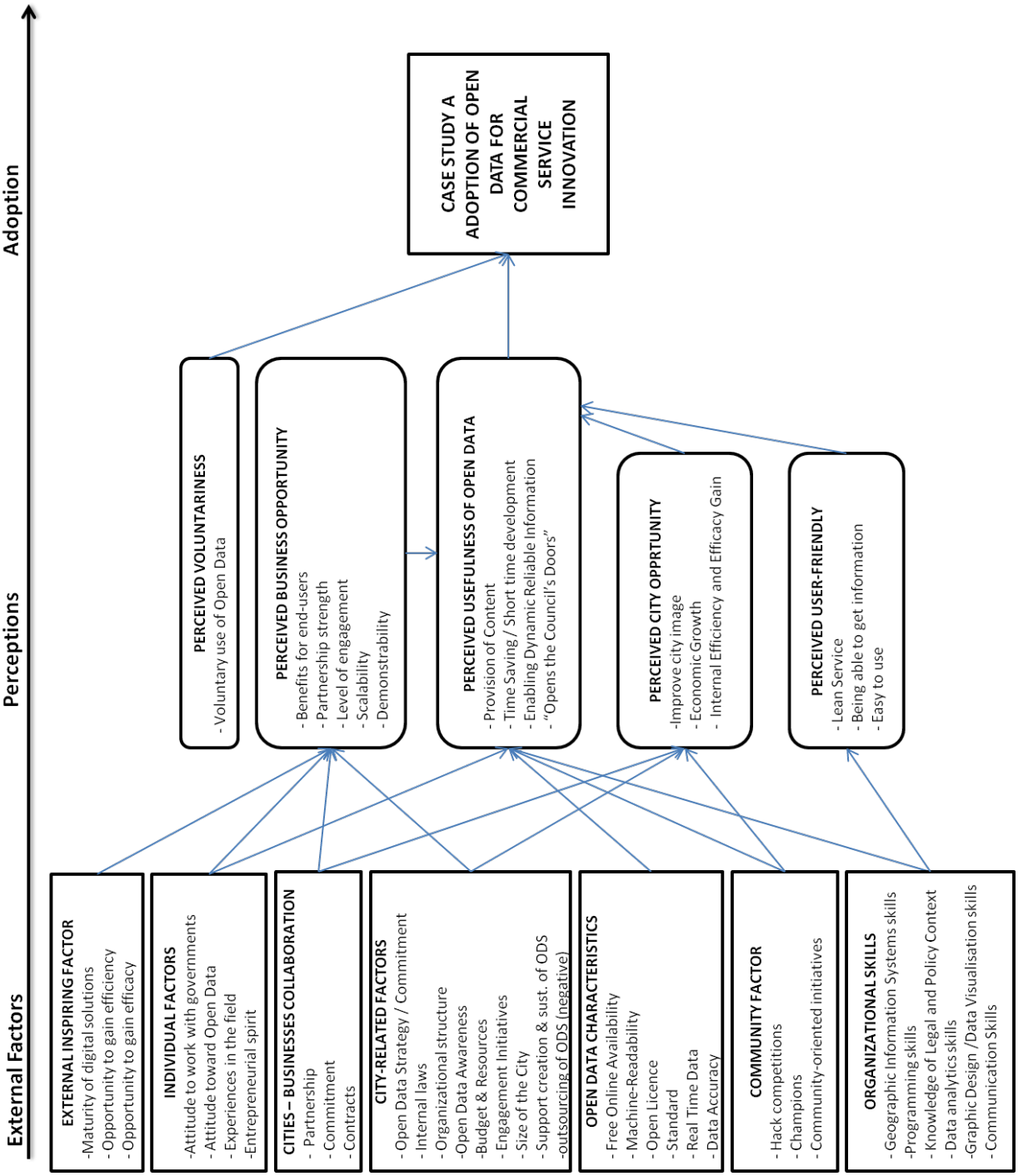

Figure 4.8: Case A adoption model

5 CASE STUDY B REPORT

This chapter is dedicated to the description of the activities and the findings related to Case B. In reporting this case study, the same structure used for Case A is adopted (see Chapter 4). In particular, after introducing the case and its open data services (section 5.1), the methodological aspects customized to this case are described (section 5.2). Subsequently, the analysis related to Case B is presented (section 5.3). Similar to the previous chapter, the fourth section is dedicated to the presentation of the findings achieved from this case study. In this section, thick descriptions are provided about each factor, variable, and link that emerged as describing open data adoption by Case B. The chapter concludes with a summary of the findings achieved from this case. This includes the formalization of Case B's open data adoption model, and the outline of its adoption stages.

5.1 Introduction to Case B

Case B represents the second case study that was carried out as part of this research. Case B integrates open data, social networking and gamification principles in order to provide tourists with a digitally interactive experience with the locals and the cultural attractions in the area.

Case B's history can be described through three main stages. Firstly, at a volunteer level, an idea emerged with the goal of improving and creating potential tourism opportunities in the local city. This area, despite the fact that several cultural attractions are in place, was at the time underexploited by the local authority and community. The idea emerged in late 2011 from a person that was previously working in this industry at a nearby location. The main aspects that a solution to these problems should have included were incentivizing both locals and tourists to visit the main attractions in the area and, as a consequence of this, improving and fostering the local economy. Furthermore, in order to achieve these goals, it was believed that the services offered should have enabled an experience that had to be challenging, interesting, and amusing. In other words, these three components were considered critical in order to engage and incentivize people to visit the city and undertake this cultural journey. With the aim of addressing these problems and achieving these goals, a mobile application was created. The main enabler for this solution was open data offered by the related council.

Case B entered the second phase of its history when it took part in a one-day hack competition organized in conjunction with the local council. Its main goal was to extract value from their open dataset as well as to demonstrate the effort that was being done towards "transparency, collaboration, participation, and fostering economic opportunities" (Local Council Report, 2012). The current CEO of Case B won the overall competition. From an early stage idea and pilot, a "minimum viable product" (CEO) was created. To the previous prototype, several additional aspects and concepts

were integrated. From a pure focus on cultural attractions, Case B's service became an integrated tool "from which all tourism/cultural sites, businesses, attractions, tours events and deals in [the area] can be marketed and updated quickly, easily and free of charge" (Company Document). Gamification principles, in which "a series of levels, rewards, and challenges [were] designed to engage and stimulate visitors" (Company Document), were added to the service. Technical functionalities were enhanced as well.

The third and current stage of the case's history identifies a significant improvement and enrichment of the service. The current application aims at improving and augmenting the visitors experience from their arrival to their departure. To do that, data about public transport, events data, "landmarks data, weather data" (among other datasets) were integrated and added to the previous configuration. Technical changes towards a much "sleeker" design and a better operability were implemented as well. Also, the CEO started to conceptualize and deploy a revenue model for this open data service.

Currently, three people are working within the company: the CEO, who is mainly responsible for project management activities, software development, and graphic design; and two more volunteers whose role is to support the technical development of the application. Due to the scalability of the solution, the company is currently expanding the service across those cities in Europe where the data needed is made available by the related local authority as open data. At the moment of conducting the case study, Case B was operating in Ireland and Germany.

The open data services offered are enabled by the following open datasets:

- Top cultural attractions: this includes information about culture and entertainment, museums and galleries, tours, zoos and other attractions. Both localizations and descriptions of the main cultural experiences in the area are provided.
- Transport: within this category data about buses, trains, trams and bicycles is used.
- Food and nightlife: localization and general information (e.g. meals, offers, and opening times) are taken about bars, clubs, and restaurants in the area.
- Outdoor locations: localization and information about scenic locations and other natural heritage sites are taken from the city council's open datasets.
- Professionals: as visitors might be professionals, information about conventions, conferences, and events in the city is used as well.
- Shops: retails, shopping centres, and information about local online shopping are subjects of this category of open data re-used.

- Accommodations: open data about hotels and apartments are integrated in order to support visitors in finding a suitable accommodation for their stays.

The ensemble of these datasets is believed to include all the information that people need when visiting the area.

The open data service still preserves one of its main initial characteristics, that is its gamification component. The application was initially conceived as a way to create a "historical challenge" (Company Document) from the city council's historical dataset. Users "are encouraged to discover historical and cultural sites in the area through an innovative find and reward Facebook app and mobile website" (Company Document). In other words, people undertake a game/challenge in which they need to find and "check into" (i.e. provide evidence that the user is at that actual location) three cultural sites. The game ends when the user is able to check into three cultural sites. If he or she has done so, a reward of a voucher for a free tea or coffee is provided. Once this destination is reached, the strategy of the company was to incentivize people to "buy a tour, scone or lunch along with the redemption of their free tea/coffee voucher" (CEO). Subsequently, the game experience was expanded into detailed "hunts". This term indicates "packages of events which a local or foreign visitor can do" (Company Document) in the area. This is supported by a mobile content platform and authoring tool, "which allows diverse organizations, from art galleries to zoos, to quickly develop and deploy mobile learning experiences, transforming fixed exhibitions into connected interactive learning challenges and games" (Company Document). By doing so, according to the CEO, "cities become interactive learning playgrounds where schools, home and educational spaces such as museums and galleries become interconnected and interactive spaces".

The service is supported by a "cloud-based cross HTML5 platform". HTML5 is an internet standard which is characterized by interoperability across internet browsers. This means that the service "runs over any browser meaning that there is no need to design individual applications for Iphones, Android, Blackberry, or Windows phones" – one of the volunteers stated. The integration of this information and the delivery of the service enable three main outcomes defined by the CEO as: (1) mobile, (2) social, and (3) analytics. The first includes all the outcomes in terms of information and experience provision to the users' mobile phones. The social component is a result of the interactivity of this open data service. In this way, the application collects "users reviews and updates", and information about preferences within the visiting experience. It also provides a platform for sharing experiences with other actual or potential visitors. Finally, at the analytics level, the CEO explained: "if we have the application we can track, we can discover patterns, we can do the enjoyment level, and where does that go? That is a government attraction; we can then analyze it at an open data level". Hence,

innovative analytics capabilities are implemented on the data stored and collected from both the usage experiences and the result of the above explained social outcome. These processes enable the discovery of "user tracking, user preferences, top patterns, top attractions, top routes, visitor enjoyment, and buying patterns" (Company Document). Part of this information is then released to the council and/or as open data to the public. However, some sections of this information are structured and form detailed knowledge about tourists' behaviours and habits. This side of the outcome became gradually part of the case's revenue model. A summary of the structure of the open data services offered is provided in Figure 5.1.

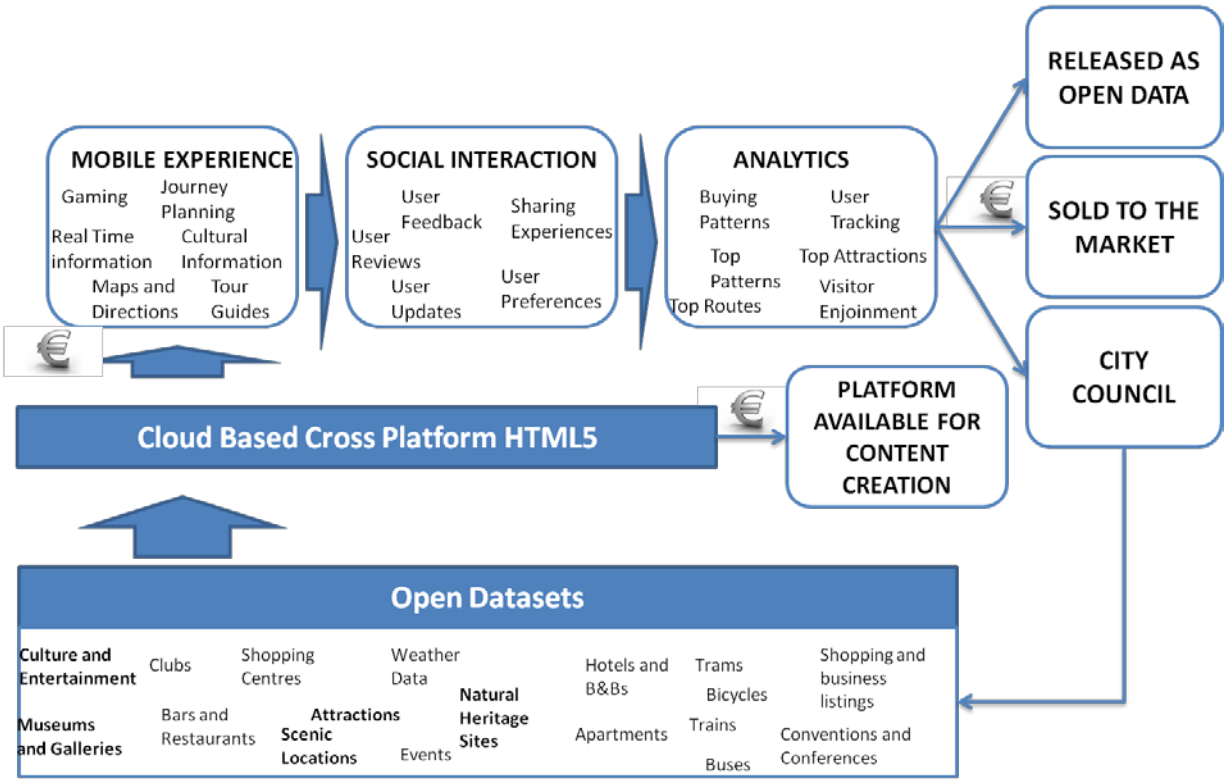

Figure 5.1: Case B's open data services

As shown in Figure 5.1, a cloud-based cross platform is in place and acts as an interoperable content management system. Open data is loaded and integrated into the system and the three main outcomes described above are enabled. As said, the outcome of the analytics stage is partially released as open data; another recipient of this output is the city council that with this detailed information is given the possibility of "taking more informed decisions, solve local inefficiencies, and design better strategies" (CEO).

It is noted that where the "€" sign is depicted, a form of revenue for the company is established. In particular, three sources of revenue are established. The first is about the possibility for local people and cultural organizations to upload content to the main platform. For example, citizens are given the

opportunity to design and sell their own tours around the city by buying the possibility of uploading content to the system. Secondly, the service is offered in a freemium version. This means that the largest portion of the functionalities of the application are available free of charge; however, some additional more advanced features can be purchased by the users. An example is the availability of "augmented reality" in relation to specific cultural attractions. Finally, the third source of revenue for the company is related to the selling of particular results of the data mining and analytics processes.

5.2 Contextualization of Methods

According to the methodology designed (see Chapter 3), three main sources of data were leveraged for this case study. At first, both internal and external documents were considered. In particular, four internal and seven external documents were collected and analysed. A description of each document, its content, the year in which it was created as well as its length is provided in Appendix 1.2. Observation was implemented through informal meetings with the CEO and, in one case, with one of the volunteers. These interactions were leveraged for initially introducing my research, and subsequently collect additional data about Case B. As part of these activities, the CEO was showing me the technical characteristics of Case B's open data services from his own computer.

Regarding the SSI process, the CEO of the company was found as the only suitable interviewee. The choice of interviewing only one person came from the fact that the other two people involved with the case are volunteers who work occasionally with the case and mainly with technical tasks.

In relation to the SSI protocol that had been previously designed (see sub-section 3.10.4), some changes were made based on the data collected. Firstly, an initial step was included in relation to the specific characteristics of the open data services offered. As opposed to Case A, Case B's services are much more complex. These are designed on top of a bigger range of open datasets and seem to evolve quickly over time. In addition, clarification was found to be needed about the strategy of embedding additional devices ("Apple Watch") in order to overcome the lack of internet access by foreigners. As part of step 2, additional questions were added in relation to main trends of open data in the tourism sector. This choice came from the fact that the subject has been actively involved in this ecosystem with several other projects. This set of questions were expected to substantially enrich the understanding of the context, as well as its main perceived trends and opportunities. As part of step 3, question 2 (*"In the choice of the specific dataset(s), which variable did you consider?*") was not found consistent with Case B's history. Similar to Case A, the idea preceded the availability of the data. A further question was added about what emerged so far as being the peculiarity of this case's open data service, i.e. the gamification component. In step 4, the three stages of the open data service development (see section 5.1) were investigated. Particular features of the pilot stage, the "minimum

viable product" one, and the journey that has been undertaken until the design and development of the current service were explored carefully. To do this, specific questions were addressed in relation to each of these stages as well as to the evolution from one to another. Moreover, the people component within the company was better explored as well; a few questions were added about the role of volunteers, including their expertise and responsibilities.

The first SSI process took place in a hotel hall. The environment facilitated a fluent and productive meeting which lasted for more than two hours. The wish of the interviewee to not be recorded probably made the overall process longer (I had to take precise notes while he was talking). However, initial interpretations were made while writing these notes, and discussed with the interviewee. The notes were structured on the same day that the interview was taken. Notwithstanding this, some aspects remained unclear or only partial information about some concepts was collected and structured. For this reason, an additional interview was scheduled with the CEO of the case with the goal of deepening my understanding of those aspects, which included:

- Open data ecosystem (EU): several times during the SSI the interviewee highlighted the importance of having in place a strategy driven by the European Commission. Specific details on how these strategies should be designed and to what extent these should influence the strategies within city councils were lacking. Thus, more clarity was sought in relation to the details of these initiatives as well as on how different stakeholders are believed to comply with such a wide-level program.
- Relationship with the council: the fact that Case B experienced interactions with the city councils that are providing the open data that is being re-used, emerged from several sections of the interview. However, no specific details about these relationships were provided.
- Business model for local open data ecosystem implementation: the last topic that remained unclear after the first round of analysis was the idea shared by the interviewee in relation to the development of a local open data ecosystem. For its development, the CEO emphasised the need of adopting a "triple helix" approach involving open data services businesses, academia, and city councils. However, no specific details were provided about the roles that each of these entities is expected to have in the resulting ecosystem. Also, the mechanisms through which this model would be put in place were not clearly described. Furthermore, the interviewee highlighted several times the importance of the "innovation network" in the overall ecosystem. Given the emphasis placed in relation to these bodies, further questions were planned to be asked to add clarity on: (a) what is meant by "innovation network" precisely; (b) if and how the case has

benefitted from engaging with these bodies; and (c) how and why these entities should be structured and implemented to positively influence adoption of open data.

All these aspects were translated into specific questions that were used as the input for the second SSI process. The second interview was undertaken three weeks after the first and took place in the same location as the previous one. These concepts were investigated together with the preliminary findings emerged from the first round of analysis. This interaction lasted for more than two hours and field notes were taken and partially structured during the meeting. The last stage of this case study was the final validation effort. This interaction lasted for two hours.

5.3 Case B Analysis

The analysis process undertaken for Case B strictly followed the method previously designed (see section 3.12). In particular, the line-by-line open review exercise conducted on the SSI transcripts led to the extraction of three hundred and thirty-eight KS from the SSI transcripts (see Appendix 4.1). As a result of the first cycle of analysis, two hundred and sixty-three KS were grouped under the external factors' category; for perception and adoption one hundred and seventy-seven and sixty-six KS arose respectively. Finally, ninety-eight KS were extracted in relation to links between factors. As done for Case A, the analysis for Case B is reported separately for external factors, perceptions, and adoption sub-factors.

5.3.1 External Factors Analysis

As a result of the first round of analysis, two hundred and sixty-three KS were found to be connected with external factors (see Appendix 4.1). These key concepts were further clustered, at a first level of detail, and five main topics could be identified. As done for Case A, one more cluster was defined and named *Other External Factors* in order to include all of those statements that did not fall within one of the other categories. The clusters identified were:

- ***Cluster A:*** open data and open dataset characteristics.
- ***Cluster B:*** organizational factors.
- ***Cluster C:*** factors related to the city. This cluster included two main themes emerged from the data collected: (1) KS about the city council and its open data program; and (2) aspects of the city that influenced the adoption process in this case. Regarding the latter, the nature of the service (i.e. mainly based on cultural attractions, sites, and experiences) makes it fundamental to consider these characteristics.
- ***Cluster D:*** open data ecosystem factors. This includes a variety of actors such as the European Union (EU) as well as "industry clusters and innovation networks" within the cities.

- ***Cluster E:*** service characteristics factors.
- ***Cluster F:*** other external factors.

It is noted that some of the KS were found to be related to multiple clusters. For example, KS 29,19 (see Appendix 4.1), i.e. "analytics enable: track, pattern discovery, enjoyment level as new knowledge created", contains information about the need of having internal skills related to data analytics and data mining (Cluster B), but also in relation to the characteristics of a service (Cluster E) which should have the ability of deriving secondary knowledge in the field.

Table 5.1 provides an overview of the coding process (more details about the coding processes are provided for each KS in Appendix 4.1). It includes, for each cluster defined, the number of supporting KS, the number of TAs developed in relation to both factors and related variables, and the number of concepts that were initially retained to be re-discussed during the final validation meeting.

	Cluster	Number of KS	Tentative Factors	Tentative Variables	Retained Variables
External Factors	A: Open Data Characteristics	72	1	8	0
	B: Organizational Factors	65	2	9	0
	C: City Related Factors	61	1	7	0
	D: Open Data Ecosystem	104	2	9	0
	E: Service Characteristics	59	1	8	2
	F: Other External Factors	12	1	3	0

Table 5.1: Case B, external factors analysis

As a result, eight tentative external factors emerged from this first stage of analysis as those that influenced Case B in adopting of open data for commercial service innovation. In total, forty-four variables arose from the data collected as a preliminary description for these factors.

5.3.2 Perceptions Analysis

From the first stage of open coding analysis, one hundred and seventy-seven KS were classified under the perceptions' category (see Appendix 4.1). As done for the external factors analysis, a number of clusters were defined in relation to the nature of these KS:

- ***Cluster G:*** perceptions about usefulness of open data.
- ***Cluster H:*** perceptions about business opportunity.
- ***Cluster I:*** perceptions about service's users, i.e. all KS related to perceptions generated by the users of the open data service and the related role these played for Case B's adoption.
- ***Cluster J:*** perceptions about the city.
- ***Cluster K****:* other perceptions.

The strategy employed to analyze these clusters was the same used for the external factors. In particular, each cluster constituted a separated and more focused unit of analysis. TAs were formulated about aspects for which strong evidence was found across the sources of data collected. A summary table describing this process is here proposed.

	Cluster	Nr. of KS	Tentative Factors	Tentative Variables	Retained Variables	Tentative Links
Perceptions	Perceptions on Usefulness of Open Data	82	1	6	1	5
	Perceptions about Business Opportunity	44	1	4	0	4
	Perceptions about the services' users	45	3	9	0	7
	Perceptions about the City	60	2	6	0	2
	Other Perceptions	6	1	1	0	1

Table 5.2: Case B, perceptions analysis

A shown in Table 5.2, eight classes of perceptions initially emerged as influencing adoption of open data for Case B. These factors were tentatively described by a total of twenty-six variables. One additional variable was retained at this stage as no sufficient evidence was found to motivate its inclusion. Finally, nineteen links emerged as connections between factors. All these preliminary outcomes for Case B were presented and re-discussed during the final validation meeting.

5.3.3 Adoption Process Analysis

With respect to the analysis that was conducted in relation to the adoption factor, a number of stages in the company's history emerged. These stages were derived from documents collected within the case, from informal discussions undertaken with the CEO and the volunteer encountered, and from sixty-six KS extracted from the SSI processes. As a result, eight steps were found to tentatively describe Case B's process from the generation of the first idea to the achievement of sustained use of open data.

	Adoption Stages	Number of Key Statements
1	Idea Generation	5
2	Open Data Adoption	8
3	Minimum Viable Product	10
4	Acquire Funding	10
5	Wider Adoption	11
6	Commercialization	8
7	Growth	25
8	Sustained Use	8

Table 5.3: Case B, adoption stages analysis

At this stage of the process, sixteen TAs were formulated in relation to sixteen factors that influence the adoption of open data for commercial service innovation for Case B. Furthermore, each of these factors was tentatively defined at a further level of detail through a set of variables. In total, seventy variables were identified. Concerning the actual adoption process, eight stages emerged and were outlined. Ultimately, three statements were found as not strongly supported by evidence. These were highlighted and some time was allocated for further reflections at the final validation stage. It is noted that the small number of retained concepts is believed to be related to the wish of the interviewee of not being recorded during the interviews. This situation enabled initial interpretation and structure of the emerging concepts. In this way, those concepts that were unclear were immediately tackled and further explained by the company's CEO. Concluding, a preliminary model for open data adoption for Case B was formulated and was brought into the final validation meeting.

5.3.4 Validation Meeting

At the beginning of this last stage of the case study, each factor and variable were carefully presented to the CEO together with the supporting evidence. The feedback received was very positive. Words like "excellent", "very good" and similar were often used by the subject at the time of presenting these concepts. As a result of this discussion only one variable was deleted. This was the case of the variable *Lean Team* within the *Organizational Factors* cluster (more details are provided in section 5.4). During the presentation of these tentative findings, no more insights regarding the definition of each factor arose. No additional information about the factors emerged from stimulating the CEO to further reflect on his personal and company's experiences, foundations, and future plans.

The focus of the second stage of the validation meeting was placed on the three TAs for which strong evidence was not available. These were re-discussed in-depth. In particular, these included: *Cloud-Based* (Cluster E), *Open Source Software* (Cluster E), and *Community Awareness* (Cluster G). The result of the discussion led to the integration of the first two variables, whereas the latter was accordingly excluded from the final findings.

The last stage was about considering the links between factors. As part of this stage, nineteen links were proposed between the previously agreed and established factors. In summary, only one of these links was found not correct by the CEO. In addition, a total of six more links has been proposed by the CEO, discussed, and subsequently established. With respect to the changes from the links proposed, important considerations were made by the case's CEO in relation to some factors (these are included in the thick descriptions of the final findings, see section 5.4).

At this stage, the validation phase was completed. Throughout the entire process questions such as "would you like to add something in relation to this aspect?" were continuously posed to the CEO. From the negative responses received it could be inferred that no more data collection was required and the final findings of Case B could be outlined.

5.4 Case B Findings

This section is dedicated to the description of the findings achieved from Case B. The strategy followed for reporting Case B's findings is the same adopted for Case A (see section 4.4). In particular, all the structured thick descriptions about factors, variables and links are provided for external factors (sub-section 5.4.1), perceptions (sub-section 5.4.2), and adoption stages (sub-section 5.4.3). All evidences supporting each component of the final adoption model is provided in dedicated tables (see Appendix 4). It is noted, that references to the documents (i.e. D1, D2, etc.) reflect the documents listed in Appendix 1.2. KS are referenced from Appendix 4.1 and where quotation marks appear ("") a quote from one of the interviews undertaken is made.

Overall, the findings demonstrate that adoption of open data for commercial service innovation for Case B is influenced by a total of eight external factors and eight perceptions. These emerged as being described by a total of seventy-three variables. Finally, twenty-four links were established between factors.

5.4.1 External Factors

From the initial six clusters identified, eight external factors were extracted and validated. A total of forty-five variables were established across these factors. One variable was finally excluded from the analysis in accordance with the feedback received at the final validation meeting. These are presented and described in the following paragraphs consistent with the format used in Chapter 4.

- ***Assertion 1: Open Data Characteristics influence the adoption of open data for commercial service innovation***

From the previous clustering exercise, seventy-two KS (see Appendix 4.1) were grouped within Cluster A, i.e. about open data and datasets characteristics. Also, all types of documents contributed to the definition of this assertion. For instance, D4 elaborates on "issues such as privacy, cleansing, accuracy, and reusability of open data". More broadly, D2 makes an implicit call to cities for more open datasets that should be accessible by machines and in near real-time. In total, eight variables were defined within this cluster. These and the related supporting evidences are summarized in the following table. Each variable is subsequently described.

ASSERTION 1					
Variables		**Key Statements**	**Int. Doc.**	**Ext. Doc.**	**Valid.**
1.1	*Amount of data available*	1,4; 2,6; 2,12; 2,13; 4,2; 4,7; 8,27; 9,11; 9,12; 9,13; 10,4; 10,5; 11,5; 11,6; 11,7; 11,10; 11,11; 11,12; 11,13; 12,1; 17,4; 21,3; 27,1; 29,2; 29,27, 31,6; 33,1; 33,2; 34,15; 36,3; 36.6; 36,9.	D1;D2; D3.	D5; D6.	Yes
1.2	*Real-Time/Timely data*	4,3; 4,5; 4,6; 5,4; 9,10; 9,13; 10,1; 10,2; 10,3; 29,24.	D1;D3; D4.	D5.	Yes
1.3	*Access and Machine Readability*	2,7; 6,24; 8,18; 8,19; 8,23; 8,24; 8,25; 11,2; 15,13; 15,15; 25,1; 29,26;.	D1;D2; D3;D4.	D5.	Yes
1.4	*Availability free of charge*	6,20; 8,24; 8,25.	D1	--	Yes
1.5	*Geo-Localization*	6,22; 6,23; 8,19; 15,7; 17,7.	D2, D3	--	Yes
1.6	*Accuracy of Data*	11,1; 15,1; 15,2; 15,3; 15,4; 15,5; 15,6; 15,8; 17,1; 17,2; 17,3.	D2, D4	D5	Yes
1.7	*Standard*	11,3.	D1	D5	Yes
1.8	*Re-usage License*	11,4.	D1, D4	D5, D9	Yes

Table 5.4: Case B Assertion 1, open data characteristics

1.1 Amount of Data Available: in order to provide a comprehensive service to the tourists in the area ("we want to own the tourists' trips"), all the data that might be involved in augmenting the availability of information about the decisions that a visitor has to take throughout his or her stay in the city could be used. Thus, the amount of open datasets released by the council is considered as a critical variable for this case. When making examples of data that could be openly available the CEO listed: maps, transport, traffic, weather, crimes, noise pollution, attractions, and landmark.

1.2 Real-Time / Timely Data: according to D4, "the most valuable open data is real-time data". This characteristic is critical for some datasets used by the case (e.g. transport data, weather data, events data), while it could be considered as a secondary need for others (e.g. natural heritage sites, scenic locations, museums and galleries). According to the CEO, having "stagnant datasets" is considered one of the main barriers that challenge the development of sustainable open data services.

***1.3 Access and Machine Readability*:** several KS were found to be related to the technical access to the dataset. The data structured suggests that "machine readability" of the data is crucial in order to develop solution "easily" and "quickly". This aspect was found as the main reason why the volunteer-based project initially started. Also, access to the data should be supported by open source-based APIs (Application Programming Interfaces) in order to facilitate mash-ups and integration of different open data sources.

1.4 Availability Free of Charge: another characteristic emphasized in the data collected was the availability of open data free of charge. The fact that open data can be freely re-used was found to be fundamental for the case to consider open data as a solution to a previously identified problem.

1.5 Geo-Localization: for Case B, having data to which localization is assigned was critical. The CEO underlined several times that "coding maps" for geo-localizing the data would be too time-consuming. From another perspective, errors in the dataset on localization were considered as one of the main issues encountered by the volunteer.

1.6 Accuracy of Data*:* another main challenge encountered by Case B was about "errors found in the open datasets". So "data cleansing activities were needed" and "double checking the data before loading it onto the system was a time-consuming process". As one of the consequences identified for inaccurate data provision, there was a final lack of trust in the data.

1.7 Standard*:* even if this variable was supported by only one KS, having a standard through which the data is exposed to the public was considered fundamental. In a hypothetical situation in which the data is supplied with different standards, the subsequent effort that would be implemented to make the data consistent with the company's system (and so with a unique standard) would represent a "significant barrier".

1.8 Re-usage License: as for the previous variable, one KS supported the importance of assigning re-usage licenses to the datasets. However, the fact that the CEO did not stress several times its importance is due to the nature of the datasets re-used by the case. In fact, "privacy and security issues are not relevant for these datasets" (e.g. scenic locations, cultural attractions). Thus, in the case's experience, a re-usage license was initially available and no major law threats were faced by Case B when re-using this data.

- ***Cluster B: Organizational Factors***

After reading several times and tentatively structuring the sixty-five KS classified under this cluster, two main categories could be distinguished. On one side, fifteen KS were found as related to the organizational network and strategy. Its importance was also supported in the Horizon2020 submission form (D1), especially in the budget and future economic business plan section. On the other hand, the remaining fifty KS were related to the internal skills that were needed to gradually transform the datasets into the open data services offered. According to the feedback received, two sub-clusters were defined and the analysis was conducted at that specific level of detail. These are: *Organizational Factors* (for the former class) and *Organizational Skills* (for the latter).

- ***Assertion 2: Organizational Factors influence the adoption of open data for commercial service innovation***

Regarding the first sub-cluster, four variables initially emerged. However, one of these was finally excluded as a result of the final validation meeting (underlined). All variables and the related supporting evidences are provided consistently with the format used so far (see Table 5.5).

ASSERTION 2					
Variables		**Key Statements**	**Internal Doc.**	**External Doc.**	**Validated**
2.1	*Lean Startup Approach*	1,7; 35,15; 35,16.	D1	D6	Yes
2.2	*Lean Team*	18,1; 18,2; 19,1; 19,2; 19,3; 19,4.	D1	--	No
2.3	*Partners Engagement*	26,9; 29,5; 29,6; 29,12; 34,10; 34,11; 34,12; 34,13; 34,14.	D1, D2, D3	D6	Yes
2.4	*Volunteers Eng.*	18,2; 26,9; 34,6;	D1	D8	Yes

Table 5.5: Case B Assertion 2, organizational factors

2.1 Lean Start-up Approach*:* immediately at the first question during the first SSI process, the CEO stated that having a "lean start-up approach" is considered both as the "right way of initiating a software business" and a strategy suitable for the open data industry. This methodology is based on the "build, measure and learn" loop where the first step in the process involved structuring the city-related problem/opportunity and the development of a "minimum viable product".

2.2 Lean Team*:* connected to the previous variable, it arose that one of the success factors of the case was the development of a "small lean team". However, this variable was criticized by the CEO during the final validation stage, who described it as "redundant". In particular, after a short discussion, the CEO and I convened that this information was already ingrained in the *Lean Startup Approach* concept.

2.3 Partners Engagement: another characteristic of the company's strategy is related to the continuous seeking of partners. According to the CEO, partners should include "academia, tourist associations, the innovation network and the business/SMEs local network". Also, "tourist attractions" were found to be relevant partners when applying for EU funding (D1). Overall, two main types of partners are identified. Firstly, those partners which engagement would enrich the final service delivered (i.e. tourist associations/attractions and business network). These are believed to enable "dynamic business and event information [that] would allow a central platform from which all businesses and community could update and contribute themselves" (D2). A second type of partners is identified in the "innovation network and academia". By "innovation network", the CEO meant organizations that operate as intermediaries between local authorities and open data users. These organizations usually operate to further establish catalogues and improve the supply process of the

open data published by the corresponding local authority. Finally, academic researchers are found as very useful sources for empirical studies as well as for volunteers' recruitment.

***2.4 Volunteers Engagement*:** this variable is considered separately from the previous because it arose that the city can be a very effective source for volunteers (not just from universities, but also for example from other community-oriented initiatives). The engagement with volunteers was a critical aspect especially within the early stages of the business.

- ***Assertion 3: Organizational Skills influence the adoption of open data for commercial service innovation***

Within the second sub-cluster of Cluster B, all the skills that were needed for the case to transform the open dataset into a sustainably delivered commercial service were derived from direct statements of the CEO as well from interpreting the skills required for delivering the service (see Table 5.6 for a summary of the analysis process).

3.1 Software Development: software development skills were the most critical aspect for this case. These skills were needed for: extracting the data (API-enabled access); ensuring that the right API is in place; integrating the data onto the system; "bounding data into tiles"; applying software codes on top of the database; enable user-service interaction; and design and run the analytics phases. These skills include mobile programming, cloud computing, data integration, API development (for enabling analytics) and usage (for accessing the open datasets), and HTML5 knowledge.

ASSERTION 3					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
3.1	*Software Development Skills*	2,7; 5,6; 5,7; 5,10; 6,7; 6,8; 8,23; 8,24; 8,25; 11,2; 11,3; 13,11; 15,9; 15,12; 15,13; 19,3; 24,3; 25,1; 29,23; 29,26; 29,8; 29,10; 29,20.	D2, D3, D4	D5, D6, D7, D8, D10	Yes
3.2	*Graphic Design and Data Visualization*	1,19; 5,1; 5,2; 5,11; 13,11; 15,10; 15,14; 15,15; 19,2;	D2, D3, D4	D6, D8, D10	Yes
3.3	*Data Analytics and Data Mining*	13,11; 24,4; 29,11; 29,9; 29,19; 29,20; 29,24.	D2, D3, D4	D6, D8, D10	Yes
3.4	*Data Cleansing*	13,11; 15,3; 15,4; 15,5; 15,6; 15,8; 15,11.	D4	--	Yes
3.5	*Project Management*	14,1; 19,1; 26,9.	D1	--	Yes

Table 5.6: Case B Assertion 3, organizational skills

3.2 Graphic Design and Data Visualization: the importance of having "sleek design" of the application was highlighted several times during the interviews. Furthermore, graphic design improvement was seen as one of the most important targets to be achieved during the company's history. This set of skills includes data visualization. These skills were described as the ability of transforming "excel tables into something easy to see and detect".

3.3 Data Analytics and Data Mining: part of the mission of the company is generating and releasing additional knowledge about the tourism context. This data is created as users use the service. To do this, data mining and analytics processes are implemented on the data that is being generated by the service's usage and through its social interaction (see section 5.1).

3.4 Data Cleansing: in the case's experience this term can be defined as identifying incorrect, incomplete, and inaccurate data and subsequently replacing, modifying, or deleting this "bad data". These skills were crucial for the case as initially "many errors were found in the open datasets". As a consequence, data had to be "fixed before loading it onto the system".

3.5 Project Management: in order to develop, establish and sustain "wide collaboration" (see variable 2.3) and relationships with the city councils, project management skills are considered to be crucial. In the company's experience, this set of skills was also critical when developing the application for Horizon 2020 funding (D1).

- ***Assertion 4: City-Related Factors influence the adoption of open data for commercial service innovation***

Within Cluster C all of the KS emerged in relation to the city were grouped together (see Appendix 4.1). Two main categories of statements could be outlined within this class: (1) statements about the local authority and (2) those about the specific characteristics of the cities where the open data service is offered. Both of these aspects were found to play a role in Case B's adoption of open data. Seven variables were identified and TAs about these variables were formulated and subsequently validated (see Table 5.7).

4.1 Open Data Program: this concept stresses the importance for the city to have an open data program in place. City authorities need to allocate budget and resources for supplying open data as part of a long-term strategy. According to the CEO, this program should be oriented towards continuously publishing and updating the data, including as much data as possible, consistently with the characteristics outlined within Assertion 1. It is also noted that the main criterion for the case to target new cities is the presence of a long-term program from the related public authority. Furthermore, only one open data initiative should be developed across departments. The final objective of the strategy should be to "build and maintain an open data ecosystem".

ASSERTION 4					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
4.1	*Open Data Program*	1,23; 2,12; 2,13; 5,5; 12,3; 13,3; 13,4; 20,7; 26,3; 27,1; 29,2; 29,17; 29,27; 31,6; 33,2; 33,4; 35,6; 36,1; 36,2; 36,4; 36,5; 36,6; 36,8; 36,9; 37,10; 37,12; 37,13; 37,14.	D1	D5, D9, D10	Yes
4.2	*Open Data Awareness / internal champions*	2,12; (2,13;) 3,2; 17,5; 27,3; 28,1; 34,4; 35,3; 35,12; 37,10; 37,11; 37,12; 37,13; 37,14; 38,7.	D2, D4	D5, D9	Yes
4.3	*Open Data Promotion*	2,12; 17,1; 17,2; 17,5; 27,6; 35,13; 35,14; 37,9; 37,12; 37,13; 37,14.	D4	D5, D9	Yes
4.4	*Underexploited areas*	1,11; 2,8; 2,9; 3,1; 3,2; 8,9; 26,10; 31,3; 31,4; 31,6; 35,12.	D2	D8	Yes
4.5	*Public free Wi-Fi*	1,10; 1,11; 8,7; 8,8; 8,9; 17,6; 33,3;	D3	D8	Yes
4.6	*Infrastructure investments*	12,2; 12,3; 13,4; 35,6.	--	--	Yes
4.7	*Local Funding Provision*	33,4; 34,18; 36,8.	D2, D3	D8, D9	Yes

Table 5.7: Case B Assertion 4, city-related factors

4.2 Internal Champions: as stated in D4, "in order for this public data to be accessible we need integrated initiatives and open data champions both on a local and EU-level" (D4). Cities need to be "pushed, encouraged, and educated" because too often, apart from some figures/entities promoting the initiative, too many people within the councils are not aware of open data. The subsequent resistances to change that arise are believed to be significant barriers for releasing open data.

4.3 Open Data Promotion: promotion of open data by the city council is meant as all the activities implemented by the authority to spread awareness outside the council itself. Promotion initiatives include "community-oriented events, free events and presentations, competitions for open data re-usage" where examples of open data services can be showcased as a demonstration of the value behind open data. As a result, more users would be stimulated to embrace open data as the source for new commercial services. Other consequences for cities include getting known worldwide for being pioneers in the open data movement and increasing "trust" from actual and potential users.

4.4 Underexploited Areas: as the first variable of the second type (i.e. related to the city's assets and characteristics), the presence of underexploited areas was found to be a critical aspect that led Case B to adopt open data. In this case, tourism in the area was previously significantly underexploited ("did you know that 95% of public attractions do not have a mobile app?", D8); no connections between

the target area and the main hubs around were previously provided ("local inefficiency"); and "local traditional culture was not promoted at all".

4.5 Public Free Wi-Fi: for Case B's services internet connection is a main requirement. On the other hand, "lack of Wi-Fi" was mentioned several times by the CEO as being a significant barrier for visitors (especially foreigners) to embrace the open data services offered. The importance of having public Wi-Fi was underlined. This issue was initially tackled by the case by adopting and giving the user the possibility of renting an additional device (Apple Watch) on which the service can be run.

4.6 Infrastructure Investments: infrastructure investments were described as enabling the development of another source of open data, i.e. from sensors. The CEO stressed the importance of having and maintaining a sensor network that is able to capture, process, and outline data from the city to be then published as open data. This source is believed to be fundamental for some types of information (e.g. "weather, noise pollution, traffic"), i.e. the one for which real-time data is needed in order to make it re-usable.

4.7 Local Funding Provision: in relation to funding, the CEO underlined how EU-driven funding should be the main contributors. However, providing awards and prizes for the "best developers" in various forms of competitions is seen as an example of small funding enabling and encouraging the design of new open data services. In addition, cities where local funding is provided are likely to be targeted as potential new locations in the "growth" process of the company.

- ***Cluster D: Open Data Ecosystem***

The inclusion of one hundred and four KS under this cluster (see Appendix 4.1) demonstrates the emphasis placed in relation to the importance of having an effective open data ecosystem to foster widespread adoption of open data. Two main categories of statements were outlined: (1) those related to an EU-driven ecosystem, (2) and those related to an ecosystem to be developed at the city level. These two categories led to the definition and validation of two different assertions.

- ***Assertion 5: EU Open Data Ecosystem Factors influence the adoption of open data for commercial service innovation***

This cluster emerged to be critical because individual cities are seen from the case as potential contexts where its services can be deployed. These contexts are seen as part of a broader ecosystem that has to be "EU-driven". "EU-driven open data" was defined as a solution to the main challenges for open data services businesses (i.e. "scale, lack of data, lack of Wi-Fi, and lack of funding") in one of the internal company reports (D3). The solution was defined in two main directions. On one side, small projects should be fostered through a "European Union open data re-use incubator". On the other hand, large projects should be leveraged through funding schemes. The following variables emerged

across the data collected and are meant to fully describe the case's view on this ecosystem as well as its relationships with it (for details about supporting evidence see Table 5.8).

ASSERTION 5					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
5.1	*EU-level strategy*	1,4; 1,22; 2,11; 13,3; 13,4; 22,5; 22,6; 23,1; 23,2; 26,5; 26,6; 26,7; 26,8; 26,9; 26,10; 26,11; 29,13; 29,17; 30,2; 30,3; 33,4; 34,1; 34,2; 34,5; 34,8; 34,9; 34,15; 34,18; 35,8; 35,9; 36,1; 36,2; 36,8; 37,1; 37,3; 37,4; 37,5; 37,6; 37,7; 37,9; 38,6.	D1, D3, D4	D6, D8	Yes
5.2	*EU-level Ambassadors*	26,2; 35,2; 35,3; 35,4; 37,8; 37,11; 37,12; 37,13; 37,14.	D1, D3, D4	--	Yes
5.3	*Triple Helix Innovation and Collaboration*	1,4; 1,21; 9,11; 12,2; 12,3; 13,1; 13,2; 13,3; 13,4; 13,5; 13,6; 23,3; 26,9; 34,10; 11; 34,12; 34,13; 34,14; 35,11; 35,12; 27,3; 38,1; 38,2; 38,4; 38,8; 38,9.	D1, D2, D3	D10	yes
5.4	*Best Practices Repository*	26,7; 35,10; 35,13; 35,14; 37,2; 37,12.	D1, D3	--	Yes

Table 5.8: Case B Assertion 5, EU open data ecosystem factors

5.1 EU-Level Strategy: according to the CEO, cities started embracing open data because of claims from EU about the related potential economic growth (see Chapter 1). As a consequence, "EU should invest". The case sees EU as the main opportunity to get funding. In addition, EU is seen as the proper driver for open data because of the following reasons: legislation (i.e. open data re-usage license) is EU-driven; EU for creating "one location for all initiatives"; and to enable knowledge-sharing and collaboration across countries. According to the CEO, this strategy would have to "supervise" the various cities and foster their open data initiatives.

5.2 EU-Level Ambassadors: referring to an example of an initiative carried out by the Swedish government, the CEO explained how "ambassadors should go around cities to educate them on how to do it" (i.e. how to effectively embrace open data). These subjects should travel across councils to "promote best practices, toolkits, and to provide examples".

5.3 Triple Helix Innovation and Collaboration*:* emphasis also emerged in relation to the key characteristic that this ecosystem needs to enable: collaboration across countries involving academia, businesses, and both local and wider governments. In addition to what is stated in terms of *Partners Engagement* (see variable 2.3), the business component of this model should be also represented by IT companies that "build and maintain" the infrastructure needed (i.e. mainly sensor networks) to generate and release real-time open data. Connected to this variable, several KS demonstrate the

growing concern about the "business model" for this collaboration. In this way, a call was made for new solutions in order to align "corporate social responsibility and commercial purposes".

5.4 Best Practices Repository: this concept is meant as a repository of the activities and operations for effectively releasing the data and for service developers. By doing this, the overall ecosystem collectively builds a structured knowledge base upon which adoption of open data can be fostered.

- ***Assertion 6: Local Community and Ecosystem Factors influence the adoption of open data for commercial service innovation***

Further five variables emerged in relation to the importance of all those initiatives at the local level "that are not too much closely related to the council". For example, in D4 community-oriented hack competitions were described as "tremendously successful" in "showcasing the true benefits and innovations that can be achieved from open data" (D4). The common underline argument across the variables in this cluster is to "spread awareness" (D2) and "helping show the value of public agencies taking the leap to open up their data" (D4). The evidences supporting the definition of these variables are provided in Table 5.9.

ASSERTION 6					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
6.1	*Local Ambassadors*	26,1; 35,2; 35,3; 35,4; 35,7.	D2, D4	D9	Yes
6.2	*Hack Competitions*	7,2; 7,3; 8,1; 22,1; 22,2; 22,3; 22,4; 34,3.	D2, D3, D4	D5, D6, D7, D9, D10	Yes
6.3	*Showcases*	8,26; 8,27; 22,1; 22,2; 26,11; 29,6; 36,7.	D2, D4	D5, D6, D7, D8, D9	Yes
6.4	*Volunteers*	34,6.			Yes
6.5	*Innovation Network*	26,3; 29,5; 29,14; 34,14;	D1, D2, D3, D4	D5, D6	Yes

Table 5.9: Case B Assertion 6, local community and ecosystem factors

6.1 Local Ambassadors: together with *EU-Level Ambassadors*, also local champions are believed to be critical in order to "spread awareness among government and consumer sides". Ambassadors are perceived as having a key role in joining together potential open data users (developers) through different means (see variables below within this cluster).

6.2 Hack Competitions: hack competitions are seen as valuable initiatives for fostering the creation of new services from open datasets. The case took part in one of these competitions (and won it). According to the CEO, this experience was fundamental to discover open data as well as to leverage the work that was being implemented. It also gave the opportunity to engage with the key people in the city council. Finally, it provided initial funding for launching the business.

6.3 Showcases: similar to the previous variable, these initiatives involve the creation of spaces to enable people to come together and to spread awareness of open data and its potential benefits. Showcases "allow effective display of open data benefits". These occasions are also seen as means through which the local council can "drive exposure of the open data service offered" (D2).

6.4 Volunteers*:* only one KS was found to support the importance of volunteers at the local level. This code was about the ability of the local ecosystem to provide volunteers (e.g. at community-oriented initiatives). During the final validation stage, confirmation about its importance from these initiatives was achieved.

6.5 Innovation Network: in the CEO's opinion and experience, a key role in this ecosystem is finally played by those new organizations of intermediaries between city authorities and the demand landscape. These are referred to as "innovation networks". The main role of these bodies is to help city councils by providing IT skills for improving the release process of existing open data as well as for creating new datasets.

- ***Assertion 7: Open Data Services' Characteristics influence the adoption of open data for commercial service innovation***

In Cluster E, fifty-nine KS related to the characteristics of the open data service offered were grouped together (see Appendix 4.1). Internal documents (mostly D2 and D3, which focus on general and technical characteristics of the service) were largely leveraged to expand and structure the variables within this cluster. In total, ten variables emerged and were validated (for details about supporting evidences see Table 5.10). Two of these were initially retained (i.e. variables 7.7 and 7.8) and their inclusion in the final findings was agreed only after the validation stage. The general concept ingrained in this assertion is that adoption of open data is stimulated if the resulting service developed has certain characteristics. Also, depending on its characteristics, an open data service is more or less adopted by end users, which in turn impacts adoption of open data in the long-term.

7.1 Scalability: having a scalable service was found to be a critical aspect for the case's success. According to the CEO, the "big work has to be done once only" and, from there, growth is enabled by customizing the first solution to other cities. In other words, once the service is designed for one city, different datasets can be loaded into the system "quickly". Critical in this way is the design of a "single content management system across datasets".

ASSERTION 7					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
7.1	*Scalability*	13,7; 13,8; 22,7; 25,1; 27,4; 27,5; 27,6; 29,1.	D1, D3, D4	D6	Yes
7.2	*Interoperability*	29,23.	D3	D6	Yes
7.3	*Incentivizing - Challenging*	1,5; 1,12; 8,14; 8,16; 24,1; 29,25; 32,1.	D1, D2, D3	D6, D8	Yes
7.4	*Enabling Interactive Experience*	2,16.	D1, D2, D3	D6, D8	Yes
7.5	*Attractive technology*	1,9; 1,10; 1,18; 1,19; 5,1; 5,2; 5,6; 5,7; 5,10; 5,11; 8,2; 8,3; 8,9.	D1, D2, D3	D6, D8	Yes
7.6	*Addressing User's needs*	1,13; 3,1; 4,7; 5,3; 8,3; 8,9; 8,10; 8,11; 8,17; 11,13; 31,5.	D1, D3	D8	Yes
7.7	*Cloud Based*	24,3.	D3	--	Yes
7.8	*Open Source Software*	2,7; 8,25.	D2	--	Yes
7.9	*Secondary Open Data Creation*	1,21; 1,23; 1,24; 2,6; 2,7; 3,2; 20,7; 24,4; 24,5; 29,18; 29,19; 29,20; 29,24; 31,6.	D2, D3, D4	D6	Yes
7.10	*Engagement with local community*	31,3; 31,4; 32,1.	D2, D3	D8	Yes

Table 5.10: Case B Assertion 7, open data service characteristics

7.2 Interoperability: this characteristic is enabled by the choice of using HTML5 language to develop the software. In the CEO's opinion, having an interoperable solution across internet browsers and devices is critical to engage several different types of end-users (e.g. mobile users, tablet users, general web users). Also, it "significantly facilitates the external content creation process" (D6), in which editing activities can be undertaken with any device using any of the available browsers.

7.3 Incentivizing and Challenging: another critical component of the service is about the nature of the experience enabled. This is related to the introduction of game theory principles to enable a challenging experience for users. In the CEO's opinion, "the gamification component and the availability of rewards are critical to incentivize users", thus for spreading adoption of the service, and therefore to achieve sustained adoption of open data.

7.4 Enabling Interactive Experience: another peculiarity of the service is its interactive component. Interactivity is achieved in three directions. Firstly, users have the chance to provide reviews of their experiences and share these opinions with other users. Secondly, the learning process of a person using the service can be considered interactive from several angles. This interactivity was described in D3 as follows: "creating interactive learning environments, create interactive tours, create interactive quizzes, and create interactive learning games" (D3). Thirdly, interactivity is also achieved with regards to the opportunity of "visitors and locals coming together" (D2).

7.5 Attractive Technology: this characteristic was included as a variable because emphasis has been put on the importance of having "fancy technology" and "design" supporting the service. This characteristic is perceived to increase the adoption of the service among users.

7.6 Addressing Users' Needs*:* the most important characteristic of the service is its ability to provide valuable information to the end-user. Case B's open data service is mainly designed for tourists. In this way, the design and study of several different daily scenarios for tourists was critical to the provision of a "tourism platform" where "information-enabled daily decisions" are displayed.

7.7 Cloud Based: this variable was initially supported by one KS only and from D3. Therefore, this TA needed to be re-discussed at the validation meeting. Thus, when further investigating its importance towards the adoption of open data, the answer was clear: "it is really important". In particular, by having a "cloud-based-system" the company avoided the development of complex databases for storing the huge amount of data uploaded, collected, and generated by the open data service.

7.8 Open Source Software: two KS emerged in relation to the importance of employing open source software as an input and outcome of the service. This characteristic is perceived by the CEO to be aligned with the open data principles, but it also reduces the company's dependency on software licenses. The latter is perceived as an action to minimize risks related to the development of this open data service. A description of this variable and a reasonable explanation on why it should be included are available. However, this variable has been underlined because more clarity was needed especially regarding why embracing open source software is particularly important for open data services businesses as opposed to other industries. Part of the answers given emerged around the "uncertainty of open data". The CEO stated "we don't know what will be open data tomorrow. And we don't want to be paying software licenses for several years that we might not need". Accordingly, this variable was included in the final findings.

7.9 Secondary Open Data Creation: the results of data analytics processes upon the data generated by the people using the service are aimed at generating new data in the context. Once structured, this data becomes knowledge that is released to the council, or as open data, or sold to the market. This step is believed to "improve local data" and "drive the ecosystem around" as well as to enable "shared value for the community". Among other consequences, local businesses are advertised and contribute to the company's revenue generation. The CEO described this process as a "crowd sourcing of valuable data for the local authority".

7.10 Engagement with Local Community: engaging with local communities is also perceived as a critical aspect for having a sustainable service and for increasing the end-users satisfaction ("when you travel the best thing you could do is to meet some locals"). By giving the opportunity to locals to

contribute to the service and/or building a further one on top of this platform (e.g. “selling their tour guides”), the company aims at playing a pivotal role in the local tourism economy.

- ***Cluster F: Other External Factors***

Within Cluster F the twelve remaining KS were grouped. The majority of these aspects were found to be related to the characteristics of the individual that developed the idea and subsequently founded and launched the business. As a consequence, an assertion is formulated outlining individual factors influencing the adoption process for this case.

Assertion 8: Individual Factors influence the adoption of open data for commercial service innovation

As shown in Table 5.11, three variables initially arose and were subsequently validated for this factor.

8.1 Attitude toward IT/Data-Oriented Business: this first variable was defined to emphasize the expertise in IT and data science of the person who initiated the business (i.e. the CEO). It was because of the subject’s passion about IT and data-oriented business that awareness of open data was achieved. This statement finds further confirmation in the introductory section of D4, D5, D8, and in D7 where the current CEO is described as “an experienced software developer” (D7).

ASSERTION 8					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
8.1	*Attitude toward IT/Data-oriented Business*	1,2; 1,3; 1,6; 2,12; 7,1; 7,2; 8,1; 23,2.	D4.	D5; D7; D8.	Yes
8.2	*Experience in the context*	2,1; 6,1; 6,2.	D1.	D5; D7; D8.	Yes
8.3	*Creativity*	24,2.		D9.	Yes

Table 5.11: Case B Assertion 8, individual factors

8.2 Experience in the Context: from the CEO’s experiences in the context, this subject developed knowledge about: what a comprehensive tourism service should encompass; what data should be included; and the market in general. Having an articulated knowledge about the tourism context highly influenced the successful design of the final open data service.

8.3 Creativity: even if this variable is supported by only one KS, the creative attitude of the founder was critical for the success of the service itself. In particular, introducing “open data games” through designing an interactive challenge is understood as the main value proposition of the company. This peculiarity was the result of the subject’s creativity.

At this stage all the data that was found to be related to external factors was thoroughly analyzed and validated. These processes resulted in the establishment of eight assertions about external factors

that influence the adoption of open data for commercial service innovation for Case B. Moreover, forty-five variables describing these factors were defined and validated.

5.4.2 Perceptions

In reporting the perceptions emerged for Case B, the strategy employed is the same used for external factors. A comprehensive view of all the evidences is provided as well as the thick descriptions of all factors and variables. At this stage, also links between factors are presented.

- ***Assertion 9: Perceptions about Usefulness of Open Data influence the adoption of open data for commercial service innovation***

Assuming that cities have open data programs in place and that this data is published consistent with the characteristics previously identified (see Assertion 1), seven tentative variables emerged describing the perceived usefulness of open data for commercial service innovation purposes. One of these was excluded after the final validation stage (variable 9.7). The evidences supporting this assertion are provided in Table 5.12.

ASSERTION 9					
	Variables	**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
9.1	*Enabling Dynamic Information*	1,14; 2,16; 4,3; 4,5; 4,6; 5,4; 6,4; 8,10; 9,4; 11,9; 17,1; 17,2; 26,10; 29,19; 29,22; 29,24; 29,25; 31,6.	D1, D3, D4	D6, D8	Yes
9.2	*Free of Charge*	2,7; 6,13; 6,20; 9,8.	D1		Yes
9.3	*Large Datasets Analyzable and Shareable*	1,11; 2,13; 4,2; 6,8; 6,11; 6,24; 6,25; 6,26.	D2, D3, D4	D8	Yes
9.4	*Short-time development*	2,14; 6,5; 6,8; 6,10; 6,12; 6,14;6,17; 6,19; 6,22; 6,23; (6,24); 6,25; 6,26; 6,31; 8,22; 9,7; 13,8; 15,6; 25,1; 29,3; 34,16.	D2	D6	Yes
9.5	*Easier Process*	5,6; 6,5; 6,8; 6,12; 9,7; 9,9; 13,8; 15,6; 15,15; 29,3.	D2	D6, D8	Yes
9.6	*Provision of content*	1,11; 1,14; 2,13; 2,15; 2,16; 3,2; 4,2; 4,3; 4,7; 6,3; 6,4; 6,6; 6,15; 6,21;6,23; 7,1; 7,2; 8,10; 8,17; 8,26; 9,2; 9,3; 11,2; 11,9; 11,13; 17,3; 21,3; 29,2; 31,5; 34,17.	D1, D2, D3, D4	D6, D8	Yes
9.7	*Community Awareness*	38,7.	--	--	No

Table 5.12: Case B Assertion 9, perceived usefulness of open data

9.1 Enabling Dynamic Information: open data is perceived to enable "dynamic information", which in turn "facilitates app developments". In particular, the supply of data that is "frequently updated" generates a "flow of information" that is leveraged by Case B into its open data service.

9.2 Free of Charge: open data is perceived as useful because it is free of charge. Having this data for free gives an additional advantage to open data compared with other sources. By using open data, "real-time experiments" and "cheap developments" can be made "at a low risk".

9.3 Large Datasets Analysable and Shareable: "open data has the ability to be collected, analysed and shared in large datasets" (D4). Thus, developers have the opportunity to analyse, combine, and re-use large datasets. Alternatively, specific datasets would need to be populated and maintained, and "achieve the same multitude of data would be almost impossible".

9.4 Short-Time Development: another perceived advantage of using open data is related to "time saving". The CEO stressed several times the importance of having solutions that can be developed "quickly". As underlined in one of the internal reports, "as more open data becomes available this can be quickly updated and turned into interactive challenges and rewards" (D8). When envisioning a situation where open data is not available, the CEO outlined a range of "time-consuming" activities that would be required. For instance, editors should be hired for populating and frequently updating the database; further activities would be needed such as coding the map on top of the data.

9.5. Easier Process: open data also "enables a much easier process". Several potential activities that developers should undertake are already ingrained in the open dataset. These include: searching and cataloguing the data; applying localization codes on top of it; and the potential need of meeting up with locals to gather information and design touristic tours. By having open data available, according to one of the volunteers, it is "easy to pull the database into the app".

9.6 Provision of Content: the variable that probably describes the best the usefulness of open data is the availability of content. Case B's open data service wouldn't have been developed without content, since it aims at providing an "information rich service" to "display information-enabled daily decisions". This is also reflected in the data that the company subsequently publishes as open data (from analytics). This data is released to "stimulate the entire ecosystem by publishing new content".

9.7 Community Awareness: only one KS supported the definition of this variable. The idea that "if all the community gains awareness of its benefits (i.e. of open data) it facilitates the process", comes from the fact that the company aims at involving the community in the delivery of the service (both businesses and ordinary people). However, more insights were needed in this way. During the validation meeting, the CEO argued that this interpretation was confusing. After an attempt to find an explanation to this variable, we decided it had to be excluded from the final findings.

Assertion 9 links

Considering links, three external factors were found to enable this set of perceptions. Firstly, open data is perceived useful for commercial service development depending on its characteristics. In other words, the characteristics outlined within Assertion 1 contribute to perceiving open data useful for commercial service innovation. A further link emerged between *Organizational Skills* and this factor. In these regards, when investigating the skills needed for transforming an open dataset into one or more commercial services, some emerged in relation to "merely access and extract the data". Following this statement, a discussion emerged during the validation meeting regarding the need of these skills for perceiving data as useful. This link has been established accordingly. Finally, given also the experience of the CEO of the case, *Local Community and Ecosystem Factors* are interpreted to have a role in perceiving usefulness of open data. These are meant to educating people about open data, facilitating its discovery and re-usage, and providing initial capabilities for designing new open data services. In the case's history, all these initiatives and experiences led to perceive usefulness of open data. On the other hand, if open data is perceived useful, it is adopted. A summary picture of these links is provided below.

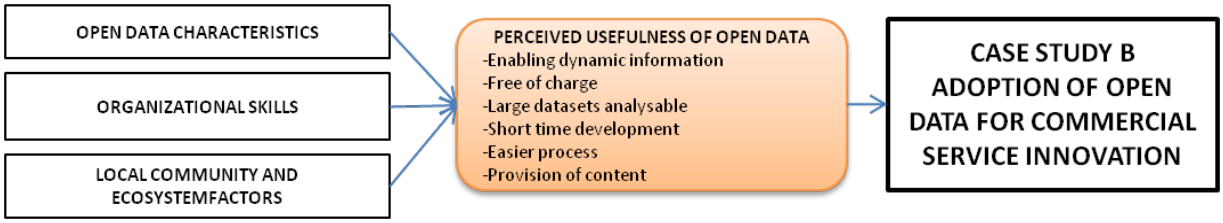

Figure 5.2: Case B Assertion 9 links

- ***Assertion 10: Perceived Business Opportunity influences the adoption of open data for commercial service innovation***

This category of perceptions derives mainly from *City-Related External Factors*, and specifically from problems and opportunities related to the city. These perceptions were observed to be strongly linked with: the user base (for generating revenues); and the ability of the company to expand its solutions to other cities (i.e. enabling growth). In this way, four variables emerged as positively influencing *Perceived Business Opportunity* (for supporting evidence see Table 5.13).

ASSERTION 10					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
10.1	*Demonstrability*	6,28; 8,20; 28,2; 29,4; 36,7; 37,2.	D1, D4	D8	Yes
10.2	*Applicability across cities*	6,8; 6,10; 6,12; 6,14; 6,31; 13,7; 22,7; 25,1; 27,2; 27,4; 27,5; 27,6; 29,1; 29,3; 29,26; 34,16; 34,17.	D1, D3, D4	D6, D8	Yes
10.3	*Funding*	20,6; 22,8; 22,9; 23,1; 23,2; 34,7.	D1, D2	D6, D7, D8, D9	Yes
10.4	*User Base*	20,5; 32,5; 35,1.	D1	D8	Yes

Table 5.13: Case B Assertion 10, perceived business opportunity

10.1 Demonstrability: this variable is believed to be fundamental for expanding the service to other cities. According to the CEO, this process is highly facilitated if the targeted city is shown the potential benefits of embracing an open data initiative. To do so, the services developed in other urban environments are used as a "demonstration for other councils of the feasibility" of the idea and to convince them to establish a relationship with the case. Presentations at local events are perceived as the best way for undertaking this process (D8).

10.2 Applicability across Cities: while the previous variable was formulated in relation to convincing other municipalities to release the data, this one is related to the service's technical characteristics. In the CEO's opinion, "scalability of the solution is highly critical" and "it is crucial for growth". By having a scalable service, new datasets can be pulled into the existing system "quickly" and "easily".

10.3 Funding: in relation to perceived business opportunity, the CEO shed light several times on the importance of funding at both the local (e.g. rewards from hack competitions) and the European level (more conspicuous fund for enabling business growth). The availability of funding is not just perceived as being connected to financial benefits, but also to other aspects such as "promotion from EU, EU-network inclusion etc.".

10.4 User Base: the last variable that describes perceived business opportunity is about the achievement of an adequate user base. Of course, having users highly influences the perceptions of a business opportunity.

Assertion 10 links

Concerning links, five external factors were found to contribute to the development of this set of perceptions. Three of these links emerged from the data extracted from the documents and the interviews. Firstly, *Individual Factors* and specifically the experience of the subject in the context were crucial for perceiving business opportunities. Secondly, the presence of an *EU-Level Strategy* and the funding opportunities within it were found and validated as enablers for this set of perceptions (see *Funding* variable). Thirdly, the ability of engaging with partners and volunteers (*Organizational*

Factors) was also interpreted as enabling beliefs about commercial opportunities in the context. In addition to these three links, two more emerged during the final validation meeting. Firstly, in relation to the characteristics of the city (see variables *Public Free Wi-Fi*, *Underexploited Areas*, and *Infrastructure Investments*), a strong link emerged with *Perceived Business Opportunity*. Secondly, the CEO highlighted the importance of the *Service Characteristics* for enabling perceived business opportunities. The rationale provided was about how having the ability of designing a valuable and scalable service is fundamental to think about starting a business in this direction. Ultimately, the presence of *Perceived Business Opportunities* was found to enable *Perceived Usefulness of Open Data.* In other words, open data is increasingly perceived useful for commercial service development if perceptions about business opportunities are in place.

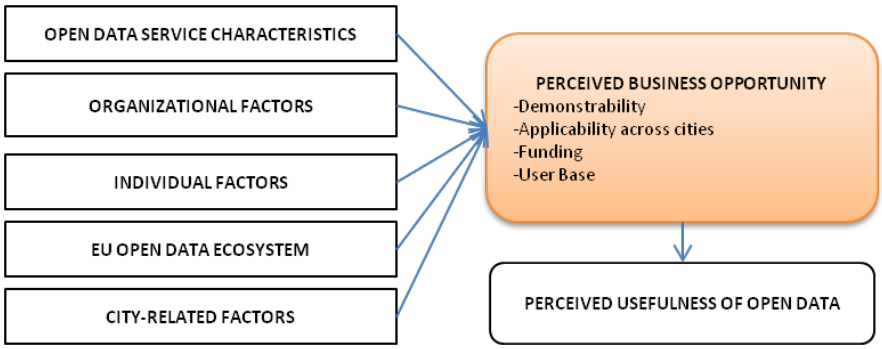

Figure 5.3: Case B Assertion 10 links

- ***Cluster I: Perceptions about Service's Users***

With respect to this cluster, three factors emerged from the analysis. These reflect the three stages of interactions between users and the service. These interactions were classified as "before" (Cluster I1), "while" (Cluster I2), and "after" (Cluster I3) using the service. In relation to the former, those KS related to the perceived user interest (as a result of incentivizing activities) in adopting the service were grouped. Secondly, KS emerged in relation to providing an "easy-to-use" experience (i.e. while using the service). Finally, the latter was formulated from a number of KS about the perceived user satisfaction enabled by actually using the service (i.e. after using the service). These variables were complemented by my personal experience using the service. The rationale provided by the CEO at the final validation meeting followed the idea that open data adoption is influenced by the ability of designing a service that enables these perceptions for users. Furthermore, the presence of these perceptions contributes to the wish of the case to continuing to use open data, i.e. achieving sustained adoption of this source.

- ***Assertion 11: Perceived Users' Interest influences the adoption of open data for commercial service innovation***

In this category, three variables (see Table 5.14) emerged in relation to actions and aspects of the service that are in place in order to stimulate potential users to adopt the service. These variables are mainly related to the integration of gamification principles into the service.

ASSERTION 11					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
11.1	*Amusement*	1,12; 2,16; 8,14; 8,15; 24,1; 29,25; 32,1.	D2, D3	D6, D8	Yes
11.2	*Rewards*	1,12; 8,14; 8,16; 32,1.	D1, D2, D3	D6, D8, D10	Yes
11.3	*Interactivity*	1,5; 2,16; 8,14; 32,1.	D2, D3	D6, D8	Yes

Table 5.14: Case B Assertion 11, perceived users' interest

11.1 Amusement: an amusing experience is designed and deployed in order to develop interest among the potential users and to stimulate their adoption of the open data services offered. This is perceived as being a concept that attracts users.

11.2 Rewards: the experience enabled by the service provides the possibility of gaining a reward for the user that successfully completes the game. This variable is believed to incentivize users to adopt the service.

11.3 Interactivity: another component emerged to be the provision of an interactive experience. Tourists are believed to be attracted by the chance of interacting with their cultural journey as opposed to be "passive observers and recipients of information". The service enables a further potential interaction between users themselves, by enabling them to share their experiences about their cultural trips. Finally, the service "allows visitors and locals to come together to chat and discover the area" (D2).

- ***Assertion 12: Perceived User-Friendly of the open data service influences the adoption of open data for commercial service innovation***

In Cluster I2, the KS that are included relate to characteristics of the service that are perceived to increase the quality of the interaction between the users and the application. Three variables emerged as describing this set of perceptions (see Table 5.15).

ASSERTION 12					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
12.1	*Easy to use*	4,8; 5,1; 5,2; 5,9.	D1, D2	D8, D10	Yes
12.2	*Interoperability across browsers*	15,15; 22,7; 29,23.	D1, D2, D3	D6, D8, D10	Yes
12.3	*Sleek Design*	1,8; 1,19; 4,8; 5,2; 5,6; 5,7.	--	D8, D10	Yes

Table 5.15: Case B Assertion 12, perceived user-friendly

***12.1 Easy to Use**:* having an easy-to-use service was critical in "adding value to it". In relation to this, the "service got cleaner and cleaner" and its "sleek design" enables users "with little or no technical skills to fully enjoy the experience".

***12.2 Interoperability across Browsers**:* the choice of HTML5 as the programming language enables running the service "on any browser and device". Thus, no specific devices and/or browsers are needed for using the service. From another angle, the availability of specific programs and/or mobile/web devices is seen as a barrier for some users to embrace the service. In addition, interoperability enabled by HTML5 is critical for enabling external people to upload their offers. "It allows any cultural attraction to be run in one single platform, but it also allows adding and editing multiple contents across multiple platforms" (D8).

12.3 Sleek Design: finally, developing a "light solution" where "data are bound into tiles" with "nice graphics" and "fancy technology" emerged as a further aspect that contributes to the successful delivery of the open data service.

- ***Assertion 13: Perceived Users' Satisfaction influences the adoption of open data for commercial service innovation***

This set of variables emerged as service characteristics enabling satisfaction of users. This is highly connected with "satisfying a want or need" through the delivery of the service. The majority of internal reports (D1, D2, D3) written by the CEO were supporting this cluster (see Table 5.16).

ASSERTION 13					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
13.1	*Gain experience efficiency*	2,9; 2,14; 2,15; 6,23; 8,4; 8,10; 8,17.	D1, D2, D3	D6, D8	Yes
13.2	*Gain experience efficacy*	1,14; 2,6; 2,9; 2,13; 2,15; 4,7; 6,23; 8,3; 8,4; 8,5; 8,6; 8,11; 8,17; 11,13; 35,1.	D1, D2, D3	D6, D8	Yes
13.3	*Learning*	1,5; 4,1; 8,3; 8,5; 8,17; 11,13; 31,5; 35,12.	D1, D2, D3	D6, D8	Yes

Table 5.16: Case B Assertion 13, perceived users' satisfaction

13.1 Gain Experience Efficiency (Time Saving)*:* this variable is mainly concerned with time saving for the user of the service. The purpose of the service, besides offering a challenging experience through cultural sites in the area, is to provide different pre-studied information-rich itineraries for tourists. These people "alternatively would spend much of their times planning and looking for attractions, and transports to reach them etc". Thus, the service offers the opportunity of significantly increasing the efficiency of their trips.

13.2 Gain Experience Efficacy (More Informed Decisions): the service offers a variety of information in relation to all the activities that tourists usually undertake (e.g. "looking for accommodation", "looking for transport", "cultural attractions in the area", "seeking to meet locals" etc.). Therefore, the service enables more informed decisions taken by the users about their visit to the city, so "increasing their stay's efficacy".

13.3 Learning: the service enables an innovative learning experience for the users. According to the CEO, "meeting with locals" is believed to be an exciting and sought touristic experience. By enabling this, the user is induced into a learning journey of the local culture that goes beyond simply visiting the main cultural attractions and heritage sites (which also implicitly entails a learning experience).

Assertions 11, 12, 13 (Cluster I) links

The presence of these perceptions (which were unavoidably developed at an advanced stage of the adoption process) was interpreted to be directly linked with (sustained) adoption of open data. Regarding the enablers, initially a unique link between *Open Data Service Characteristics* and these three factors emerged from the data collected. This idea relied on the fact that a service is perceived to be interesting, satisfactory, and user-friendly, based on its basic characteristics. However, whereas two of these links were immediately confirmed by the CEO, the one between *Open Data Service Characteristics* and *Perceived User-Friendly* was deleted at the validation stage. According to the CEO, perceptions about the service being user-friendly can be seen more as a consequence of *Organizational Skills* as opposed to the characteristics of the service itself. In other words, while the first two perceptions are seen as a consequence of the service features, the third was interpreted as a more technical-based connection. Concluding, the next figure summarizes the links emerged for these three factors defined within Cluster I.

Figure 5.4: Case B Assertions 11, 12, and 13 links

- ***Cluster J: Perceptions about the City***

This cluster encompasses all the KS describing perceptions about benefits and opportunities for the city. The CEO emphasized how, in order to adopt open data in the long run, the outcome of the final service should entail benefits for the city in which it is delivered. These benefits are perceived to motivate the councils in releasing open data and "welcoming new businesses". Among other findings, it could be distinguished a number of concepts related to potential economic growth of the area. Given this specific focus, this category was considered separately (see Assertion 15 below).

- ***Assertion 14: Perceived City Opportunity influences the adoption of open data for commercial service innovation***

In this assertion most of the variables that emerged are related to perceived opportunities by the city from the new open data that is created and released by the company as a result of running the service among an "adequate number of users". As stated in D2, the service "also allows tracking and analysing visitors' numbers and patterns through an innovative analytics engine". One further reason for Case B's objective of fostering open data is "because of its potential to aide accountability and transparency" (D4). In particular, open data "allows public participation and collaboration [variable 14.5] in reviewing government data, thereby allowing for greater transparency [variable 14.3] and promoting efficiency [variable 14.1] and effectiveness [variable 14.2] within government" (D4). The definition of the five variables that emerged in this sub-cluster is here provided (for details about evidences see Table 5.17).

ASSERTION 14					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
14.1	*Solving local inefficiencies*	1,11; 2,8; 3,1; 9,4.	D1, D2, D3, D4	D8	Yes
14.2	*Greater Efficacy in the context*	2,6; 2,8; 2,10; 3,1; 3,2; 4,1; 8,5; 8,26; 9,5; 9,6; 20,7; 26,10; 28,1; 28,2; 29,4; 29,7; 29,22; 29,24; 31,6.	D1, D2, D3, D4	D8	Yes
14.3	*Transparency*	1,21; 1,22, 1,23; 2,6; 2,7; 5,5; 7,1; 7,2; 8,26; 29,18; 29,22.	D4	D5, D9	Yes
14.4	*Competitiveness*	1,11; 2,8; 2,10; 3,1; 5,5; 8,5; 9,5; 13,5; 13,9; 26,11; 37,9; 37,10.	D4	D9	yes
14.5	*Participation and Collaboration*	1,21; 6,28; 29,19; 29,20; 29,21; 29,22; 29,24; 35,12.	D1, D2, D3, D4		Yes

Table 5.17: Case B Assertion 14, perceived city opportunity

***14.1 Solving Local Inefficiencies*:** a range of benefits for the city was described as "solving local inefficiencies". Significant emphasis was put on this variable from the potential information acquired by the city council from the data that is generated and subsequently released by the company. This information is structured from collecting and analysing data created from the interaction between users and the experiences offered by the service (including for instance, review of cultural attractions, feedback on their experiences, suggestions for improvement) as well as from the mere usage of the service (e.g. visitors' patterns, buying patterns). This is believed to better informing new actions from the city authority in this context.

***14.2 Greater Efficacy in the Context*:** in relation to this variable, it is reminded that the service was ideated and designed starting from a city-related "structured problem formulation" (e.g. the presence of underexploited cultural sites and attractions). By providing this service, Case B aims at addressing these issues, so at increasing the value of tourism in the city. Results are believed to augment efficacy of tourism initiatives, policies, and directives from the city authority.

***14.3 Transparency**:* in relation to the "secondary" open data published by the company, the city is believed to have as a result its open data portal expanded with new detailed datasets. In turn, this aspect is perceived to contribute in increasing the "city's transparency" in the context.

***14.4 Competitiveness*:** another critical aspect for the city emerged to be related to the "competitive nature of the city" itself. This variable can be decomposed in two main concepts: (1) competitiveness from open data perspective, and (2) competitiveness from the context perspective. While the latter can be inferred (i.e. providing augmented tourism experiences is expected to attract more tourists than before and so to improve competition with other cities), regarding the first a reflection is needed. According to the CEO, "cities want to state that they are embracing open data" and they "compete for being pioneers" in this field. Similarly, D4 reports that "cities want to be at the forefront of global open data initiatives" and "be positioned as a leader in the worldwide open data movement" (D4). The reasons why so much importance is given to open data in this way include attracting new businesses and investments to exploit this "huge opportunity".

14.5 Participation and Collaboration: this variable is about the need for cities to establish commitments to collaborate with the open data services business. In other words, given that the company aims at improving the local open data program, "an effort from the city-side is needed" in order for these impacts to happen.

Assertion 14 links

Of a total of five links defined for this cluster, three were added as a result of the final validation effort. Starting from those emerging from the data, the first link is defined as *City Related Factors* enabling *Perceived City Opportunity*. The CEO agreed that if an open data program is established and awareness

is spread within the council, perceptions about opportunities for the city are generated. These are mainly related to “improved transparency and competitiveness”. The second link that arose from the data collected sees these perceptions enabling *Perceived Usefulness of Open Data*. The fact that cities perceive these opportunities is seen as a key for the authorities to implement and continuously improve their open data programs. In this way, open data is perceived as a “sustainable and reliable dynamic source of information”, and so useful for being adopted for commercial service development. Regarding those links emerged during the final validation meeting, the first highlights the role of *Open Data Service Characteristics* as enablers for these perceptions. Of particular attention was the variable *Secondary Open Data Creation*. The creation of new data about the culture and entertainment context was described during the interviews as “additional detailed knowledge in this domain”. Part of this knowledge is shared with the council with the goal of enabling improvements in the internal operations in the context. Finally, *Local Community and Ecosystem Factors* as well as the establishment of an *EU Open Data Ecosystem* are thought to generate perceptions of opportunities for the city. For instance, initiatives outside the boundaries of the city council, as well as the presence of a long-term EU-driven strategy, are seen as means to “educate councils” and “to make them embracing this movement”. This is achieved through stimulating and promoting the demonstration of the potential value that can be generated through open data. A summary picture of the links to and from this factor is provided below.

Figure 5.5: Case B Assertion 14 links

- ***Assertion 15: Perceived Potential Local Economic Growth influences the adoption of open data for commercial service innovation***

The second class of variables in this cluster is dedicated to perceptions about potential local economic growth enabled by the service offered by the company and the secondary data that it produces and releases. Citing one of the internal documents collected, “we hope users will then buy a tour, scone or lunch at one of our [associated] businesses. We will continue working with them to track these upsells” (D2). In other words, businesses and people can make their offers in the application directly to the users. The final objectives are: to promote existing businesses; create new jobs (as a result of those users who want to enjoy the opportunity to be guided by a local); and stimulate the overall

economy (through the releasing of additional more-detailed open data for new service creation). Further evidence for this cluster was found in D4: open data released by Case B "can drive significant economic value in terms of job creation [variable 15.2] and new start-ups [variable 15.3]" (D4) (for details about evidences supporting this assertion, see Table 5.18).

ASSERTION 15					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
15.1	*Promote existing businesses*	1,21; 2,6; 2,10; 3,2; 6,30; 6,32; 6,33; 8,5; 9,6; 26,10; 31,1; 31,3; 31,4; 32,2; 32,3; 32,5.	D1, D2, D3	D8, D10	Yes
15.2	*Job creation*	1,21; 6,29; 6,30; 8,5; 11,9; 12,4; 29,24; 31,2; 31,3; 31,4; 32,2; 32,3; 32,5; 36,7.	D1, D2, D3, D4	D6, D8, D9	Yes
15.3	*Stimulate local economy*	1,21; 1,22; 1,23; 1,24; 2,6; 2,10; 3,2; 6,30; 8,5; 9,5; 9,6; 11,9; 26,10; 26,11; 29,6; 29,7; 29,18; 29,19; 29,20; 29,21; 30,1.	D1, D2, D4	D5, D6, D8, D9	Yes

Table 5.18: Case B Assertion 15, perceived potential local economic growth

15.1 Promote existing Businesses: the first variable emerged in relation to the business lists added to the service and exposed to the user. "Businesses are encouraged by the user base to embrace the service". The expected result is that businesses would sell more products and services. Therefore, a positive contribution to the local economy is expected to be achieved.

15.2 Job Creation*:* this variable is connected to the characteristic of the service to include the possibility of "engaging with locals and give them touring opportunities". In other words, locals can join the service and "design and sell their own tour-guides". A critical enabler for this variable is found in the unique and integrated content management system, "that allows other people and organizations to quickly and easily upload their offers" (D8).

15.3 Stimulate Local Economy*:* this variable encompasses all the expected consequences of contributing to the overall open data ecosystem by publishing new open data as a "shared value for the community". "New start-ups can jump in" and leverage the new open data that is being created. "This will be vital to generating economically viable start-ups either based on this open data or based on the advanced analysis of this data" (D4). In addition, by achieving a "valorisation of tourist attractions", the existing economy is expected to benefit as well.

Assertion 15 links

Regarding links, it is noted that all the variables in this cluster are enabled in first place by the specific characteristics of the open data services offered by Case B. The peculiarity of the service of *Engagement with Local Community* was derived from two different features: (1) the service enables

locals to upload their offers for additional tour guides; and (2) the service includes "local business listings". Regarding point (1), the service can potentially enable new job creation in the area (variable 15.2). The second, instead, is related to promoting existing businesses (variable 15.1). Finally, this set of perceptions was found to contribute to the *Perceived City Opportunity*. In other words, through the potential creation of new jobs and promotion of existing ones, and the potential for new businesses to "enter the local market", a clear opportunity for the city is perceived to be in place as a result. All these links were presented and approved without any specific objection at the validation meeting. These are proposed in the following figure.

Figure 5.6: Case B Assertion 15 links

- ***Cluster K: Other Perceptions***

With respect to this last cluster, only six KS were needed to be analysed (see Table 5.19). The common underlying topic was found to be about perceptions on the voluntary involvement in the open data landscape. Given that initially Case B was conceived as a voluntary project, these perceptions played an important role in the company's processes of adopting open data.

Perceived Voluntariness					
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Validated**
16.1	*Voluntary Use of Open Data*	1,6; 2,2; 18,2; 21,2; 32,4; 34,6.	--	D10	Yes

Table 5.19: Case B Assertion 16, perceived voluntariness

***16.1 Voluntary Use of Open Data**:* this aspect was considered to be crucial at the beginning of the case's activities when revenue models had not been established yet. Case B was initially conceived as a volunteering project. Thus, a voluntary use of open data for new service creation was in place.

Therefore, the following assertion was formulated and validated:

Assertion 16: Perceived Voluntariness influences the adoption of open data for commercial service innovation

Assertion 16 links

Perceived Voluntariness was found to be directly linked with the adoption factor. This reflects the case's experience in which open data was initially adopted at the "volunteer level".

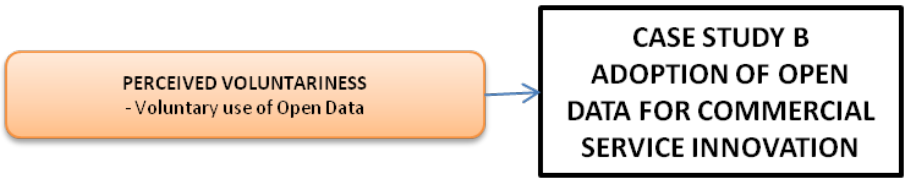

Figure 5.7: Case B Assertion 16 links

As a result of the perceptions' analysis, eight factors emerged to be critical in the adoption process of open data for commercial service innovation. A total of twenty-eight variables and twenty-four links were established across these factors.

5.4.3 Adoption Stages

The last and final category to be analysed was about the actual adoption stages the company went through from the initial generation of the idea to achieving sustained use of open data for its services' delivery. The process followed was the same used for external factors and perceptions. In total, eight adoption sub-factors were identified from the data collected, structured during the qualitative analysis, and validated at the final validation meeting (details about supporting evidence are provided in Table 5.20). Probably because these stages were already discussed during the interview, no specific comments emerged during the final stage.

ADOPTION STAGES					
Stage		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
1	*Idea Generation*	5,8; 6,2; 6,3; 6,16; 7,1.		D6, D7, D8	Yes
2	*Open Data Adoption*	1,20; 6,6; 6,14; 6,15; 6,16; 7,2; 7,3; 8,1.	D2, D4	D6, D7, D8	Yes
3	*Pilot Implementation*	1,1; 1,15; 1,18; 5,8; 5,9; 6,17; 6,18; 8,13; 8,20; 8,21.	D1, D2	D6, D7, D8	yes
4	*Acquire Funding*	20,6; 21,1; 22,6; 22,8; 22,9; 30,3; 34,1; 34,5; 34,7; 34,9.	D1	D6, D8	Yes
5	*Wider Adoption*	1,16; 1,17; 1,18; 2,4; 5,8; 6,33; 8,3; 22,8; 29,5; 34,15; 35,1.	D1, D2, D3		Yes
6	*Commercialization*	2,3; 8,2; 20,1; 20,4; 20,5; 22,8; 32,5; 34,7.	D1, D4	D6	Yes
7	*Growth*	1,4; 1,20; 1,23; 2,5; 6,28; 8,20; 13,7; 13,8; 20,7; 21,1; 22,7; 22,8; 27,1; 27,4; 27,5; 27,6; 28,2; 29,3; 29,4; 29,6; 34,4; 34,15; 34,16; 34,17; 36,7.	D1, D3	D8	Yes
8	*Sustained Use*	1,16; 1,22; 1,23; 1,24; 13,8; 22,7; 29,3; 36,7.	D1, D3		Yes

Table 5.20: Case B, adoption stages

Stage 1, Idea Generation: according to the CEO, the "idea of developing a platform for tourists came first, but the lack of data was the main issue". In other words, as opposed of a situation in which the person was facing the data available and thinking about a potential solution using this data, the idea was generated from the previous work experience in the field. The initial idea was formulated as a "dream for a solution" for the highly inefficient touristic strategy of the local council. As a result of this step a basic service platform was designed.

Stage 2, Open Data Adoption: once a basic version of the service was developed, the CEO of the company started to engage with the council and to "look for availability of data" needed to enable the service. The importance of data from the council was conceived as a form of "enabling quick development of an already existing idea" – the CEO stated. At that time (2011), a local organization organized a hack competition with the goal of exploiting local open data. According to the CEO, "the hack competition was fundamental to discover open data" and it was "found as the best way for implementing the idea and leverage the work previously done". Hence, the subject took part in this competition and won the award for the best application developed. In summary, open data was discovered, adopted and integrated into the previously defined idea, when the subject participated in the local open data hack competition.

Stage 3, Minimum Viable Product: the success of the hack competition experience led the subject to push the idea forward. The reward received was perceived as a stimulus to develop the idea further. However, given that this was not his primary work activity, and no revenues at that time were supporting the work undertaken on this initiative, "only a small amount of time could be dedicated to it". The time and money available were leveraged to develop a "minimum viable product" in which the basic application was turned "from a rough and ready to a much cleaner and easy to use service". These activities were further described by the CEO as "re-bounding the entire website and re-bounding information into tiles". The improved service was used as a demonstration for the local city authority of the potential value of providing open data to Case B.

Stage 4, Acquire Funding**:** once the pilot was implemented and tested in the local area, the company needed further funding for embedding new features that would have enabled the conceptualization of revenue models. The CEO of the case underlined several times how funding were seen as fundamental to establish the project "as a commercial one" and to expand it through new city environments. The main source for funding was found in the European Union. Hence, a key step in the process from generating the idea to achieving sustained adoption of open data for the case studied was related to applying and receiving funding that were needed for fostering "wider adoption, commercialization, and expansion across cities".

***Stage 5, Wider Adoption**:* at this stage the main goal for the company became to develop and achieve a conspicuous user base that would have enabled thinking about commercialization. To do so, Case B integrated into the service several additional open datasets in order to shift from a basic game experience into a "fully integrated tourism platform". The goal was to reach a "massive scale".

Stage 6, Commercialization: as the user base increased, the perceived commercial opportunity increased subsequently. In other words, "businesses were encouraged by the user base to embrace the service". The integration of new or existing businesses was the cornerstone to "establish B2B revenue models" on top of the open data service offered. Interestingly, the perceived commercial opportunity emerged "later in the process". This step was critical for the subsequent growth plan and for achieving sustained use of open data, which probably wouldn't have happened if the service had not the potential of producing incomes for the company.

***Stage 7, Growth*:** during the SSIs, one of the main aspects tackled was about current and planned growth. With initial funding available, and with a commercialization aspect in place, the main goal for the company became to expand the service across different cities. "To do that we need to exploit a super niche. Our super niche is public cultural attractions in European cities where more than one hundred thousand do not have currently a mobile application" (D8). One of the main characteristics that were crucial for achieving this status was the scalability of the initial service. In other words, "the big work was done once only". Another aspect that influenced this step was about "demonstrability". In the case's experience this was crucial because it was actually through short presentations and discussions with the various councils that those were convinced to release their data and commit in the long-term. "City councils need to be encouraged" and "demonstrability was crucial to convince them that it is feasible". The main variable considered by the case when targeting new cities is the availability of open data; "with the data it can be done everywhere".

***Stage 8, Sustained Use*:** the last step in this process is the achievement of sustained use of open data to feed the continuously delivered service. At this stage, the crucial factor is to ensure sustainable provision of valuable data by the related councils. The key here is to establish long-term "relationships" in which the company commits to provide those secondary open data emerging from the usage of the service, in exchange for valuable open data released on a "timely-enough basis".

5.5 Case Study B Conclusions: Adoption Factors Variables and Links

After refining the results achieved through several cycles of qualitative analysis at the last validation meeting, the case study was concluded. Overall, the findings demonstrate that sixteen factors, seventy-three related variables, and twenty-four links between factors are the most relevant in the process of adopting open data for commercial service innovation for Case B. Figure 5.8 (next page) depicts the final adoption model emerged for Case B. All factors, related variables, and links between

factors are represented. In addition, eight sub-factors emerged as describing Case B's adoption process, from the initial generation of the service idea, to the achievement of sustained use of open data for delivering the service (Figure 5.9).

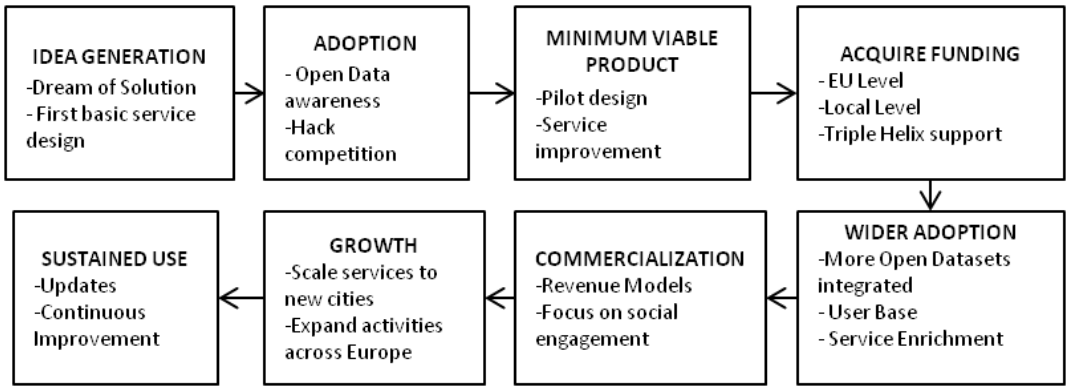

Figure 5.8: Case B adoption stages

These findings concluded the analysis of Case B. Together with the findings from the other two cases studied in this research, these represented the input for the cross-case analysis. As done for Case A, the evidences collected and leveraged were stored. These were recursively re-captured when developing cross-case assertions.

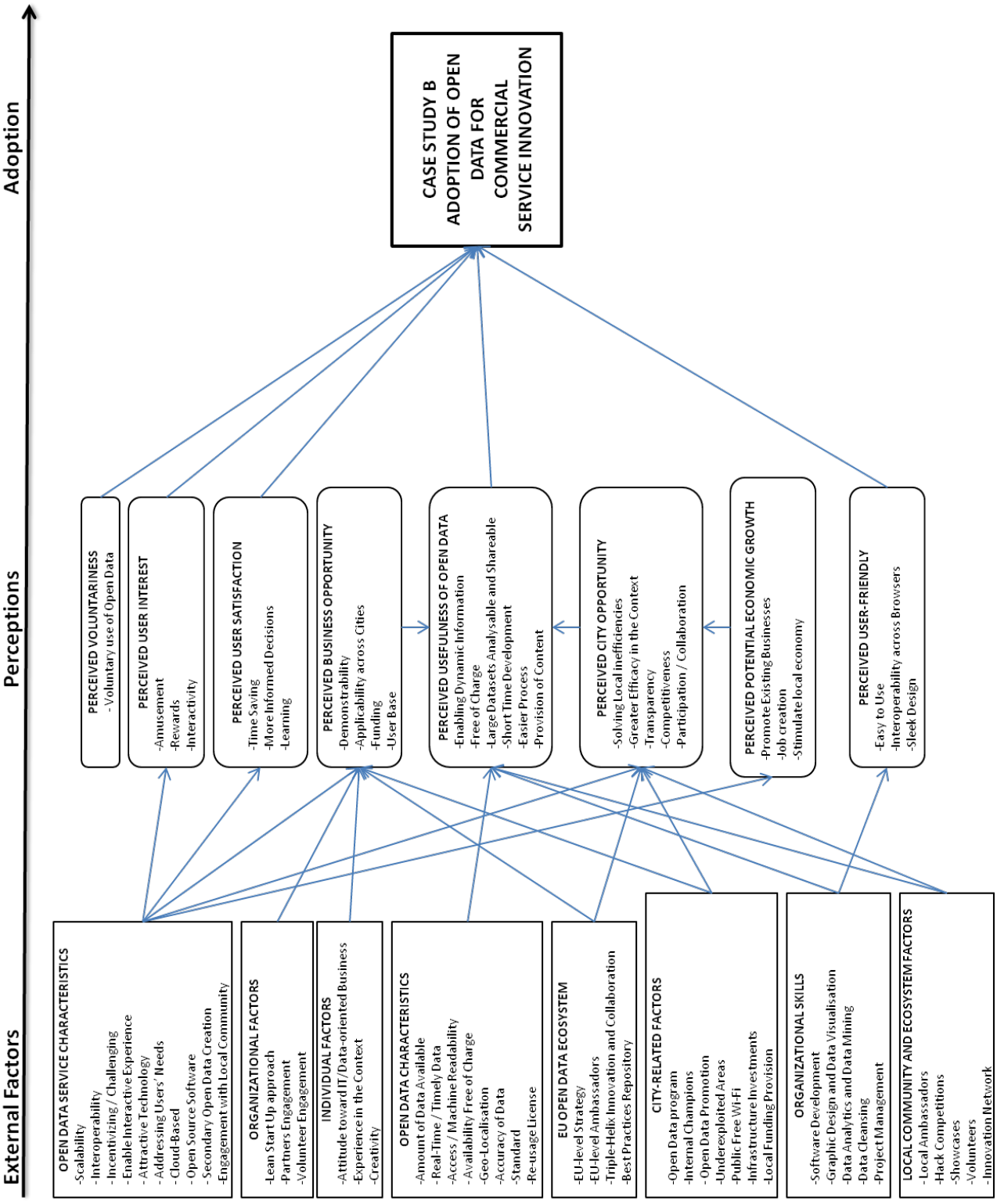

Figure 5.9: Case B adoption model

6 CASE STUDY C REPORT

This chapter is dedicated to the description of the activities and the findings related to Case C. In reporting this case study, the same structure adopted for Case A and Case B is used (see Chapters 4 and 5). In particular, after introducing the case, its history and its open data services (section 6.1), the methodological aspects customized to this case are described (section 6.2). Subsequently, the analysis related to Case C is presented (section 6.3). The fourth section is dedicated to the presentation of the findings achieved from this case study. Here, thick descriptions are provided about each factor, variable, and link that emerged as describing open data adoption by Case C. The chapter concludes with a summary of the findings achieved from Case C, i.e. the formalization of the case's open data adoption model and the outline of its adoption stages.

6.1 Introduction to Case C

Case C is the third case study that was carried out as part of this research project. This open data services business provides a range of services in the domain of parking. These focus on facilitating parking in the city by making these processes easier and quicker.

Case C was officially launched around the end of 2012. However, the two co-founders had been working on this project since the year before. The current CEO generated the idea at an open data hack competition in 2011. After this experience, the service was gradually enriched, improved, and revenue models in relation of the different features of the service were established. Currently, three people are working full-time within the company. According to the CEO, "my own background was software development not long time ago, but for the last couple of years I have been business developer commercial". The co-founder is responsible to all activities related to product design, "like service ecology, helping people using the product", he stated. He also has an "IT background". Furthermore, an expert in communication is involved full-time in the case. This person is responsible "of all the coms and social media", the CEO stated. Ultimately, two "interns" are helping on an occasional basis to undertake technical tasks.

The company operates in Europe and Australia. However, the data that is being re-used outside Europe is not consistent with the definition of open data considered in this study. In particular, some of the data re-used comes from private entities and the data taken from the city councils is not available to the public but obtained as a result of negotiations. Thus, all activities and services delivered by the company in Australia were not considered for this research. In Europe, the company offers its services in two major cities in Ireland and Spain. However, growth is pursued by the case through scaling the services across different urban environments.

The main feature of the service is meant to help drivers in the city to find the optimal parking location based on both prices and closeness. It visualises the cheapest parking space in proximity of the user. The service is enabled by a number of datasets taken from the city council and made available as open data through a dedicated online portal. These include:

- Parking locations: datasets that outline disabled and generally accessible parking bays in the city.
- Tariffs: this dataset provides structured information about tariffs for each of the parking slots in the city. Also, it indicates if clearways are in operation (and so if parking is no longer permitted).
- Location of meters: these datasets include location information of all public meters in the city.
- Revenue data from the meters.
- Multi-storey car park feeds: this dataset is described in the city council's website as a "service [that] gathers information from car parks to show you what parking spaces are available".
- Trips data: data about traffic flows in the city is described in the online portal as a service in which the "city council extracts traffic volumes from sensors located at 1100 road links across the city, which are listed in the spreadsheet and are also assigned to road centre lines in a GIS [Geographic Information System] map. Data is collected in separate excel spreadsheets".

The integration of these datasets into one content management system identifies a structured knowledge created about parking in the city. Both public and private parking spaces are included. The service is delivered to both drivers and parking operators. Overall, the open data service offered to drivers is composed of six main features (Company Website):

- Parking nearby: by sharing his or her location the user is able to find the closest parking option. Details about parking include location, category (disabled-general), tariff, and location of closest pay and display machine. When using the app, users can decide to highlight best parking rates available or to have an "ordered list by distance" (Company Document).
- Before you travel: this option is meant to enable users to search for parking close to her or his final destination prior to their departure.
- Simple payments: the application includes an "easy to use mobile payment system" (Company Document) leveraged into the private parking operators' systems.
- Parking history: parking expenditures are automatically tracked and stored by the service. It avoids users to keep receipts to track their parking expenditures.
- Directions: the service also provides "turn to turn" directions to the selected parking spot.
- Reminders: personalised reminders are available as "push notification" to the users (e.g. "if the parking time is expiring soon").

In addition, a critical part of the company's mission is to facilitate parking operators' activities by giving them the opportunity to share their parking locations and availability. This service is free of charge. However, more advanced premium listings are also offered. The first of these cases sees a situation in which the company supplies an e-payment system for the parking operator. This is believed to be a further advantage for parking businesses in terms of avoiding the development of their own applications and having an automated secure payment process. The revenue model for this open data service is flexible and depends on the actual features of the operator itself (Company Document). Tiered pricing is available to suit both large and small parking operators. For some a transaction fee is taken, for others a monthly subscription fee is charged. Furthermore, the company offers the opportunity for parking operators to promote their parking spaces among drivers. In particular, parking operators can be listed in the application for free. However, an advanced version (premium) is applied if the parking operators want to be featured in a previously defined area of the city. This feature was described as "advertising targeted by location" (Company Document). The CEO of the company explained this step as follows: "if anyone searches within those (radios) or if anyone opens the application within that area, their premium listing is featured". The company also offers the opportunity for parking operators to develop proprietary apps including the features listed above. This "white label solution" is meant to increase the operators' user bases by having their own branded app that can be also customised for their specific business needs. This solution also includes a comprehensive mobile-friendly management dashboard as well as the opportunity to enforce parking by entering the vehicle's license plate in the app.

Finally, secondary data is generated from data analytics processes undertaken upon the data extracted from the usage of the service itself. In other words, the service offers the opportunity to collect and subsequently analyse data about "how drivers behave in the city" – the product manager said. Significant results in terms of new knowledge creation from the discovery of drivers' patterns in the city pushed the company to enlarge its revenue model. This new knowledge is structured and "sold to the market" (CEO). A summary of the services offered by Case C is provided in Figure 6.1.

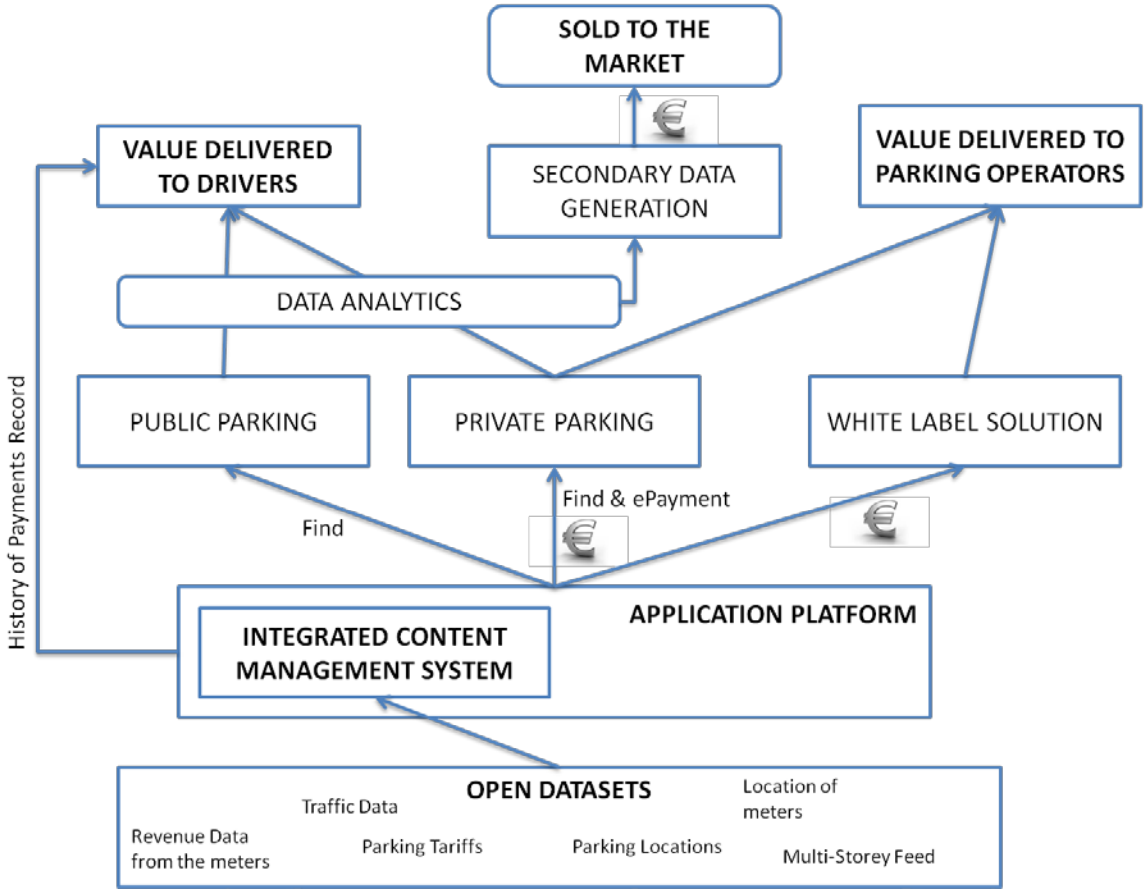

Figure 6.1: Case C's open data services

It is noted that where the "€" symbol appears, a form of revenue model is in place.

6.2 Contextualization of Methods

As done in reporting the previous cases, this section is dedicated to the contextualization of methods employed when carrying out this specific case study. According to the methodology designed for this research (see Chapter 3), three sources of data were leveraged. These are: documents, observation, and Semi-Structured Interviews (SSIs).

With respect to documents, three internal and eight external ones were used as sources of evidence. Appendix 1.3 provides a description of each document, its content, its length, and the year in which it was created.

Observation activities were implemented in the company's offices. In addition to informal conversations across the different stages of the single case protocol (see section 3.9), other activities included the display by both of the co-founders of examples of technical activities undertaken. These were critical in understanding the complex nature of the open data services offered, as well as specific procedures for extracting and managing open datasets. According to the methodology designed, this data was stored in field notes.

Regarding SSIs, both the CEO and the product manager (and co-founder of Case C) agreed to be interviewed and were found as relevant sources of information for studying the case. The third person currently working for the case was not found as a valuable source of data for the purposes of this research. This subject is responsible "only" for marketing and social media activities.

Based on the data gathered about the case before entering the SSI processes, the related protocol (see sub-section 3.10.4) was updated and augmented accordingly. Regarding step 2, I believed more clarity was needed in relation to the actual open data services offered; the actual actors involved (i.e. drivers, parking operators) and the related revenue models applied were not clear. Also, in relation to the component of the revenue model where the data generated by the application is extracted and "sold to the market" (see Figure 6.1), more clarification was needed. "What is the market? Who buys this information? And what are examples of knowledge that is derived from the open data service?" were added as further questions to the initial protocol. Regarding step 3, from the data collected it emerged that at the time of adopting open data, the current CEO of the case was already working at a technical solution for the taxi industry in the city. Then, open data about parking was adopted and the solution was contextualized to this new domain. Thus, it became relevant to explore why parking-related datasets have been chosen among others. As part of step 4, one of the aspects that emerged from several documents is about the importance for the case to have a flexible revenue model. One more question was addressed in this way in order to add clarity on if and how a particular revenue model formulation can play a key role in achieving sustained adoption of open data for commercial service development. Finally, questions were added to step 5 to explore the experience of the case in having "strong roots with the open data community and [being] regular contributors on open data policy and initiatives" (company website). The main purpose was to investigate if this experience is being leveraged by the case in some way, and if so, how. As a result of this stage, the SSI protocol was updated. This was used twice (i.e. with the CEO and the product manager).

The first two SSIs lasted for shortly more than one hour (time recorded). After conducting a first round of analysis of the transcriptions of these interactions, some aspects remained unclear. As a consequence, these were highlighted and new questions were shaped accordingly. These were the focus of the second round of interviews. These aspects included:

- The company, beside open data available through a city council's online portal, has been leveraging other datasets provided by the local authority itself for the "last couple of years". According to the CEO, these datasets "haven't been published until two or three weeks ago". In relation to this aspect, further investigation was needed about the mechanisms through which

these datasets were obtained. New information was sought about the people/departments that were approached and how these interactions were conducted.

- The case has been using two types of open data so far. Open data obtained from the city council was leveraged together with datasets released by an intermediary organization. The latter was described as a body that "facilitates the publication of data from the council". Hence, it emerged the need to explore the differences existing between the data released by these two entities.
- Although it was considered a priority for the case during its lifetime so far, the relationship with the council seems to be at the same time the biggest barrier in its experience. The cause is seen in relation to commercial partnerships previously established in the parking context. No specific questions on how these partnerships work and if/how the case has tried to pursue this way were addressed in the data collection process so far. Thus, understanding these characteristics was believed to be a cornerstone to understand the case's activities and future plans.
- Finally, in relation to the intermediary organization, it emerged that "instead of being an integrative part of the council, these should challenge the government and be detached. Like an NGO". The reasons why these statements were made were not provided by the interviewees.

These topics were addressed at the second SSI cycle. The second interviews took place in public spaces and lasted for one hour and thirty minutes (with the CEO) and one hour (with the product manager). All sessions were found to be highly productive by me and the interviewees.

At the end of this process I was able to structure preliminary findings from Case C as the input for the last stage of this case study, i.e. the validation meeting. This validation meeting was undertaken twice. First, preliminary results were presented and discussed with the CEO. Subsequently, after revisiting the feedback gained from this interaction, a validation effort has been conducted with the product manager as well. These meetings lasted for two hours and one hour respectively.

6.3 Case C Analysis

The analysis process implemented for Case C strictly followed the method previously designed (see section 3.12). A total of four hundred and forty-four KS was extracted from the SSIs (see Appendix 5.1). Similar to the other cases, a significant amount of statements was found to be relevant to more than one adoption category. For example, "open data in machine readable format enables quick time development" (KS 4,1, see Appendix 5.1) contains information about: external factors (open data characteristics, i.e. open data needs to be machine readable); perceptions (i.e. open data is perceived useful because it enables short time development of services); and about the link between the two factors that were expected to emerge in this way. As a result of the first clustering exercise: three

hundred and forty KS were categorized under external factors; two hundred and fifty-nine emerged in relation to perceptions; and sixty-two for adoption stages. Ultimately, one hundred and seventy-two KS were categorized as containing information about links between external factors, perceptions, and adoption (see Appendix 5.1).

6.3.1 External Factors Analysis

As a result of the first round of analysis, three hundred and forty KS were found to be related to external factors. The analysis of Case C's external factors proceeded with reading several times these KS and finding underlying common topics that could potentially represent clusters of analysis. This exercise took more time than compared with previous cases. As a result, ten different clusters were defined (details about the allocation of each KS to the various clusters are provided in Appendix 5.1):

- ***Cluster A:*** open data and datasets characteristics.
- ***Cluster B:*** city council related factors (independently from the open data services business).
- ***Cluster C:*** intermediary organization factors.
- ***Cluster D:*** service characteristics.
- ***Cluster E:*** skills and resources.
- ***Cluster F:*** events / community factors (outside the city council).
- ***Cluster G:*** city factors.
- ***Cluster H:*** individual factors.
- ***Cluster I:*** city council - business relationship.
- ***Cluster L:*** other external factors.

As done for the previous cases, a cluster named *Other External Factors* (Cluster L) was defined to encompass all those KS that were found not relevant within the other clusters. Through this process, more focused units of analysis were available at this stage. Also at this level, several KS were found to be related to more than one cluster. For example, from the following extract, information in relation to at least two clusters emerge: "a big challenge was that different departments [in the city council] had different datasets in terms of: content, format, structure, and frequency of update". Firstly, it can be interpreted that having a unique open data format, structure and frequency of update is crucial for Case C (Cluster A). However, also information supporting the need for the council to have an integrated open data strategy across departments is contained in this statement (Cluster B). The coding exercise proceeded separately for these categories. The following table gives an overview of the clusters emerged, the number of tentative factors and variables, and the number of retained variables for each cluster.

	Cluster	Nr. of KS	Tentative Factors	Tentative Variables	Retained Variables
External Factors	Open Data and Datasets Characteristics	113	1	8	0
	City Council-related Factors	120	1	5	0
	Intermediary Organization Factors	36	1	4	0
	Service Characteristics	50	1	5	1
	Skills and Resources	49	1	4	1
	Events / Community Factors	42	1	2	1
	City Factors	22	1	2	0
	Individual Factors	5	1	2	0
	City Council - Business Relationship	43	1	3	0
	Other External Factors	36	1	3	0

Table 6.1: Case C, external factors analysis

As a result, ten tentative external factors emerged from these steps of the analysis as influencing open data adoption for commercial service innovation for Case C. A total of thirty-eight variables were tentatively defined as descriptions of these factors. Ultimately, three variables were retained at this stage as additional clarity was needed. An in-depth discussion about the relevancy of these aspects was undertaken as part of the final validation meetings.

6.3.2 Perceptions Analysis

At the perceptions' level of analysis, two hundred and fifty-nine KS were initially grouped (see Appendix 5.1). Moreover, those KS that were found to contain information about links were also considered here. After several attempts, seven clusters were defined. These arose as areas with a certain degree of homogeneity where a significant number of concepts extracted from the SSIs, the field notes, and the documents could consistently fit. These are:

- ***Cluster M****:* perceptions about usefulness of open data.
- ***Cluster N***: perceptions about business opportunity.
- ***Cluster O****:* perceived ecosystem awareness.
- ***Cluster P****:* perceived user-satisfaction.
- ***Cluster Q****:* perceptions about commitment/trust from city councils.
- ***Cluster S:*** perceptions about city opportunity.
- ***Cluster T:*** other perceptions.

The process continued taking these clusters as more focused and manageable units of analysis. KS within each of these clusters were again grouped based on their similarity. When sufficient evidence was found, TAs about the case were formulated. Those concepts for which only weak evidence was available were highlighted, retained at this stage, and re-discussed during the final validation meetings. A summary of this process is provided in the following table.

	Cluster	Nr. of KS	Tentative Factors	Tentative Variables	Retained Variables	Tentative Links
Perceptions	Perceptions about Usefulness of Open Data	95	1	5	0	3
	Perceptions about Business Opportunity	67	1	3	0	4
	Perceived Ecosystem Awareness	48	1	2	1	2
	Perceived User-Satisfaction	12	1	4	0	3
	Perceptions about Commitment/Trust from City Councils	94	1	0	1	4
	Perceptions about City Opportunity	29	1	3	1	3
	Other Perceptions	45	1	2	0	5

Table 6.2: Case C, perceptions analysis

As shown in Table 6.2, seven perceptions factors emerged and were established as part of the preliminary findings from Case C. A total of nineteen variables arose as tentative descriptions for these factors. Furthermore, twenty-four inbound and outbound tentative links were defined at this stage. Finally, three variables were retained to be further discussed during the final validation stage. With respect to Cluster Q (i.e. perceived commitment/trust from city councils), ninety-four supporting KS were extracted from the data collected. However, no tentative variables emerged at this stage. Only one retained variable (supported initially by four KS) was identified. In fact, the remaining ninety KS were found to be related to links to this cluster. Specifically, perceived commitment and trust from the city council was found to be a cornerstone for achieving sustained adoption of open data for Case C. Therefore, a significant number of KS was related to links (in total four links emerged at this stage for this cluster) representing: (1) external factors that (if consistently in place) lead to this set of perceptions; and (2) further perceptions enabled by this factor.

6.3.3 Adoption Process Analysis

The final activity at this stage of the analysis was related to the investigation of the adoption process. A number of sub-factors was sought to explain the adoption stages that saw Case C undertaking the journey from initially generating the idea, through adopting open data, to achieving its sustained use. Evidence for these sub-factors was found in a total of sixty-two KS derived from the SSI processes (see Appendix 5.1), as well as from the documents analysed. In total, eight stages were tentatively established from the data collected and analysed (see Table 6.3).

	Adoption Stages	Number of Key Statements
1	Idea Generation	5
2	Open Data Adoption	13
3	Market Analysis	7
4	Private Beta Release	11
5	Public Beta Release	7
6	Commercialization	3
7	Growth	14
8	Sustained Use	2

Table 6.3: Case C, adoption process analysis

As a result, seventeen TAs were formulated in relation to seventeen factors that influence Case C's adoption of open data for commercial service innovation. These factors were found to be connected by a total of twenty-four links. Furthermore, each of these factors was tentatively defined at a further level of detail through a set of variables. In total, fifty-seven variables were identified. Regarding the adoption process, eight stages emerged and were outlined in relation to the different milestones the case went through from generating the idea to sustainably delivering a commercial open data service. Ultimately, six variables were retained as not strongly supported by evidence. Therefore, a preliminary model for open data adoption for Case C was formulated and was the main input for the final validation meetings.

6.3.4 Validation Meetings

This effort took place twice with the two main representatives of Case C. The process followed for both sessions was the same used for cases A and B. Three main stages were undertaken: (1) presentation of the overall preliminary findings; (2) review and discussion of retained concepts; and (3) review of links between factors.

In the first stage of each meeting, an overview of all factors and variables was given to the subjects. All the TAs generated from the data previously collected and structured were tackled one by one. When approval for each factor was achieved, questions like "do you think this set of variables fully describes this factor?", "is there something you would like to add, correct or delete from this factor?" were asked to the subjects. Unlike the previous cases, no additional factors or variables emerged from this stage of the interaction. Both of the subjects used congratulatory statements in relation to the analysis conducted. However, a few reflections were added to the final findings and augmented the general understanding of some factors. For example, both of the subjects emphasised several times how "the big key of the overall thing" is represented by perceived commitment from the city council and perceived credibility and sustainability of the overall ecosystem. Notwithstanding this, significant approval was gained for all factors and variables emerged during the analysis. Also, the fact that both

the subjects specifically requested to be able to see the findings that will emerge across cases (I inferred they believe that Case C can benefit from these outcomes), was a further demonstration of the goodness of the analysis as well as the serious involvement of these participants in this research.

The main focus of the second stage of these validation meetings was about reviewing those six variables that emerged from the analysis but for which strong evidence was not found. These were tackled one by one and a decision of whether each of these should be included or excluded from the final adoption model was collectively taken. As a result of this effort, all six retained concepts were included in the final findings of Case C. Details about these discussions are provided in the next section, i.e. when presenting the findings from this case. In the third and last stage, all the links between factors that emerged from the data collected were carefully explained to the subjects. In both meetings, this stage was the longest. In summary, twenty-two of the twenty-four proposed links were approved and confirmed by both subjects. The remaining two links were deleted and replaced with links to other factors. Finally, one more link emerged; it was proposed in first stance by the CEO during the first meeting, and agreed by the second subject subsequently.

Through undertaking this validation effort, all emerging concepts were presented and reviewed by the two co-founders of the case. As a last step the CEO proposed two personal contacts within the city councils as potential interviewees for the later steps of this research (i.e. validation across cases, see section 7.3). Given that no specific insights emerged suggesting future data collection, the case study was considered concluded and the findings about Case C were established.

6.4 Case C Findings

This section is dedicated to the description of the findings achieved from Case C. The strategy followed is the same adopted for Case A (see section 4.4) and Case B (see section 5.4). In particular, all the structured thick descriptions about factors, variables and links are provided for external factors (sub-section 6.4.1), perceptions (sub-section 6.4.2), and adoption stages (sub-section 6.4.3). All evidences supporting each component of the final adoption model is provided in dedicated tables (see Appendix 5). It is noted, that references to the documents (i.e. D1, D2, etc.) reflect the documents listed in Appendix 1.3. KS are referenced in Appendix 5.1, and where quotation marks appear (“”) a quote from one of the interviews is made.

Overall, the findings demonstrate that Case C’s adoption of open data for commercial service innovation is influenced by ten external factors and seven perceptions. Twenty-three links were established between factors. These factors emerged as being described by a total of sixty-three variables.

6.4.1 External Factors

All the initial ten clusters defined for this category (see Table 6.1) were translated into actual external factors that influenced Case C in adopting open data for commercial service innovation. The thirty-eight tentative variables initially emerged were confirmed at the validation meetings. From these interactions, three more variables were integrated. The findings are described in the following paragraphs, consistently with the format used for the previous cases.

- ***Assertion 1: Open Data and Datasets Characteristics influence the adoption of open data for commercial service innovation***

The first cluster encompassed one hundred and thirteen KS emerged from the SSIs in which aspects related to the actual datasets re-used, were tackled. The evidence that emerged from the data collected allowed outlining a range of eight variables explaining the open data and datasets characteristics required for Case C. The evidences supporting each variable are provided in Table 6.4.

ASSERTION 1					
Variables		**Key Statements**	**Int. Docs.**	**External Docs.**	**Validated**
1.1	*Availability Online*	2,1; 2,3; 2,10; 2,14; 2,16; 3,1; 3,2; 12,4; 17,13; 35,2; 36,6; 36,8; 38,4; 41,1; 41,8; 45,2; 45,4; 45,5; 45,10; 45,11; 46,2; 46,4; 47,5; 47,6; 47,7; 47,8; 48,2; 50,1; 51,2; 51,4; 54,1.	D4.	D6; D8.	Yes
1.2	*Amount of Data Available in the Context*	2,2; 2,6; 2,8; 2,9; 2,13; 2,18; 12,5; 12,6; 14,3; 14,4; 14,5; 14,7; 15,1; 21,2; 21,6; 21,10; 21,15; 23,2; 27,1; 33,11; 38,4; 41,1; 41,8; 45,2; 45,4; 45,5; 45,10; 45,11; 46,2; 46,4; 47,5; 47,6 ; 47,7; 47,8; 48,2; 50,1; 51,2; 51,4; 54,1.	D4.	D6; D8.	Yes
1.3	*Structure – Meta Data*	2,4; 2,5; 15,5; 16,3; 16,4; 35,1; 45,6; 45,7; 47,1; 47,2; 47,4; 50,3; 51,1; 51,3; 54,7;	--	D8.	Yes
1.4	*Frequency of Update*	2,11; 17,3; 17,4; 21,1; 21,9; 45,8; 45,9; 54,5; 54,11; 64,1.	D1.	D7.	Yes
1.5	*Machine Readability*	4,1; 4,3; 47,2; 54,2; 54,6.	--	D8; D10.	Yes
1.6	*Standard / Format*	4,2; 15,2; 15,3; 15,4; 15,6; 15,7; 16,4; 33,7; 34,4; 34,9; 45,4; 45,6; 45,7; 47,1; 47,2; 47,3; 47,7; 47,8; 48,2; 50,1; 51,4; 54,3; 54,4.	--	--	Yes
1.7	*Data Accuracy*	16,1; 16,2; 16,5; 38,1; 52,1; 53,2; 54,5.	--	--	Yes
1.8	*Open License*	7,1; 9,14; 24,1; 24,3; 24,4; 24,5; 24,6; 24,7; 24,8; 24,12; 50,2; 61,1; 61,3.	--	D6.	Yes

Table 6.4: Case C Assertion 1, open data and dataset characteristics

This result was largely expected given the enabler role that open data plays in the service offered by the case. However, supporting evidence from the documents analysed was low. This fact reflects the

company's hesitation to highlight open data as the main enabler of the service delivered. This aspect is taken into account and further explained within a specific perception (i.e. Assertion 17).

***1.1 Availability Online*:** this variable outlines the importance of having open data freely available through online portals. Two different bodies supplied different datasets that are currently leveraged by the case. Some datasets are taken directly from two departments of the city council (urban planning and transportation). However, the majority of datasets re-used are released by an intermediary organization that takes "raw data" from the council and facilitates its releasing processes.

***1.2 Amount of Datasets Available in the Context*:** this variable played a crucial role since the beginning of the case's history for several reasons. The idea of the service was generated from the open data that was available at a hack competition. At that time, an idea for an application supporting the taxi's industry was in the mind of the current CEO of Case C. However, the fact that multiple parking-related datasets were available led the subject to shift from assisting people in the city to find a taxi to "facilitate parking in the city". According to both of the founders of the company, "the majority of data was about this context" and since "open data provides the content for our service, the biggest dataset was chosen because of the amount of information within it". The case is currently pushing the council to release additional data about parking. These include "data that can be gathered from existing systems" (e.g. parking meters) as well as data that can be captured by investing in "sensors to be placed around the city".

***1.3 Structure and Metadata*:** a structured dataset is seen by the case as something that can be understood and navigated without "large time consumed to figure out what is what". A dataset is structured if it includes "metadata to describe and guide the exploration of the dataset" as well as the descriptions where "all the standards" are displayed (e.g. format, location's standard used).

1.4 Real-Time / Frequency of Update: for the case, data needs to be "current" in order to avoid the provision of a non-effective service. One of the service's goals is to "empower dynamic pricing and trigger targeted events such as promotions based on real-time payment trends" (D7). However, for some datasets (e.g. location of meters) "real-time is not fundamental, but once a year the data has to be cleaned". For others (e.g. "revenue data from the meters"), the higher the frequency of update is, the greater value is perceived in relation to that specific dataset.

***1.5 Machine Readability*:** for enabling the open data services offered by the case, the datasets have to be readable by machines. The company uses datasets in KML/csv format and "supported by APIs".

1.6 Standard / Format: in the past the company encountered "significant difficulties" from collecting datasets that had different (and not clearly described) standards. For example, in relation to geo-location of data, some datasets re-used were available through the "common longitudes-latitudes system", while for others a local standard system was used.

1.7 Data Accuracy: accuracy of data influences directly the quality of the services offered by the case. If errors are in place in the dataset these would be reflected in the data displayed by the application (e.g. an error in data representing a parking space would lead to the display to the drivers of a parking space that doesn't exist). This aspect was tackled by the case by "extensive and time-consuming data cleaning processes" and "lot of field work was done to improve the dataset". However, inaccuracy of data was "an issue but not a major one, because it could be fixed".

1.8 Open License: concerning the datasets re-used by the company, the majority of these were covered by the European Public Sector Information (PSI) license. According to the interviewees, "restrictions should be limited to privacy concerns", and "commercial openness is critical to sell data on the back of open data". In these regards, the interviewees underlined Creative Commons as the optimal solution, and specifically its "Attribution Formulation".

- ***Assertion 2: City Council Factors influence the adoption of open data for commercial service innovation***

Cluster B emerged from one hundred and twenty KS. These were connected to the structure, the people, and the activities of the city councils where open data is taken. Therefore, several variables emerged in relation to characteristics that city councils need to have in order to ensure a long-term provision of open data, which is translated in enabling sustainable delivery of the open data services. The concepts that were categorised under this cluster relate to variables within the council that should be in place independently from the case for perceiving open data as useful. Interestingly, major evidence emerged from the SSIs, while only little information about how city councils should operate was found within the documents analysed (more details are provided in Table 6.5).

2.1 Non-Fragmented Strategy**:** city councils need to have an established, structured, and long-term open data program. Some recommendations on how these strategies should be structured emerge from the case's experiences. One of the challenges faced by the company was in relation to the data made available from different departments. As a result of work conducted within different areas of the council, the final open datasets were diverse in terms of: "access, contents, standards, formats, quality, and associated licenses". Thus, people within the case (also based on experiences with other cities) suggest the need for councils to have a unique open data strategy that has "a central fund body across departments" with a "unique open data committee".

ASSERTION 2					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
2.1	*Non-Fragmented Strategy*	1,2; 1,4; 2,1; 2,3; 2,6; 2,7; 2,10; 3,2; 17,3; 17,13; 21,2; 24,9; 34,5; 34,6; 35,2; 35,3; 36,6; 36,7; 37,3; 39,1; 41,1; 41,4; 41,5; 41,7; 44,1; 44,2; 44,3; 45,2; 45,4; 45,5; 45,6; 45,7; 45,8; 46,2; 46,3; 46,4; 48,2; 51,4; 52,2; 55,3; 63,2; 63,3; 63,4; 63,5; 63,6; 65,1; 66,1; 66,2; 66,3; 66,4; 66,5; 66,6; 66,7; 66,8.	--	--	Yes
2.2	*Fostering Engagement*	1,3; 12,2; 28,1; 40,3; 40,6; 42,7; 2,8; 42,9; 45,11; 46,1; 60,2; 60,3; 60,4; 60,5; 60,8; 60,9; 60,10; 60,11; 63,1; 65,4; 65,5; 65,6.	D4.	D6; D9.	Yes
2.3	*Open Data Awareness*	12,1; 28,2; 35,4; 35,5; 35,6; 38,3; 38,4; 38,5; 40,3; 40,6; 42,7; 42,8; 42,9; 45,11; 60,2; 60,3; 60,4; 60,5; 60,7; 60,8; 60,9; 60,10; 66,10.	--	D6.	Yes
2.4	*Investments*	13,3; 21,6; 21,7; 22,1; 40,5; 42,5.	--	--	Yes
2.5	*IT Skills / department*	21,4; 21,5; 35,6; 46,1; 66,9.	--	--	Yes

Table 6.5: Case C Assertion 2, city council factors

***2.2 Fostering Engagement*:** another aspect in the company's experience is related to events and workshops organised by the local city council. Differently from events organised outside the authority (see Assertion 6), in this case councils should focus on opportunities and current situations about specific public-related-industries (such as parking). The final objective should be "stimulating the emergence of valuable ideas in these contexts". In summary, periodic meetings should be organised by the city council as effective ways of promoting open data, fostering engagement with actual and potential users, and getting aware of what the user's needs are.

***2.3 Open Data Awareness*:** for city councils it is also critical to be aware of open data in terms of what it is, its potential, how it should be managed and released, and who and how is using it. These concepts should be in place at all stages when users interact with open data. Councils should be able "to understand what are the needs of open data users" for implementing an idea as well as for sustainably deliver an already available open data service.

***2.4 Investments*:** the CEO highlighted the importance of having "sensed data" as a further source of open data. This class of data would enable more detailed near real-time knowledge about the specific context. This entails substantial "infrastructure investments from the council" for enabling the collection of data from these sensors.

2.5 IT Department / IT Skills: councils are found to positively influence adoption of open data if internal IT skills are in place. These skills are critical in taking raw data and transforming it into a dataset with the characteristics outlined above (see Assertion 1). The presence of an established IT

department is believed to be important for ensuring that the most appropriate software is used in this process. Finally, the presence of such a function enabled the company to have an informed and effective interlocutor when approaching open data. In other words, the case approached "the city CIO, or the head of IT if a CIO was not in place" in order to be able to effectively communicate technical aspects, which otherwise would have to be carefully explained.

- ***Assertion 3: Intermediary Organization Factors influence the adoption of open data for commercial service innovation***

Cluster C was populated with thirty-six KS related to the role that "intermediary organizations" play in terms of provision of open data to Case C and in relation to the overall ecosystem. These organizations' purposes are taking open data released by the council and developing portals through which these datasets are made available to the public. "More work is done on the data if compared to the one released by the council". Similarly, to what emerged in relation to the need of a structured initiative from the council (see Assertion 2), variables in this cluster indicate the need for these organizations to have knowledge about open data, to foster engagement, and to have long-term plans. However, an interesting aspect emerged also in relation to the nature of these entities and particularly on the type of connection with the city council. Table 6.6 provides an overview of the variables formulated for this cluster and the related supporting evidences.

ASSERTION 3					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
3.1	*Long-Term Structured Plan*	1,2; 1,4; 3,1; 23,2; 36,6; 37,3; 41,1; 41,4; 41,8; 57,9.	--	D8.	Yes
3.2	*Workshops / Level of ecosystem engagement*	1,3; 22,2; 40,3; 45,11; 57,1; 57,7; 57,8; 59,2; 59,3; 65,5.	D4.	D6; D8.	Yes
3.3	*Level of Engagement / Independence with the City Council*	22,2; 22,3; 22,4; 22,5; 22,6; 23,1; 40,3; 45,11; 57,2; 57,5; 57,6; 57,7; 57,8; 59,2; 59,3; 65,5.	--	--	Yes
3.4	*Internal Open Data Knowledge / Skills*	50,1; 50,2; 50,3; 51,1; 51,2; 51,3; 52,1; 53,2.	--	D8.	Yes

Table 6.6: Case C Assertion 3, intermediary organization factors

3.1 Long-Term Structured Plan: similarly, to the need of establishing long-term commitment in releasing open data from the council, also these intermediary organizations should operate upon structured strategies. This was found as a critical aspect for perceiving open data to be a sustainable source for the company's offer.

3.2 Workshops / Level of Ecosystem Engagement: as also stated for city councils (variable 2.2), engagement with the overall ecosystem should be enabled and facilitated by these organizations.

These initiatives should have as a central goal to "educate the councils" and "to stimulate the council to do more work" based on the actual needs of demanders. Based on Case C's history, in the case of events and interactions led by the councils, the actual contents of these meetings were based on activities and results that local authorities had previously undertaken and achieved. Differently, for initiatives led by the intermediary organization, "challenges were posed to the council" in terms of stimulating the authority to release data previously retained for some unclear reasons.

***3.3 Level of Engagement with the Council*:** another variable emerged for these bodies in relation to their level of dependency from the related council. This understanding was achieved by the CEO from experiencing several of these entities across Europe. The CEO underlines the importance for these organizations to be "independent NGOs in order to be able to act as a counterpart for the city councils". This is believed to be critical not just for pushing local authorities to implement effective open data programs, but also to give them the ability of being "more vocal". In this way, issues and challenges in place within the authorities (which might be "voluntarily omitted by the councils") are expected to emerge and being addressed. A reflection on "motivations of the city councils" and how this relates with challenges posed to these authorities is provided as part of Assertion 17.

3.4 Internal Open Data Knowledge and Skills: the goal of these organizations should be "refining the open data provided by the council" and fostering its re-usage. By acting as intermediary between the council and the open data service developers, this body is expected to have a "360 degrees understanding of the overall ecosystem". Achieving this status allows to "effectively find the balances between suppliers of the data and demanders".

- ***Assertion 4: Service Characteristics influence the adoption of open data for commercial service innovation***

Cluster D emerged from fifty KS extracted from the SSI transcripts as well as from all the documents analysed. All this information was about characteristics of the open data services influencing Case C's adoption of open data. In other words, open data was found to be sustainably adopted based also on the actual features of the service designed. A set of six variables was defined within this factor (details about supporting evidence are provided in Table 6.7). Variable 4.6 (underlined in the following list) was included only after the validation meetings.

***4.1 Valuable Service in the Context*:** one aspect that fosters adoption of open data is intrinsically related to the ability of developing a valuable service in the context. The services offered by Case C are meant to be valuable for drivers and private parking operators (see Figure 6.1). In summary, for open data to be re-used for commercial service innovation, this service needs to deliver some value to a range of end-users.

ASSERTION 4					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
4.1	*Valuable Service in the context*	2,2; 5,1; 6,1; 7,1; 17,1; 17,2; 17,3; 17,5; 17,8; 24,6; 27,1; 33,11; 61,1;	D1; D3; D4;	D6; D7; D8; D9; D10;	Yes
4.2	*Secondary Data Generation*	2,11; 2,13; 2,15; 2,18; 17,10; 17,12; 21,15; 24,5; 30,3; 31,2;	D1; D3; D4;	D6; D7; D8; D9; D10;	Yes
4.3	*White Label Solution*	5,2; 8,1; 8,2; 8,4; 8,5; 8,6; 33,8;	D1; D3; D4;	D6; D7; D8; D9;	Yes
4.4	*Flexible Revenue Model*	5,3; 5,4; 5,5; 6,2; 7,2; 7,3; 24,5; 30,5;	D1; D4;	D8; D9;	Yes
4.5	*Unique Content Management System*	8,1; 8,4; 8,5; 21,15; 33,3; 33,4; 33,5; 33,6; 33,7; 33,9; 33,10; 37,2; 41,2;	D1; D2; D3; D4;	D8; D9; D10;	Yes
4.6	*Complexity*	27,2; 36,8; 47,8; 56,4;	--	--	Yes

Table 6.7: Case C Assertion 4, service characteristics

***4.2 Secondary Data Generation*:** another aspect was related to the importance of designing a service that enables "generating new knowledge on the back of open data" (or "modelling new knowledge about how cities are used" (D7)). This knowledge consists of discovering previously unknown patterns on how drivers behave in the city. For example, by aggregating information on "how long people stay in a specific parking, how many people look for parking in certain areas and so on", patterns about drivers' behaviours can be extracted. The knowledge that is gradually created by aggregating this data is a critical part of the case's activity in two main directions. On one side it can be "sold to retailers". On the other hand, it can contribute to the council's operations through enabling a more efficient and effective management of the parking context.

***4.3 White Label Solutions*:** a critical section of the case's offer is the "white label solution" that is proposed for private parking operators. This is designed from the company's central platform in order to ensure "quick and prompt development" as well as "to protect our brand". This personalised solution maintains the main structure of the case's application and its logo, with limited additional graphical features that can be added at the discretion of the customer.

***4.4 Flexible Revenue Model*:** Case C applies three different revenue models. Beside the development of a personalised solution (which is a fully-premium service), the company's income also comes from "business analytics from the data collated through the app" (D9). Finally, the company applies a "flexible revenue model" from the payments enabled at private operators' parking facilities. The term "flexible" is being used because the revenue model in place "is tier-based in relation to the client's dimension". Overall, having a flexible revenue model in the open data service offered is believed to be an important aspect for the case to be able to sustainably use this source.

4.5 Unique Content Management System: this technical characteristic was found to be fundamental for the case (among others that emerged as not-particularly relevant). Having a unique content management system is crucial to develop a "cross-platform" (D1) upon which "branded apps" (D1) or personalised solutions are built. Also, having a single system increases the "scalability of the service". This characteristic is believed to be important for expanding the offer to other cities.

***4.6 Complexity*:** four KS emerged in relation to the importance of having a "complex service". However, some aspects of this variable remained unclear and no sufficient evidence was initially found. As a consequence, this variable was established only after being re-discussed during the validation stage. In general, it is perceived that open data intrinsically has a risk, that is: "it is ready". This is perceived by the CEO as a risk that the overall company's activities might face in two main directions. Firstly, if no major work is conducted on the original open data, "other businesses might easily replicate the service". Secondly, especially at the time in which the company was looking for investors, the value of the service delivered would be under-estimated. An example was proposed of an American company involved in parking-related open data services. This business was promoting its service without mentioning that open data was re-used within it. In fact, these people were describing their database as being built by field work and creative ways of collecting data about parking spaces and locations. The results were the same as the ones from Case C (in terms of being able to detect parking spaces and meters in the city). However, by making the service more complex an increased value of the company was perceived among the community. These reflections suggest that complexity of the service is believed to positively affect the open data services business success.

- ***Assertion 5: Organizational Skills and Resources influence the adoption of open data for commercial service innovation***

When analysing the SSIs' transcripts, forty-nine KS emerged in relation to the skills that were needed within the case to transform the open datasets into the final services. Further sets of skills were found to be relevant in managing the relationships with the overall ecosystem (particularly with the council and the intermediary organization). Other important sources from which this cluster was generated came from D1 and D9 (for more details see Table 6.8). In these documents the people working at the case and their roles are outlined as well as their previous experiences. It is noted that variable 5.5 was included only after the final validation stage.

ASSERTION 5					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
5.1	*Data Analytics and Management Skills*	2,4; 2,5; 4,2; 9,15; 9,16; 9,18; 15,8; 16,2; 17,10; 29,1; 29,2; 29,6; 29,7; 29,8; 29,9; 36,8; 37,2; 38,1; 47,1; 47,7; 52,1;	D1; D3;	D7; D8; D9; D10;	Yes
5.2	*Programming and Graphic Design Skills*	8,1; 8,6; 9,5; 9,9; 9,12; 9,13; 15,7; 17,10; 29,3; 29,4; 29,5; 29,9; 47,6; 47,8;	D1; D3;	D7; D8; D9; D10; D11;	Yes
5.3	*Lean Start Up Approach / Lean Team*	9,1; 9,2; 9,3; 9,4; 9,17; 11,1; 32,1; 32,2; 32,3; 43,2;	D1; D2;	D8; D9;	Yes
5.4	*Business Development and Communication Skills*	9,6; 9,10; 33,9; 52,2;	D1; D2;	D8; D9;	Yes
5.5	*Legal Knowledge*	9,14;	D4;		Yes

Table 6.8: Case C Assertion 5, organizational skills and resources

***5.1 Data Analytics and Management Skills*:** this set of skills was found to be critical in several stages of the company's activities. General knowledge about data and databases was leveraged to understand the actual datasets. Then, in order to solve "inaccuracy issues", data cleaning skills were employed on the original dataset. Subsequently, data integration and database management skills were required to load the open datasets and integrate these into the content management system. From there, data visualisation skills were needed to graphically display this data. Finally, data mining and data analytics skills are leveraged to extract knowledge from "data collated through the app" (D9) and so to enable *Secondary Data Generation*.

***5.2 Programming and Graphic Design Skills*:** in order to develop a mobile application, programming skills are needed. These include both iOS and Android languages (D11). Two more sets of skills are included in this variable. Firstly, "graphic and product design" were critical in designing "an application that is easy to use and where information is easy to detect". Secondly, without the internal availability of GIS knowledge, the open data service couldn't have been designed. In this way, a standard could be applied on the "geo-codes, which were not in the form latitude-longitude". To cope with accuracy issues, field work was undertaken "and geo-codes were reversed to understand how many spaces were actually available". Thus, extensive GIS knowledge was required.

***5.3 Lean Start-up Approach*:** another variable that resulted to be relevant for this cluster is represented by the importance for the case to have a "lean start-up approach". The idea is to design the team based on the process and empower people based on their skills. Three people are currently

working in the company with three distinct sets of skills and subsequent responsibilities. Volunteers and interns are engaged (mainly from academia) to "help out with technical stuff".

5.4 Business Development and Communication skills: according to the CEO of the case, beside "business development skills", also communication skills were required especially in establishing and maintaining the relationships with the city council, the intermediary organization, and the industry partners. Also, effective communication of the actual activities undertaken by the company was needed for convincing investors about the value of the service proposed.

***5.5 Legal Knowledge*:** from the data collected it could be inferred that legal knowledge about the actual re-usability of open data was required. In particular, the case had to be aware of the commercial constraints associated with open data re-usage. Also, legal knowledge is leveraged into the service itself. This means complying with road safety legislation (e.g. "do not use the app while you are driving") and in terms of information protection on the data collected on the back of the service's usage. Moreover, as stated in D4, the company also acts as a "contributor to the open data policy making process" (D4). In other words, this knowledge is used to canalise policy makers' activities towards fostering the development of a sustainable open data services ecosystem.

- ***Assertion 6: Events and Community Factors influence the adoption of open data for commercial service innovation***

In this cluster, forty-two KS underline the importance of having initiatives in place that are led and organised by actual and potential open data users as well as by other organizations or consortia (i.e. outside the city council and the intermediary organization). Councils "should be invited" and be educated about wants and needs of re-users, as opposed to be active promoters of their open data initiatives. Of the three variables emerged, one was initially retained (variable 6.3). This was included as a result of further discussions undertaken with the case's people. Details about supporting evidences for this assertion are provided in Table 6.9.

ASSERTION 6					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
6.1	*Workshops / Engagement Opportunities*	1,1; 1,3; 24,1; 34,7; 35,4; 35,5; 35,6; 36,3; 38,3;38,4; 38,5; 40,3; 60,1; 60,2; 60,3; 60,4; 60,5; 60,6; 60,7; 60,8; 60,9; 60,12; 60,13; 60,14; 60,15; 60,16; 60,17; 60,18; 63,1;	D2; D4;	D6; D7; D8;	Yes
6.2	*Hack Competitions*	12,1; 12,2; 12,3; 12,4; 12,5; 14,3; 14,4; 14,5; 14,7; 32,4; 32,5; 36,3; 41,3; 41,7;	D2; D4;	D6; D7; D8; D9;	Yes
6.3	*Champions*	39,3;	--	--	Yes

Table 6.9: Case C Assertion 6, events and community factors

***6.1 Workshops / Engagement Opportunities*:** according to the product manager, organising "meetings and workshops outside the councils is much more effective"; "cities can be met" and initial engagement can be enabled. More emphasis was however placed in describing these initiatives as ways of "educating and stimulating the councils". To do this, actual and potential users present their ideas and needs. Subsequently, the open data needed is investigated and asked. In this way, "face to face interactions are much more effective" than other forms. In addition to events organised by open data users, also other events fall into this category. For example, as observed from the case's experience, by joining initiatives led by local enterprise associations, the "company was helped in establishing milestones and motivated to achieve them" (D2).

***6.2 Hack Competitions*:** another key role within this cluster is played by hack competitions. This experience was leveraged by the case in two main ways: (1) "to know and get aware of open data"; and (2) to "receive buckets of training, mentoring and support with the added benefit of some seed funding to help get the project off the ground" (D2). However, both of the interviewees stressed that "hack competitions alone are not a viable option for spreading open data".

***6.3 Champions*:** one KS emerged to be related to the key role that should be played by "external champions", i.e. from outside the city councils. The idea behind this thought is that "internal champions could help us to some extent, but their role was limited". External champions are believed to overcome the issues that "champions within councils face in terms of ensuring fair competition" and simultaneously support open data services businesses.

- ***Assertion 7: City-Related Factors influence the adoption of open data for commercial service innovation***

Across the data collected significant evidence was found in relation to city's characteristics influencing the development of commercial services based on open data. This statement can be further demonstrated by the fact that, as part of its "growing process" (see sub-section 6.4.3), the company targets new cities for expanding its services. Together with the availability of open data in these cities (see Assertion 2), two more variables are considered in this choice independently from the activities of the city councils (see Table 6.10). For example, cities in which parking is not found as a time-consuming process and/or where payment for parking is not required (and so parking meters are not in place) would significantly limit the value provided by Case C's open data service.

***7.1 Market Size*:** when describing future activities and related growth opportunities across cities, one of the main variables considered is market size (D9). Before adopting open data, "extensive market analysis is conducted on potential cities to be targeted". The successful delivery of the open data service is a consequence of a city-related problem and opportunity. These include "traffic and parking

congestions", and so relative short number of parking spaces compared to the number of cars driven in the city. Additional aspects include "smart-phone penetration" and general "socio-economic status of the city".

ASSERTION 7					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
7.1	*Market Size*	11,5; 18,3; 18,5; 18,6; 31,3; 33,3; 33,10; 44,6; 44,7; 44,8; 44,9; 44,10; 44,11; 44,12;	D1;, D2;	D9;	Yes
7.2	*Existing Infrastructure*	18,4; 31,6; 34,5; 34,6; 44,1; 44,2;	D3;		Yes

Table 6.10: Case C Assertion 7, city-related factors

***7.2 Existing Infrastructure*:** in relation to this variable, an observation was made about the need for cities to have in place a parking payment system in order to be targeted for the delivery of the open data service. Hence, cities need to have parking plans as well as a network of meters in place.

- ***Assertion 8: Individual Factors influence the adoption of open data for commercial service innovation***

Cluster H encompasses all the KS emerged in relation to the characteristics of the individuals that actually founded the business. These characteristics have been observed and interpreted as playing a role in the adoption of open data. Strong evidence about all variables in this cluster was found in the documents that were collected and analysed (see Table 6.11).

ASSERTION 8					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
8.1	*Entrepreneurial Spirit*	9,7; 11,1;	D1; D2; D4;	D5; D8; D9;	Yes
8.2	*Attitude Toward IT Business*	9,8; 24, 1; 20,4;	D1; D2;	D5; D8; D9;	Yes

Table 6.11: Case C Assertion 8, individual factors

8.1 Entrepreneurial Spirit: both of the co-founders of the company have highlighted their passion for entrepreneurship. According to them, without this passion and huge dedication no major results could have been achieved. These statements are also demonstrated by the active role that both of these people are still playing in the local start-up ecosystem.

8.2 Attitude towards IT Business: the second concept was also crucial to lead the subjects to consider the development of digital-related solutions as the major focus of their working lives. According to the CEO, "without this attitude we wouldn't have even known about open data".

- ***Assertion 9: Council-Business Relationship Factors influence the adoption of open data for commercial service innovation***

The fact that working "closely with the city council" (D4) was a cornerstone for the company is the common underlying assumption made under this cluster. These topics were emphasised during the SSI processes. According to the CEO, "an established relationship is needed if you want the data". Moreover, important considerations emerged in relation to the characteristics of these interactions as well as the specific mechanisms that made this collaboration successful. Forty-three KS were included in Cluster I (more details about supporting evidence are provided in Table 6.12).

ASSERTION 9					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
9.1	*Long-Term Commitment on Open Data Provision/Usage*	2,3; 2,6; 2,7; 2,10; 2,11; 2,13; 3,2; 17,12; 17,13; 38,7; 39,1; 55,3; 56,4;	D1; D4;	D6; D8;	Yes
9.2	*Collective Learning*	1,1; 20,5; 26,2; 38,2; 38,3; 38,4; 38,5; 40,6; 44,3; 55,1; 60,7;	--	D8;	Yes
9.3	*Structure of Relationship / Partnership*	19,1; 20,1; 20,2; 20,3; 20,6; 20,8; 25,2; 26,1; 26,8; 26,9; 38,7; 40,4; 40,5; 40,7; 42,1; 42,10; 44,3; 55,1; 55,9; 55,10; 56,1; 56,2; 56,3; 60,6;	--	--	Yes

Table 6.12: Case C Assertion 9, council-business relationship factors

9.1 Long-Term Commitment on Open Data Provision / Usage: this variable highlights the need for the company to ensure that open data will be sustainably available. To do this, a relationship with the council needs to be in place. According to the case's experience, in order to achieve long-term commitment from councils, the open data service business needs to provide some value in return. The company leveraged the potentialities for enabling more efficient and effective management in the context from the secondary data generated by the service. Hence, in exchange for the ensured provision of open data from the city council, the case commits to deliver new valuable knowledge about the context. How this knowledge can actually improve the council's management of the parking industry emerged and was established as a *Perceived City Opportunity* (see Assertion 16). Among other benefits, a positive impact is found to be in place also in relation to investors. At the questions initially asked when presenting the idea to investors such as "what are the guarantees that this data will be available in the future?", no satisfactory answer was given. Thus, establishing a relationship with the council helped to overcome this critical issue.

***9.2 Collective Learning*:** this aspect emerged as an important feature that should be an active part of this interaction. Emphasis was put on the need for councils and the open data services business to

"learn from each other" based on both the needs and the constraints that are in place for both parties. A few examples were given. One was about the experience in which the case was invited by the council for a presentation of problems that the city was facing in the context of parking. From there, both parties thought about new potential solutions in a "sort of brainstorming session". Another example was about the need for the case to "educate the councils" especially in relation to technical capabilities required to efficiently and effectively release open data. Hence, this interaction should enable collective learning for both of the parties involved.

9.3 Structure of Relationship / Partnership: having an established relationship with the council emerged as "a cornerstone for success". However, this interaction also represented the main challenge faced by the case. At the time of generating the idea, the main goal was "to establish a commercial relationship with the council". This form of partnership was built on the "supply of commercial services to the council" in which the company was responsible of managing the parking context and the council was giving it the data to do so. However, a commercial relationship in the context was already in place. Together with impeding the possibility for the company to supply commercial services to the council, this situation created "insecure perceptions" about the future ability of delivering the open data services. These were described by the CEO of the case as the risk that the existing business partner of the council would have replicated the service since the service is mainly enabled by open data that "is ready". Hence, from being the final goal, the commercial relationship became a "big barrier" for the case. To cope with these issues, Case C initially adopted a strategy in which the fact that open data was being re-used was hidden from all internal documentation and from the website. In addition to this, complexities in various forms were associated with the service when describing or presenting it (see also variable 4.6). However, these resulted (so far) to be just perceptions which did not become part of the company's reality. The case was able to establish a partnership with the existing council's partner. The interviewees concluded that commercial partnerships negatively influence the adoption of open data for commercial service innovation. In fact, these might result into "lock-in situations" where innovative open data services are precluded to become sustainable business activities. In proposing potential solutions for this problem and at the same time harness the possibility of engaging with the council, the CEO of the case pointed out some interesting scenarios. Briefly, the possibility of establishing "small contracts with the council for piloting a solution" and/or putting in place "shared investments for the development and initial deployment of a pilot solution" in the context were proposed.

- ***Cluster L: Other External Factors***

The last cluster of external factors included thirty-six KS that did not fall consistently in the clusters previously tackled. After several attempts, these KS were grouped into three main categories. In order

to define a common factor to encompass all these TAs, the name *Ecosystem Factors* was generated and subsequently validated. This includes three variables in relation to (supporting evidence is provided in Table 6.13): (1) the importance for the case to have established partnerships with other businesses; (2) the ability of the overall ecosystem to provide volunteers; and (3) the need of a broad EU-driven strategy leading open data programs across cities.

ASSERTION 10					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
10.1	*Partnership with other Businesses*	1,1; 11,3; 11,4; 11,6; 17,6; 24,13; 33,5; 33,6; 33,7; 33,8; 33,9; 33,10; 34,9; 34,10; 44,4; 55,9; 55,10; 65,1; 65,4; 65,5; 65,6;	--	D6; D8; D9;	Yes
10.2	*Volunteers Engagement*	9,3;	D2;	D8; D9;	Yes
10.3	*EU-Level Strategy and Funding*	22,1; 32,5; 36,2; 36,5; 44,5; 63,1;	--	--	Yes

Table 6.13: Case C Assertion 10, ecosystem factors

Assertion 10: Ecosystem Factors influence the adoption of open data for commercial service innovation

10.1 Partnerships with other Businesses: Case C successfully overcame the "commercial relationship with the council" barrier by "relying on the council's existing partner". According to the subjects interviewed, "a partnership was established with competitors" and "that was crucial to survive". Furthermore, since the company doesn't have internal multiple language skills, partnerships with specialised businesses where leveraged to "localise both language and currencies" across cities. Also, a partnership with the national parking association was established and "contributed significantly to our growth". Partnerships with payment providers in the city were also crucial in the company's experience (D6, D8). All in all, the CEO summarised the importance of establishing partnership in the open data landscape by saying: "you need collaboration and co-creation in this world".

***10.2 Volunteers Engagement*:** only one KS was found to support this variable. However, the role that volunteers played with the case is clear. Hence, a variable emphasizing the importance for the overall ecosystem to provide volunteers is found to be relevant. The role of volunteers (and interns) is also highlighted in D2 where the importance of having a "lean team" is reported.

***10.3 EU-Level Strategy and Funding*:** by having common focuses regarding open data across Europe, the scaling process for growing and delivering the open data service in different cities would be facilitated. The importance of an EU-driven strategy was also highlighted in relation to the need of

funding for emerging open data businesses. An integrated strategy from the European Commission is interpreted by the product manager as “a further push for councils to embrace open data”.

At this stage, ten assertions about the case were formulated in relation to ten external factors influencing the adoption of open data for commercial service innovation. Forty-one variables were found to fully describe the external factors emerged.

6.4.2 Perceptions

In reporting the perceptions emerged, the same strategy used for external factors is employed. At this stage, also links between factors are outlined. Thus, when defining each perception two types of links are highlighted: (1) the external factors that influence the generation of these perceptions; and (2) the link between each perception with either the adoption factor, or another set of perceptions.

- ***Assertion 11: Perceived Usefulness of Open Data influences the adoption of open data for commercial service innovation***

The first class of perceptions emerged in relation to the usefulness of open data. From the initial ninety-five KS, five variables could be defined within this factor. A list of these variables with the related evidence supporting their definitions is provided in Table 6.14. It is noted that the content of one of these variables (11.5) was substantially enriched during the final validation meetings undertaken.

ASSERTION 11					
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Valid.**
11.1	*Provision of Content / Knowledge about the context*	2,2; 2,6; 2,7; 2,9; 2,12; 2,14; 2,15; 2,16; 2,17; 9,16; 11,5; 12,1; 12,2; 12,3; 12,4; 12,5; 12,6; 14,2; 14,3; 14,4; 14,5; 14,6; 15,1; 16,5; 17,3; 17,8; 17,9; 17,10; 17,12; 18,; 21,2; 21,11; 21,12; 21,13; 21,14; 21,15; 27,1; 27,3; 27,4; 30,2; 30,3; 30,4; 30,6; 30,7; 31,2; 33,3; 34,5; 34,6; 34,11; 41,1; 41,6; 54,8; 54,9; 54,11; 54,12;	D1; D3; D4;	D6; D7; D8; D10;	Yes
11.2	*Quick development / time saving*	2,4; 2,5; 8,5; 12,8; 16,2; 36,9; 38,1; 47,1; 47,2; 47,3; 47,4; 47,7; 50,1; 51,1; 51,2; 51,3; 53,1; 53,2; 56,2;	--	--	Yes
11.3	*Resource Saving*	12,7;	--	D8;	Yes
11.4	*Discoverability / Understandability*	15,5; 16,3; 16,4; 47,4; 47,7; 51,1; 51,2; 51,3; 52,1; 52,2; 53,1; 53,2;	--	D8;	Yes
11.5	*“Badge of Honour”*	20,2; 21,11; 21,12; 21,13; 27,3; 27,4; 30,7; 35,3; 36,1;	--	--	Yes

Table 6.14: Case C Assertion 11, perceived usefulness of open data

***11.1 Provision of Content / Knowledge about the Context*:** open data usefulness is found to depend on the amount of content provided within it. The first reason why open data is perceived useful is related to the provision of detailed content about the context of interest. For this reason, open data is defined as "a main requirement for the service". At the time when the actual adoption of open data was undertaken (i.e. at the hack competition), the specific parking-related dataset was "selected because of the amount of content within it". Further datasets were later integrated into the service as these were perceived to be useful in structuring the knowledge about the context.

***11.2 Quick Time Development / Time Saving*:** the time that is employed in re-using the dataset is found to influence this factor. This refers to the "work that has to be done on the data". It includes: "understanding the dataset, its standard, content, and format", "cleaning the data" and integrating it into the platform. As this time decreases, perceptions about the usefulness of open data increase. However, in relation to this, it should be also considered that open data is believed to intrinsically carry a risk with by the fact that "it is ready" (aspect considered in variable 4.6).

***11.3 Resource Saving*:** having open data available avoids all the processes that would be undertaken to collect and structure this content. As part of these processes, significant resources would need to be employed.

***11.4 Discoverability and Understandability*:** by having a "discoverable dataset", data doesn't have to be interpreted and, more generally, "re-use is facilitated". This characteristic is enabled by the presence of clear metadata describing the dataset and supporting its "navigation" as well as by the clear "display of all the standards for all the characteristics of the dataset".

11.5 "Badge of Honour": this variable emphasizes the perceived usefulness of open data for improving the image of the company in some regards. Using open data was described by the CEO as a "badge of honour". In other words, "open data was leveraged to talk to the council" and to "enter the context". In the case's experience it was actually because open data was re-used that the company could actually engage with the city council. However, both of the interviewees agreed that "this value can't be easily measured".

Assertion 11 links

As shown in Figure 6.2 below, three external factors were found to influence perceptions about the usefulness of open data. Firstly, open data is perceived useful depending on its characteristics (Assertion 1). In relation to *Events and Community Factors*, the CEO highlighted the importance that the hack competition had in his experience for perceiving open data as useful. As part of this experience, open data was discovered and perceived as a valuable source for new commercial service creation. An additional link to *Perceived Usefulness of Open Data* was established as a result of the

validation discussions from *Organizational Skills and Resources*. Both of the people encountered as part of this stage underlined how without this set of skills, open datasets wouldn't be easily accessible. Since data is provided in machine readable format, with little information about its structure, and with different formats, certain IT skills are required for exploring, understanding, and navigating the datasets. It is noted that two more perceptions (i.e. *Perceived Business Opportunity* and *Perceived Ecosystem Awareness*) were found to be linked with this factor. These links will be described in the dedicated paragraphs of this section. On the other hand, the data collected suggested that if open data is perceived useful, it is adopted for commercial service innovation purposes.

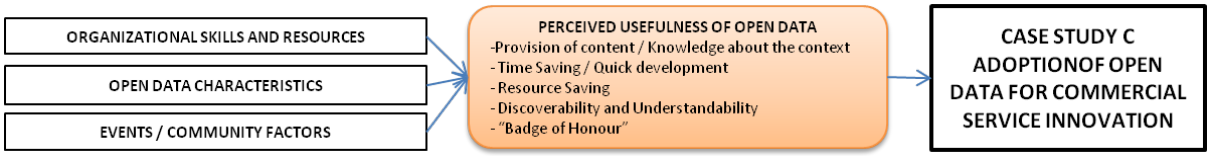

Figure 6.2: Case C Assertion 11 links

- ***Assertion 12: Perceived Business Opportunities influence the adoption of open data for commercial service innovation***

Cluster N was defined from sixty-seven KS in which information about perceived business opportunities were contained. Three variables were defined from the data collected and subsequently confirmed by both of the people approached (details are provided in Table 6.15).

ASSERTION 12					
Variables		**Key Statements**	**Internal Docs.**	**Ext. Docs.**	**Validated**
12.1	*Knowledge potentially derived*	2,12; 2,15; 2,17; 11,5; 12,3; 13,1; 17,9; 17,10; 21,13; 21,14; 21,15; 30,3; 30,4; 30,6; 31,1; 31,2; 31,4;	D1; D3; D4;	D7; D9;	Yes
12.2	*Scalability*	8,1; 8,2; 8,4; 8,6; 33,1; 33,4; 33,5; 33,6; 34,10; 41,2; 44,1; 44,4; 44,5; 44,6; 44,7; 44,8; 44,9; 44,10; 44,11; 44,12; 44,13; 55,4;	D1; D3;	D8; D9;	Yes
12.3	*Market Opportunities*	14,6; 31,3; 31,5; 33,3; 44,4; 44,5; 44,6; 44,7; 44,8; 44,9; 44,10; 44,11; 44,12; 44,13;	D1; D2;	D9;	Yes

Table 6.15: Case C Assertion 12, perceived business opportunities

***12.1 Knowledge Potentially Derived*:** this variable represents a source for the company to increase perceived business opportunity from the data collated through the app (D9). This is related to the *Secondary Data Generation* process (see variable 4.2). According to the interviewees, "the main business opportunity is from the data you can extract"; "a lot of information can be captured" and the knowledge created "can be sold to retailers to enable more informed decisions on their side".

***12.2 Scalability*:** perceptions about business opportunities are tightly related to the possibility of expanding the service in different cities. The platform is developed once only, and new data pertaining to different cities is "loaded in". In the case's experience, the importance of having a scalable solution was also "strongly advocated by potential investors".

***12.3 Market Opportunities*:** perceptions about business opportunities reflect which are found to be the actual context-related market opportunities. The bigger the market is, the more business opportunities are perceived. These can be described through two main concepts: market size and market opportunities. With respect to the former, it emerged that the size of the market is estimated from a number of variables including: "smart-phone penetration", "socio-economic status of the city", "number of cars driven in the city", and "relative number of parking spaces". On the other hand, opportunities are seen in relation to "presence of investors in the area", "opportunities for establishing partnerships", and based "on traffic congestions" and difficulties in finding a parking space. Finally, existing competition in the context is also found to influence this variable.

Assertion 12 links

The first link that emerges as describing contributors for this set of perceptions is from *Service Characteristics*. As demonstrated by the revenue models in place for Case C, business opportunities are partially related to the additional knowledge that is being created in the parking context. A further reason supporting the definition of this link is found in relation to the importance outlined above of having a scalable service. This is connected with one of the characteristics of the service, i.e. developed on a *Unique Content Management System*. The second and third inbound links defined are also intuitive. As confirmed by both of the co-founders of Case C, their attitude towards entrepreneurship contributed to augment beliefs about business opportunities. On the other hand, mostly in relation to variable 12.3, perceived business opportunities were found to depend on the actual characteristics of the city in which open data is taken. Clearly, the presence of a market opportunity is connected to the size of the market and the existing level of maturity of the parking infrastructure and related solutions (see Assertion 7). If these perceptions are in place, a contribution towards perceiving open data as useful for commercial service innovation is generated (see Figure 6.3).

Figure 6.3: Case C Assertion 12 links

- ***Assertion 13: Perceived Ecosystem Awareness influences the adoption of open data for commercial service innovation***

In this cluster it was found that *Perceived Ecosystem Awareness* positively influences the adoption of open data. The variables emerged reflect the different parties, whose awareness of open data is found to be beneficial for adopters. These are: city councils, investors and the overall community. Forty-eight KS were found to support this assertion (see Table 6.16). It is noted that variable 13.3 was initially retained from the preliminary findings, and added after further data was collected.

ASSERTION 13					
Variables		**Key Statements**	**Internal Docs.**	**External Docs.**	**Validated**
13.1	*Community Awareness*	12,1; 14,1; 22,2; 22,3; 22,4; 22,8; 23,1; 39,3; 40,3; 42,8; 42,9; 43,1; 44,4; 45,2; 50,1; 55,9; 57,5; 57,6; 57,7; 57,8; 57,9; 58,1; 58,2; 58,3; 59,1; 59,2; 59,3; 60,3; 60,5; 60,7; 60,8; 60,9; 60,12;	D2; D4;	D6; D7; D8;	Yes
13.2	*City Council Awareness*	13,2; 22,2; 28,2; 40,3; 42,8; 42,9; 60,1; 60,4; 60,6; 60,9; 60,13; 60,14; 60,15; 60,16; 60,17; 66,10;	D4;	D6; D8;	Yes
13.3	*Investors Awareness*	36,3; 44,5; 60,7;	--	--	Yes

Table 6.16: Case C Assertion 13, perceived ecosystem awareness

***13.1 Community Awareness*:** following the interviewees' opinions, as more people in the local community are aware of open data, the more effective the related ecosystem is expected to be. In this way, community awareness is seen as a key "to stimulate councils to evolve" and to "provide guidance on how to do it". The overall result is described by the CEO as "enabling co-collaboration".

***13.2 City Council Awareness*:** when tackling the issues around council's awareness, much has been already stated in relation to the need for councils to be educated through different forms (see variables 2.2, 2.3, 3.2, 3.3, and 6.1). In this variable the attention is placed on the importance for councils of "being aware of who and how is using open data" as the source for business activities. Case C's people believe that councils must be effective in reporting results of their open data initiatives. This is expected to subsequently increase trust among the community, especially among investors.

***13.3 Investors Awareness*:** this variable highlights the importance for the investors' community of being aware of open data, its benefits, and its potentialities. If this is achieved, a substantial push (in terms of funding) would be given to the overall adoption process. This aspect emerged in relation to the case's experience in which at the time of presenting the idea to potential investors, these resulted to be reluctant because of their "poor knowledge on open data".

Assertion 13 links

As depicted in Figure 6.4 below, only one link has been identified as direct contributor for this set of perceptions. In particular, as the main consequence of implementing and having in place *Events and Community Factors*, spread awareness in the overall ecosystem is expected to be achieved. Indeed, the main purpose of workshops and engagement initiatives, hack competitions, and of having effective champions in place, is to spread awareness of open data and fostering engagement. This achievement is captured within the definition of the link between Cluster F and *Perceived Ecosystem Awareness*. On the other hand, if these perceptions are in place, *Perceived Usefulness of Open Data* was found to increase subsequently.

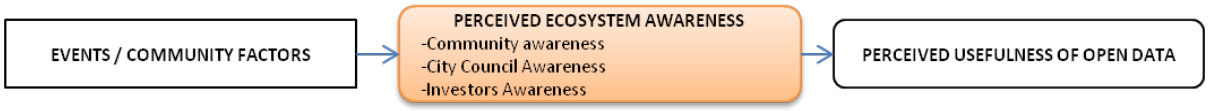

Figure 6.4: Case C Assertion 13 links

- ***Assertion 14: Perceived User Satisfaction of the open data service influences the adoption of open data for commercial service innovation***

Cluster P encompasses all the data collected about the perceptions of users of the open data service. Four variables arose from the analysis (Table 6.17). It is noted that specific information about the service were collected before the actual SSIs. Hence, no supporting evidence for variables 14.1 and 14.2 is coming from the actual KS extracted from the SSIs' transcripts. However, both of the co-founders emphasised the importance of these variables at informal meetings as well as during the validation stage. These aspects are believed to be a key for the service to diffuse among users (and so to increase the user base), which in turn positively impact the likelihood of achieving sustained adoption of open data for commercial service delivery.

ASSERTION 14					
Variables		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
14.1	*Time saving*	--	D1; D3; D4;	D6; D7; D8; D9; D10;	Yes
14.2	*Resource Saving*	--	D1; D2; D3; D4;	D6; D7; D8; D9; D10;	Yes
14.3	*Personalized Service*	5,2; 6,2; 7,1; 7,2; 7,3; 8,1; 8,2; 8,6; 34,11;	D1; D3; D4;	D6; D7; D8; D9; D10;	Yes
14.4	*Easy to Use*	18,7; 30,1; 30,2;	D1; D2; D3; D4;	D8; D10;	Yes

Table 6.17: Case C Assertion 14, perceived user satisfaction

***14.1 Time Saving*:** the service's main objective is to enable users to find a parking spot easily and quickly. For example, the service's option "parking nearby" (D4) avoids waste of time of the drivers by indicating the nearest parking available. The features "ordered list by distance" and "direction with maps" (D1) are in place to minimise the time the users spend in looking for parking.

***14.2 Resource Saving*:** by giving the opportunity to the drivers to find "best parking rates" (D2), the user's expenses for parking can be potentially minimised. Also, one of the company's documents (D1) indicates that "30% of emission from traffic in the city is caused by drivers searching for a parking spot" (D1). Thus, the service is designed to have an overall negative impact on emissions.

***14.3 Personalised Service*:** this variable can be considered relevant for both of the types of targeted users: drivers and parking operators. Concerning the first, by logging in into the account, the drivers can access a "payment history" file, where all the transactions are stored and reported by date, place and time (D1). In addition, personal notifications, such as "reminders and alarms" as well as "loyalty and rewards mechanisms" (D1), are supported by the service. The second class, instead, has the opportunity to upgrade the coverage of the specific parking place into an "own brand parking app" (D1), i.e. a fully personalised application.

14.4 Easy-to-Use: developing an open data service that is easy to use and where information sought is easy to detect is believed to be crucial for its success. Internal *Organizational Skills and Resources,* and specifically graphic design and data visualisation skills, are critical factors for enabling this variable.

Assertion 14 links

With respect to Cluster P, two main inbound and one outbound links were identified and confirmed by the case's people. Firstly, it emerged that developing a service that is perceived to be valuable by the end users significantly affects the achievement of sustained adoption of open data. However, this link was agreed to be mediated by *Perceived Usefulness of Open Data*. In other words, following the reasoning provided, if the perceived ability of providing a service that satisfies end-users is in place, open data is perceived to be useful for service innovation. Secondly, these perceptions are directly connected with the characteristics of the service deployed and particularly with the variables *Valuable Service in the Context, White Label Solution,* and, "to some extent" *Flexible Revenue Model*. Finally, the fact the service is perceived to be *Easy to Use,* was interpreted as being a consequence of internal *Organizational Skills and Resources* (especially on *Graphic Design and Data Visualization*). These links are summarized in Figure 6.5.

Figure 6.5: Case C Assertion 14 links

- ***Assertion 15: Perceived Commitment of City Councils influences the adoption of open data for commercial service innovation***

Initially, ninety-four KS were classified under Cluster Q. However, the majority of these KS accounted for evidence of links between factors. In addition, a considerable number of KS was connected to the aspects already tackled in Cluster O (*Perceived Ecosystem Awareness*) in relation to the importance of *Councils Awareness.* Therefore, only one variable was found to fit into this cluster. This variable was initially retained as no sufficient clarity was found when undertaking the first cycles of data analysis (see Table 6.18 for details). Therefore, reflections in this way were brought to the final validation meetings where its importance was confirmed and its description enriched.

ASSERTION 15					
Variables		**Key Statements**	**Int. Doc.**	**Ext.Doc.**	**Valid.**
15.1	*Account for responsibility on the data released*	16,5; 18,1; 55,3; 56,3;	--	--	Yes

Table 6.18: Case C Assertion 15, perceived commitment of city council

15.1 Account for Responsibility on the data released: this aspect initially emerged at a question that was asked to the CEO: "in a potential situation in which there is an error in the dataset, for example a place where parking seems to be permitted is no longer available and a driver decides to park as indicated by the app, who is responsible for this?" The CEO, underlined that the council is not responsible for mistakes in the data and that the company doesn't want to be either. The solution that was found was to "put it on the back of the drivers, with notifications like 'follow road signs first'". "It worked; we had no major complains so far" – the CEO stated. However, this issue still remains when presenting the idea to potential investors. During the validation stage, this aspect was found to be relevant because if there is responsibility from the council on the "goodness of open data", its strategy and the overall ecosystem are perceived credible and sustainable (see Assertion 17). Emphasis has been put towards the importance of perceiving commitment from the city councils in order to develop secure perceptions about the availability of open data in the long-term. These perceptions are interpreted not to affect the case in question only, but also investors, to which great importance has been given by Case C's people.

Assertion 15 links

According to both of the interviewees, two main contributors for this set of perceptions are identified. First, commitment from the city council depends on the characteristics of the council itself. In other words, commitment is increasingly perceived if: an integrated *Open Data Strategy* is in place supported by extensive *IT skills*; awareness is spread within it; and if engagement initiatives and investments are made towards this program. The second contributor is the council-business relationship factor. Clearly, if a *Long-term Commitment on Open Data Provision and Usage* is established by both parties, both the open data services business and investors increasingly perceive long-term commitment of the council. Finally, a further link emerged between this factor and *Perceived Credibility and Sustainability*. The latter will be introduced as the result of the analysis conducted within Cluster T (see Assertion 17 below).

Figure 6.6: Case C Assertion 15 links

- ***Assertion 16: Perceived City Opportunities influence the adoption of open data for commercial service innovation***

Another factor that contributed to Case C's adoption of open data consists of a set of perceptions generated in relation to opportunities perceived for the city. Through pursuing these opportunities, municipalities are stimulated to effectively embracing open data to enable these opportunities. A total of twenty-nine KS were classified under this cluster, and these complemented significant evidence found in the documents collected and analysed (see Table 6.19). Variable 16.4 was initially retained and subsequently added to the findings as a result of the validation discussions.

ASSERTION 16					
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Valid.**
16.1	*More informed decisions in the context*	2,12; 17,8; 17,9; 17,10; 17,11; 17,12; 21,11; 21,13; 21,14; 21,15; 5,1; 25,2; 26,3; 30,3; 30,4; 30,6; 30,7; 31,1; 33,3;	D1; D3; D4;	D7; D8; D9; D10;	Yes
16.2	*Increased efficiency*	6,1; 12,7; 12,8; 17,8; 17,10; 17,11; 21,11; 21,12; 21,15; 24,11; 25,1; 25,2; 26,3; 30,3; 30,4; 30,6; 33,3;	D1; D3; D4;	D7; D8; D9; D10;	Yes
16.3	*Increase Open Data outcomes*	25,1; 25,2; 26,3; 35,3; 49,1; 49,2;	--	--	Yes
16.4	*Attract New Businesses*	33,2; 60,8;	--	--	Yes

Table 6.19: Case C Assertion 16, perceived city opportunity

***16.1 More Informed Decisions in the Context*:** the potentialities of new data generated on the back of the service have been highlighted in previous clusters. Here, the focus is on the use that councils can make of this data to create value. This value is understood as "enabling more informed decisions for the parking infrastructure management" from the "additional detailed knowledge created" about this context (e.g. "discovery of drivers patterns in the city"). An important source in this way is the dataset about "revenue from the meters". One of the outcomes of the analytics processes is described as "revenue forecasting analytics" (D1). With this "reliable information", more effective management in the context is enabled for the city council.

16.2 Increased Efficiency: similarly, to the previous, this variable represents the potential of delivering value to the council from the open data service. The value is considered in terms of augmenting efficiency in managing the parking context. For instance, the mobile-payment option provided to the drivers was described during one interview as a "service provided to the council" that increases the efficiency of the parking-payment-related processes.

***16.3 Increased Open Data Outcomes and Transparency*:** the initial process for re-using a dataset is about refining the dataset itself. Adjusting the standards, the formats, and the errors in the dataset are the mainly required activities. This "improved dataset" can be "shared back to the council" and it can replace the original one. Also, in the interviewees' opinion, the fact that a company is re-using open data is leveraged by the council as an opportunity to claim transparency.

16.4 Attract New Businesses: a last variable emerged in relation to the increasing attractiveness of the city for new businesses willing to re-use open data as a source for service innovation. Although this statement finds confirmation in the case's experience when targeting new cities (cities that have effective and relatively mature open data programs in place are targeted by the company), no sufficient evidence was initially found in the data. Thus, its integration was agreed only after the final validation meetings.

Assertion 16 links

The importance of this factor within Case C's adoption process becomes clearer when looking at the links generated. As shown in Figure 6.7 below, the fact that cities perceive this set of opportunities is "reasonably" leading to an increased "motivation" of the council to pursue open data objectives, which is translated into increasingly perceiving commitment from the city authority. In relation to this cluster, a change in the links was made as a result of the validation meetings by deleting the previously established link between *City-Related Factors* and *Perceived City Opportunity*. The rationale behind the establishment of this link was about cities perceiving opportunities in relation to their natures (e.g. bigger cities with a certain infrastructure are expected to have greater perceptions about opportunities related to the delivery of open data service). However, it emerged that these

opportunities are enabled mainly by the nature of the service and its ability to provide more detailed knowledge about the context. Therefore, the previously established link has been replaced with a connection with *Service Characteristics*. A final link emerged with *City Council Related Factors*. The motivation for establishing this connection is found within the concept of a potentially increased local economic growth (see variable 16.4). This was inferred by the process through which Case C selects and targets additional cities for scaling its open data services. This can be seen as the exploitation of the opportunity to *Attract new Businesses* from the perspective of the cities selected. As mentioned within Assertion 2, one of the key requirements for a city to be targeted is related to the council's characteristics.

Figure 6.7: Case C Assertion 16 links

- ***Cluster T: Other Perceptions***

This last cluster was generated from all the KS that did not fall in any of the previously defined categories. A common theme across these statements has been interpreted to be around perceived trust in open data and the overall related ecosystem as well as in the possibility of sustainably deliver an open data service. In order to encompass the meaning of both of the variables in this cluster, the resulting perception factor was named, in accordance with both of the subjects interviewed, *Perceived Credibility and Sustainability.*

Assertion 17: Perceived Credibility and Sustainability influence the adoption of open data for commercial service innovation

Throughout its history, Case C faced several stages in which questions arose in relation to the perceptions of open data to be a useful source for the service in the long-term. These questions were relevant for both people within and outside the case. The latter were mainly represented by investors to which the idea was presented. Overall, two variables were defined as part of this cluster (see Table 6.20).

ASSERTION 17					
Variables		**Key Statements**	**Int. Doc.**	**Ext. Doc.**	**Validated**
17.1	*Challenges to the Council*	22,8; 23,1; 57,2; 57,3; 57,4; 58,2; 59,1;	--	--	Yes
17.2	*Replicability of the Open Data Service*	26,5; 26,6; 26,7; 26,8; 36,8; 36,9; 37,1;	--	--	Yes

Table 6.20: Case C Assertion 17, perceived credibility and sustainability

***17.1 Challenges to the Council*:** the first of these aspects is about the actual challenges that are posed to the council. A similar variable was defined for describing the role that *Intermediary Organizations* should play in this context. This is understood as the need of establishing parties or mechanisms that act as "counterparts for the council". For better explaining this concept it is useful to mention an extract of an interview undertaken with the product manager:

"In a lot of government's departments, if you do something wrong, you tend not to get sacked, you get moved somewhere else. Also, the job description in their contracts is so tight and strictly defined. So this slows things down and they are not motivated to do things well. If they can do their jobs like okay, it is fine. They lack motivation. So if you have a third party that pushes, that adds credibility. These are mainly [the local *Intermediary Organization*] at the local level and the EU at the wider level". Therefore, challenges to the council should be in place to overcome "motivational problems" of those assigned to open data activities within the councils. The CEO explained that if no challenges are posed to the councils, this will result into "complacency" where "the councils just do something to state they are doing open data". Therefore, open data initiatives, in order to develop trust across the ecosystem, should be critically assessed and subsequently stimulated.

17.2 Replicability of the Open Data Service: this variable emerged from "worries" that investors highlighted in relation to the replicability of the open data service by other parties. The fact that "open data is ready" is indeed seen as a risk factor perceived by the case. Considering that the critical enabler of the service is something that is publicly available to everyone, the more value is added to this data, the less chance is in place that the service can be replicated. Actions undertaken in this way by the case were: "keep key ideas and processes internally", "hide open data from our website and from presentations to investors", and focus on the value of *Secondary Data Generation.*

Assertion 17 links

For describing the three contributors to this set of perceptions, it is useful to consider the two variables defined above separately. In particular, variable 17.1 stresses the importance of having in place entities that continuously challenge the council's operations. As demonstrated by the extract proposed within its description, two main sources of challenges should be in place: (1) from the EU

level, and (2) from the *Intermediary Organization*. These are believed to be different but equally important. The former is believed to have a key role in relation to legal aspects, and general stimulation for embracing open data. The latter is expected to play a role at a greater level of granularity and more on "ground-based tasks". Given that these bodies are actively involved in daily activities with the councils, the effect in this way is expected to be bigger and "more systematic". Therefore, two contributors to this factor emerged to be *Intermediary Organization* and *Ecosystem Factors*. Considering variable 17.2, the "insecure perceptions" generated for Case C were already tackled when defining the importance of designing a service with a certain degree of *Complexity* (see variable 4.6). Therefore, having an open data service that is complex is found to reduce its potential replicability, which led to consistently establish a link between *Service Characteristics* and this factor. The fourth contributing factor was found to be the *Perceived Commitment of the City Council*. This connection was already outlined within Cluster Q. Now that this cluster has been described it becomes clearer that if perceptions about the commitment of the council are in place, perceptions about the credibility of the overall ecosystem are subsequently generated. On the other hand, concerning the outbound links from this factor, an initial link had been defined between *Perceived Credibility and Sustainability* and the adoption factor. In this regard, during the final validation stage both subjects emphasised how having these perceptions in place increase the basis upon which perceived business opportunities are created. From another perspective, opportunities for implementing a commercial open data service are perceived if credibility and sustainability of the overall ecosystem and the open data service proposed are believed to be in place. These links are represented in the following figure

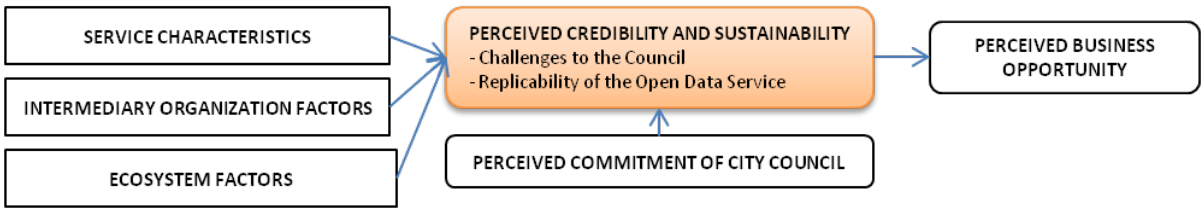

Figure 6.8: Case C Assertion 17 links

At this stage, all the KS related to perceptions were rigorously structured and analyzed. As a result, seven factors, twenty-two variables, and twenty-three links were derived, validated, and established as part of Case C's adoption model of open data for commercial service innovation.

6.4.3 Adoption Stages

The last and final category analysed was related to the actual adoption stages the company went through from the initial generation of the idea to the achievement of sustained use of open data for its services' delivery. In total, eight adoption sub-factors were identified from the data collected,

structured during the qualitative analysis, and validated at the final validation stage (see Appendix 5.2). Very helpful in this way was found to be the company website (D4). In particular, a section with "key milestones achieved" by Case C is provided. This was a very useful starting point from which these eight stages were finally derived.

Stage 1, Idea Generation: as the first step for the case, the current CEO generated the service idea. Initially, the idea was thought for the taxies' industry. The actual plan of delivering a service for facilitating parking in the city was developed "from the data available". However, the technical side of the service was kept from its previous formulation (i.e. for taxies). Critical in this way were the experiences of the co-founders in IT-related organizations as well as their entrepreneurial spirits.

Stage 2, Adoption*:* the second stage sees the leveraging of the previously generated idea through the adoption of open data. It was at the time when the current CEO participated in the locally organised hack competition, that awareness of open data was gained. The CEO reported that: "a previously designed solution was leveraged with open data for parking". Thus, open data was adopted to extend the existing technical application to the parking industry.

Stage 3, Market Analysis**:** the hack competition experience and the prize won stimulated the CEO to further develop the business idea. The actual centric phase in the design of the application was informed by an extensive market analysis. "Completing the commercial side of the business plan and market research" were undertaken once open data has been adopted (D2). Results from market analysis generated perceptions about potential business opportunities.

Stage 4, Private Beta**:** once knowledge about the targeted market had been acquired, the next step for the company was to release a "private beta version of the application". The idea was to take a sample of the dataset to test and improve the initial solution. In this way, "a group of one hundred drivers in the city was selected" (D2). These initial users were given a basic version of the application and were asked to use it while involved in parking in the city. The actual testing of the dataset and of the solution as well as the collection of feedback from these users, contributed significantly to the development of the application.

Stage 5, Public Beta**:** as a result of the analysis undertaken during the private-beta period, the case detected strengths and weaknesses of both the datasets and the service. "A lot of field work has been done to improve the data". During these processes, the case started to get involved in the overall ecosystem. This enabled the co-founders to develop knowledge about the councils, open data, the related legislation, and of the parking context. As a result, new open data was successfully asked to the council (and the intermediary organization) and subsequently integrated within the service's systems. The content management system has been adapted to accommodate these new datasets

and the final service was enriched with several additional features. This moment is considered as the official launch of the open data service.

Stage 6, Commercialization: after a public beta of the open data service was delivered, the business model for the company was starting to be established. The first actions taken in this way by the case were to "enable the up-sell of the white label solution" from the existing open data service, and to integrate it with the existing payment infrastructure. While growing, the company achieved a "considerable user base". It was only from there that it could think about generating "knowledge on the back of open data". Thus, the three main components of the company's revenue model (see section 6.1) were conceptualised and established.

Stage 7, Growth: from this stage of development, the first step undertaken by the case was to promote the application developed. To do that, a partnership with two major local newspaper companies was established and leveraged to substantially increase the user base in the initial city. From there, the case's growth was represented by "expanding the service to different cities". Both US and Europe were targeted at that time. However, so far only European cities have been successfully included (also Melbourne, Australia, but it was ignored in this research as the data supporting this application was not consistent with the definition of open data). "Approaching the council is always the first activity" undertaken when targeting a new city.

Stage 8, Sustained Use: the last stage in the case's adoption process is represented by the achievement of sustained use of open data for the service development and delivery. Continuous improvements are made on the actual service. Critical at this stage is having open data that is updated according to the needs of the case.

6.5 Case Study C Conclusions: Adoption Factors Variables and Links

Concluding, from Case Study C, seventeen assertions were formulated in relation to seventeen factors influencing the adoption of open data for commercial service innovation. Each of these factors is described by a set of variables, and links between factors were derived as well. The resulting adoption model for Case C is depicted in Figure 6.10 (next page). In addition, eight main stages emerged to be critical in describing the process through which Case C went from the initial generation of the idea for the open data service, until sustained use of open data for the delivery of the service was achieved. These stages are represented in Figure 6.9.

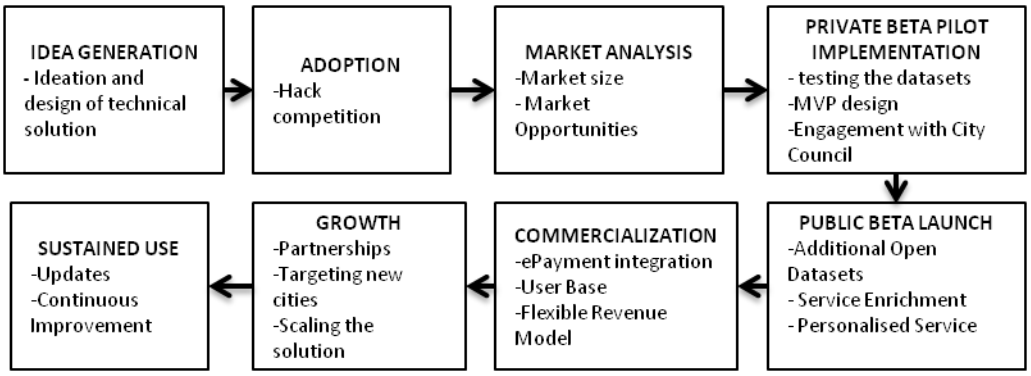

Figure 6.9: Case C adoption stages

These findings concluded the analysis of Case C. Together with the findings from the other two cases studied in this research, these represented the input for the cross-case analysis.

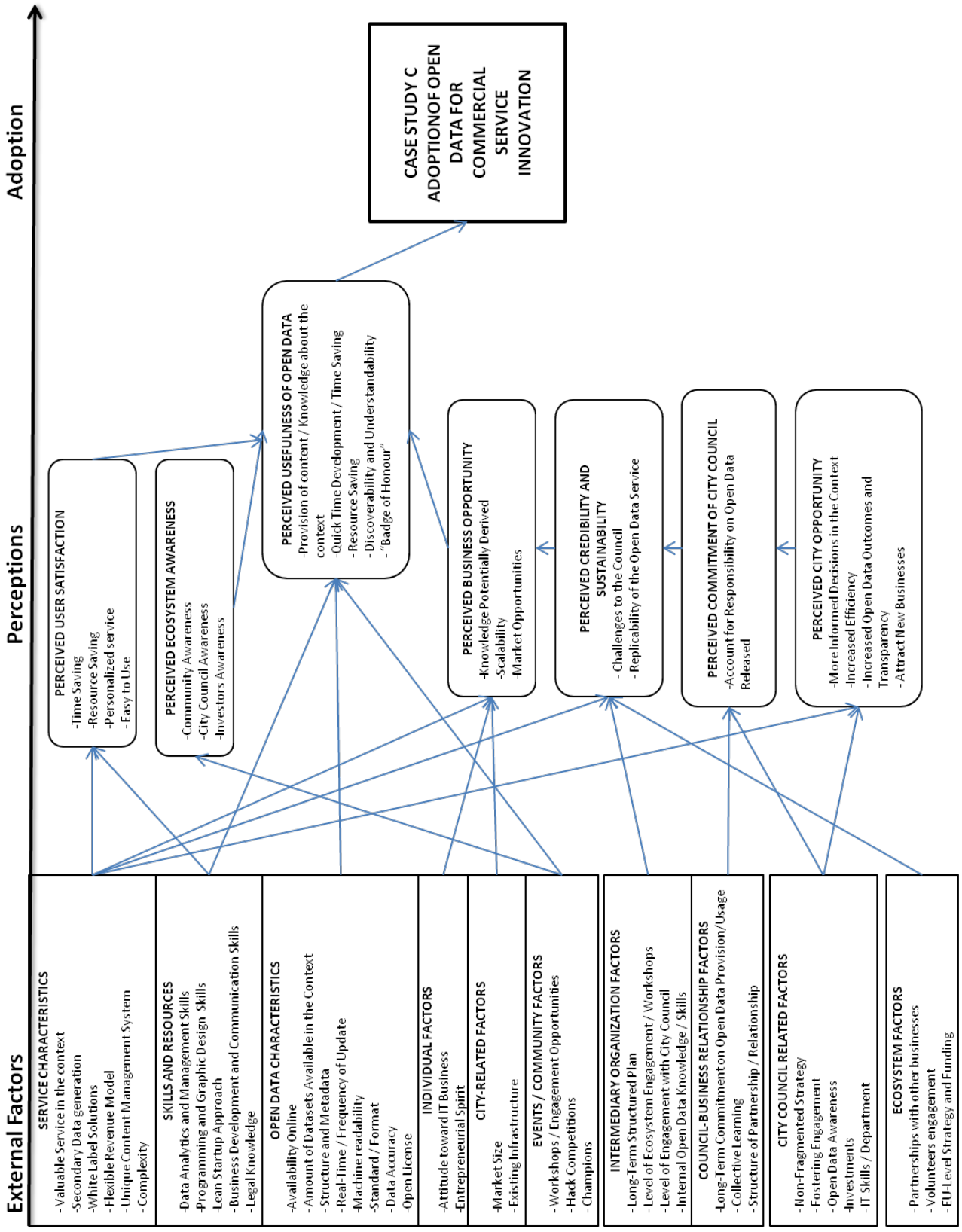

Figure 6.10: Case C adoption model

This chapter concludes the segment of this dissertation dedicated to the findings achieved from the single cases. The next chapter focuses on the findings achieved from the qualitative analyses conducted across these three cases.

7 CROSS-CASE FINDINGS

This chapter is dedicated to the reporting of the findings achieved from the cross-case analysis and the validation across cases. These represent the main outcome of this research from which contributions were formulated (Chapter 9).

In reporting these findings, a similar structure to the one adopted for the single cases is used (see sections 4.4, 5.4, and 6.4). First, the cross-case analysis process is reported (section 7.1). Subsequently, three separate sections are dedicated to the three main categories of IS adoption, i.e. external factors (section 7.2), perceptions (section 7.3), and adoption sub-factors (section 7.4). Within each of these, a paragraph is dedicated to each Cross-Case Assertion (CCA) established. Each assertion is presented starting with an introduction to the factor and its supporting evidence. All evidences supporting each CCA are provided (as done for reposting the single cases findings, for space reason, one table including supporting evidence is provided for each category; the remaining tables are provided in Appendix 6). Differently from the single cases, tables of evidences are provided for each CCA including: each variable defined; its occurrence across cases; if it was initially retained or not; and whether the specific variable was finally validated or not. The reporting continues with the description of each variable defined within the specific CCA (i.e. the specific factor). The thick descriptions generated for each variable are provided. Within the perceptions' section also links between factors are presented. For each CCA two additional paragraphs are included. Firstly, the outcome of the final validation stage is also integrated. Secondly, an additional paragraph is added describing those variables that were retained from the specific CCA. It is noted that in reporting outcomes from the validation across cases, quotations from these interactions are related to Table 7.13 (e.g. S1, S2 etc., see sub-section 7.3.1). The chapter concludes with the outline of the cross-case adoption model generated and the adoption sub-factors derived (section 7.5).

Overall, the findings demonstrate that sixteen factors described by a total of sixty variables are the most relevant in influencing adoption of open data for commercial service innovation in cities across the three cases studied. A total of twenty-two links between factors emerged from this analysis. Furthermore, seven sub-factors were established across cases as describing the adoption process of open data from the initial generation of the idea, to achieving sustained usage of this source for the delivery of commercial services.

7.1 Cross-Case Analysis Process: CCTA Development

The following sub-sections tackle the specific activities related to Level 1, 2, and 3 of the analysis implemented for external factors, perceptions, and adoption sub-factors respectively. Links between factors are outlined within the perceptions' section.

7.1.1 External Factors CCTA Development

The following table represents a list of external factors identified for all the cases studied. Based on each factor's occurrence across the three cases, Level 1, 2, or 3 of the analysis is highlighted.

EXTERNAL FACTORS	CASE A	CASE B	CASE C	LEVEL
Open Data Characteristics	x	x	x	1
External Inspiring Factors	x			3
Individual Factors	x	x	x	1
City-Business Relationship	x		x	2
City and City Council Related Factors	x	x	x	1
Local Community Factors	x	x	x	1
Organizational Skills	x	x	x	1
Service Characteristics		x	x	2
EU-Driven Open Data Ecosystem		x		3
Organizational Factors		x		3
Intermediary Organization Factors			x	3
Ecosystem Factors			x	3

Table 7.1: External factors across cases

As shown in Table 7.1, a total of twenty-five external factors emerged across cases. Of these, five were found relevant for all the three cases studied; two were derived for only two of the cases; and five emerged in first approximation for one case only. Based on this categorization, factors were allocated across the three different levels of analysis explained above.

- ***Level 1: external factors across all the cases***

The focus of the first level of the cross-case analysis was about merging those factors that emerged for all the three cases studied. As shown in Table 7.1, five factors were included in this cluster. However, when undertaking the actual analysis process, one factor was divided into two components. This was the case of *City and City Council Related Factors* (see Table 7.1). This was found to contain information about both the city where the service is implemented and the city councils from which open data is actually released. This factor emerged as an individual construct for Case A and Case B, while two different factors emerged for Case C. However, for the first two cases the variables defined within them could be clearly divided between these two aspects. Only one variable was considered for both. This was *Infrastructure Investments* (Case B). Indeed, this variable, as emerged from Case B,

contained information about: the need for councils to invest in generating capabilities for deriving and subsequently releasing additional open data; and the need for cities to have in place these infrastructures in order to be perceived as attractive urban environments where open data services can be deployed.

In relation to the analysis process, the merged factors were tackled separately. Similar concepts from similar variables across cases were merged. When strong evidence across cases was found in relation to the variables, CCTAs were initially formulated. On the other hand, those variables which occurrence emerged from one case only were retained at this stage. These retained variables were stored and considered in future steps of this analysis. Table 7.2 gives a summary of this process.

LEVEL 1 EXTERNAL FACTORS CROSS-CASE ANALYSIS					
Merged Factors / CCTAs	**Variables**				
	Case A	**Case B**	**Case C**	**CCTAs**	**Retained**
Open Data Characteristics	6	8	8	8	1
Individual Factors	4	3	2	3	2
Local Community Factors	3	5	3	3	2
Organizational Skills	6	4	6	6	2
City-Related Factors	1	3	2	2	2
City Council-Related Factors	8	5	5	5	3

Table 7.2: External factors level 1

As shown in Table 7.2, six cross-case tentative external factors were formulated from this stage of the analysis. A total of twenty-seven variables emerged as critical in describing these factors, while twelve were retained at this stage. These were taken into account for Level 2 and Level 3 of this analysis.

- ***Level 2: external factors across two cases***

The cross-case analysis continued with the focus on those factors that emerged from two of the three cases studied. These included *City-Business Relationship* (Case A and Case C) and *Service Characteristics* (Case B and Case C). The strategy was to consider these clusters and reflect upon these concepts in relation to the "missing case". Similarly, to what done for Level 1, the following table is proposed as a summary of this stage of the analysis.

LEVEL 2 EXTERNAL FACTORS CROSS-CASE ANALYSIS					
Merged Factors / CCTAs	**Variables**				
	Case A	**Case B**	**Case C**	**CCTAs**	**Retained**
City-Business Relationship	3	--	3	2	1
Service Characteristics	--	10	6	2	12

Table 7.3: External factors level 2

As shown in Table 7.3, two tentative factors emerged from Level 2 of the cross-case analysis. For both these factors, additional information was found regarding the "missing case" from other case-related

factors and/or from revisiting the original data collected. It is also noted that thirteen variables were retained at this stage. Of these, twelve were retained from *Service Characteristics*. Although these were subsequently considered within Level 3, it was clear already at this stage how certain specific features of the different open data services were tightly related to the context in which these are delivered. For example, from Case B it emerged the importance of delivering an open data service that enables interactive learning experiences; it was clear at this stage how this concept was strongly related to the culture and tourism context. Finally, none of the retained variables at this stage was found related to factors emerged from Level 1, nor those retained at that stage were found to be relevant here.

- ***Level 3: remaining external factors***

In this last level of the external factors' cross-case analysis, the first step followed was to list all the remaining variables across the factors that emerged in relation to one case only. This list was enriched with those variables that were retained from the previous levels. This process was undertaken in order to ensure that the variables that did not find considerable confirmation across cases, were not considered within other factors from different organizations. A total of forty-two variables were considered at this stage. Of these, seventeen were variables that emerged in specific cases in relation to the five single cases' factors initially clustered within this class. The remaining twenty-five variables were those imported from the previous levels. At this stage all variables were clustered and combined based on their similarities. The process followed was about looking at the original thick description generated for each variable within the specific case. Based on these definitions, similar variables were grouped together aiming at finding considerable evidence across cases for the definition of additional CCTAs. Three possible situations were encountered: (1) the emerging concept was defined as a new factor and a new CCTA was formulated; (2) the emerging concept was defined as a variable of an already emerged factor (if there was similarity with factors emerged from Levels 1 or 2); and (3) no sufficient evidence was found to support the emerging concept which was therefore object of further reflections.

After several attempts, three main topics emerged: (1) variables emphasizing the importance of the city-related opportunity; (2) variables emphasizing the importance of having place and being involved with an open data ecosystem; and (3) variables about the importance of having in place EU-level commitment and a related strategy. The first emerging topic was finally established as an actual variable related to the already tentatively defined *City Related Factors* (see Level 1). Details about the variables merged from the single cases for its definition are provided in Table 7.4.

Tentative Variable	Variables Integrated	Original Factors	CASE
City Related Opportunity	*Maturity of Digital Solutions*	*External Inspiring Factors*	A
	Opportunity to gain Efficacy	*External Inspiring Factors*	A
	Opportunity to gain Efficiency	*External Inspiring Factors*	A
	Underexploited Areas	*City-Related Factors*	B
	Valuable Service in the Context	*Service Characteristics*	C

Table 7.4: City related opportunity, level 3

Regarding the two additional clusters, these were translated into CCTAs. Table 7.5 provides an overview of the emerging factors, the variables merged for their tentative definition, and the number of tentative cross-case variables and retained ones.

LEVEL 3 EXTERNAL FACTORS CROSS-CASE ANALYSIS			
Merged Factors / CCTAs	**Variables integrated across cases**	**New tentative variables**	**Retained**
Open Data Ecosystem	11	2	1
European Level Commitment and Strategy	4	1	0

Table 7.5: External factors level 3

As shown in Table 7.5, two additional factors were tentatively defined from this stage of the analysis. Furthermore, as part of this clustering exercise, one additional variable emerged in relation to the already tentatively defined *City Related Factor* (level 1).

Concluding, twenty-five variables from the single cases were retained at this stage. Therefore, more data was needed to be collected in these regards in order to motivate either their inclusion into the final cross-case findings, or their allocation to the specific contexts in which these were defined. These decisions were taken as part of the final validation across cases stage (see section 7.3). Furthermore, ten cross-case tentative external factors were defined described by a total of thirty-five variables.

7.1.2 Perceptions Factors CCTA Development

Table 7.6 represents the list of perceptions identified for all the cases studied. Based on the occurrence of each factor across the three cases, Level 1, 2, or 3 of the analysis is highlighted.

As shown in Table 7.6, from a total of twenty perceptions emerged across cases: three were found to be relevant for all the three cases studied (i.e. for Level 1); three were highlighted as relevant for only two of the cases (i.e. for Level 2); and five emerged in first approximation for one case only (i.e. for Level 3). Based on this distinction, the three different levels of analysis were undertaken.

PERCEPTIONS FACTORS	CASE A	CASE B	CASE C	LEVEL
Perceived Voluntariness	x	x		2
Perceived Business Opportunity	x	x	x	1
Perceived Usefulness of Open Data	x	x	x	1
Perceived City Opportunity	x	x	x	1
Perceived User-Friendly	x	x		2
Perceived User-interest		x		3
Perceived User-Satisfaction		x	x	2
Perceived Potential Local Economic Growth		x		3
Perceived Ecosystem Awareness			x	3
Perceived Commitment by City Council			x	3
Perceived Credibility and Sustainability			x	3

Table 7.6: Perceptions across cases

- ***Level 1: perceptions across all the cases***

This first option encompassed three perceptions factors that emerged as influencing the adoption of open data for commercial service innovation across all the cases studied. In relation to the analysis process, the merged factors were tackled separately. Similar concepts from similar variables across cases were merged. When strong evidence across cases was found in relation to the variables, CCTAs were formulated. On the other hand, those variables which occurrence emerged from one case only were retained at this stage. Table 7.7 gives a summary of this process.

LEVEL 1 PERCEPTIONS CROSS-CASE ANALYSIS					
Merged Factors / CCTAs	Variables				
	Case A	Case B	Case C	CCTAs	Retained
Perceived Usefulness of Open Data	4	6	5	5	4
Perceived Business Opportunity	5	4	3	5	4
Perceived City Opportunity	3	6	3	3	2

Table 7.7: Perceptions level 1

As shown in Table 7.7, three CCTAs were generated from this phase. These factors were enriched by a total of thirteen tentative variables. Ten variables were retained at this stage. In addition, for each CCTA, a number of inbound and outbound links were tentatively defined. In particular, a total fourteen links were tentatively merged across these factors.

- ***Level 2: perceptions across two cases***

As part of Level 2, the cross-case analysis continued with the focus on those factors that emerged from two of the cases studied. These included *Perceived Voluntariness* (Cases A and B), *Perceived User-Friendly* (Cases A and B), and *Perceived User-Satisfaction* (Cases B and C). The strategy here was to consider these clusters and reflect upon these concepts in relation to the "missing case". Similarly, to what done for Level 1, a table is proposed as a summary of the analysis process.

LEVEL 2 PERCEPTIONS CROSS-CASE ANALYSIS					
Merged Factors / CCTAs	**Variables**				
	Case A	**Case B**	**Case C**	**CCTAs**	**Retained**
Perceived Voluntariness	1	1	--	1	0
Perceived User-Friendly	3	3	--	2	2
Perceived User-Satisfaction	--	3	4	3	2

Table 7.8: Perceptions level 2

As shown in Table 7.8, the three factors emerged as part of this stage were all translated into CCTAs. Evidence suggested that where the contribution of a specific case was missing, a reasonable explanation could be provided based on either the peculiarity of the context of the "missing case", or from other factors and variables within the case-specific adoption models. It is noted that one of the variables retained from *Perceived User-Friendly* (i.e. *Being Able to Get Information* from Case A, see sub-section 4.4.2) was re-considered as additional evidence supporting *Perceived User-Satisfaction*. Similarly, the fact that an open data service needs to be *Easy to Use* emerged from Case C within the related *Perceived User Satisfaction* factor. This was subsequently moved adding evidence to the newly emerged CCTA regarding *Perceived User-Friendly*. Therefore, at the end of this process two of the initially four retained variables were "re-allocated". Concluding, only two variables were retained at the end of Level 2 of this stage of the cross-case analysis. Ultimately, concerning the analysis of links to and from these factors, five were tentatively defined at this stage.

- ***Level 3: remaining perceptions***

In this last class of the cross-case analysis related to perceptions, the first step was to list all the remaining variables across the factors that emerged in relation to one case only. A total of twenty-four variables were considered at this stage. Twelve of these were variables emerged in specific cases in relation to five perceptions emerged for one case only. The remaining twelve variables were those retained form Levels 1 and 2 of the perceptions' cross-case analysis. At this stage, all these variables were listed and combined based on their similarities. As a result of the analysis of the content of these twenty-four variables, one tentative factor emerged across cases (including five variables) and four variables were found related to CCTAs that were previously generated. Table 7.9 (next page) gives a summary of this process. The last column on the right indicates whether each specific variable: was considered for the additional factor (indicated as "new CCTA"); was included in a previously defined CCTA (indicated as "integrated"); or if it was retained after this stage and its inclusion was re-considered during the final validation across cases stage. From the analysis of the content of the remaining variables, it emerged significant information related to the importance for the council and the overall ecosystem to gain credibility among the community of users and investors. Two specific factors were originally defined in this way for Case C (i.e. *Perceived Commitment of City Council* and

Perceived Credibility and Sustainability). Regarding Case A and Case B, other variables that emerged within different factors were found to contain similar information. Therefore, additional evidence was available for tentatively asserting the importance of perceiving credibility and sustainability of the open data release as well as of the long-term commitment of the city council. An overview of the variables included in this new CCTA is given in Table 7.10.

#	VARIABLE	ORIGINAL FACTOR	CASE	STATUS
1	*Rewards*	*Perceived User Interest*	B	Retained
2	*Amusement*	*Perceived User Interest*	B	Retained
3	*Interactivity*	*Perceived User Interest*	B	Retained
4	*Promote Existing Businesses*	*Perceived Potential Economic Growth*	B	Retained
5	*Job Creation*	*Perceived Potential Economic Growth*	B	Retained
6	*Stimulate Local Economy*	*Perceived Potential Economic Growth*	B	Integrated
7	*Community Awareness*	*Perceived Ecosystem Awareness*	C	Retained
8	*City Council Awareness*	*Perceived Ecosystem Awareness*	C	Retained
9	*Investors Awareness*	*Perceived Ecosystem Awareness*	C	Retained
10	*Account for Responsibility on Open Data Release*	*Perceived Commitment from City Council*	C	New CCTA
11	*Challenges to the Council*	*Perceived Credibility and Sustainability*	C	New CCTA
12	*Replicability of the Open Data Service*	*Perceived Credibility and Sustainability*	C	Retained
13	*Large Datasets Analysable and Shareable*	*Perceived Usefulness of Open Data*	B	Retained
14	*Discoverability and Understandability*	*Perceived Usefulness of Open Data*	C	Retained
15	*Partnership Strength*	*Perceived Business Opportunity*	A	New CCTA
16	*Level Of Engagement*	*Perceived Business Opportunity*	A	New CCTA
17	*Funding*	*Perceived Business Opportunity*	C	Retained
18	*Revenue Model Structure*	*Perceived Business Opportunity*	C	Retained
19	*Participation / Collaboration*	*Perceived City Opportunity*	B	New CCTA
20	*Interoperability Across Browsers*	*Perceived User-Friendly*	B	Retained
21	*Learning Opportunity*	*Perceived User-Satisfaction*	B	Retained
22	*Free of Charge*	*Perceived Usefulness of Open Data*	B	Integrated
23	*Easier Process*	*Perceived Usefulness of Open Data*	B	Integrated
24	*Detailed Knowledge Creation*	*Perceived City Opportunity*	B	Integrated

Table 7.9: Perceptions level 3, list of variables considered

LEVEL 3 PERCEPTIONS CROSS-CASE ANALYSIS			
New CCTA	**Variables Merged**	**New Tentative Variables**	**Retained**
Perceived Credibility and Sustainability	*Account for Responsibility on Open Data Release* (Case C) *Challenges to the Council* (Case C) (retained) *Partnership Strength* (Case A) *Level of Engagement* (Case A) *Participation / Collaboration* (Case B)	1	1

Table 7.10 : Perceptions level 3, emerging factor

Concerning this emerging factor, the tentative findings also included the inbound contribution of three external factors, and one perception, as well as one outbound link to another class of perceptions. Therefore, a total of five links were tentatively defined at this stage for this emerging CCTA.

On the other hand, as shown in Table 7.9, four variables were found to contain additional evidence for already defined CCTAs. Table 7.11 gives an overview of where these variables were re-allocated in terms of factors (central column) and of which specific variable (right column).

LEVEL 3 PERCEPTIONS CROSS-CASE ANALYSIS – REALLOCATION		
Reallocated Variables	**Factors**	**Variables**
Stimulate Local Economy	*Perceived City Opportunity*	*City's Competitiveness*
Free of Charge	*Perceived Usefulness of Open Data*	*Resource Saving*
Easier Process	*Perceived Usefulness of Open Data*	*Time Saving*
Detailed Knowledge Creation	*Perceived City Opportunity*	*Internal Efficiency and Efficacy Gain*

Table 7.11: Perceptions level 3, reallocation

At this stage, the first two phases of the cross-case analysis regarding perceptions were concluded. As a result, seven tentative factors emerged described by a total of twenty-one variables. Furthermore, a total of twenty-one links were tentatively defined. In addition, sixteen variables were retained at this stage as no sufficient evidence was found across cases for motivating their inclusion. These were carefully re-discussed and related additional data was collected during the final validation across cases stage.

7.1.3 Adoption Stages CCTA Development

The third unit of cross-case analysis was about the actual adoption stages through which the companies went from the initial generation of the idea for the service, through the actual adoption of open data, to achieving sustained use of this source for the service delivery. The main goal at this stage was to analyse the content of each stage and compare it with the remaining cases. Hence, the first

activity conducted was about listing all the adoption sub-factors emerged from the three cases and merging those that are similar in their contents across cases. Table 7.12 represents this first merging exercise.

#	ADOTION SUB-FACTORS	CASE A	CASE B	CASE C	STATUS
1	*Idea Generation*	x	x	x	Merged
2	*Open Data Adoption*	x	x	x	Merged
3	*Market Analysis*	--	--	x	Retained
4	*Pilot Implementation*	x	x	x (Private Beta) x (Public Beta)	Merged
5	*Acquire Funding*	--	x	--	Retained
6	*Wider Adoption*	--	x	--	Retained
7	*Commercialization*	x	x	x	Merged
8	*Growth*	x	x	x	Merged
9	*Sustained Use*	x	x	x	Merged

Table 7.12: Adoption stages cross-case analysis

As shown in Table 7.12, a total of nine different stages emerged across the three cases. Interestingly, six of these sub-factors were identified as similar for all the cases studied. Therefore, six new CCTAs were formulated. Furthermore, three sub-factors emerged in relation one case only. These were retained at this stage and their inclusion was re-discussed during the final validation across cases.

At this stage of the process, external factors, perceptions and adoption sub-factors were merged from the single cases findings. In total, seventeen CCTAs were formulated in relation to ten external factors and seven perceptions that influence the adoption of open data for commercial service innovation across cases. These were comprehensively enriched with the definition of fifty-six tentative variables, and twenty-one links between factors. Also, six adoption stages were tentatively defined as the most relevant in explaining the adoption process across the three cases studied. These results were aggregated and a tentative adoption model was generated subsequently. This constituted one of the inputs of the last stage of this cross-case analysis, i.e. the validation across cases. Finally, forty-one variables and three adoption sub-factors from the single cases were retained after this process because no sufficient evidence was found to support their integration. Therefore, more data needed to be collected in relation to these aspects. This was conducted as part of the final validation effort.

The next sections are dedicated to the reporting of the cross-case findings. The thick descriptions constituting factors, variables, and links between factors influencing open data adoption are outlined. Discussions of these findings and their positioning with the extant literature on open data services and IS adoption are presented in Chapter 8.

7.2 External Factors Cross-Case Assertions

In this section the nine external factors established as influencing open data adoption for commercial service innovation across cases are presented. Those defined from Level 1 of the cross-case analysis (see section 7.2) are tackled first. Regarding those emerged from Level 2 and Level 3, further reflections about the importance of the specific aspect in relation to the "missing case(s)" are provided.

- ***CCA1: Open Data Characteristics influence the adoption of open data for commercial service innovation***

The first factor that emerged from Level 1 of the analysis focuses on the characteristics of the open data and the dataset that is re-used for commercial service innovation purposes. In other words, for all the three cases these characteristics emerged as critical for adopting this source for designing commercial open data services. A total of nine variables were defined (see Table 7.13). It is noted that one of these was integrated into the final findings only after the validation stage (1.9).

CROSS-CASE ASSERTION 1						
Variables		**Case A**	**Case B**	**Case C**	**Retained**	**Validated**
1.1	*Open Data Online Availability Free of Charge*	x	x	x	No	Yes
1.2	*Amount of Datasets Available*		x	x	No	Yes
1.3	*Frequency of Update / Real-Time Data*	x	x	x	No	Yes
1.4	*Access / Machine Readability*	x	x	x	No	Yes
1.5	*Re-usage Open License*	x	x	x	No	Yes
1.6	*Standard / Format*	x	x	x	No	Yes
1.7	*Data Accuracy*	x	x	x	No	Yes
1.8	*Geo-Localisation*		x		No	Yes
1.9	*Structure / Metadata*			x	Yes	Yes

Table 7.13: CCA1 open data characteristics

CCA1.1 Open Data Free Online Availability: this variable could be considered ingrained in the definition of open data assumed for this study. All the cases studied use open data that is being delivered by the related local authorities (and by intermediary organizations that are in place to help city councils in the releasing process) free of charge through online portals.

CCA1.2 Amount of Datasets Available in the Context: the cases suggest that as more open data is provided about a context, the more likely it is that new open data services will be developed for the specific domain. The amount of open datasets made available defines the amount of information about the specific context to be re-used.

CCA1.3 Frequency of Update: one of the main barriers for re-using open data that emerged across the three cases is the availability of stagnant datasets. The cases studied suggest that the required

frequency of update varies across the datasets. For instance, while information about the status of a planning application should be updated "on a daily basis" (for Case A), data providing location of parking meters in the city (Case C) needs to be "cleaned once a year". Overall, data needs to be current in terms of representing realistic knowledge about the specific domain of the city.

CCA1.4 Machine Readability: datasets need to be machine-readable, i.e. that can be processed by technologies. Data in KML and CSV formats (Case B and Case C) as well as open data supported by APIs (Case A, Case B, and Case C) is being re-used by the cases. In relation to the development of APIs, it emerged that mesh-ups of different datasets would be highly facilitated by the development of a single platform for accessing and extracting all open data from one city authority.

CCA1.5 Re-Usage Open License: another recurring theme across all the cases highlights the need of open data to be legally re-usable for commercial purposes. This was found to be critical in two cases (cases A and C). Its importance is related to the actual nature of the data. Regarding Case B, no particularly sensitive data is re-used, meaning that no specific needs to address legal concerns are in place (e.g. outdoor cultural attractions). All the cases had undertaken activities in their early stages on research about legal aspects related to the datasets. The solution proposed by the cases is to adopt Creative Commons license in its Attribution formulation.

CCA1.6 Standard / Format: another variable that was common for all the cases was about the importance of having one clear standard for all data. From a technical perspective, all the cases developed software that enables the extraction of data from one or more databases. If the data is released in different formats, an effort needs to be undertaken to make it consistent with the software codes and so with the company's systems. It also emerged that the standard should be clearly explained and displayed as information attached to the related dataset. In order to develop a common standard, evidence from Case A suggests that this should be developed in cooperation between actual and potential users and the city council.

***CCA1.7 Data Accuracy*:** in summary, data needs to be exact, real, and precise as much as possible. No perfect data is expected by any of the cases, and all the companies are involved to some extent in data cleaning activities to fix minor errors in the datasets.

***CCA1.8 Geo-Localisation*:** this variable was established for Case B only. However, all the cases actually use data that is assigned with a location (where relevant). Examples include parking locations and location of meters (Case C), outdoor cultural attractions (Case B), and planning application information (Case A). Furthermore, the data collected suggests that one standard should be used (longitudes and latitudes) for providing the dataset with this feature.

***CCA1.9 Structure / Metadata*:** the need for having a structured dataset provided with metadata was highlighted by the people within Case C. In this way, the datasets need to be discoverable and

navigated without “time-consuming processes to figure out what is what”. However, no evidence was found from the other two cases. As a consequence, this aspect was initially retained and its inclusion was motivated from the outcomes of the final validation stage. As noted by S5, “discoverable, accessible, interoperable and re-usable are the key for the portal” (and for the open datasets). This idea was strengthened by the developer and the business representatives (S1, S2, and S7). All these reported their experiences navigating the datasets and “not understanding what these datasets were about” (S7). Thus, the importance of having structured datasets provided with metadata was included as the ninth variable within this factor.

Open Data Characteristics validation

This cluster was one of the most discussed and all subjects commented on this factor. However, significant validation of the CCA and its variables was achieved. The importance of releasing data through online portals was confirmed by all the participants. “The portal is at the heart of our open data strategy, and the key to provide a single access of structured data in an open format” (S5). However, as noted by S3, “the portal is just a catalogue”. Thus, data needs to have certain characteristics in these portals. In this way, the importance of having updated data, data that is accurate and geo-localised, and of focusing on the quantity of data published was acknowledged by all subjects. One of the most discussed statements was about machine readability and the subsequent formats of the datasets. While this is found as a requirement for potential re-users to adopt open data, this characteristic was argued to preclude non-IT-skilled people to access the data (S4, S8). An example was made by S8 about the provision of real-time data about public transports. According to this participant, providing this data in other forms (e.g. pdf) would be technically impossible as “by the time in which the format is completed, the data results not in real-time anymore” (S8). However, for transparency reasons (one of the main objectives of open data strategies by cities – as noted by S6), these datasets must be readable by anyone. Notwithstanding this, the re-usage of open data for these scopes goes beyond the focus of this study. All subjects agreed that, for commercial re-use of open data, APIs are better ways to give access to the datasets. A potential issue in this way arose from the interaction with S5, and was described as “amount of downloads versus greatness of the file”. In relation to this, the subject said: “an advantage of bulk downloads is that datasets can be downloaded at any time by as many people as needed, while for APIs the portal collapses if too many people try to access the same file at the same time” (S5). To address this issue the council decided to publish through APIs those files that are smaller than twenty megabytes. In relation to APIs, the developers were according on JASON to be the optimal one as it is “highly re-usable due to its interoperability and the usage of a unique standard” (S7). For spatial data, instead, the city council uses KML which, however, is considered “horrible for large datasets” (S1, S2). Emphasis was also placed on open data

licenses. As noted by S5, “one of the main goal is enabling perceptions of no-commercial restrictions for users” (S5). Creative commons and the attribution license is the suggested scheme according to S1, S3, S5, and S8.

- ***CCA2: Individual Factors influence the adoption of open data for commercial service innovation***

Across the three cases, the characteristics of those individuals that actually started the businesses were found as an additional factor that influences the overall adoption process (details are provided in Appendix 6.1). Overall, five variables emerged in this factor across the three cases studied. Of these, two were finally excluded from the cross-case findings.

CCA2.1 Experience in the Context: before actually adopting open data, the experiences of the CEOs of Case A and Case B had given these subjects deep knowledge about the context. This was leveraged in several ways, including: being aware of the current situation of the context and, subsequently of the market opportunities; knowing what data is needed, where it is situated, and who owns it; and exploiting the network of people and companies developed during these experiences (Case A). Case C demonstrates that this aspect is not necessarily a strict requirement. Nonetheless, evidence demonstrates that experience in the context can significantly facilitate the adoption process of open data for commercial service innovation.

CCA2.2 Attitude towards Open Data and IT: significant evidence was found across the cases in relation to the importance of individuals’ attitudes towards IT. All the founders and co-founders of the cases leveraged their knowledge in the field of software development and data science for adopting open data.

CCA2.3 Entrepreneurial Spirit: this variable appeared in two of the three cases studied. However, it was found that all the founders (besides being entrepreneurs given their open data services businesses) are playing roles in their local start-up ecosystem.

Individual Factors validation

The presentation of this factor did not lead to significant discussions. Participants agreed that having *Experiences in the Context* where open data is wanted to be re-used, and being IT-oriented entrepreneurs strongly facilitates open data adoption. Only in one case (i.e. S8) an observation has been made about the fact that “open data is driven by IT people” and about related consequences. However, this comment was mainly addressed for people within city councils.

Retained variables

Attitude to Work with Governments: this concept emerged from Case A only and has not been acknowledged as a requirement for re-using open data for service development. Only S1 argued that "this might help in putting in place relationships with the councils".

Creativity (Case B): this variable was in part believed as something that is not specifically related to open data, and in part as an aspect that is relevant to the tourism context (e.g. "in tourism you need to develop something creative, because there is already too much stuff out there", S2).

- ***CCA3: City-Related Factors influence the adoption of open data for commercial service innovation***

An initial cluster was defined encompassing the factors that emerged across the three cases about specific aspects of cities and the related city councils influencing the adoption of open data. For cases A and B one unique factor was defined encompassing both of these concepts. This was different in Case C where two different factors were formulated. At this stage it was rational to divide these variables between two different clusters: (1) *City-Related Factors;* and (2) *City Council-Related Factors* (see CCA4). Within the first cluster, all variables emerged across the cases in relation to the nature of the city were listed and analysed (see Appendix 6.2). This factor emerged from the fact that all the cases base their growth upon expanding the open data services across different cities. These cities are targeted based on several different aspects including the nature of the cities themselves. Three variables were established as influencing the adoption of open data for commercial service innovation. One more arose from Case B only (i.e. *Public Free Wi-Fi)*. This was initially retained and re-considered at the final validation stage. These interactions further supported its exclusion.

CCA3.1 Size of the City / Market Size*:* in cities were the context-related market opportunities are higher, the chances for new companies to re-use the related open data increase. For example, in cities where parking is easily accessible for free, or where a few constructions and building applications are submitted, the open data services offered by the cases A and C wouldn't be valuable as these wouldn't match an actual need to be solved. Evidence also suggests that bigger cities are places where open data is more likely to be re-used.

CCA3.2 Infrastructural Characteristics: a particularly important contribution to this variable was given by cases B and C. The focus was on innovative infrastructures for generating additional data from observing the city at a micro level, i.e. from sensors. In this way, a range of new datasets can be created with reliable data in real-time. Therefore, if cities have these infrastructures in place, these are more likely to be places where open data services are developed. However, from the case's experiences it arose that cities need to find and establish business models for the creation, and most importantly,

the maintenance and sustainability of these projects. It is noted that for Case C this aspect also included the presence of a certain parking infrastructure (e.g. “presence of meters”, “payment for parking” etc.). The absence of a contribution from Case A was interpreted as being motivated by the peculiarity of the urban planning context. No specific additional infrastructures are believed to improve the delivery of Case A’s open data services.

CCA3.3 City-Related Opportunity: this variable was defined as part of Level 3 of the cross-case analysis, and subsequently integrated within this factor. It encompasses five variables from across the three cases (see Table 7.4, sub-section 7.2.1). The common underlying assumption upon which these variables emerged is the presence of general opportunities within the cities targeted by the cases. The three cases were able to deliver their open data services because these were designed to address an existing problem and/or to harness a specific opportunity. These concepts emerged in Case A in relation to the low maturity of solutions addressing the process of getting urban planning information, and in the related opportunities of gaining efficiency and efficacy in these processes. In relation to Case B, these opportunities arose in terms of *Underexploited Areas* in the tourism sector. Finally, a similar concept emerged for Case C when highlighting its need to provide a *Valuable Service in the Context* (originally part of the *Service Characteristic*s factor). This aspect entails an opportunity for improving drivers’ experiences when involved in parking activities, and for providing new solutions to other existing operators in the context.

City-Related Factors validation

Similar to the previous CCA, this factor was well accepted by all participants and no significant feedback was received. All the subjects approached agreed that efforts to design and deliver commercial open data services depend reasonably on the *Size of the City* and its market potential. According to S3 and S8, a viable way to deliver open data services on “small and medium cities” is if open data is provided at the national level as the aggregation of all cities-related datasets. Also, infrastructural characteristics of cities were argued to be relevant, especially if these are related to the development of technological capabilities (e.g. “an integrated sensors’ network”, S7) that enable the extraction of additional data and its release on a near real-time basis.

Retained variable

Public Free Wi-Fi: this aspect arose from Case B where the targeted users of its open data services (i.e. visitors) often face “roaming problems”. In other words, foreign visitors usually do not have internet connection which in turn precludes them from using Case B’s open data service. Hence, it was considered critical for cities to have *Public Free Wi-Fi* for re-using its open data. As a result of the

validation discussions this was excluded and found to be relevant merely for Case B's context (i.e. tourism).

- ***CCA4: City Councils-Related Factors influence the adoption of open data for commercial service innovation***

In all the cases, an assertion has been formulated in relation to the role played by city councils. Several features regarding how these entities are structured and operate were found to influence open data adoption. From another angle, all the cases base their growth upon extending their activities across different cities. When targeting a new city, together with looking at its specific characteristics (see CCA3), the first set of activities aims at ensuring the sustainable availability of open data. To do this, city councils are approached and a variety of aspects emerged to influence the choice of adopting (or not) their open data. A total of eight variables initially emerged across the cases for this cluster (see Appendix 6.3). However, three of these were retained at the first stage as only one case contributed to their definitions. As a result of the final validation interactions, one of these three variables was integrated in the final findings (4.6).

CCA4.1 Open Data Strategy / Commitment: all the cases studied shared the idea that city councils should have in place long-term strategies and commitment for releasing valuable open data. These strategies need to be unique as opposed to fragmented across internal departments. For instance, Case C experienced a situation where the open data needed was coming from different internal functions (i.e. urban planning and transport). This resulted in data provided with different formats and licenses and updated at different times. Ensuring long-term commitment for its provision is also required for generating perceptions about the opportunity of creating a sustainable commercial business based on open data (both among re-users and potential investors).

CCA4.2 Open Data Awareness: this variable emerged in relation to the challenges that the cases faced in terms of "educating the councils". City councils should be aware of open data, what are its benefits, and of who and how is using it, in order to understand the needs of the users and react accordingly. This concept includes also awareness of the activities, technologies, stakeholders, and processes that need to be in place to effectively release open data. Case B, in this way, emphasised the importance for city councils to have internal champions.

CCA4.3 Organizational Structure / IT Skills and Department / City CIO*:* the importance for councils to have the right IT skills in place and an effective IT department established was emphasized in cases A and C. Furthermore, the presence of a city CIO is believed to significantly increase the quality and the effectiveness of the open data initiative. This feature enabled the companies to approach the city through an effective and informed interlocutor (i.e. "the city CIO or the head of IT", Case C). By doing

this, the company's needs were outlined in a language that was understandable by the councils (i.e. "technical language").

CCA4.4 Budget and Resources: another aspect emerged in relation to the need for councils to have budget and resources allocated to open data ("across departments", Case C). One of the most common responses from those councils that rejected the idea of embracing open data was "lack of budget and resources" (Case A).

CCA4.5 Engagement Initiatives: as stated by Case A's CEO: "just releasing open data and wait that something happens is not enough". The three cases agree that it is crucial that city councils engage with actual and potential users through different mechanisms. The forms through which these engagements should be undertaken emerged to be "workshops", "showcases", "periodic meetings" and other types of events. The objective of these initiatives should be twofold. On one side, city councils should highlight their needs in the contexts and promote the related open data that is, or can be, released. On the other hand, actual and potential users should educate the councils in terms of what data they need, in which format, and how the current data can be improved.

CCA4.6 Outsourcing of Open Data Services: this concept initially emerged from Case A only. It arose that one of the barriers for new companies to develop open data services is represented by some situations in which the city councils outsource the creation of these services to "big companies". Further confirmation about the role of these activities was sought during the validation stage. Surprisingly, this idea found considerable confirmation from both of the business representatives approached as well as from S4 and S8. According to these participants, city councils seek external development of applications to gain legitimacy of the goodness of their open data programs. Big corporations and other bodies (such as "the central statistic office", S4) are hired for developing open data services. Although benefits emerging from these approaches were highlighted in some ways, participants agreed that these initiatives are obstacles for new start-ups that want to develop new open data services. Accordingly, this aspect was included as the sixth variable within this factor.

City Council-Related Factors validation

Together with CCA1, this factor was the most discussed; significant feedback has been received. In summary, considerable validation was achieved for this CCA. All subjects agreed that city councils should have long-term strategies in place and awareness should be widespread across departments. S5 highlighted the need of an open data strategy that "must be underpinned by structured governance and a technical framework developed across departments". Furthermore, "another objective is to present the concept of open data to as many people as we can". An open data governance board needs to be established to drive the strategy, "because without focused committee we won't succeed" (S5).

Another interesting point of view emerged from a representative of a city council that hasn't embraced open data yet (S6). This subject highlighted the issues that the council is facing in deciding whether implementing an open data strategy or not. These include: (1) "ownership! Often the local authority does not own the data in total"; (2) the local authority is "conservative" in its nature and big internal resistances to change are in place; and (3) "it is hard to understand the value behind it". Challenges perceived by users (S1, S2) include the lack of champions and/or people responsible of open data for a specific sector/department/area. These subjects also underlined the importance of having an IT department and a related officer as the body that can be approached by people willing to re-use open data. Connected to this aspect, an interesting discussion emerged from S8. According to this participant, having "IT people" driving open data within councils is seen as negatively contributing to open data outcomes beyond service innovation. S8 highlighted this aspect by criticizing the absence in these findings of a reflection about the "epistemic community" involved: "you don't talk about the epistemic community; who is that is involved in open data in the city? Who is the champion in the city? If you look at their characteristics they are all with an IT background, and the leaders are always the CIO or CTO. They are looking at open data from an IT perspective. So data becomes an object and they don't think substantially what is in the dataset" (S8). Therefore, interpreting this reflection, having IT-oriented people driving open data within councils is seen as beneficial only for other IT people (as asserted within *Individual Factors*). On the other hand, this is seen as precluding other aspects of the councils' open data strategies. It is noted that other uses of open data (e.g. as a community engagement tool) are beyond the scope of this research. Notwithstanding this, I acknowledge how a heterogeneous group of people working at open data might be beneficial for open data strategies in general.

The importance of engaging with actual and potential users was commented as follows by S5: "for us it would be important to know who is using the data and how is using it. By knowing this we can help, we can focus our efforts towards something that is valuable, as opposed to roughly estimate what can be potentially re-used. We were doing this and it was as navigating in the dark". This subject continued saying: "we need to bridge the technical skills gap, by engaging with the huge amount of stakeholders including start-ups and businesses, as well as researchers. Every month a meeting is scheduled with stakeholders with the goal of presenting the progresses and collect feedback which will be inputs for the work conducted in the following month". Therefore, the open data services' ecosystem should be "demand pulled as opposed to supply pushed" (S5). Informative sessions and workshops are believed to be the best ways to stimulate this phenomenon. Also, "a support desk for users is available" (S5). Furthermore, with initiatives like public meetings were groups are formed and discuss how to move forward with the program, "we want to show that we are responsibly taking this strategy forward"

(S5). This last aspect strengthens the idea of city-council related characteristics influencing the creation of *Perceived Credibility and Sustainability* of the overall ecosystem (see CCA16). Finally, S5 highlighted how important it is "to allocate budget and resources internally to both the data collection and the open data dissemination processes".

Retained variables

Support Creation and Sustainability of Open Data Services (Case A): initial funding provision and promotions from the councils were the main focuses of this variable. While the option of co-funding open data services was excluded ("open data is at a too early stage for this" – Case A's CEO stated), examples included prizes for best projects or other similar forms of support. These were leveraged by Case A as acknowledgments when targeting new councils to demonstrate that the solution proposed was valuable somewhere else, and so to "push them to come on board". This variable was presented to the city councils and business representatives (S5, S6 and S1, S2 respectively). This aspect was generally interpreted as being already integrated in the need of engaging with actual and potential users already tackled within another variable (CCA4,5). All in all, no considerable confirmation was achieved for this variable, which was therefore excluded from this CCA.

Internal Laws: this variable was supported by an experience of Case A in which one city council in US was releasing open data "by default" (i.e. internal laws were in place regarding open data). During the validation stage, this aspect was criticized by the city council's manager (S5). This subject's critiques were oriented towards the impossibility of establishing such laws "at the current level of maturity of open data". Hence, given that no supporting evidence was collected for establishing this variable, this was excluded.

- ***CCA5: Local Ecosystem Factors influence the adoption of open data for commercial service innovation***

As stated in section 7.3.2 of this dissertation, this CCA represents an external factor that was refined according with the feedback received during the final validation interactions. Two CCTAs were previously formulated for two different factors named *Local Community Factors* (from Level 1 of the analysis) and *Open Data Ecosystem* (from Level 3 of the analysis). This step was undertaken as the content of the original CCTAs presented was found to be similar. Both of these factors underline the fact that entities, initiatives, and local actors (different from the actual open data services businesses and the city council) influence the adoption of open data for commercial service innovation.

Regarding *Local Community Factors*, the term "local" had been added to this cluster in order to distinguish this factor from others similar in nature but that have a focus on a larger scale (i.e. *EU Open*

Data Ecosystem Factors from Case B, and *Ecosystem Factors* from Case C). Regarding the second CCTA integrated (i.e. *Open Data Ecosystem*), eleven variables were initially found to be related to the need of establishing actors and respective roles within an ecosystem. These variables, and the related factors and cases from which they were taken, are listed in Table 7.14.

Variable		Factor	Case
1	*Innovation Network*	*Local Community and Ecosystem Factors*	B
2	*Triple-Helix Innovation*	*EU-Driven Open Data Ecosystem*	B
3	*Volunteers Engagement*	*Organizational Factors*	B
4	*Partners Engagement*	*Organizational Factors*	B
5	*Long-Term Structured Plan*	*Intermediary Organization Factors*	C
6	*Level of Ecosystem Engagement / Workshops*	*Intermediary Organization Factors*	C
7	*Relationship (Independence) with the Council*	*Intermediary Organization Factors*	C
8	*Internal Open Data Knowledge / Skills*	*Intermediary Organization Factors*	C
9	*Partnerships with other businesses*	*Ecosystem Factors*	C
10	*Volunteers Engagement*	*Ecosystem Factors*	C
11	*Volunteers*	*Local Community Factors*	B

Table 7.14: Level 3, open data ecosystem CCTA

After, re-visiting the content of these variables across the two cases in which they emerged (cases B and C), three themes emerged (see Table 7.15). Firstly, five variables were related to the need of establishing an intermediary organization that facilitates the process of releasing the data from the city council (this entity was described as "innovation network" by the CEO of Case B). Secondly, the cases' experiences show the importance of establishing partnerships with other businesses in the local area. Finally, volunteers' engagement has emerged as relevant for both cases.

New Variables	Case B	Case C	Variables integrated*	Retained
Intermediary Organization	x	x	1; 2; 5; 6; 7; 8;	No
Partners Engagement	x	x	4; 9; 2;	Yes
Volunteers Engagement	x	x	3; 10; 11;	No

Table 7.15: Open data ecosystem CCTA, variables integrated
*numbers are related to Table 8.2

As a result of the validation stage, these two CCTAs were integrated. Appendix 6.4 provides an overview of the variables emerged, the original CCTAs in which these were included, the cases in which each variable occurred, and whether these variables were initially retained and subsequently validated (or not). As a result, one variable related to the need of engaging in partnerships (outside the city council) was excluded as no sufficient evidence and confirmation was achieved.

CCA5.1 Hack Competitions: this experience represented the moment in which all the cases adopted open data for developing an idea previously generated. The different cases participated in different

competitions organised by different entities (i.e. national councils and business associations). Across the cases, these were seen as experiences that were fundamental to discover open data and to get initial funding for stimulating further developments of the related open data services. However, “hack competitions alone are not a viable option for spreading open data” (Case C).

CCA5.2 Champions: this variable has been placed under this cluster because of the emphasis emerged in relation to the “limited role” that champions from inside the city councils can play in this ecosystem. The idea is that champions from outside the authorities can operate with flexibility without facing issues in relation to ensuring fair competition and at the same time support open data service businesses. According to the cases, their roles should include “spread awareness among government and consumer sides” and join together the overall ecosystem’s stakeholders through community-oriented initiatives (see next variable).

CCA5.3 Community-Oriented Initiatives: this range of initiatives highlights the need of stimulating engagement opportunities outside the city councils. Differently from city council-related initiatives (CCA4,5), councils shouldn’t lead these events in order to enable actual and potential open data re-users to present their ideas and their needs. This is seen as a “more effective way for educating and stimulating the councils”. The formats proposed to implement these events include: workshops and periodic meetings, showcases, and coder-dojo. Community organizations, enterprise associations, and other NGOs are believed to be the entities that should organise these initiatives.

CCA5.4 Intermediary Organizations: this variable emerged from Case B and Case C. The main role of these bodies is to help cities in the open data release process by providing IT skills and creating new datasets. Similar to city councils, given that these intermediaries are publishing open data through online portals for being re-used, these should have: internal open data knowledge and skills; a long-term structured plan; the ability to foster engagement with other parties in the ecosystem (Case C). The findings from Case C suggest that the level of relationship that these bodies have with the related city council needs to be considered. In particular, if too closely related to the council, these entities do not act in a way that stimulates and challenges their operations. Thus, the CEO of Case C highlighted the importance for these organizations to be NGOs (he mentioned Open Data Institute in UK as the “perfect example”) in order to act as the counterpart of the council. No insights emerged in this way from Case A due to the fact that it does not take any open data from these organizations.

***CCA5.5 Volunteers Engagement*:** both Case B and Case C leveraged the possibility of engaging with volunteers and interns. These played a critical role during the first stages of the cases’ developments. Technical tasks were assigned to volunteers/interns in both cases. Regarding Case A, the contribution of volunteers was not observed. However, the CEO underlined their importance based on his experience in community-oriented initiatives (outside the company).

Local Ecosystem Factors validation

Significant validation was found in the interpretation of the cases and the resulting five variables established across the two CCTAs. Firstly, great importance was given to the hack competitions as initiatives that substantially foster the diffusion of open data for new service creation. In this way, S3 highlighted: "we used the hack competitions as a way to demonstrate the value of open data to both the council and the public. The great support we had from media helped dramatically to spread awareness among both councils and researchers". The importance of having in place champions and initiatives from outside the council was further supported from the data collected in relation to the previous cluster (see "this ecosystem should be demand pulled"). As reported by S5, "we want to hear more from businesses, we need to know a lot more from them; there is too much data to be published all. We all create data continuously. If you use open data, we encourage you to be part of our development plan and to tell us your story". This discussion continued with a confirmation of the importance of intermediary organizations in the overall process. "The help from intermediaries is needed and crucial" – S5 stated. The Open Data Institute in UK is again taken as an example for the effectiveness of posing challenges to the council (aspect underlined by S1 and S3). Related to this, S6 highlighted how today the Open Knowledge Foundation acts as a sort of stimulator for these initiatives, but internationally and across councils. Intermediary organizations, on the other hand, act at the local level. As confirmed by one of its representatives (S3), the focus of these bodies should be on single cities and tackling one department at a time. "We are strongly focusing on timeliness of data and we do it at the individual dataset level" – S3 added. Finally, the role of volunteers (emerged in both of the original CCTAs) was found to be relevant from the experiences of the subjects with other community-oriented initiatives. S5 described his surprise in "seeing the contribution from volunteers" emerging from these initiatives.

Retained variable

Partners Engagement: in both cases B and C the establishment of partnerships with other businesses contributed to the successful deliver of their open data services. Regarding Case B, two types of partners are identified: (1) entities that are targeted to enrich and improve the open data service (e.g. tourism associations); and (2) partners within the open data field. Concerning Case C, two types of partners emerged as those operating in the parking industry (e.g. payment providers) and those that are not (e.g. newspaper agencies for advertisement). During the validation phase, all subjects interpreted this aspect as important depending on the context. No specific connections with this variable were underlined for open data services businesses (as opposed to other forms of business).

- ***CCA6: Organizational Skills influence the adoption of open data for commercial service innovation***

Another recurring factor across all the cases was related to *Organizational Skills*. In the three cases studied, the presence of certain internal skills was found to be a critical aspect for businesses to re-use open data. Interestingly, an almost homogeneous set of skills emerged across these three companies. When aggregating the variables from the different cases (see Appendix 6.5), one simplification has been made: the variables *Data Analytics Skills* (Case A), *Data Analytics and Data Mining Skills* and *Data Cleansing Skills* (Case B), and *Data Analytics and Management* (Case C), were integrated under the newly defined *Data Analytics and Data Management Skills* (CCA6.2).

CCA6.1 Programming Skills: all the cases offer a range of digital services delivered through mobile and web applications. Thus, internal programming skills (i.e. iOS and Android for mobile and HTML5 for web applications) are needed to actually design and implement such applications. This was described by Case A's CEO as their "core skill". All the three founders and the vast majority of their employees, interns, and volunteers, have software development skills and backgrounds.

CCA6.2 Data Analytics and Data Management Skills: all the cases went through a certain set of activities for extracting the data and visualising it onto the final application. These activities included: inspecting and understanding the datasets, data cleaning processes, modelling the data, developing the standard, integrating the data, and loading it into the actual content management system. Within this variable, also data mining and analytics skills are included. A peculiarity that is common across the three cases is that these analyse and extract knowledge "on the back of the open data services" (Cases A and B) or, in other words, "from data collated through the app" (Case C). The outcomes of these processes are parts of their value propositions.

CCA6.3 Geographic Information Systems Skills: all the cases operate with data that is localised geographically. This variable did not emerge for Case B; however, GIS components are ingrained in the open data services offered by this organization. For example, localisation of cultural attractions and businesses (among other elements) is active part of the service offered. With respect to both Case A and Case C, these have faced issues in this way. The CEO of Case A stated that "in Europe it's very difficult to geo-localise information" referring to the scarcity of these capabilities within public authorities. Related to this, Case C faced a situation in which data was localised inaccurately and using different standards. Thus, in both cases internal GIS skills were leveraged to re-use open data.

CCA6.4 Graphic Design and Data Visualisation Skills: by looking at the services offered by the three cases, it can be inferred that graphic, web, and mobile design skills were essential for the design and delivery of these services. These skills were leveraged in Case B "to transform excel tables into

something easy to see and detect" within an application "with a sleek design". Emphasis was placed from all the cases on the importance of achieving a sleekly designed application. These are believed to lead to a service that is "attractive for users" (Case B), "easy to use" (Case C), and "user-friendly" (Case A).

CCA6.5 Project Management and Communication Skills: the fact that the companies grow by scaling their services across cities, requires specific skills to approach these councils. These skills were /are leveraged to convince city authorities on the value of the proposed solution, and to "establish and maintain" relationships. For cases B and C this set of skills was needed to establish and sustain local collaboration also with other public and private organizations (intermediary organizations and industry partners). Finally, Case B required this internal knowledge when involved in the "complex funding application process".

CCA6.6 Knowledge about Legal and Policy Context: this set of skills was identified to be critical in cases A and C. Particular importance is given in terms of the need for open data services businesses to: acquire knowledge about what data can be re-used and what can't be; and in relation to how this data can be legally published and sold to the market. The legal licenses associated with the open datasets need to be understood. Often, policy issues reflect privacy concerns. Moreover, companies should have specific legal knowledge about their own contexts.

Organizational Skills validation

Regarding this CCA, general confirmation was achieved from the interactions at this stage. No major discussions were undertaken. The majority of subjects referred to these skills as "IT skills". The participants confirmed the interpretations around the importance of generating secondary data from the service and the role of data analytics skills for undertaking this process ("data analytics is the key" (S1)). The need of having in place these IT-related skills was interpreted from the nature of subjects that are generally interested or actually involved in open data re-usage processes. For example, both of the business' representatives (S1 and S2) have advanced IT-related background, experiences, and interests. Data analytics is in both cases the main component of their open data service business ideas. Also, people from the councils (S5 and S6) confirmed that those that manifested interest in re-using open data for new commercial service development are mainly "software developers and data scientists". When looking at the skills presented, a significant confirmation was provided also in relation to the need of knowing the policy context in terms of what and how open data can be re-used for commercial activities. Finally, no major objections emerged when describing the interpretations about the importance for the cases studied of *Project Management and Communication Skills*.

The cross-case analysis continued with the focus on those factors that emerged from two of the three cases studied (i.e. Level 2 of the analysis, see section 7.2). These are *City-Business Collaboration* (cases A and C) and *Service Characteristics* (cases B and C). Differently from the CCAs presented so far, these also include reflections in relation to the "missing case".

- ***CCA7: City Council – Business Relationships influence the adoption of open data for commercial service innovation***

The findings from the cross-case analysis suggest that having a relationship with the city council is a key for developing a successful open data services business. This is reflected in a range of perceptions around business opportunities, which, in turn, led the cases to perceive open data as a valuable source for commercial service innovation. The common underlying assumption supporting this CCA was that for both Case A and Case C "working closely to the city council" has been a cornerstone for their successes. Open data services companies need to ensure that the open data available today (consistently with CCA1) will be sustainably provided in the future. To ensure this long-term commitment from the councils, two different elements were leveraged by the cases. Firstly, both cases designed open data services that benefit the council's internal operations in the context. Secondly, Case A established commercial contracts with the councils. Moreover, Case C acknowledges the importance of enabling *Collective Learning* for fostering long-term collaboration. Therefore, three variables were initially considered for this cluster (see Appendix 6.6). One of these was initially retained and finally excluded (i.e. *Collective Learning*).

CCA7.1 Win-Win Relationship: the release of open data is often perceived as a process that "increases the workload" within the councils. One aspect through which both Case A and Case C stimulated the councils to provide "precise, current, and useful data" on a continuous basis is ingrained in the open data services offered. Both cases leveraged the fact that the services developed were designed to have a positive contribution to the councils and the related management of the respective contexts. Case A's open data services minimize the internal resources allocated to the management and administration of urban planning information. Similarly, Case C, through the data analytics processes, is able to extract significant knowledge about the parking domain. This is in turn leveraged by the councils to substantially improve internal management of the context. An important consequence of these activities is reflected in the companies' growing processes. When targeting other cities, both cases approach the councils leveraging relationships already established in other places as a demonstration of the value provided by the proposed solution. The name *Win-Win Relationship* (Case A) was assigned to this variable.

Win-Win Relationship - Missing Case: although this variable didn't emerge from Case B, its open data services are designed to bring additional benefits for the city council. According to the CEO, by running analytics on the services' usage, Case B does "crowd sourcing of valuable data for the local authority". Unlike the other two cases, no specific evidence was found about leveraging these outcomes to establish a relationship with the councils. In a smaller scale, this was leveraged when approaching new councils for having the data needed. The reason why this variable did not emerge for Case B can be traced in the nature of open data re-used. The service is currently based on open data for which high frequency of update is not required (i.e. outdoor locations, scenic locations, nightlife, accommodations, maps etc.). This reflection further explains the words used by the CEO when he stated: "if the data is available we make it […] and we are getting to a situation in which all of this will be fully scalable. So then it is just a matter of getting the data". Thus, it can be concluded that these are the reasons why no emphasis has been placed towards establishing relationships with the councils. It is however rational to think that this need will emerge for Case B in the coming future (the case is working at integrating new "near real-time datasets" such as "events", "weather data", and "queue times at cultural attractions" etc.). It is also noted that, regarding Case B, a set of perceptions emerged to be critical in relation to city opportunities from the open data services offered. These perceptions are understood as a stimulus for the council to embrace open data, which is strongly connected to what stated in relation to Case A and Case C within this variable.

CCA7.2 Structure of the Relationship: another variable emerged in relation to the nature of this relationship. Regarding Case A, two different situations were encountered. In Europe, it was sufficient to establish the above mentioned *Win-Win Relationship* to ensure the provision of data in the long-term. Differently, in US, contracts were stipulated with the local authorities. Regarding, Case C, "establishing commercial partnerships with the council" (i.e. stipulating contracts) was the initial objective of the company. However, when adopting open data in the first city, another private company was responsible for managing the parking context. This situation led the case to develop insecure perceptions about the ability of delivering the service in the future for potential "replication threats" from the existing partner. These perceptions, if becoming true, would have impeded Case C "to exist". This discussion led the people within the case to conclude that commercial partnerships might result into "lock-in situations" where innovative open data services developed by emerging businesses are inhibited to become sustainable business models. The CEO proposed a solution to this issue. He believes that these challenges could be solved by establishing "small contracts with the council for piloting a solution" and/or putting in place "shared investments for the development and initial deployment of a pilot solution" in the context.

***Structure of the Relationship - Missing Case*:** similar to what stated for the previous variable, this is expected to become relevant for Case B in the near future (i.e. when the nature of the open data required will be reflected in the need of establishing relationships with the related council).

City Council – Business Relationship validation

When tackling this factor all the subjects acknowledged the importance of establishing relationships between the open data services businesses and the councils that supply the datasets. The reason why this aspect is considered a cornerstone in this landscape is found in the need of ensuring the long-term commitment of the council in providing open data (which in turn is believed to be fundamental for perceiving business opportunities). However, how to establish these collaborations remains unclear. S1 proposed the idea of having "service level agreements" with the councils. According to the councils' representatives (S5 and S6), the solution of enabling win-win relationships is found as the most suitable, because the "idea of providing open data as a service to the businesses is not part of the thinking within public authorities at the moment" (S5). The evidence from the cases that supported this statement found considerable interest from the businesses' representatives (S1 and S2). This aspect stimulated reflections on the future activities planned for their businesses. Notwithstanding this, an optimal structure for these relationships "has not been discovered yet" (S5). Commercial contracts between the two entities might be beneficial for one case but would constitute a barrier for new innovative services and related businesses to emerge. As noted in CCA4, having internal regulations that force council to publish open data by default is currently not achievable; "we are not at that level of maturity yet" (S5). Overall, the importance of this variable has been extensively acknowledged and all the subjects agreed that new mechanisms should be studied and established. As part of this, no shared agreement was found on whether these should be context-specific or to be applied across different domains. For instance, in the opinion of S8, "contracts could work for some domains like urban planning in your study but might be a barrier for others, like the parking-related service".

Retained variable

Collective Learning: this variable emerged from Case C. It was excluded from the cross-case analysis as its content was already considered as part of two other variables established. Arguments about the need for councils to engage with businesses and explore together context-related opportunities were already ingrained in CCA4,5. On the other hand, in relation to the need for businesses to educate the councils (e.g. on the data needed), these concepts were already tackled as part of *Community-Oriented Initiatives* (CCA5,3).

- ***CCA8: Service Characteristics influence the adoption of open data for commercial service innovation***

The second and last external factor emerged from two of the three cases studied highlights the importance of the characteristics of the open data services offered for the successful development of the businesses. The nature of this cluster (i.e. specifically related to the service offered) suggests that most of the concepts emerged within each case represent peculiarities of the specific context in which each case operates. For example, the fact that for Case B the service had to be designed in order to provide an "interactive learning experience to the users" does not bring any relevant argument for studying the adoption of open data across contexts. Overall, it emerged that two specific characteristics are in place for both Case B and Case C. Further reflections led to consider these variables as relevant for Case A as well. On the other hand, twelve variables were initially retained as the interpretation of the evidence suggested that their contribution was limited to the specific contexts from which these aspects emerged. However, when undertaking Level 3 of this analysis (i.e. all remaining and previously retained variables were considered as the unit of analysis), two of these twelve variables were re-allocated to other factors. This was the case of *Valuable Service in the Context* (Case C, see CCA3,3 above) and *Addressing Users' Needs* (Case B, see CCA11,1 below). In summary, from the fourteen variables emerged across cases B and C, two were established and validated here; two were re-allocated to other factors; and ten were finally excluded as their contribution was found to be relevant in the specific contexts in which these were defined (see Appendix 6.7). Overall, the findings demonstrate that the two variables included in this CCA play a crucial role for: (1) facilitating growth across cities (CCA8,1); (2) providing a valuable service for both the council and the market (CCA8,2); and, as a consequence, (3) fostering the establishment of *Win-Win Relationships* with the local authorities.

CCA8.1 Unique Content Management System: this variable was defined from the need for cases B and C of having a scalable service. Scalability is achieved by having a service designed with a unique content management system. For both cases achieving scalability is crucial as part of the growing process. By having a scalable service, "the big work has to be done once only" (Case B) and "new datasets can be loaded easily and quickly from different cities" (Case C). The idea of developing a unique content management system as a technical aspect of the service relies on the fact that it provides a single interface for loading, editing, updating, deleting and maintaining the content. The availability of this solution significantly facilitates the integration of new datasets (both from already integrated cities and from new ones).

***Unique Content Management System - Missing Case*:** with respect to Case A, the importance of having a scalable solution to facilitate growth across cities emerged within the *Perceived Business Opportunity* factor. However, a focus of major importance in this way was about the data standard. This might seem dissimilar and not related to the need for Cases B and C to have a unique content management system. However, data format management is one of the features of such systems. Therefore, confirmation about the importance of this variable is also found in Case A.

CCA8.2 Secondary Data Generation: the second variable that emerged across the two cases is about the importance of having additional data generated from the actual open data service. Both cases ingrained in their services data analytics processes to explore, derive and structure previously unknown knowledge about their specific contexts. This knowledge is subsequently leveraged in two different ways. Firstly, both companies deliver part of this knowledge generated to the councils. This concept has been already underlined as crucial in establishing *Win-Win Relationships* (CCA7,1). Secondly, this knowledge constitutes in both cases a relevant part of the revenue model associated with the open data services. These processes were commonly understood by both cases in terms of the importance of generating new knowledge "on the back of open data".

Secondary Data Generation - Missing Case: as expected, when re-visiting the nature of the service offered by Case A, it emerged that this characteristic is ingrained within its open data services too. In particular, Case A, besides improving councils' internal operations, provides the councils with additional knowledge extracted through data analytics processes. This knowledge includes "trend analysis" in the context for "assisting future development plans" and "locating inefficiencies".

Service Characteristics validation

Also regarding this cluster no significant objections were made in relation to the data interpreted from the three cases. Also, no significant discussions emerged. A common belief around the importance of the data generated "on the back of the service" (S2) was observed. In this way, "data analytics is the key" (S1). Thus, services provided need to have a form of data generation process on the back of the service enabled by analytics. Furthermore, all subjects acknowledged the importance of providing a scalable service and that the design of a single content management system is "the right way of doing it" (S7).

Retained variables

As expected, the majority of variables emerged across the cases were finally excluded. When discussing the variables initially retained at the validation interactions, no specific arguments emerged against these interpretations. The only aspect that was briefly discussed was in relation to *Complexity*

(Case C). Both of the business representatives (S1 and S2) initially agreed about the importance of reducing the probabilities of "someone else doing the same thing" (S1), i.e. of replicability of the open data service. However, when trying to match these thoughts with the other contexts, no further evidence was found (or arguments provided) supporting its inclusion in the final cross-case findings. The importance of adding complexity to the service designed was interpreted to be relevant only for the peculiarities of the parking context (i.e. high level of maturity of solutions, high competition, and extremely regulated business environment). Therefore, confirmation was achieved about the exclusion of all the ten variables initially retained.

- ***CCA9: EU-Level Factors influence the adoption of open data for commercial service innovation***

This factor represents the last external factor emerged. Its definition was achieved by merging similar variables across cases as part of Level 3 of the analysis. Emphasis was placed from Case B (and in a minor scale from Case C) on the importance of having an EU-level strategy and commitment. Such an approach is believed by the CEO of Case B to provide a common way of releasing data which would improve subsequent scalability of open data services across cities. These strategies should also include "funding schemes". A total of four variables were established across the two cases as shown in the following table.

#	Variable	Factor	Case
1	*EU-Level Strategy*	*EU-Driven Open Data Ecosystem*	B
2	*EU-Level Champions*	*EU-Driven Open Data Ecosystem*	B
3	*Best Practices Repository*	*EU-Driven Open Data Ecosystem*	B
4	*EU-Level Strategy*	*Ecosystem Factors*	C

Table 7.16: EU-level factors, variables integrated

These were integrated in a single concept named *European Level Commitment and Strategy*. This concept was further validated during the final validation stage. Given that the variables integrated from the single cases did not include Case A, further reflections in this direction are provided below.

CCA9.1 European Level Commitment and Strategy*:* the importance of having an EU-level approach was highlighted in Case B for the intrinsic role that the European Commission plays in the open data context. Open data gain momentum across European cities following claims that the commission made around the potential of open data. Furthermore, legislation (i.e. open data re-usage licenses) is EU-driven. A call is made to the commission to act as a showcase and best practices repository, thus creating "one location for all initiatives" (and "not fragmented projects"). This is believed to enable knowledge-sharing and collaboration across countries. As part of these initiatives it emerged the key role that champions (called "ambassadors") should play to foster the overall ecosystem by: "going around cities and educate them", "providing examples of successful services", "exporting toolkits",

and “promoting best practices”. From Case C, on the other hand, the focus was on funding opportunities and the importance of having a common approach across European cities to facilitate scalability of solutions (i.e. to enable growth). Moreover, as part of Case C, the role of EU has been extended to “challenging the operations of the city councils”. The importance of having challenges to the councils emerged as a specific variable (see CCA16,2) contributing to the perceived credibility and sustainability of the overall ecosystem.

Regarding Case A, no emphasis on these aspects was observed. This was interpreted as motivated by the fact that the company is operating in both EU and US. However, it can be inferred that if all cities in Europe would adopt the same approach in releasing open data, significant benefits would be achieved by Case A. These benefits are expected to address the main barriers and challenges that the case faced throughout its history so far, i.e. educating the councils and promoting the standard. Ultimately, also Case A uses open data covered by EU-based legislation. Hence, to some extent, the importance of an EU-level commitment can be inferred as a result.

EU-Level Factors validation

When presenting this factor during the validation stage, no major comments emerged. The idea of having an established commitment at the European Commission level was commonly shared by all subjects. People from the city authorities emphasised more the importance of policies for commercial re-usage of public sector information. From the business side more importance was given to potential investments for fostering an open data ecosystem. According to S2, additional funding opportunities should be in place and driven by the commission.

CCA9 represents the last external factor established from the cross-case analysis. As a result, nine CCAs were formulated in relation to those external factors that were found to be relevant across the three cases studied, and validated across all the participants, in influencing the adoption of open data for commercial service innovation in cities. A total of thirty-seven variables emerged within these factors. Ultimately, nineteen variables were excluded and their role was agreed to be relevant for those specific contexts in which these emerged.

7.3 Perceptions Cross-Case Assertions

In this section all the seven perceptions factors established across cases as influencing open data adoption for commercial service innovation are presented. Similarly to what done for reporting external factors, for each CCA defined the thick descriptions merged from the single cases are provided. Also, additional data collected during the validation interactions is included as well as

reflections on those variables that were excluded from the cross-case findings. Finally, information about links between external factors, perceptions, and adoption is presented for each CCA.

- ***CCA10: Perceived Usefulness of Open Data influences the adoption of open data for commercial service innovation***

This CCA emerged across the three cases in relation to the role played by perceptions about the usefulness of open data in influencing its adoption. A total of fifteen variables arose from the cases. These were grouped together based on similar underlying topics. As a result, nine different variables were listed. The next stage of the analysis led to the formulation of five CCTAs in relation to variables influencing this cluster across cases; four additional concepts were retained as no sufficient evidence was found across the cases studied. When discussing these concepts at the final validation stage, only one (10.6) was ultimately included in the cross-case findings. Details about these processes are provided in Appendix 6.8.

CCA10.1 Provision of Content: open data is perceived to be useful for service innovation mainly because it provides the content for designing the service itself. This perceived characteristic of open data emerged across the three cases. This content is understood as "knowledge about the context".

CCA10.2 Time Saving / Short Time Development: this variable emerged in relation to the usefulness of open data in terms of enabling significant time saving (as opposed to alternative sources of information). Open data is understood across the cases as giving the opportunity of avoiding the manual collection and structuring of the data used for designing and delivering the services. In addition, the time that is actually employed in re-using the dataset is found to influence this factor. This relates to the process undertaken to understand the dataset, its standard, content, and format, clean the data and integrate it into the companies' systems. As this time decreases, perceptions about the usefulness of open data increase.

CCA10.3 Resource Saving: similar to the previous variable, here the focus is placed on the resources that would be needed to actually develop the datasets through other means (manual data collection and storage) if open data wasn't available.

***CCA10.4 Enabling Dynamic Reliable Information*:** if certain characteristics of open data are in place (CCA1), all the companies perceived it as a dynamic (i.e. frequently updated, "real", and "current" data) and reliable (i.e. "accurate" data complying with legislations) information. It was this actual perception that led all the cases to implement further efforts to design and deliver the services.

***CCA10.5 Opens the Councils' Doors / Badge of Honour*:** this variable emerged from cases A and C. It highlights the ability of leveraging the fact of re-using open data to approach and establish relationships with the related city council (see CCA7). All the cases agreed that councils "are

desperately looking for examples" (Case A) of open data re-usage. This is perceived to be relevant because it enables the authorities to claim legitimacy of their open data initiatives and to claim, at a wider scale, concepts around open government and transparency (see CCA12). In relation to Case B, similar considerations to those made for CCA7 can be done (i.e. in relation to the currently not critical relationships with the local councils, but the expected need of establishing collaborations in the future). In summary, the findings suggest that, because of the desire of the councils to showcase open data re-usage, the use of open data is perceived as improving access to these entities. This, in turn, augments the perceived usefulness of open data for commercial service innovation.

CCA10.6 Discoverability and Understandability: this variable emerged from Case C only. It relates to an increased perceived usefulness of open data based on the ability of discovering and understanding the various datasets. This variable is connected to the *Open Data Characteristics* and specifically to the availability of metadata. In other words, open data is perceived to be useful if it doesn't require efforts to "navigate" the datasets and to interpret the data. However, no confirmation was found across the cases. Thus, this variable was initially retained for further analysis. As a result of the validation discussions, and as a consequence of including the variable *Structure and Metadata* within CCA1, a final decision of including this variable was made.

Perceived Usefulness of Open Data links

With respect to the analysis of the links connected to this factor, no major issues emerged during this process. Three inbound and one outbound links were re-connected to all the three cases and were subsequently validated. These are proposed in Table 7.17.

CCA10 Links	Case A	Case B	Case C	Val.
Organizational Skills → Perceived Usefulness of Open Data	x	x	x	Yes
Local Ecosystem → Perceived Usefulness of Open Data	x	x	x	Yes
Open Data Characteristics → Perceived Usefulness of Open Data	x	x	x	Yes
Perceived Usefulness of Open Data → Adoption	x	x	x	Yes

Table 7.17: Perceived usefulness of open data links

All the cases agreed that *Perceived Usefulness of Open Data* is directly connected with *Adoption*. Moreover, in order to actually access, understand, navigate, extract, and re-use open data certain technical skills must be in place. This aspect was emphasised also by S8 during the validation stage. In this way, the subject underlined several times how the scope of open data today is limited to the "IT skilled community". Hence, without IT skills open data can't be understood and therefore can't be perceived as useful. Furthermore, based on the experiences of all the cases, further perceptions around the usefulness of open data were developed as a result of participating in *Hack Competitions, Community-Oriented Initiatives*, and as a consequence of interactions happened with local champions.

Finally, for all the cases studied these perceptions were enabled by the actual open data and dataset's characteristics. These links are summarised in Figure 7.1.

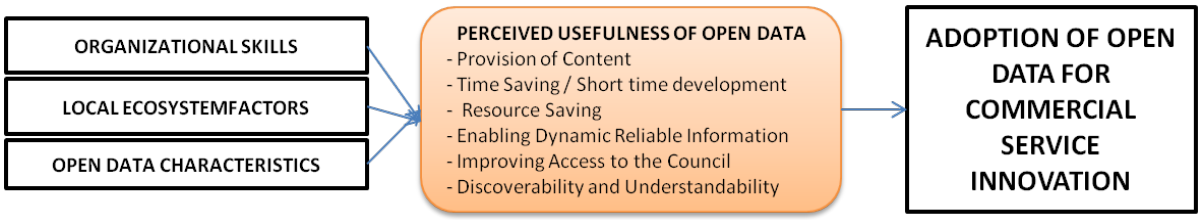

Figure 7.1: Cross-case assertion 10 links

Perceived Usefulness of Open Data validation

One of the tactics (besides the presentation of all the evidences from the cases studied) used for achieving confirmation about this cluster was to pose questions to both S1 and S2 (i.e. actual re-users of open data) and S7 (i.e. a person considering to re-use open data) on the reasons why they decided to explore open data as the source for new service development. According to these subjects, the primary role of open data is to provide information about a domain. All subjects acknowledged that open data enables significant time and resource saving as opposed to alternative ways of developing this knowledge. S1 and S2 indicated that stagnant dataset were perceived as the major cause for non-re-usage of open data. This was explained by saying that "open data is useful if it is an actual dynamic source of content" (S2); and this "has to be provided on a near-real-time basis to be of some value" (S1). Interestingly, when approaching S4, the discussion led to one of the retained variables (i.e. *Discoverability and Understandability*) before this was introduced. Indeed, when discussing about the usage of APIs, S4 distinguished between usability and discoverability of open data. In this way, S4 highlighted that "the introduction of metadata to APIs [CCA1,9] is also believed to be a potential source for increasing interpretability of the machine readable datasets". Aligned with this thought, S2 stated: "datasets need to be understandable to be re-usable".

Retained variables

Free of Charge: the fact that open data is free of charge emerged as a variable from Case B only. However, as also noted by some of the subjects approached, this condition is ingrained in the definition of open data assumed for this research. Thus, this variable was excluded from the cross-case analysis process.

Large Datasets Analysable and Shareable*:* Case B perceives open data as a way that enables the possibility of analysing and sharing large datasets. This variable is enabled by open data being provided with a common standard and in machine-readable format (CCA1). During the validation stage, no significant feedback was received for motivating its inclusion. The majority of the participants found

this variable unclear and did not see any additional aspect than those previously introduced within this factor. The only observation emerged when presenting this factor was in relation to the possibility of developing solutions in a "relatively short time" (S2). However, this aspect was already integrated in the *Time Saving / Short Time Development* variable.

***Easier Process**:* similar to what stated in terms of the previous variable, this concept emerged from Case B only, and it was already ingrained within the *Time Saving* and *Resource Saving* aspects.

- ***CCA11: Perceived Business Opportunity influences the adoption of open data for commercial service innovation***

The second class of perceptions that emerged across all the cases studied is about the role that perceived business opportunities played for pushing the cases to adopt open data. A range of variables emerged as contributors to this factor. These were grouped based on their similarities and a total of nine variables where imported for this stage of analysis. Of these, five were tentatively established and four initially retained. At the validation stage, confirmation has been achieved for six concepts (the five tentatively defined and one initially retained). For a summary of this process see Appendix 6.9.

***CCA11.1 Benefits for End-Users*:** this variable emerged only for Case A. However, having an open data service that entails a range of benefits for the final end-users was a fundamental aspect for developing a value proposition for all the three cases studied. All the open data services studied are based on a significant change of how operations in their contexts are undertaken. This change, in all the cases, is meant to improve the experiences of those people involved in these contexts (i.e. end-users), thus providing benefits to them. Therefore, providing open data services that entail benefits for the end-users contributes to the creation of *Perceived Business Opportunities*.

CCA11.2 Scalability across Cities: all the cases shared the idea that open data services businesses' growth is mainly related to the scaling of the service across different cities. This ability considerably contributes to the creation of perceptions about business opportunities. The main enabler for this variable is seen across cases to be the design of a *Unique Content Management System* (CCA8,1).

CCA11.3 Demonstrability: during the growing process, the strategy for having the city councils to "come on board" (Case A) (i.e. to release their data in an appropriate way and to establish a long-term commitment in doing so) is undertaken by approaching the councils and "ask for the data". In this process, Case A and Case B emphasised the importance of demonstrating the value of the open data service proposed. By doing so, the organizations try to convince the authorities to release their data and to establish relationships. Thus, the assumption made here is that the more the value of an open data service can be demonstrated, the greater is the chance that new city councils will deliver the data

needed, i.e. will enable the possibility for the company to deliver its service in an additional city. This variable did not emerge in this cluster for Case C. However, similar considerations can be inferred from the *Growth* sub-factor in its adoption stages. "Approaching the councils is always the first activity" undertaken as part of this process. In this process, Case C presents its open data service to convince the council to embrace the idea, and to release its data.

CCA11.4 Funding: emphasis emerged from Case B about the importance of funding for developing perceptions about business opportunities. These encompass funding at both the local and the European levels. Regarding the former, the CEO of Case B underlined the case's experience at the hack competition. All the cases went through this experience and received some funding as a result. Nonetheless, in cases A and C this initial financial support was not highlighted as influencing the *Perceived Business Opportunities*. Regarding the European level, instead, this aspect emerged for Case B only. Thus, no sufficient evidence was initially found for postulating a related CCTA. Hence, this variable was re-discussed during the final validation stage. Here, S1, S2, and S3 emphasised how as funding (at the local and European levels as well as from other forms of private investments) grow, so does the perceived business opportunity. It is also noted that the role of private investors was highlighted several times within the data collected from Case C.

CCA11.5 User Base: this variable emerged from Case B. However, differently from the previous, the achievement of a considerable user base is a major requirement for the cases to extract *Secondary Data* (CCA8.2) from the usage of the services. For instance, no significant trend analysis (for Case A), or favourite journeys for tourists (for Case B), or patterns of drivers in the city (for Case C) can be derived if the user base is not "big enough". Given that these processes are fundamental part of the cases' revenue models, it can be inferred that as the service's user base increases, *Perceived Business Opportunities* increase subsequently.

CCA11.6 Knowledge Potentially Derived: according to Case C's interviewees: "the main business opportunity is from the data you can extract; a lot of information can be captured and sold to retailers to enable more informed decisions on their side". This information (see CCA8,2) is seen as new knowledge that can be potentially structured about the city. This concept is at the base of the revenue models of Case B and Case C. Also regarding Case A, the production of additional knowledge (e.g. trends in the urban planning domain) is part of its perceived business opportunity. Indeed, this knowledge is leveraged to convince the councils (in US) to establish commercial contracts.

Perceived Business Opportunity links

Similar to the previous factor, no particular issues emerged when merging links for this set of perceptions. Three inbound and one outbound links were defined. All these links were observed and established for all the three cases studied. An overview of this process is provided in Appendix 6.10.

The first link that emerged highlights the contribution of *Service Characteristics* to perceiving business opportunities. Both of the variables defined within CCA8 are found to be related with two of the variables within this CCA. Firstly, designing an open data service with a *Unique Content Management System* (CCA8,1) is found to enable *Scalability* of the service across cities (CCA11,2). Secondly, the ability of the service to generate additional knowledge about the context is linked with these perceptions. Indeed, in all the cases this additional data generated is part (to some extent) of their revenue models. The second link shows that characteristics of the individuals influence the creation of perceived business opportunities. The focus here is on the variables *Experience in the Context* (CCA2,1) and *Entrepreneurial Spirit* (CCA2,3). All the cases agreed that these were significant drivers for enabling these perceptions. In addition, both cases A and B leveraged the individuals' previous experience (in urban planning and tourism respectively) to generate the idea of a service that entails a range of *Benefits for End-Users* (CCA11,1). As the third link, the characteristics of the city in which the service is aimed to be delivered were found to influence the creation of these perceptions across the cases. Cities to be targeted are selected in relation to the potential business opportunities. The evidence collected showed that bigger cities (CCA3,1), where a certain infrastructure is in place (CCA3,2), and where a market opportunity exists (CCA3,3) are places where open data services' business opportunities are expected to be generated. Finally, according with the findings of all the cases, these perceptions contribute to the *Perceived Usefulness of Open Data*.

Figure 7.2: Cross-case assertion 11 links

Perceived Business Opportunity validation

Regarding this CCA, only business representatives (S1 and S2) provided additional reflections on the variables presented. Both S1 and S2 agreed that perceiving business opportunities is a fundamental step in the decision of adopting open data for commercial purposes. Considering the variables arose within this factor, significant confirmation was found in relation to the importance of providing a service that entails a range of benefits for the end-users. In addition, as the knowledge potentially derived increases, so do perceived business opportunities. No major observations emerged also when presenting the statement around the positive contribution of user base development to this set of perceptions. S1 and S2 agreed that once the system is able to ingrain open data from one city, then *Scalability* of that system becomes critical to perceive growth potential. Connected to this,

Demonstrability has been acknowledged as something beneficial in order to convince the councils to release the data on a sustainable basis. To this variable, additional aspects were also added by the subjects from the city councils. In particular, S5 stated: "case studies that demonstrate successful implementation of open data services are crucial"; S6 reported: "we face a situation in which we proposed to the city CFO to undertake an open data program. His answer was 'what is the return on this?' We tried to explain him that there is a chance that new economic activities can be enabled and the CFO said: 'do you want me to invest on a chance?' He prefers investments with zero risk but that at least can guarantee a return". Interpreting this statement, demonstrations to support these arguments are critical for engaging with new cities, i.e. for enabling growth.

Retained variables

Partnership Strength: this variable emerged in relation to Case A, in which the importance of having a relationship with the council was emphasised. According to the feedback received at the validation stage, this concept was interpreted as redundant as a related variable was already established (see CCA16,1).

Level of Engagement: similar considerations to the ones provided for the previous variable were made with respect to the role that the *Level of Engagement* with the council played in the creation of perceptions about business opportunities for Case A. As noted by S4 and S8, this concept is more related to the *Perceived Credibility and Sustainability* factor (see CCA16) and its contribution to the *Perceived Business Opportunity* is already ingrained in the link between CCA16 and this factor.

Revenue Model Structure: the ability to have an open data service that allows establishing flexible revenue models is perceived by Case C to increase the related business opportunity. This aspect is believed to accommodate the needs of a wide range of customers based on their characteristics. However, no insights in this way emerged from the remaining cases. When re-discussing this variable with the subjects selected for the final validation, no insights were found to suggest its importance across domains for the creation of perceived business opportunities.

- ***CCA12: Perceived City Opportunity influences the adoption of open data for commercial service innovation***

Another set of perceptions that emerged across cases as influencing the adoption of open data for service innovation is related to perceived opportunities by the city. From the cases studied it arose that these perceived opportunities result in actual commitment and efforts in fostering the open data service ecosystem. The general understanding behind this factor relies in the fact that cities benefit from the creation of the cases' open data services in two main directions. Firstly, the actual services provide the council with better ways of managing the contexts. Secondly, the perceived opportunities

by the councils include the ability in terms of allowing to improve their images (i.e. transparency claims are legitimated) and to attract new businesses and investments. These benefits are perceived as generating "motivations" (Case C), which in turn are perceived as adding credibility to the overall ecosystem. Four different variables emerged across the cases (see Appendix 6.11). Three were established and validated within this CCA. The fourth (i.e. *Participation and Collaboration*, Case B) was taken into account as additional evidence from the "missing case" when defining the importance of City-Business Relationships (CCA7). Therefore, also according with the feedback received during the final validation phase, this variable was excluded from this cluster.

***CCA12.1 Improve City Image / Transparency*:** the underlying assumption related to this variable is that cities want to claim to be at the "forefront of the open data movement and pioneers of transparency" (Case B), and open government more generally. Thus, the fact that there is a business that is successfully re-using the city's open data for delivering a commercial service is seen by the city council as the opportunity to state these claims with perceived legitimacy. More observable claims can be stated around the improved outcomes of the open data program. Indeed, from both Case B and Case C new open data is delivered back to the council.

CCA12.2 City Competitiveness: the concept of urban competitiveness emerged across the three cases as an aspect related to open data in terms of attracting new businesses by giving them a source for service creation. This concept is also indirectly strengthened by the approach taken by the cases when targeting new cities. In doing so, the cases aim at delivering their open data services in those cities where a mature open data program is in place. Finally, as part of Case B, this concept emerged also concerning the competition between cities to be positioned as leaders in the worldwide open data movement.

CCA12.3 Internal Efficiency and Efficacy Gain: this variable is related to the fact that all the open data services studied provide the cities with an overall better management of the specific contexts. This was observed as an improved efficiency and efficacy of the operations in domains in which the services are delivered. This was found to constitute a perceived opportunity for the city that contributed to the councils' decision to undertake efforts towards sustainably embracing open data. This aspect was already partially covered in CCA7,1, where the importance of providing benefits to the council was underlined for establishing *Win-Win Relationships* with the authority.

Perceived City Opportunity links

In total, three inbound and one outbound links were defined for this CCA. Given that the outbound link is related to CCA16, for clarity, its definition is provided within the paragraph dedicated to CCA16.

The remaining three links established and validated are shown here (details about the process are provided in Appendix 6.12).

Firstly, a link is established between *Local Ecosystem Factors* and this set of perceptions. This link was observed in Case A originating from *Community Factors*, in Case B from *Local Community and Ecosystem Factors*, and for Case C from *Events and Community Factors* (it is noted that the original link from the single case model was established with the factor *Perceived Ecosystem Awareness*, and the specific variable *City Council Awareness*; this variable was integrated as further evidence for this factor). The main focus of this link is placed in the variables *Hack Competition* (CCA5,1) and *Community-Oriented Initiatives* (CCA5,3). These initiatives are in place to "educate the councils" (Case A) about the data needed and to make them "embracing open data" (Case B). From Case C, this specific connection emerged in relation to the ability of participants and organizers of initiatives (outside the city council) to be more vocal and educate the authorities on how to move forward. Secondly, it emerged across the three cases that these perceptions are generated if certain characteristics of the city council are in place (CCA4). For example, if an open data strategy (CCA4,1) is not undertaken and awareness (CCA4,2) is not spread across the council's departments and operators, *Perceived City Opportunities* are precluded to be generated. Also, businesses aim at scaling their services in those cities where certain characteristics of the councils are in place. Therefore, this argument further strengthens the relationship between *City Council-Related Factors* and *City Competitiveness*. Thirdly, *Internal Efficiency and Efficacy Gains* are enabled by *Service Characteristics* for all the three cases studied. A representation of these links is provided in Figure 7.3.

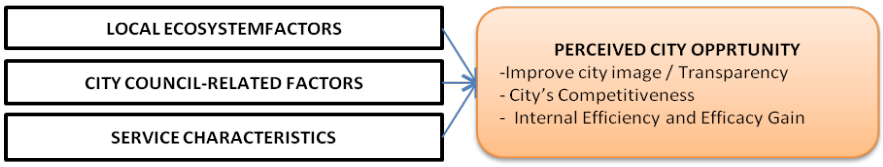

Figure 7.3: Cross-case assertion 12 links

Perceived City Opportunity validation

The interpretations about *Perceived City Opportunities* as a factor that influence the adoption of open data were agreed by all the subjects when presenting evidences from the cases. Confirmation about the fact that cities need to perceive a range of opportunities in order to effectively implement an open data strategy was achieved. Those external factors that (if correctly in place) contribute to the creation of this set of perceptions are generally seen as *Local Ecosystem Factors, City-Council Related Factors,* and the *Secondary Data Generation*, thus confirming the links to this factor previously presented.

Regarding the variables, significant confirmation came from S5 and from the subject from the intermediary organization (S3). The latter sees potential benefits in different directions: (1) as an increased "internal efficiency and less phone calls and less inquiries"; (2) "time and cost saving"; and (3) "a significant image improvement for increasing city competitiveness worldwide and also within the city for providing information to the public in a sleek-way with fancy technologies". S5 also highlighted the main goals for the local authority. "We want to be transparent and open government is our first goal. Then pursue economic benefits". Finally, "I encourage the business sector to engage with this initiative to explore how open data can drive new services and products. Public bodies need to treat data as a corporate asset" (S5).

- ***CCA13: Voluntary Use of Open Data influences the adoption of open data for commercial service innovation***

This CCA was defined from Level 2 of this perceptions' cross-case analysis. In Case A, this aspect emerged during the final validation meeting. From the experience that the CEO has in the open data ecosystem outside the case, reflections were undertaken in relation to the possibility of people that voluntarily re-use open data. Furthermore, initial activities for Case B were undertaken by the current CEO as a "volunteering project". Hence, a set of perceptions regarding the voluntary use of open data was established and subsequently validated (see Appendix 6.13). No specific variables, however, emerged from this factor across the two cases. Regarding Case C, no significant evidence arose in this direction. However, following the arguments of Case A's CEO, it can be reasonably accepted that "someone can use open data voluntarily for creating a new commercial service".

Perceived Voluntariness links

This factor arose as non-linked with any of the external factors in both cases in which it originally emerged (i.e. cases A and B). Therefore, its original formulation, i.e. directly linked with the adoption factor was kept in this cross-case analysis.

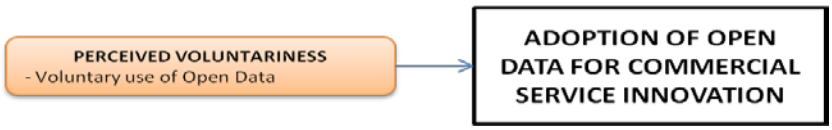

Figure 7.4: Cross-case assertion 13 link

Perceived Voluntariness validation

This factor was presented to all subjects in a very short time by saying: "it emerged that people from the cases believed that someone using open data voluntarily for creating a new service can influence the adoption process of open data for commercial service development". No objections or additional

information were provided by any of the subjects encountered. Those who provided a comment were aligned in saying: “well, someone could start at a volunteer level [as for Case B] and then study business models. Said that, yes I agree” (S3).

- ***CCA14: Perceived User-Friendly of the Open Data Service influences the adoption of open data for commercial service innovation***

Emphasis was placed in both Case A and Case B about the importance of designing an open data service that is user-friendly. The main assumption made within this cluster is that open data is adopted for commercial service innovation if the ability of designing a service that is easy-to-use is perceived. Also, in order for open data services to be sustainably delivered (and so for open data to be sustainably adopted), users need to use these services. As shown in Appendix 6.14, two variables were established for this cluster. One more variable (i.e. *Being Able to get Information,* Case A) was initially retained and subsequently re-allocated as constituting additional evidence for CCA15. Similarly, the variable *Easy to Use* from Case C emerged for the related factor *Perceived User Satisfaction* was moved to this cluster adding evidence from the “missing case”.

CCA14.1 Lean Service / Sleek Design: across Case A and Case B the need of developing a lean open data service was explained as the need of highlighting in the service what adds significant value (to the user) by reducing everything else. Case B added that this feature of the service is enabled by designing a “light solution” achieved through employing a “sleek design” and “fancy” technologies.

Lean Service / Sleek Design - missing case: two sources of evidence allow relating this concept to Case C as well. Firstly, the importance of providing a service where the information sought is “easy to detect” was highlighted. Secondly, further confirmation came from personally experiencing the service. All the relevant information is displayed with no redundancies. Thus, the open data service is considered lean consistently with how this term is understood by cases A and B.

CCA14.2 Easy to Use: all the open data services that were studied were designed in order to enable users to fully enjoy the experience without the need of significant technical skills. In other words, the service needs to be easy to use. This concept ingrains the availability of a service where the information or experience sought by the end-users is easy to obtain.

Easy to Use - missing case**:** regarding Case C, this variable emerged within another cluster (i.e. *Perceived User Satisfaction*). Therefore, its content was integrated into this variable and contributed to its establishment (due to the achievement of confirmation across all the cases studied).

Perceived User-Friendly links

As shown in Appendix 6.15, two links were established for this factor. Across the two cases, these perceptions are seen as the result of particular graphic design and data visualization techniques. Further evidence came from Case C for which a link was established between *Organizational Skills* and *Perceived User Satisfaction* (and specifically underlying the contribution of these skills to the development of an *Easy to Use* open data service). Therefore, a significant contribution to the achievement of these perceptions is acknowledged to be from *Organizational Skills Factors*. On the other hand, a conflict regarding these links was in place between cases A and C, and Case B. All the cases agreed on the fact that the design and delivery of an open data service that is user-friendly contributes to the achievement of sustained adoption of open data. Nonetheless, in cases A and C this connection was found to be mediated by *Perceived Usefulness of Open Data*. When interpreting this "inconsistence", the reasoning provided by these two cases was validated. The fact that for Case B a direct link with the adoption factor was established was interpreted to be in relation to the initial non-commercial purpose of the service. Therefore, because the project was initially at the "volunteering level only", this link was found to be relevant when Case B faced the actual adoption of open data. A summary of the two links established is proposed in Figure 7.5.

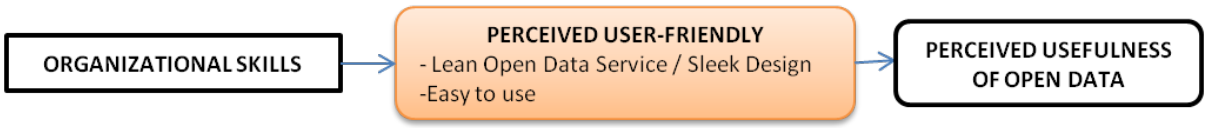

Figure 7.5: Cross-case assertion 14 links

Perceived User-Friendly validation

As expected, the interpretations about this factor were confirmed by all the participants. No objections arose from any of the interactions. Also, no additional comments were added.

Retained variable

Interoperability across browsers: significant importance was given to this aspect of the service by the CEO of Case B. Interoperability of the service is enabled by the usage of HTML5 as the programming language used for its design. Its importance was related to overcoming a potential barrier for users, i.e. the need of having specific programs or devices to be able to use the service. However, no further confirmation about this feature was initially found within the two remaining cases. When discussing this variable during the validation stage, no major feedback was collected to motivate its inclusion. Indeed, this was interpreted as a non-necessary and "minor" aspect of a service (S1).

- ***CCA15: Perceived User-Satisfaction of the Open Data Service influences the adoption of open data for commercial service innovation***

CCA15 originally emerged for cases B and C. This cluster encompassed concepts around the importance of satisfying users in order to sustainably deliver the service (i.e. to adopt open data on a sustainable basis). Five variables were defined within this cluster across the two cases. However, one of these (i.e. *Easy to Use*) was already ingrained in the previous cluster (see CCA14.2, "missing case"). On the other hand, as already highlighted above, a variable emerged from Case A was added to this cluster and formed the basis for the evidence coming from the "missing case". Regarding the remaining four variables, two found considerable reflections in Case A as well, and one more (i.e. *Personalised Service*, initially established for Case C only) was included after the validation efforts. The remaining variable was finally excluded (details are provided in Appendix 6.16).

***CCA15.1 Gain Experience Efficiency / Time Saving*:** both cases B and C's services aim at significantly reducing the time spent by users in their contexts. Case B's open data service gives the opportunity for tourists to save significant time by providing information-rich itineraries that are pre-studied consistently with different users' characteristics. Case C provides a service that increases the overall efficiency of the process of searching for a parking space in the city.

Gain Experience Efficiency / Time Saving - Missing Case: although the overall cluster did not emerge in relation to Case A, it is clear the fact that the service provided enables users to have a significantly improved experience when looking for urban planning information in terms of time saved. This variable has been defined as an additional component of the *Perceived Business Opportunity* for Case A. However, in this case the focus of these perceptions is placed on users. Hence, the variable defined in relation to addressing a market opportunity (i.e. inefficiency in the usual process of managing and delivering urban planning information) can be also considered in terms of providing an increased efficiency from the point of view of final end-users of the open data service.

CCA15.2 Gain Experience Efficacy / Resource Saving: this concept emerged in relation to the importance (for Case B and Case C) of providing a service that enables users to save resources if compared to the previous ways of conducting the related activities in the context. This emerged in Case B as providing tourists with more informed decisions (and so in turn increase the efficacy of their journeys). In Case C, this aspect was related to both enabling users with the ability to find the best parking rates, and to save consumptions when driving in the city.

Gain Experience Efficacy / Resource Saving - Missing Case: regarding Case A, similar considerations to the ones made for the previous variable, can be applied here. In particular, the importance of providing a service that entails benefits for end-users is supported by significant evidence (e.g.

CCA11,1). These benefits were outlined for Case A and included: (1) enabling users for an improved efficacy in the process (i.e. enable more informed decisions in the urban planning domain); and (2) decrease resources that were previously employed to get the required information.

***CCA15.3 Personalised Service*:** the importance of having a personalised service was interpreted as influencing the final users' satisfaction from Case C only. This concept was connected to the service's characteristics in which users get personalised information (e.g. parking payment history) and, if requested, personalised "push notifications" (e.g. reminders and alarms). Regarding the other two cases, this variable did not clearly emerge within the related findings. However, as also discussed during the validation interactions, both Case A and Case B provide an open data service where personalised features are in place. Considering the former, the "Get Notified" feature of the service (see section 4.1) is in place. In Case B, the journeys proposed to the different users are generated based on personal characteristics of these subjects (this feature was ingrained into the *Interactive Experience* variable).

Perceived User Satisfaction links

Appendix 6.17 shows the links emerged for this factor from the single cases. As a result of the cross-case analysis, two links were established. Firstly, a link was defined between *Service Characteristics* and this factor. For both cases B and C, *Perceived User Satisfaction* was established as a consequence of the different features of the open data services offered. Secondly, regarding its outbound connection, a similar situation to the one encountered for the previous factor (see CCA14) was found here. Whereas for cases A and C the link with *Adoption* was found to be mediated by the *Perceived Usefulness of Open Data*, in Case B a direct link was established. The interpretation undertaken was similar to the one presented above. In brief, Case B's initial "volunteering" characteristic is interpreted as the reason for this inconsistence. Therefore, the assumptions made for cases A and C were kept across cases. A summary of these two links is provided in Figure 7.6.

Figure 7.6: Cross-case assertion 15 links

Perceived User Satisfaction validation

Similar to the previous factor, no major comments emerged when presenting the evidences from the cases supporting this CCTA. The importance of this factor was interpreted by the subjects consistently with my personal interpretation of the data collected. In particular, in order for open data to be

adopted on a sustainable basis as the source of a commercial service, this service needs to be adopted by users. This contributes to the achievement of a considerable user base that triggers *Secondary Data Generation* (which in turn enables perceived business and city opportunities).

Retained variable

Learning Opportunity: this variable was defined for Case B. It highlights the importance of providing a service that entails a learning opportunity. It can be inferred that this concept is aligned with the specific context of Case B, i.e. tourism. All subjects approached during the validation stage agreed on its importance to be connected with the specific tourism context.

- ***CCA16: Perceived Credibility and Sustainability of the Council's Open Data Strategy influences the adoption of open data for commercial service innovation***

This CCA was formulated in relation to the critical importance of perceiving the credibility of the city councils-related open data strategies in order to develop beliefs about open data services business opportunities. This factor is believed by the majority of the subjects encountered as the cornerstone of the overall adoption process. This CCA was generated from Level 3 of the perceptions' cross-case analysis (see section 7.2.2). In this process, five variables were grouped together across the cases studied. These were found to have an underlying common topic in relation to the importance of ensuring a perceived credibility and sustainability of open data. While for Case C two specific factors emerged in this direction, similar concepts were ingrained in other factors for the remaining cases. As a result of the analysis, two variables were finally established to support this cluster (see Appendix 6.18). The first is related to the level of commitment of the city council (topic covered across all the three cases). The second emerged from the analysis of Case C in relation to the importance of having in place challenges to the councils' operations. As a consequence, a CCTA was formulated in relation to the first, while the second was initially retained and re-considered during the validation stage. After these interactions, a final decision to include this variable was made. Arguments supporting its inclusion are provided below. It is noted that the variable *Replicability of the Open Data Service* emerged from Case C has been excluded from the analysis at this stage. Similar reflections to those provided for excluding the importance of having complex open data services are applied here (see CCA8).

CCA16.1 Level of Commitment: in order to perceive business opportunities from open data, re-users (and potential investors) need to perceive the sustainable availability of this source. This means that the councils' open data programs must be perceived as initiatives in which open data will be provided (consistently with CCA1) in the long-term. Regarding Case C, a specific factor named *Perceived*

Commitment of City Councils was defined. In relation to Case A, the importance of ensuring commitment of the council was highlighted as a relevant concept constituting *Perceived Business Opportunities* through two variables, i.e. *Partnership Strength* and *Level of Engagement*. The former stresses the importance of establishing strong partnerships between the council and the business to generate perceptions about the credibility and sustainability of the open data supply. The latter sheds light on the contribution of the councils' engagement with the entire ecosystem (e.g. through hack competitions, workshops etc.), towards perceiving an increased commitment. Concerning Case B, this concept emerged as the need for the council to establish commitments in participating and collaborating with the open data services businesses. This was found to lead to *Perceived City Opportunities.* These concepts were clustered together providing a general understanding of the importance of perceiving the commitment of councils for perceiving open data as the basis for sustaining a service business in the long-term.

CCA16.2 Challenges to the Council: this variable emerged from Case C in relation to the importance of having entities in place that challenge the councils' activities. This is understood as critically assessing its performances (towards open data) and stimulating subsequent improvements. This set of activities is believed to avoid "complacency" situations where the councils rely on "stagnant datasets" to state claims around transparency and open government. According to Case C's CEO, these challenges should be posed primarily by the intermediary organizations (and from the European Commission at the broader level). If compared to other bodies, intermediary organizations are believed to be "more vocal". In other words, entities that operate outside the council can be more objective and effective in assessing current ways of providing open data. However, no further insights arose in this way from cases A and B. Hence, its inclusion was re-discussed as part of the validation stage. During these interactions, S4 emphasised its importance. The role of the Open Data Institute in UK was taken again (it emerged also from Case C) as an example for the effectiveness of posing challenges to the council. This is believed develop perceptions on "the goodness and sustainability of open data". One of the city council's representatives (S6) highlighted how today the Open Knowledge Foundation acts "as a sort of stimulator" for these initiatives but internationally and across councils. This participant sees this set of activities as "pushing the city councils to do better" (S6). Overall, interpreting the feedback received, posing challenges to the council is acknowledged as a way to increase perceived credibility of the council's open data strategy.

Perceived Credibility and Sustainability links

With respect to the links, the fact that this factor originally emerged from Case C only motivated the need of further interpretations. After carefully reviewing all links of all single cases related to the variables integrated in this cluster, three inbound and one outbound links were initially defined.

Appendix 6.19 provides a summary of this analysis. The first link that emerged acknowledges the connection between *City Council-Related Factors* and these perceptions. Regarding Case C, *Level of Commitment of the City Council* (which is integrated in this CCA) was found to be partially enabled by the characteristics of the council itself. The evidence for cases A and B came from a reflection on the actual role of the *Perceived City Opportunity* in the related models. In both cases, the importance for cities to perceive opportunities from releasing open data to the cases was established because of its ability to generate perceptions about future availability of this source. From another angle, the role of this factor is reflected in the links emerged for both cases A and B in the form: *City Council-Related Factor*s → *Perceived City Opportunity* → *Perceived Usefulness of Open Data*. In this way, the underlying argument provided was about the importance for cities to have certain characteristics in place (CCA4) in order to perceive open data services-related opportunities. According to these reflections, a link was established between *Perceived City Opportunity* and *Perceived Credibility and Sustainability*. The rationale behind this connection can be summarized as follows: if cities perceive an opportunity from releasing open data, this contributes towards perceiving open data as a credible and sustainable source for commercial service innovation.

Another link was established from the factor *Cities-Business Collaboration*. The focus of this link is on the variable *Level of Commitment*. For Case C this link was originally established. In Case A, a link was defined between *Cities – Business Collaboration* and *Perceived Business Opportunity*. This was describing the importance of having relationships in place for perceiving sustainability of the open data availability in order to perceive actual business opportunities. In this way, the establishment of this CCA is meant to add importance to the link originally emerged from Case A through the definition of an actual factor that mediates this connection. Hence, following this reasoning, the fact that the open data strategy from the council is perceived as credible and sustainable was found to enable *Perceived Business Opportunities* (this link explicitly emerged from Case C).

Furthermore, one variable (i.e. *Challenges to the Council*) was integrated only after the validation efforts. Therefore, giving the primary role of intermediary organizations and the European Commission in posing these challenges, two further links were consistently established: (1) from *Local Ecosystem Factors;* and (2) from *EU-Level Factors*. A summary of all links established to and from this factor is provided in Figure 7.7.

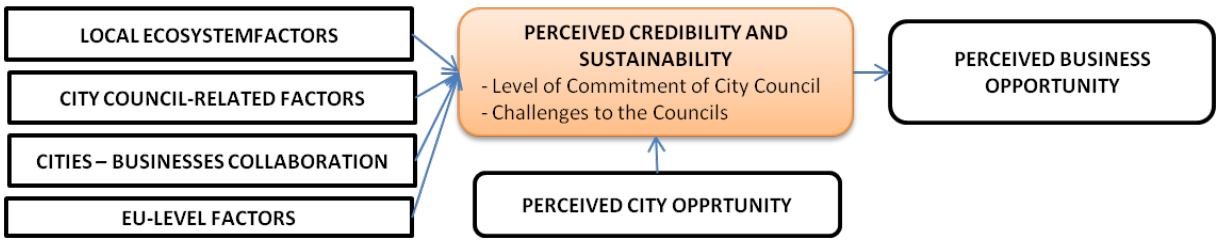

Figure 7.7: Cross-case assertion 16 links

Perceived Credibility and Sustainability validation

Perceiving commitment from the city council and subsequently credibility and sustainability of the local open data ecosystem was interpreted by the majority of the participants as the cornerstone of the overall process. According to S1 and S2 this class of perceptions is "the most important one". A strong focus has been placed by both of these participants upon the importance of perceiving credibility of the council-related open data program. "We need to be sure that the data will be available in two years from now" (S1). These perceptions need to be in place in order to allocate time and resources and towards open data (S2). A different and interesting point of view was provided from those participants from the city councils (especially S5), when asking the following question: "what does the city council do for helping SMEs and start-ups to re-use open data?" The answer provided included: "trying to release as much data as we can"; organize events where re-usage is stimulated; and having in place "local enterprise officer" (S5). As a response to this comment, I described the desire emerged from businesses of having "open data published by default" (Case A). However, mechanisms are not in place to achieve this status yet – S5 argued. This subject also added that the city council doesn't "know who uses it [i.e. open data] for commercial purposes and who does not" and "one of the main goals is enabling perceptions of no-commercial restrictions for users". The city council's commitment was explained as follows: "with initiatives like public meetings where groups are formed and discuss how to move forward with the program, we want to show that we are responsibly taking this strategy forward" (S5). The concept of service level agreement between the open data service business and the council was re-taken into account (it was mentioned as part of the validation process for CCA7). No positive feedback was collected in this way in terms of the feasibility of this option. The solution of enabling *Win-Win Relationships* is found as the most suitable at the moment, because the "idea of providing open data as a service to the businesses is not part of the thinking within public authorities at the moment" (S5). In particular, as noted by S6, "SLAs are a big challenge, because there is no clarity on the returns of open data". Thus, this is not currently perceived (by the councils' representatives) as a feasible solution for ensuring sustainability of open data provision.

At this stage, the perceptions' cross-case analysis was concluded. As a result of combining and merging similar topics across the three cases, seven perceptions were established and validated as influencing the adoption of open data for commercial service innovation. These are described by a total of twenty-three variables and connected through twenty-two links. Ultimately, for a total of seventeen variables no significant overlap was found across cases during the analysis and the validation phase. Thus, these variables were excluded from the cross-case findings.

7.4 Adoption Stages across Cases

The third and last phase of the cross-case analysis focused on the adoption stages. The main goal at this stage was to analyse the content of each sub-factor and compare it with the remaining cases. Hence, the first activity conducted was about listing all the adoption sub-factors across cases and merge those that are similar in their contents. As shown in Appendix 6.20, a total of nine different stages emerged across the three cases. Interestingly, six of these sub-factors were identified for all the cases studied. Three more emerged in relation to one case only. However, for one of these (i.e. in relation to the process of acquiring funding), confirmation was achieved from re-visiting the single case reports and during the final validation stage. The remaining two sub-factors were finally excluded from the cross-case findings. Thus, seven adoption stages emerged to be relevant in describing the process from the initial idea generation to the sustained use of open data.

Stage 1, Idea Generation: as a first stage for all the cases, the initial idea about the service was generated. However, this process was undertaken in two different ways across the three cases. On one side, Case A and Case B's current CEOs leveraged their previous experiences in the context. These experiences enabled them to identify a relevant problem that needed to be solved, and subsequent opportunities to be targeted. On the other hand, in Case C the subjects who generated the idea did not have any experience in the parking industry (apart from being citizens that actually drive in the city). Their entrepreneurial spirit and their attitude towards developing an IT-related business, led these people to design a technical solution for the taxi's industry. What was common across the three cases is that the idea preceded the actual awareness of open data and so, its adoption. Nonetheless, all these ideas faced a challenge in terms of lack of data for their actual implementations.

Stage 2, Open Data Adoption: after the initial idea for the service was developed, all the cases faced a stage in which open data was actually adopted for enabling the development of the service itself. For Case A and Case B this process was different than the one followed by Case C. In particular, the former cases started to look for the data by engaging with the councils. Both the cases had in mind what data was needed for the development of the previously generated idea. Then, it was by engaging with the councils that awareness of open data was gained. This initial relationship was also leveraged

by both these cases to get to know the different local hack competitions that were planned by the city councils and other organizations in the area. Regarding Case C, this business did not approach the councils before participating in the hack competition. Thus, the three cases adopted open data at a hack competition. At this stage, while people within cases A and B had clearly in mind what data was needed for the development of their ideas, this situation was different for Case C. The current CEO of Case C modified and extended the previously designed solution to the parking industry. This choice was made based on the data available ("it seemed to be all about parking"). All the cases won their competitions and, besides the prize won, a strong push to develop the solution further was perceived. As a result of this step all the three cases had the first basic version of the service designed.

Stage 3, Pilot Implementation: after designing a first basic solution, all the cases entered a stage in which the first pilot was implemented and delivered in the city where the hack competitions had been undertaken. This pilot was described as a "minimum viable product" across cases. Common activities conducted at this stage included: testing the datasets and their suitability for the delivery of an effective and valuable service; improving the service, its design and features; and engaging with the councils aiming at shaping the idea of a *Win-Win Relationship*. Already at this stage, initial perceptions arose within the cases in terms of the need of ensuring the availability of this data in the long-term. This concept was considered to be a critical requirement for the cases in order to develop the solution further. However, some peculiarities emerged in this stage across the different companies. In particular, Case A placed major attention on the data itself. The main focus at this stage was about working on the standard, the cleaning, and the geo-localization of the datasets. Case B, developed the original service further in terms of improving its usability and design. The main goal of designing a pilot at this stage was to have access to funding needed to further develop the solution. For Case C two different stages emerged as part of its pilot implementation. These were labelled as *Private Beta* and *Public Beta*. As part of the first, the initial pilot developed was tested among a limited number of users. Results were shown to the council to leverage the opportunity of establishing a partnership. During these processes, all the cases started to get closer to the overall ecosystem (e.g. by attending public meetings and workshops on open data and other community-oriented initiatives). As a conclusion of this stage, the three cases officially launched the companies.

Stage 4, Acquire Funding: this stage emerged from one case only. A major importance of the availability of funding emerged from Case B. This phase was perceived to be fundamental by the CEO for moving from a volunteering project to a commercial one. To do this, the main goal was seen as expanding the service with open data across Europe. Thus, EU-based funding schemes were targeted. The reason why this sub-factor was not excluded although it emerged for one case only, relies in the fact that at this stage of development also Case C sought funding for supporting future activities.

However, as opposed to Case B, Case C was oriented towards private investors. The absence of a contribution to this concept from Case A can be generally understood from the fact that this business leveraged the establishment of commercial contracts with the city councils. Hence, neither specific funding schemes nor investors were needed to be targeted. Confirmation about this interpretation was also achieved as part of the final validation stage.

Stage 5, Commercialization: after implementing the first pilot and (for Case B and C) after receiving external funding, the open data services – related revenue models were conceptualized and established. For Case A and Case C perceptions about the business opportunity related to their open data services were previously generated (in Case A from the initial *Idea Generation* and in Case C from its *Market Analysis* stage). Case B developed this perception while funding was sought. All the three cases established revenue models on additional characteristics added to the original pilot. For example, Case A added to the pilot an offer for purchasing push notifications and specific personalised features; similarly, Case B offered augmented reality of cultural attractions as an additional premium feature of the application; Case C integrated ePayment processes as a further possibility for drivers involved in parking. Ultimately, both cases B and C established revenue models based on the analytics of the "data collated through the app" (i.e. from the *Secondary Data Generation* processes, see CCA8,2).

Stage 6, Growth: at this point, the three cases had developed their open data services for one initial city. These services and the related open datasets were tested and improved accordingly. The next stage was about seeking growth opportunities. Interestingly, growth was meant across the cases as the ability of delivering the service in other cities (Case A also emphasised the goal of covering the entire national territory of the initial city). The key for this process was to have in place a scalable solution. All the cases shared the idea that for this ecosystem the "big work" has to be done for the first city only. Once the technical solution had been developed and delivered in the first city, other cities were (and currently are) targeted based on some relevant variables (see CCA3). In addition, the cities that are found to be attractive by the cases are those in which: an open data program is in place at the related authority (CCA4,1); awareness of open data is widespread (CCA4,2); an IT department is responsible for open data with an allocated budget (CCA4,3, CCA4,4); and where room for partnership with the council and possibility of engaging with the overall ecosystem are in place. All the cases agreed that when targeting new cities, "approaching the council is always the first activity". This happens because of the need to establish long-term *Win-Win Relationships* with these entities to perceive a sustainable locally-related business opportunity. When approaching the councils, it is critical to effectively demonstrate the value that can be potentially achieved through the proposed

open data service. A critical step in this set of activities emerged to be about "educating the councils". The CEO of Case A named this overall process as "making cities to come on board".

Stage 7, Sustained Use: as the last stage, all the companies achieved sustained use of open data for the continuous delivery of their commercial services. These services are also continuously improved and enriched based on emerging opportunities in the various contexts. At this stage, data is provided with the required characteristics (see CCA1) and is being perceived useful as the source for the various commercial services offered (see CCA10). Perceived sustainability of the overall ecosystem is in place (see CCA16) and it is mainly enabled by the establishment of partnerships and relationships at various levels with the councils (see CCA7). Thus, perceived business (see CCA11) and city (see CCA12) opportunities continue to be in place and foster the effective development and maintenance of these relationships. Concluding, the three cases are seeking growth by continuously monitoring the open data landscape across cities in terms of who is implementing an effective open data program (see CCA4), i.e. which city can be potentially targeted.

Adoption Sub-Factors Validation

When presenting these adoption stages to the participants of the validation efforts, no significant comments were provided to the single stages other than mere approval. The only feedback emerged from two of the subjects approached, i.e. S4 and S8. This was related to the general formulation of these adoption stages. These were presented as sequential process made of separated stages. This "linear" approach was the subject of the critiques arose by S4 and S8.

Retained sub-factors

Market Analysis: this stage emerged within Case C's model only. No specific insights about the importance of undertaking market analyses were found from the remaining cases. Also during the validation stage, no further confirmation was achieved. Both of the business representatives (S1 and S2) agreed on the importance of doing market analysis, without, however, placing emphasis about its importance within the open data services field.

Wider Adoption: this sub-factor emerged for Case B only. It is related to that stage in which Case B aimed at including more open datasets. These were about the information that tourists need and look for during their visits. However, since this stage was observed for one case only, it was initially retained and re-presented at the final validation stage. During these interactions, no further motivation was collected to motivate its inclusion. Therefore, this sub-factor was retained as its relevancy was found to be limited to Case B's specific context.

Concluding, seven sub-factors emerged across the cases studies to be relevant in describing the adoption stages.

7.5 Cross-Case Conclusions: Adoption Factors Variables and Links

This chapter presented the findings achieved from the cross-case analysis. These findings demonstrate that sixteen factors described by a total of sixty variables and connected by twenty-two links, are the most relevant in influencing adoption of open government data for commercial service innovation in city contexts across the three cases studied. The resulting adoption model generated from the aggregation of the CCAs formulated is depicted in Figure 7.9 (next page). This model represents the main outcome of this study.

In relation to the second research question for this study, seven adoption stages across cases were defined as the most relevant in describing the process through which the organizations went from the initial generation of the service idea, to achieving sustained use of open data. These stages are summarized in Figure 7.8.

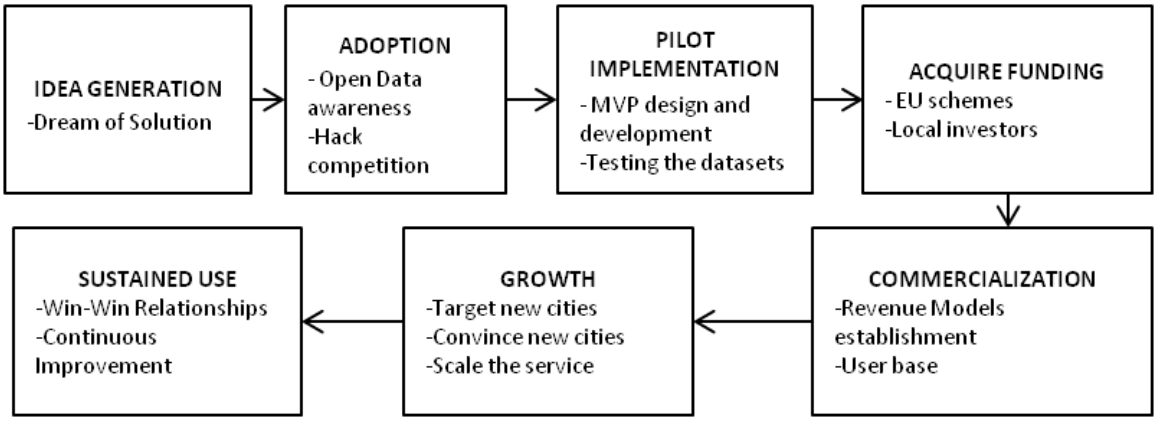

Figure 7.8: Adoption stages across cases

Ultimately, a total of thirty-six variables and two adoption sub-factors were finally retained (a full list is provided in Appendix 7) and their contribution was acknowledged to be limited to those specific contexts from which these concepts actually emerged.

The next chapter focuses on the last stage undertaken in this inductive study, i.e. positioning the cross-case findings within the literature. The thick descriptions generated for each component of the cross-case adoption model were therefore located within both the open data services and the IS adoption literatures. This exercise led to the formulation of specific contributions to these two fields.

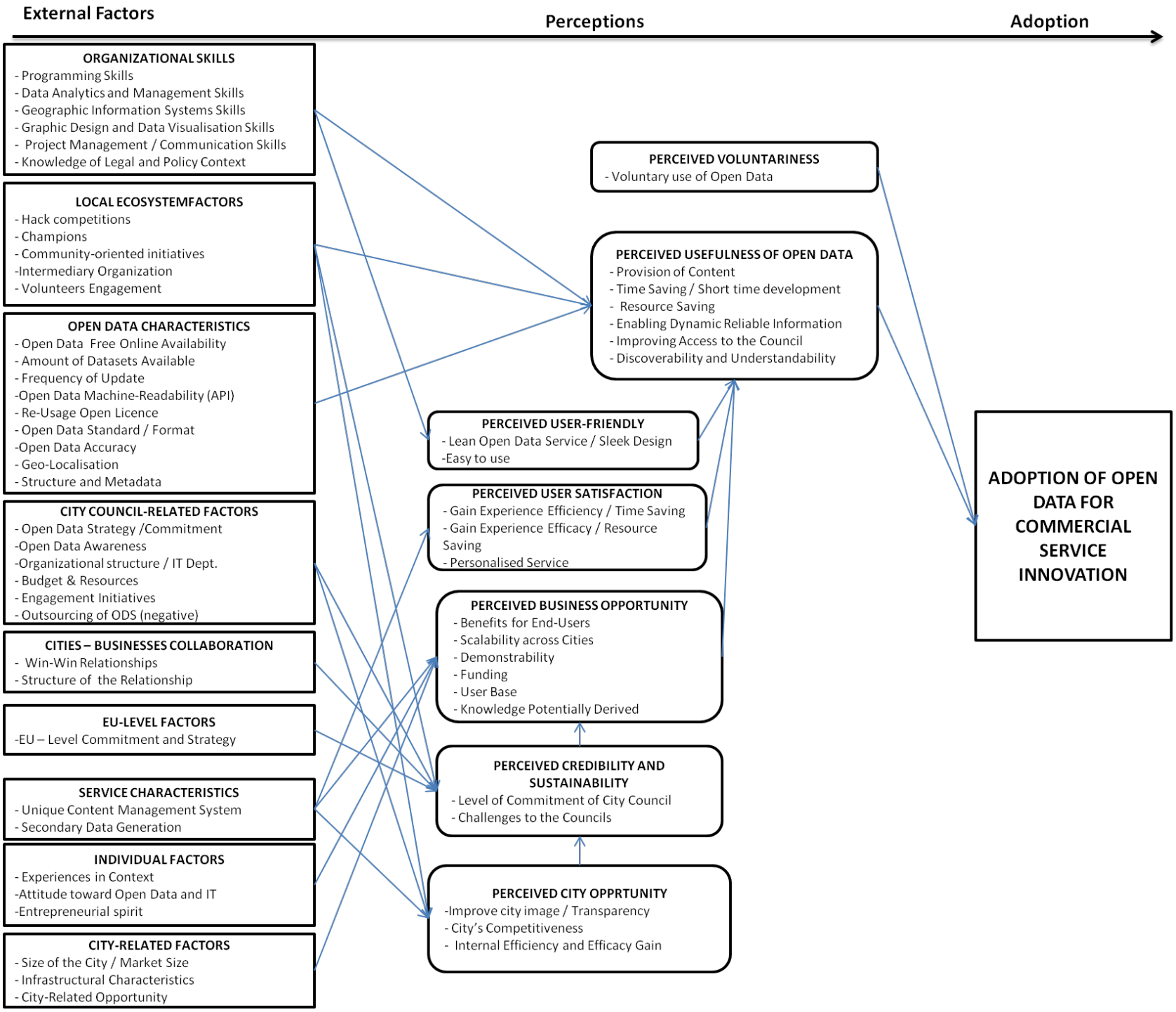

Figure 7.9: Adoption model across cases

8 POSITIONING THE FINDINGS WITHIN THE LITERATURE

This chapter focuses on the last step of this research. According to the methodology designed for this study (see Figure 3.2, section 3.7), this stage was about positioning the inductive findings achieved from the multiple case study investigation (see Chapter 7) within the existing literature. This process has been also defined as "enfolding literature" (Eisenhardt, 1989). It involves "the comparison of predicted patterns and/or effects [i.e. from the existing literature] with the ones that have been empirically observed [i.e. inductive results of this multiple case study], and the identification of any variances or gap" (Baskarada, 2014, p.16). The methods followed for this exercise are explained in section 3.18. Results of these processes enabled the formulation of the contributions of this research to the existing knowledge base. This set of activities has been undertaken in relation to the two main bodies of literature tackled by this research: (1) the IS open data services academic debate and (2) the IS adoption literature. Therefore, specific contributions to each of these domains are highlighted in the following sections. Section 8.1 is dedicated to the positioning of the findings within the open data services literature. On the other hand, section 8.2 focuses on IS Adoption.

8.1 Positioning Findings within the Open Data Services Literature

The choice of adopting an inductive reasoning and a qualitative approach enabled the emergence of factors, variables, and links (enriched by thick descriptions) influencing open data adoption for commercial service innovation from the point of view of re-users. Their aggregation led to the formulation of a model in which drivers of open data adoption are outlined as well as the relationships between them (see Figure 7.9). The inductive investigation of the demand-side of this ecosystem and the specific focus on commercial open data services are new to this academic conversation. However, as expected, some of the aspects emerged were already tackled (or mentioned) in the broader open data literature (within and beyond open data services). For example, the needs of releasing open data that is frequently updated and assigned with a re-usage license were already well-acknowledged in the existing literature. Thus, in accordance with the methodology employed, all the papers selected for the SLR study (see Chapter 2) were re-taken into account at a greater level of granularity. By comparing the findings from this research with the insights emerged from this process, more specific contributions to this academic conversation arose. As discussed in sub-sections 8.1.2-4 below, some of the factors emerged from this study were previously mentioned in the literature (mostly through thin descriptions). Here, an extension of these concepts is proposed from the thick descriptions emerged from this research. On the other hand, some of the factors established were neglected in research so far, and are subsequently proposed as enabling a comprehensive understanding of the predictors of open data adoption for commercial service innovation. The following sub-section

proposes the overall contribution of this research to the open data services academic conversation. Subsequently, each of the CCA established (see Chapter 7) is discussed in relation to the existing literature.

8.1.1 Positioning External Factors

This sub-section focuses on positioning CCAs related to external factors within the existing open data services literature. A paragraph for each external factor defined is proposed.

CCA1: Open Data Characteristics

This research asserts that certain open data characteristics need to be in place in order for this source to be perceived useful for commercial service innovation. This factor has been extensively acknowledged in the broad open data literature to-date. With respect to the actual variables emerged, the importance of having data freely available online through portals that act as single access points to the city's open data (CCA1,1) is a widely researched topic (Stephenson *et al.* 2012) (Rittenbruch *et al.* 2012) (Lakomaa and Kallberg, 2013) (Lindman *et al.* 2013b) (Zuiderwijk *et al.* 2015). Some previous studies were specifically focused on how these portals should be designed (Palka, 2013) (Palka *et al.* 2013), i.e. as a "platform that offers a one-stop-experience for data seekers by providing a data catalogue" (Tcholtchev *et al.* 2012, p.862). This data should be free to use and re-use with little to no requirement of authentication or approval (Oh, 2013). The findings from this inductive study also demonstrate that the chances of open data to be re-used for commercial service development increase together with the amount of datasets available (CCA1,2). In relation to this, the literature is contradictory. A portion of papers stresses the importance about "making all public data available" (Matheus *et al.* 2012) (Feng *et al.* 2012) (Hjalmarsson *et al.* 2014, 2015) and that entrepreneurs' perceptions about the usefulness of open data increase with the number of datasets released (Lakomaa and Kallberg, 2013). However, Janssen *et al.* (2012) argue that only the data for which benefits are clear should be published. These authors motivate this statement through three main reasons: (1) "why spending tax payers' money for something that yields no benefits? (2) Risk for privacy and security; and (3) the resulting information quality might be too low". Also aligned with the findings of this study (CCA1,3), open data should be provided "as quickly as necessary to preserve the value of the data" (Matheus *et al.* 2012, p.23), because "applications based on old data are worthless, they only build customer dissatisfaction" (Kuk and Davies, 2011, p.8). All the cases emphasized the importance of having data readable by machines (CCA1,4), in accordance with the majority of the literature (Aoyama and Kojima, 2013) (Wilde, 2010). As also emerged across the cases and during further interactions, APIs should be in place to facilitate access (Herschel and Manolescu, 2012) (Stephenson *et al.* 2012) (Palka *et al.* 2013) (Fox, 2013) (Frosterus *et al.* 2012) (Guo and Kraines, 2010)

(Hannemann and Kett, 2010) (Jurisch *et al.* 2015) and to "provide more stable platforms to develop new services on" (Lindman *et al.* 2013, p.1244). The usage of APIs is believed to help councils in monitoring actual reuse of the data (Foulonneau *et al.* 2014). Moreover, where relevant, geo-localization of data should be provided (CCA1,8). These formats and the standard used (CCA1,6) should be clearly displayed in the portal through the usage of structured metadata to enable individuals to easily search and consume data (CCA1,9). These features are found to improve the accessibility of open data "for others beyond the primary data providers by describing, locating, and retrieving the data efficiently" (Zuiderwijk *et al.* 2014a, p. 8). If not in place, this might lead to considerable amount of "time spent trying to understand and manipulate the data into the required format" (Bonina, 2013 p.26). The absence of a standard and the need to rigorously define the format were found to be two barriers for open data re-usage (Janssen *et al.* 2012). Also, in order to keep consistency with the definition of open data, formats should enable readings without the need of proprietary software (Kuk and Davies, 2011). As expected, data accuracy and information quality (CCA1,7) are also considered crucial in the literature (Herschel and Manolescu, 2012) (Zuiderwijk *et al.* 2014a). Finally, a strong focus in the literature is being placed on open data licensing. As the findings of this research suggest (CCA1,5), Creative Common licenses (Bellini and Nesi, 2013) "by Attribution" (Oh, 2013), should be applied. According to these schemes, "open data should be licensed in a way as to enable re-use in commercial cases, whilst at the same time acknowledging the key contributor that provide, maintain, and improve the data" (Tcholtchev *et al.* 2012, p.861).

CCA2: Individual Factors

In relation to *Individual Factors* influencing adoption of open data, some related considerations were found in the current literature. Particularly explanatory is one of the myths related to open data outlined in (Janssen *et al.* 2012): "the dream is that everyone can make use of the data that is available and that anybody can use the data directly" (p.19). Oh (2013) outlines three types open data re-users: (1) "casual users", i.e. consumers of data "as is" who appreciate human readable formats; (2) "semi-technical users", i.e. those that might manipulate the data with simple mechanisms but still appreciate data in a human-readable format (e.g. journalists); and (3) "technical users", i.e. those with the ability of designing valuable services by using "script-based approaches to data manipulation to handle very large volumes of data" (p.5). In general, open data is believed to be for "IT people" (Hielkema and Hongisto, 2012) (Hjalmarsson *et al.* 2014, 2015) or "IT entrepreneurs" (Lakomaa and Kallberg, 2013) (Kuk and Davies, 2011). Zuiderwijk *et al.* (2015) argue that "without IT skills open data cannot be found, curated, scrutinized, processed, visualised, and used" (p.2). The findings of this research confirm these statements by postulating that individuals with a positive attitude toward IT (CCA2,2) and with entrepreneurial spirit (CCA2,3) are more likely to re-use open data. Also, this research asserts

that for individuals that have experiences in the context (CCA2,1) re-usage of related open data is facilitated. This aspect was found to be relevant in generating perceived business opportunities from open data adoption.

CCA3: City-Related Factors

With respect to cities' characteristics as a further factor that influences open data adoption, only a few studies shed light on this issue. Some contributions highlight the importance for cities to have in place a certain infrastructure (CCA3,2), "in order to integrate data gathered from sensors, information provided by public institutions and social information created by social application" (Rittenbruch *et al.* 2012, p.2). On the other hand, the findings from this multiple case study suggest that bigger cities (and with a related bigger market size) are places where open data is more likely to be re-used for commercial service development (CCA3,1 and CCA3,3). The literature acknowledges that different cities have different opportunities (Hielkema and Hongisto (2012) described this variable as "demand conditions") and that "multifaceted market conditions and uncertain product demand" are variables influencing the perceptions about a business opportunity from an already developed open data service (Hjalmarsson *et al.* 2014). Nonetheless, no specific considerations about the actual size of the city were found.

CCA4: City Councils – Related Factors

Together with the characteristics of open data, some attributes of city councils were of the most researched in this field to-date. According to the findings of this research, city councils must have in place strategies across departments for the delivery of open data (CCA4,1). In alignment with this research, the literature acknowledges that developing such strategies requires "considerable investment into reverse-engineering, documenting data elements, data clean-up, schema mapping, and instance matching" (Cyganiak *et al.* 2010, p.1). Hence, resources and budget should be allocated (CCA4,4) merely to open data initiatives (Tcholtchev *et al.* 2012). From another perspective, budget and resource constraints are seen as the two economic barriers for open data release (Barry and Bannister, 2013) (Janssen *et al.* 2012). As a basis for this advocated strategy, awareness of open data should be spread throughout the city authority (CCA4,2). Barry and Bannister (2013) promote the idea that un-awareness of councils "serves as an excuse for inaction" (p.13). Awareness was also found in this research to include two further aspects: (1) awareness of IT; and (2) awareness of the requirements for actual and potential re-users. The former is reflected in the need for councils to have an established IT department and a city CIO in place (CCA4,3). The latter was stressed within the variable *Engagement Initiatives* (CCA4,5). In the extant literature, open data is found to be mainly a data provision movement (Foulonneau *et al.* 2014). This "one-way communication" is argued to be a

barrier for open data re-usage (Janssen *et al.* 2012). This research proposes the implementation of various programs to enable the councils to spread awareness of open data among potential re-users. These initiatives, on the other hand, are meant to enable re-users to educate the councils in terms of what their needs are. The last aspect emerged from this research within this CCA is related to the practice of outsourcing the design of open data services (CCA4,6). Bonina (2013) acknowledges that these practices are in place, with local governments actually investing in the co-creation of services. However, no specific negative implication is assigned to these activities. The findings of this research suggest that these practices negatively affect perceived business opportunities for re-users. Concluding, the literature also suggests that city councils should promote open data services developed from their datasets on the city authority's website (Chan, 2013). This aspect emerged from Case A and no significant confirmation for its establishment was found in later steps of this research.

CCA5: Local Ecosystem Factors

In this research, *Local Ecosystem* factors were defined as elements that should be in place in the local environment outside the city council. *Hack Competitions* (CCA5,1) were critical for all the cases in the adoption process. These are seen as "intermediaries for open data innovation" (Juell-Skielse *et al.* 2014). Hjalmarsson *et al.* (2015) define these as "events in which third-party developers compete to design and implement the most firm and satisfying digital service prototype, for a specific purpose, based on open data" (p.2). The same authors propose a measurement model for assessing and subsequently managing open data hack competitions (Ayele *et al.* 2015). However, also in alignment with the findings of this research, these competitions are not found to be directly linked with the market entry of the developed solutions (Hjalmarsson *et al.* 2015). It is noted that in some cases (e.g. (Juell-Skielse *et al.* 2014)) these initiatives are considered to be part of the city council's strategy. The reason why this variable was included in this cluster relies on the fact that the cases analysed leveraged this form of engagement from competitions organised by other associations or authorities outside the city councils.

The importance of *Intermediary Organizations* (CCA5,4) is also acknowledged in the open data literature (Bonina, 2013). Tammisto and Lindman (2011) when proposing a framework of actors involved in the open data services' landscape, highlight the importance of intermediaries named "linked data producers". According to the authors, their goal is to take raw data from city authorities and develop linked data for application developers. The general idea is that these bodies are meant to help the city councils (Stephenson *et al.* 2012) (Bonina, 2013) (Barry and Bannister, 2013) in effectively releasing open data. These propositions differ from what emerges from this research. One of the key aspects for these entities was found to be their level of independence from the city councils.

The situation in which these organizations do not work closely with the authority is believed to "challenge the city council's initiative". This results in a strong stimulus for the council to improve open data outcomes and subsequently in an increased credibility of the overall open data movement, among both re-users and potential investors. Based on the same concept (i.e. "the importance of being vocal towards the councils and their initiatives" - Case C), this research proposes that community-led initiatives should be also undertaken (CCA5,3). Differently from those emerged in relation to the councils' activities, these were found to be more effective in providing input and feedback on the council's open data initiative. The current literature takes into account various forms of programs including living labs (Hielkema and Hongisto, 2012) and training programs for re-users from third parties (Foulonneau *et al.* 2014). The central idea emerging from this research is that city councils should participate without leading these initiatives. Some further examples of these initiatives are "Code for America" (http://www.codeforamerica.org/) and those designed on the same principles in other countries worldwide (e.g. Code for Ireland, http://codeforireland.com/). These initiatives, among other outcomes they achieve, have the potential to provide volunteers (CCA5,5) that contribute to the development of open data services. Finally, the role of *Champions* (CCA5,2) from outside the city authorities was highlighted. These subjects are perceived as more effective than those within the municipalities. The latter are believed to have a limited role in relation to the trade-off between ensuring fair competition and at the same time supporting open data services businesses.

CCA6: Organizational Skills

The investigation of what skills are needed to develop a service from open data is original in this academic conversation. Lindman *et al.* (2013b) highlight the need for IS researchers to answer the following research question: "what are the skills and knowledge needed in developing high quality open data services?" (p.1245). This research proposes a set of six skills identified across the three cases studied. These emerged as being required from taking raw open data to achieving sustained adoption of this source for the delivery of commercial services. *Data Analytics and Data Management* (CCA6,2), *Programming* (CCA6,1), *Geographic Information Systems* (CCA6,3), and *Graphic Design and Data Visualisation Skills* (CCA6,4) emerged to be fundamental throughout all stages of open data usage, i.e. from accessing the dataset to the delivery of the final service. On the other hand, knowledge about the legal and policy context (CCA6,6) should be present in order to comply with the current heterogeneous licenses addressing open data re-usage for commercial purposes. Finally, this research stresses the importance of developing *Project Management and Communication Skills* (CCA6,5). These emerged to be critical especially for approaching the councils. The importance of establishing relationships with the council where open data is actually released was identified for one critical reason, i.e. achieving credibility and sustainability of the open data provision.

CCA7: City Council – Business Relationship

The main challenge that emerges from the open data services businesses studied is related to ensuring the future availability of the data. This concept was found to be critical and a main condition to perceive business opportunities. This aspect remains substantially unexplored in the literature to-date. In their studies aiming at identifying inhibitors for open data services to "enter the market", Hjalmarsson *et al.* (2014, 2015) identify a barrier named "hard to interact with data providers". A possible reason can be traced in (Hielkema and Hongisto, 2012) when the authors underline the fact that a "mismatch [exists] between governmental strategy implementation and SMEs expectations on development and commercialization" (p.193). Diverse motivations between city councils and open data service developers are also highlighted in (Lindman *et al.* 2013b) and left as an open issue for future IS studies. Similarly, Chan (2013) advocates the need for IS researchers to explore "what are the strategies to foster participation and collaboration in open data initiatives", without, however, specifically mentioning the actors and the goals for establishing participation and collaboration. The findings of this research suggest that establishing relationships and partnerships with city councils is a cornerstone for businesses to sustainably adopt their open data. Two aspects emerged to be critical towards building these relationships. Firstly, establishing *Win-Win Relationships* (CCA7,1) can be seen as a solution towards aligning objectives of the two parties. In this way, businesses should, through the design of their open data services, provide an outcome that is beneficial for the council, in exchange for the sustainable provision of open data. The second variable under this cluster is related to the structure of these relationships (CCA7,2). The cases indicate that no optimal solution is yet found for this field. In fact, this research suggests that the establishment of contracts between the two parties can be beneficial for the business in question (e.g. Case A), but results in precluding other businesses to innovate in this field. The term "lock-in" emerged from the interaction with Case C to explain these situations in which new businesses are discouraged to re-use open data in a specific context because of the presence of an existing partner. Therefore, initial interpretations suggest that optimal solutions for establishing these relationships might exist, but are strongly dependent on the characteristics of the specific context in which open data is re-used.

CCA8: Service Characteristics

CCA8 emerged in relation to two main characteristics of open data services: (1) *Unique Content Management System*; and (2) *Secondary Data Generation*. With respect to the latter, Lakomaa and Kallberg (2013) showed that a large proportion of open data entrepreneurs "intend to use data indirectly" (p.561). However, this aspect is emphasized regarding a potential underestimation of the total impact of open data release. Ferro and Osella (2012) acknowledge to some extent that in order

to provide a commercial open data service, open data should be an "ingredient of the final service rather than the final good in itself" (p.15). From this inductive study, the importance of deriving secondary knowledge through data analytics processes emerged as being twofold: (1) this knowledge can be a fundamental part of the open data service's revenue model; (2) this knowledge was found to be the key for leveraging the possibility of establishing relationships with the council. In other words, the cases stressed the importance for the services to generate more detailed understanding of the context in question. This previously unavailable knowledge represents the "offer" that open data services businesses make to the councils in exchange for the sustainable release of open data. The final implication of this outcome is about enabling more informed decisions as well as improving efficiency in managing that particular area of the city. This is believed to contribute to perceived opportunities from the city (CCA12). The latter was found as a main condition for businesses to perceive the availability of open data as credible and sustainable.

The second variable within this CCA emerged in relation to the importance of having a service that is designed using a *Unique Content Management System* (CCA8,1). This feature was found to enable scalability of the service. This characteristic is perceived to be a critical component for growth. Indeed, growth is conceived by the three cases studied as scaling the service across different cities selected consistently with the city councils' and cities' characteristics outlined above.

CCA9: EU-Level Factors

This CCA stresses the importance of having EU-level commitment in terms of fostering an integrated strategy across cities, and pushing towards the development of innovative policies and licenses. These processes are believed to facilitate commercial re-usage of open data. In relation to these aspects, two main topics are covered in the literature. On one side, Foulonneau *et al.* (2014) describe the importance of EU-related policy strategies as follows: "to ensure the lowest possible barriers to reuse and in particular that re-users will not have to address heterogeneous reuse conditions when reusing multiple datasets in conjunction". According to the cases, the focus of EU should also include promoting best practices for open data release and re-usage. By undertaking these activities, the development of "one location for all initiatives" (Case B) can be achieved. On the other hand, EU is expected to play a key role in co-funding the development of services that reuse open data (Hjalmarsson *et al.* 2014, 2015). Ultimately, this research suggests that champions should be in place, and actions from EU should aim at educating councils as well as challenging their operations. The latter is perceived to add credibility to the city-related open data program.

8.1.2 Positioning Perceptions

Regarding perceptions emerged from this research, little (i.e. from thin descriptions derived from deductive studies) or no information was previously available in the literature from the point of view of those re-using open data. However, some studies focusing on the supply-side of this ecosystem attempt to predict what perceptions for users might be. These concepts are usually stated as opinions of the authors in relation to the outcomes of solutions proposed for the effective and efficient open data release. These studies are included in this exercise, although it is noted that the inductive investigation of these factors from the point of view of the demand-side of this landscape is new to this academic conversation. As done for external factors, one paragraph is dedicated to each CCA established.

CCA10: Perceived Usefulness of Open Data

When adopting open data for commercial service development, the cases were found to perceive open data useful across six main variables. Firstly, and most intuitively, open data provides content for the service to be developed (CCA10,1). Chan (2013) noted that "developers do not need to jump through all these hoops now" referring to the possibility of avoiding time-consuming processes related to searching for the agency owning the data and to asking for permissions to use it. These ideas are aligned with the assertions made in relation to *Perceived Usefulness of Open D*ata in terms of saving time and resources (CCA10,2 and CCA10,3). Beyond these characteristics, in order for open data to be adopted, this needs to be perceived as a reliable and dynamic source of information (CCA10,4). The literature includes this concept as being enabled by certain open data characteristics (Ijima *et al.* 2011). However, the findings of this research enable further reflections on how this perception is developed. In fact, *Perceived Usefulness of Open Data* is found to be in place only if perceptions about credibility and sustainability of open data are present. These, in turn, are generated from the presence of several external factors, such as the establishment of a relationship with the council (CCA7) and the presence of certain characteristics within these authorities (CCA4).

Open data is also increasingly perceived useful as *Understandability and Discoverability* (CCA10,6) of the datasets increase. In alignment with this research, this characteristic is believed to be enabled by the thoughtful development of metadata within the open datasets (Zuiderwijk *et al.* 2014a). From another perspective, "current heterogeneity of metadata is an obstacle to the dataset's discovery" (Foulonneau *et al.* 2014, p.3). Finally, this research suggests open data is perceived as useful in terms of improving access to the councils (CCA10,5). In this way, open data is found as a "badge of honour" (Case C) that "opens the councils' doors" (Case A). Therefore, the fact of using open data is found to

be valued by councils that are “desperately looking for examples of re-usage of their data to gain legitimacy for their programs” (Case A).

CCA11: Perceived Business Opportunity

Perceptions about business opportunities were found to play a role in influencing open data adoption. Ferro and Osella (2012, 2013) significantly contribute to this discourse when identifying business models for open data re-use. In addition to this, this study suggests that there are several concepts involved in the generation of these perceptions.

CCA11,1 stresses the need for the service to match a market opportunity. This opportunity depends on the nature of the city. It is identified from the knowledge that individuals have about the specific context. In the literature it is acknowledged the fact that open data services are developed in response to the popular demand of the market (Kuk and Davies, 2011). Identifying a market opportunity is one of the main barriers for the market entry of open data services (especially at early stage of development) (Hjalmarsson *et al.* 2014, 2015), which in turn demonstrates its importance. Furthermore, the service’s potential to derive secondary knowledge from the usage of the service itself has important implications (CCA11,6). These outcomes constitute part of the revenue models of the cases analysed. The bigger the user base is and the greater is the potential value of this knowledge (CCA11,5). The existing literature confirms that perceived business opportunities increase in parallel with the number of service’s users. Halb *et al.* (2010) argue that “revenues largely depend on the number of users consuming content and indirectly also on their session length” (p.1). However, this statement is based on the assumption that revenues from open data services come from advertisements. No specific relationship with data potentially structured from the actual service is made by the authors. This research also suggests that perceived business opportunities are related to the perceived ability to scale the open data services to other cities beyond the one where the service was initially designed (CCA11,2). For achieving scalability, two critical components were identified. Firstly, the service needs to be developed on a *Unique Content Management System*. Secondly, it is important to demonstrate the value of the service in order to convince city councils to embrace the idea and subsequently establish a relationship with the business (CCA11,3). This aspect provides a partial answer to the question proposed in (Lindman *et al.* 2013b): “how to convince data providers to open their data, if not already available?” (p.1245). Finally, this CCA suggests that perceived business opportunities increase with the availability of funding (CCA11,4). Different funding schemes and sources were leveraged by the cases studied.

CCA12: Perceived City Opportunity

A further CCA states that perceived opportunities from the city influence the adoption of open data for commercial service innovation. The idea behind this CCA is that: if cities perceive opportunities, perceptions are generated about additional motivations affecting the councils that lead to the development of valuable, consistent, and long-term open data strategies. The findings of this research suggest that the following opportunities exist for city councils: (1) opportunities from the actual design and delivery of the open data service (CCA12,3), (2) opportunities in terms of improving their images (CCA12,1); and (3) in terms of the perceived ability to enable economic growth (CCA12,2). Improving the city's image is understood as increasing the perception about the city being "open and transparent" (Zuiderwijk *et al.* 2014a). In relation to point (2), the term "city competitiveness" emerged from the interactions undertaken in this research. Hielkema and Hongisto (2012) mention "city competitiveness" as a major goal for European cities. According to the authors, open data is "becoming a new impulse for competitiveness for the city providing the data" (p.201). In relation to point (3), it emerged to be critical for cities to perceive opportunities in terms of internal efficiency and efficacy gain. This aspect is believed by the cases studied to be a significant push for cities to engage in relationships with open data re-users. In particular, this research suggests that the fact that additional knowledge can be delivered to the councils (i.e. secondary knowledge created from the delivery of the open data services) leverages these perceptions by adding additional value generated on the back of open data (as opposed to from the data itself). In the literature, Lindman *et al.* (2013) shed light on how open data services innovation can lead to "efficiency gains in administration [that] can be more significant than the loss of revenues" (p.1244). Establishing *Win-Win Relationships* arose to be the key for generating these beliefs. On the other hand, enabling these perceptions emerged as critical for re-users for perceiving long-term commitment of the city councils.

CCA13: Perceived Voluntariness

The concept of voluntariness in relation to open data services is not new (Bonina, 2013). When talking about who is involved in re-using open data for service innovation, Kuk and Davies (2011) describe this process as "voluntaristic hacking" (p.5). This particular situation was encountered in relation to Case B, where an individual started a project at the volunteer level, and perceived commercial opportunities emerged later in the process. It is noted that the role of voluntariness is different across contexts of adoption (i.e. within and beyond open data for service innovation). For example, Zuiderwijk *et al.* (2015) tackle a different population and demonstrate that "voluntariness of using open data negatively influences open data use and acceptance" (p.8). The authors justify this statement as follows: "the more compulsory, required and demanded the use of open data becomes,

the more the behavioural intention to use open data increases" (p.8). However, their study "focused on the use of open data for the purpose of research, scrutinizing data, and obtaining new insights" (p.9).

CCA14 and CCA15: Perceived User-Friendly and Perceived User Satisfaction

These two factors emerged from the cases in relation to perceptions generated for the open data services' users. These aspects are believed to be widely researched in the service innovation literature beyond the open data services domain. In this specific academic conversation, Juell-Skielse *et al.* (2014) highlight the importance of enabling these perceptions for users when describing evaluation criteria for selecting winners of open data hack competitions. Further confirmation comes from (Hielkema and Hongisto, 2012). This paper states that those who will succeed are open data services that are "user-friendly, relevant and attractive" (p.202). Lindman *et al.* (2013) underline the fact that in order to make open data services sustainable there need to be customers, i.e. "users that are willing to pay for the services". When discussing this issue, the paper proposes the following research question: "what impacts customer willingness to pay for open data services?" (p.1245). This research, in addition to *Perceived User-Friendly*, proposes that characteristics that foster end-users' adoption of open data services include: (1) the ability of the services to enable time saving (CCA15,1) and (2) resource saving (CCA15,2), and (3) their level of personalization (CCA15,3). In summary, the cases studied in this research suggest that adoption of open data is influenced by the ability of designing services that enable perceived ease of use and satisfaction among users.

CCA16: Perceived Credibility and Sustainability

The last CCA was believed by all the cases to be the key for enabling a number of perceptions, and for achieving sustained adoption of open data for commercial service innovation. Indeed, perceiving credibility and sustainability of the provision of open data enables *Perceived Business Opportunities*; these, in turn, enable individuals to perceive open data as useful for commercial service development. This argument is new to the academic conversation on open data services. Some could interpret that Bonina (2013) have touched on this aspect when arguing that one of the characteristics of open data services today is that they "rely on ad-hoc development without expectations on long-term economic profit" (p.16). Nevertheless, no specific descriptions are provided for this statement. On the same line, the fact that governments publish their data "does not automatically translate to trust and confidence from the users" was also mentioned in (Lakomaa and Kallberg, 2013). However, as in the previous case, no solutions are proposed to address this problem. The definition of this CCA and its (inbound and outbound) links provide an answer to the research questions proposed for IS researchers in (Lindman *et al.* 2013b): "what are the motivations and incentives needed for the data providers to

maintain the data and ensure its continued accuracy?" When looking at the broader open data literature, however, some (contradictory) insights seem to emerge in this way. For example, Barry and Bannister (2013) also highlight the importance of trust in governments, however from the point of view of transparency (i.e. releasing open data leads to increase transparency which in turn leads to increased trust). Their conclusion is that "the data released is not found to increase trust in governments" (p.27). The findings from the cases studied acknowledge the absence of this relationship also in relation to open data services (there is no direct link between *Open Data Characteristics* and *Perceived Credibility and Sustainability*). Nonetheless, the cross-case findings demonstrate that this set of perceptions is augmented if: (1) intermediaries are in place and operate independently from the councils and, together with an EU-related effort, provide "challenges to these initiatives" (CCA16,2); and (2) the level of commitment of the council is perceived as credible and sustainable (CCA16,1). In relation to the latter, the findings suggest that city councils-related characteristics influence the creation of these perceptions. Also, and more importantly, the establishment of a relationship with the council is the key driver for generating these beliefs.

8.1.3 Positioning Adoption Stages

In relation to the actual adoption factor, seven sub-factors emerged across the three cases as the description of the process undertaken from the initial generation of the idea to the achievement of sustained use of open data (Figure 7.8). Within both the open data services and the broader open data literatures, the investigation of this process represents a novel contribution.

As the first stage for all companies, the initial idea about the service was generated (*Stage 1 – Idea Generation*). What is common across the three cases is that the idea preceded the actual awareness of open data and so, its adoption. All these ideas faced a challenge in terms of "lack of data" for their actual implementations. Subsequently, all the cases entered a stage in which open data was actually adopted for enabling the development of the service itself (*Stage 2 - Adoption*). The three cases adopted open data at a hack competition. After each case designed a first basic solution as a result of these events, the first pilot of the open data service was implemented and delivered in the city where the hack competition was undertaken (*Stage 3 – Pilot Implementation*). Common activities conducted in this stage included: testing the datasets; improving the service, its design and features; and engaging with the councils aiming at shaping the idea in a way that would enable the establishment of a *Win-Win Relationship*. Already at this stage, initial perceptions arose within the cases in terms of the need of ensuring the availability of this data in the long-term. Thus, the cases started to get closer to the overall ecosystem (e.g. by attending public meetings and workshops on open data and other community-oriented initiatives). The fourth stage followed the official launch of the companies, and

was about the searching of funding opportunities (*Stage 4 – Acquire Funding*). Commercial contracts with the council (Case A), public funding schemes (Case B), and private investors (Case C) were the sources successfully targeted. At this stage, revenue models of the open data services were conceptualized and established (*Stage 5 - Commercialization*). The concept of an increased user base that leads to an increased perceived business opportunity emerged to be particularly important. Indeed, both cases B and C established revenue models also based on the analytics of the "data collated through the app" (i.e. from *Secondary Data Generation* processes). Moreover, all the three cases established revenue models on additional characteristics added to the original pilot. At this stage, all open data services were fully available in the original city. Interestingly, for all the cases, growth (*Stage 6 - Growth*) is meant to be a consequence of the ability of delivering the service in other cities. The key for this process is to have in place a scalable solution. The three cases agreed that when targeting new cities, "approaching the council is always the first activity" (Case C). This is because of the need to establish long-term relationships with these entities to perceive a sustainable locally-related business opportunity. As the last milestone, all companies achieved *Sustained Use* of open data for the continuous delivery of their commercial services. At this stage data is provided with the required characteristics and is being perceived useful as the source for the various commercial services offered. Perceived sustainability of the overall business is in place and it is mainly enabled by the establishment of partnerships and relationships at various levels with the authorities. Thus, perceived business and city opportunities continue to be in place and foster the effective development and maintenance of these relationships. Concluding, all the cases are seeking growth by continuously monitoring the open data landscape across cities in terms of who is implementing an effective open data program, i.e. to identify cities that can be potentially targeted for scaling the services.

8.2 Positioning Findings within the Literature on IS Adoption

According to the role of theory in this inductive study (see section 3.6), the last stage of this research involved the enfolding process of theories related to adoption of Information Systems (IS). As argued in sub-section 3.6.2, the starting point for this phase of the research was the Unified Theory on Acceptance and Use of Technology (UTAUT) (Venkatesh *et al.* 2003). An overview of the UTAUT model is proposed in Figure 8.1.

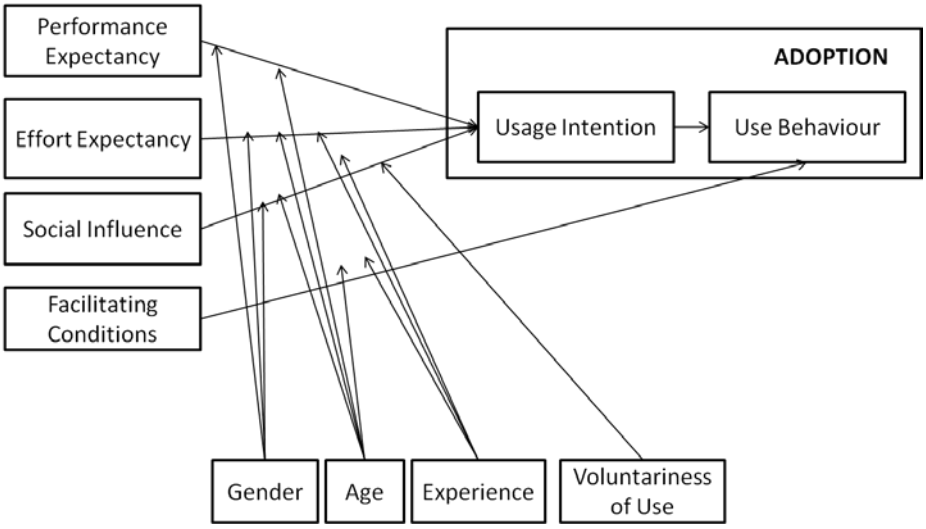

Figure 8.1: UTAUT model (Venkatesh *et al.* 2003)

It is noted that those blocks depicted at the bottom of Figure 8.1 represent moderating variables. The purpose of considering these variables in the theory refers to increasing the predictive validity of the model itself (i.e. increasing the variance explained by considering gender, age etc.). Given that the identification of these moderating variables was beyond the scope of this inductive study, these were not taken into account as part of this exercise. In fact, moderating variables are typically identified through quantitative studies involving large samples (Hartwick and Barki, 1994) (Venkatesh and Morris, 2000). Therefore, the investigation of the potential role of moderating variables such as gender and age will be proposed as a scope for future research, i.e. when leveraging the findings of this research to quantitatively investigating this phenomenon (see section 10.3).

To be consistent with the methodology adopted for this stage of the research (see section 3.18), the strategy that was followed for this exercise was to first compare the cross-case findings with those constructs defined within the UTAUT configuration. If for a specific CCA a match with one or more of the UTAUT's constructs was not found, an additional attempt was made to position its content within other constructs defined across all the models integrated in the UTAUT. It is noted that the authors of the seminal paper proposed an extension of the original UTAUT model (i.e. UTAUT2) (Venkatesh *et al.* 2012). This was taken into account as a further source of "matchable" constructs.

Therefore, the following sub-sections focus on describing all the single constructs defined in the UTAUT theory within which one or more factors emerged from this research could be positioned. The definition of each UTAUT's construct is provided together with an overview of those constructs that were integrated from the eight theories that were combined for the development of the UTAUT model. In integrating these constructs, this exercise was also conducted at the item's level. These items represent the scales used in each of the theories to measure each of the constructs. Therefore,

an overview of the items used in each construct is also provided. After describing those elements of the theories that were found to be relevant for this research, reflections are proposed on the CCAs positioned within each of these constructs.

8.2.1 Performance Expectancy

Performance Expectancy is defined as the "degree to which an individual believes that using the system will help him or her to attain gains in job performance" (Venkatesh *et al.* 2003, p.447). This factor was drawn from Perceived Usefulness (Davis, 1989) (Venkatesh and Davies, 2000), Relative Advantage (Rogers, 2010), Outcome Expectations (Compeau and Higgins, 1995), Extrinsic Motivation (Davis *et al.* 1992), and Job-fit (Thompson *et al.* 1991). An overview of the constructs combined is provided in the following table. Subsequently, an overview of the items used across the combined constructs for investigating this factor is also provided (Table 8.2).

Construct	Reference	Definition
Perceived Usefulness	(Davis, 1989) (Davis *et al.* 1989)	The degree to which a person believes that using a system would enhance job performance.
Extrinsic Motivation	(Davis *et al.* 1992)	The perception that users will want to perform an activity because it is perceived to be instrumental in achieving valued outcomes that are distinct from the activity itself, such as improved job performance.
Job-fit	(Thompson *et al.* 1991)	How the capabilities of a system enhance an individual job performance.
Relative Advantage	(Rogers, 2010)	The degree to which using an innovation is perceived better than using its precursors.
Outcome Expectations	(Compeau and Higgins, 1995)	Includes performance and personal expectations as a consequence of the adopting behaviour.

Table 8.1: Performance Expectancy, root constructs; adapted from (Venkatesh *et al.* 2003)

By looking carefully at the definitions of Performance Expectancy and the integrated constructs, three factors from the cross-case findings of this research were found to consistently fit. These are: *Perceived Usefulness of Open Data* (CCA10), *Perceived Business Opportunity* (CCA11), and *Perceived City Opportunity* (CCA12).

The first immediate correspondence was found with *Perceived Usefulness of Open Data*. In the constructs integrated in Performance Expectancy, *Perceived Usefulness of Open Data* finds a significant match with the following: Perceived Usefulness (Davis, 1989), Extrinsic Motivation (Davis *et al.* 1992), Relative Advantage (Rogers, 2010), and Outcome Expectations (Compeau and Higgins, 1995).

Construct	Items	UTAUT
Perceived Usefulness (Davis, 1989)	-Using the systems enables me to accomplish tasks more quickly. -Using the system would improve my job performance. -Using the system in my job would increase my productivity. -Using the system would enhance my effectiveness on the job. -Using the system would make it easier to do my job. -I would find the system useful in my job.	PE1: I would find the system useful in my job
Extrinsic Motivation (Davis *et al.* 1992)	Taken from (Davis, 1989), see cell above.	
Job-fit (Thompson *et al.* 1991)	-Using the system will have effect on the performance of my job. -Use of the system can decrease the time needed for my important job responsibilities. -Use of the system increases the quality of output on my job. -Use of the system increases the effectiveness of job tasks. -Use can increase the output for the same amount of effort. -Considering all tasks, the general extent to which use of the system could assist on the job (different scale used).	PE2: Using the system enables me to accomplish tasks more quickly
Relative Advantage (Rogers, 2010)	-Using the system enables me to accomplish tasks more quickly. -Using the system improves the quality of the work I do. -Using the system makes it easier to make my job. -Using the system enhances my effectiveness on the job. -Using the system increases my productivity.	PE3: Using the system increases my productivity
Outcome Expectations (Compeau and Higgins, 1995)	If I use the system: -I will increase my effectiveness on the job. -I will spend less time on routine job tasks. -I will increase the quality of output of my job. -I will increase outputs for the same amount of effort. -My co-workers will perceive me as competent. -I will increase my chance of obtaining a promotion. -I will increase my chance of getting a raise.	PE4: If I use the system I will increase my chances of getting a raise

Table 8.2: Performance Expectancy, items; adapted from (Venkatesh *et al.* 2003)

Referring to Table 8.2, the variables emerged within this cluster can be related to three of the four UTAUT's items defined for Performance Expectancy. In particular, PE1 (i.e. "I would find the system useful in my job") relates to this study to the variable *Provision of Content*. Indeed, for performing commercial service innovation (in PE1 "my job"), open data (i.e. "the system") is useful in terms of *Provision of Content* for the design of the service itself. Similar considerations (in relation to PE1) can be made in relation to open data perceived as understandable and discoverable enabling dynamic information and useful in terms of improving access to the council. The latter can be considered as "useful in my job" (PE1) given the importance emerged about achieving access and establishing relationships with the council. The second item (PE2) is related to open data enabling *Time Saving*. Finally, this research postulates that open data is perceived useful in terms of *Resource Saving* which clearly relates to "using the system increases my productivity" (PE3).

Performance Expectancy also emerged from this research as *Perceived Business Opportunity* and *Perceived City opportunity*. In particular, the performance expected by adopting open data is related to expected opportunities for creating a successful business and opportunities for the city in which open data is actually re-used. Furthermore, this research postulates that these two perceptions (the former directly and the latter indirectly) enable the main aspect considered in this cluster, i.e. *Perceived Usefulness of Open Data*. The overall *Perceived Business Opportunity* can be considered as part of PE4 (if correctly interpreted). Item PE4 postulates: "if I use the system I will increase my chances of getting a raise". Therefore, by interpreting this item as "if I use open data I will increase my chances of creating profit", a match between these two aspects is found. On the other hand, *Perceived City Opportunity* is meant as expected performance achieved by the city if use of open data for commercial service innovation is undertaken. These performances were detailed in terms of: improving city image and competitiveness, and internal efficiency and efficacy gain.

8.2.2 Effort Expectancy

Effort Expectancy is defined as the "degree of ease associated with the use of the system" (Venkatesh *et al.* 2003, p.450). This concept was covered by three constructs across the theories combined: Perceived Ease of Use (Davis, 1989) (Venkatesh and Davies, 2000), Complexity (Thompson *et al.* 1991), and Ease of Use (Rogers, 2010). A summary table including the definitions of these constructs is provided in Appendix 8**.** The items used in each specific construct and those integrated in the UTAUT model are presented in Appendix 8.1. The theory postulates that this factor's influence on adoption is "more likely to be important in mandatory usage settings" (p.469).

Unlike most of the theories combined for the development of the UTAUT, the findings of this inductive study did not include a specific factor about perceived ease of use of open data influencing its adoption. However, aspects and items integrated in this construct overlap with variables defined across different external factors. For example, part of the emerged Effort Expectancy's role in influencing open data adoption emerged within the *Open Data Characteristics* factor. This can be clearly related to Roger's (2010) conjecture about innovation characteristics (including a variable about complexity in using the innovation) influencing adoption and diffusion of the innovation itself. Therefore, this additional construct (i.e. Roger's (2010) Innovation Characteristics) was considered as the most reasonable match with *Open Data Characteristics* (see sub-section 9.2.5 below). However, aspects related to Effort Expectancy also emerged as part of other external factors. Therefore, further reflections were needed in this way. In Zuiderwijk *et al.* (2015) the definition of this construct was related to open data as "the extent to which a person or an organization believes that using open data technology will be free of effort" (p.3). This general definition makes interpretations difficult. For

example, establishing relationships with the councils was found to be a cornerstone for the cases that achieved sustained adoption of open data; a substantial effort was undertaken to establish these relationships, which makes this aspect comparable to Zuiderwijk *et al.* (2015)'s definition of Effort Expectancy. Notwithstanding this, Venkatesh *et al.* (2003) acknowledge that not all models include a specific construct related to Effort Expectancy (e.g. in the case of Theory of Planned Behavior (Ajzen, 1991)). Therefore, in relation to this specific construct, this research's findings are more "comparable" with other theories than UTAUT. One of the consequences of not including this construct is explained by Venkatesh *et al.* (2003) as follows: "if effort expectancy is not present in the model, then one would expect Facilitating Conditions to become predictive of intention" (p. 453).

Overall, factors and variables that can be matched with Effort Expectancy in the UTAUT model were related to: (1) *Open Data Characteristics*; and (2) other external factors. Therefore, the first CCA was matched with the Innovation Characteristics construct as proposed by Rogers (2010). Regarding point (2), according to what suggested by Venkatesh *et al.* (2003), these considerations were addressed and are outlined as part of Facilitating Conditions (sub-section 8.2.4).

8.2.3 Social Influence

Social Influence is defined as "the degree to which an individual perceives that important others believe he or she should use the new system" (p. 451). This construct integrates several aspects from the combined theories including: Subjective and Social Norms (Davis, 1989) (Venkatesh and Davies, 2000) (Fishbein and Ajzen, 1975); Social Factors (Thompson *et al.* 1991); and Image (Rogers, 2010). This construct is found to be relevant to compliance in mandatory contexts. Compliance is related in the theory to the social influence and/or the normative pressure. As a further confirmation of this, when studying open data adoption, Zuiderwijk *et al.* (2015) relate Social Influence to the case in which "open data use is urged by supervisor, managers, teachers, or other influential persons" (p.3). This case is clearly not an example of voluntary use.

This research, consistently with the UTAUT's considerations (i.e. this specific construct is found to be relevant when adoption is mandated, as opposed to voluntary) does not suggest this factor as being critical in the adoption of open data for commercial service innovation.

8.2.4 Facilitating Conditions

Facilitating Conditions are defined as "the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system" (Venkatesh *et al.* 2003, p.453). This factor captures three constructs across the theories combined: Perceived Behavioural Control (Ajzen, 1991) (Taylor and Todd, 1995), Facilitating Conditions (Thompson *et al.* 1991), and

Compatibility (Rogers, 2010) (see Table 8.3). The underlying assumption across these models is that aspects of the technological and/or organizational environment are in place to lower the barriers for using a technology.

Construct	Reference	Definition
Perceived Behavioural Control	(Ajzen, 1991) (Taylor and Todd, 1995)	Reflects perceptions of internal and external constraints on behaviour and encompasses self-efficacy, resource facilitating conditions, and technology facilitating conditions.
Facilitating Conditions	(Thompson *et al.* 1991)	Objective factors in the environment that observers agree make an act easy to do.
Compatibility	(Rogers, 2010), (Moore and Benbasat, 1991)	The degree to which an innovation is perceived as being consistent with existing values, needs, and experiences of potential adopters.

Table 8.3: Facilitating Condition, root constructs; adapted from (Venkatesh *et al*. 2003)

At a greater level of granularity, the items proposed for each construct, and those selected for the UTAUT model are presented in Table 8.4.

The exercise conducted as part of this construct led to the inclusion within Facilitating Conditions of six external factors emerged from this inductive study: (1) *Organizational Skills* (CCA6); (2) *City-related Factors* (CCA3); (3) *City Council Related Factors* (CCA4); (4) *Local Ecosystem Factors* (CCA5); (5) *City Council – Business Relationships* (CCA7); and (6) *EU-Level Factors* (CCA9).

Construct	Items	UTAUT
Perceived Behavioural Control (Ajzen, 1991) (Taylor and Todd, 1995)	-I have control over using the system. -I have the resources necessary to use the system. -I have the knowledge necessary to use the system. -Given the resources, opportunities and knowledge it takes to use the system, it would be easy for me to use the system. -The system is not compatible with other systems I use.	FC1: I have the resources necessary to use the system. FC2: I have the knowledge necessary to use the system. FC3: The system is not compatible with other systems I use. FC4: A specific person (or group) is available for assistance with system difficulties.
Facilitating Conditions (Thompson *et al.* 1991)	-Guidance was available for me in the selection of the system. -Specialized instruction concerning the system was available to me. -A specific person (or group) is available for assistance with system difficulties.	
Compatibility (Rogers, 2010), (Moore and Benbasat, 1991)	-Using the system is compatible with all aspects of my work. -I think that using the system fits well with the way I like to work. -Using the system fits into my work style.	

Table 8.4: Facilitating Conditions, items; adapted from (Venkatesh et al. 2003)

The first and most immediate correspondence was found with *Organizational Skills*. This set of variables relates to the FC2 item (i.e. I have the knowledge necessary to use the system, see Table 8.4) contained in the Facilitating Conditions dimension. Indeed, this factor describes, through the definition of six variables, the skills and the knowledge that were observed across the cases to be required for using open data for commercial service innovation.

Regarding the remaining external factors that were positioned within this construct, partial overlap with actual UTAUT items was found only for *Local Ecosystem Factors* and, to some extent, for *EU-Level Factors*. This conclusion was achieved by interpreting item FC4 (i.e. a specific person (or group) is available for assistance with system difficulties). In this way, the role of *Hack Competitions, Champions, Community-Oriented Initiatives, Intermediary Organization, Volunteers*, and the presence of an overarching strategy led by the EU, were all interpreted as being aspects related to specific persons or group of people / organizations / entities that (if properly in place) were found to foster adoption of open data for commercial service innovation. Indeed, all these variables were established with different roles in terms of assistance, guidance, and help. These include: (1) discovering, analysing, and thinker with the open datasets; (2) helping the council and the business with realizing more consistent and reusable datasets; (3) helping the overall ecosystem by providing "one location for all initiatives" (Case B) with aspects such as spreading awareness, taking the lead on the policy side, and establishing best practices repository (*EU-Level Factors*); (4) educating and challenging the councils; and (5) helping the businesses, especially at early stages of developments, with (usually) technical-related tasks. Furthermore, in relation to points (1) and (4) above, also *Engagement Initiatives* from the city council can be related to the FC4 UTAUT's item. However, it was mainly by considering aspects that were integrated in this constructs from other theories (and for which no specific items were defined as part of the UTAUT model) that the remaining three external factors were positioned within Facilitating Conditions. These constructs are: Perceived Behavioural Control (Ajzen, 1991) (Taylor and Todd, 1995) and Facilitating Conditions (Thompson *et al.* 1991). The former is defined as "perceptions of internal and external constraints on behaviour and encompasses self-efficacy, resource facilitating conditions, and technology facilitating conditions" (Ajzen, 1991) (Taylor and Todd, 1995). The latter was described as "objective factors in the environment that observers agree make an act easy to do" (Thompson *et al.* 1991). Considering *City-Related Factors* and *City Council Related Factors*, these were interpreted as "objective factors in the environment" that facilitate and foster the re-use of open data. In this way, open data adoption for commercial service innovation was found to be influenced by: (1) characteristics of the city in which open data is actually taken in terms of its size, infrastructure and related market opportunities; and (2) the presence within the councils of an established IT skilled workforce and department, and of structured open data

strategies where awareness of open data is spread across departments and budget and resources are allocated.

The match between *City-Business Collaboration* and Facilitating Conditions was established by interpreting Ajzen's (1991) formulation of the Perceived Behavioural Control construct. With respect to its definition, this correspondence was found in relation to "resource facilitating conditions", and to two of the items defined (i.e. "I have control over using the system", and "given the resources, opportunities and knowledge it takes to use the system, it would be easy for me to use the system" (Ajzen, 1991) (Taylor and Todd, 1995)). In this research, the importance of establishing relationships with the council was defined to emphasize the need of businesses to rely on open data as a sustainable source for their service delivery. All the cases designed open data services that benefit the management of the related contexts from the council to leverage the possibility of establishing *Win-Win Relationships.* It is also noted that establishing relationship with the councils also fits into Thompson *et al.*'s (1991) definition of Facilitating Conditions by being an "objective factor in the environment" that facilitates adoption of open data for commercial service innovation.

After considering the four constructs included in the UTAUT theory, nine of the sixteen CCAs were positioned within the model, i.e. their definition was found to correspond with its theoretical constructs. In order to reflect on the remaining seven CCAs, those theories that were combined for the formulation of UTAUT were considered consistently with the strategy employed. As a result, the remaining CCAs were found to be related to the following constructs: Innovation Characteristics (Rogers, 2010), Experience and Attitude (Ajzen, 1991) (Venkatesh *et al.* 2012), Output Quality (Venkatesh and Davis, 2000), and Voluntariness of Use (Moore and Benbasat, 1991). Finally, reflections about *Perceived Credibility and Sustainability* are provided as no specific construct from the combined theories was found to clearly match its definition.

8.2.5 Innovation Characteristics

In most of the studies on adoption and diffusion of IS innovations, the characteristics of such innovations were found to influence these processes. Also in this research, open data adoption was found to be influenced by its characteristics (CCA1). In his seminal contribution to this field, Rogers (2010) defines five different factors that influence adoption of the innovation based on the attributes of the innovation itself. These are: relative advantage, compatibility, complexity (reverse), observability, and trialability. Aspects related to relative advantage of open data emerged in this research as part of *Perceived Usefulness of Open Data* (which was considered within the Performance Expectancy section).

The importance of Roger's (2010) construct is further motivated by the fact that the claimed importance of Effort Expectancy partially emerged as part of the *Open Data Characteristics* (CCA1). Rogers' (2010) study included this aspect through the variable "Complexity" within the Innovation Characteristics' construct. In this research, the following concepts that emerged within CCA1 can be seen as aspects meant to decrease the complexities involved in using open data (i.e. the innovation): *Free Online Availability* (CCA1,1); availability of accurate (CCA1,7) and structured datasets provided with metadata (CCA1,9); data assigned with geo-location information (CCA1,8); data released in machine readable format (CCA1,4); and data with a unique standard and format clearly displayed (CCA1,6). Compatibility emerged in relation to the importance of the re-usage open license (CCA1,5). More intrinsic characteristics that contribute towards perceiving this innovation as useful emerged as *Amount of Datasets Available* (CCA1,2) as well as their *Frequency of Update* (CCA1,3).

It is noted that the establishment of this construct supports the call for future research made by Zuiderwijk *et al.* (2015). In particular, after deductively leveraging the UTAUT model for studying open data adoption for research purposes, the authors concluded: "the large part of the variance is not explained" and "there is a need for open data specific theories that address the idiosyncratic nature of open data, including aspects such as data quality" (Zuiderwijk *et al.* 2015, p.9).

8.2.6 Experience and Attitude

When reflecting on *Individual Factors* (CCA2), different constructs were found to be related to these aspects. When considering the UTAUT's constructs, correspondence was observed between the variables *Attitude towards Open Data* and *IT* and *Entrepreneurial Spirit* and the construct initially defined by Venkatesh *et al.* (2003) as Attitude toward Using Technologies. For its formulation, four specific constructs from other relevant theories were combined. These are: Attitude toward Behaviour (Fishbein and Ajzen, 1975) (Ajzen, 1991) (Venkatesh and Davies, 2000), Intrinsic Motivation (Davis *et al.* 1992), Affect toward Use (Thompson *et al.* 1991), and Affect (Compeau and Higgins, 1995). After concluding their analysis, this construct was "dropped from the model" (Venkatesh *et al.* 2003, p.461). However, the paper states that "the attitudinal constructs are significant only when specific cognitions – in this case constructs related to Effort Expectancy – are not included in the model" (p. 455). Therefore, since UTAUT considers Effort Expectancy extensively important, these constructs were hypothesized to not have a direct influence with adoption. As outlined above, Effort Expectancy emerged from this research through the definition of several variables across several factors (as opposed to a unique one). Therefore, positioning *Individual Factors* within the Attitude toward Using Technologies construct was considered relevant and appropriate. Examples of items related to the variables *Attitude towards Open Data and IT* and *Entrepreneurial Spirit* are (among others): "I

dislike/like the idea of using the system", and "using the system is unpleasant/pleasant" (Ajzen, 1991) (Taylor and Todd, 1995) (Davis *et al.* 1989); "I look forward to those aspects of my job that require me to use the system" (Compeau and Higgins, 1995); and "the system is okay for some jobs, but not the kind of job I want (reverse)" (Thompson *et al.* 1991).

When tackling the variable *Experience in the Context*, it became useful to consider Venkatesh *et al.*'s (2012) contribution to the original UTAUT model. In the original UTAUT model, experience is acknowledged to moderate relationships between different constructs and adoption (Venkatesh *et al.* 2003). As part of the study that extended the UTAUT model (Venkatesh *et al.* 2012), a construct named Habit and Experience was added to its original formulation as a predictor of adoption of IS. In this way, the definition of "experience" was taken from (Venkatesh *et al.* 2003) and (Kim and Malhotra, 2005) as reflecting "an opportunity to use a target technology typically operationalized as the passage of time" (Venkatesh *et al.* 2012, p.161). "Habit", on the other hand, is defined as "the extent to which people tend to perform behaviours because of learning" (Venkatesh *et al.* 2012, p.161). Clearly, the variable *Experience in the Context* can be matched with this category.

In summary, the *Individual Factors'* CCA was positioned within a construct named Experience and Attitude, defined from the combination of the two following constructs: (1) Attitude towards Using Technologies (Venkatesh *et al.* 2003); and (2) Experience and Habit (Venkatesh *et al.* 2012). It is noted that in relation to point (1), Attitude Toward Behaviour (Ajzen, 1991) (Taylor and Todd, 1995), (Davis *et al.* 1989), Intrinsic Motivation (Davis *et al.* 1992), Affect Toward Use (Thompson *et al.* 1991), and Affect (Compeau and Higgins, 1995) were also considered.

8.2.7 Output Quality

Output Quality is defined in (Venkatesh and Davis, 2000) as "how well the system performs those tasks" the system itself is capable of performing (p.191). This concept is theorized to have a positive effect on perceived usefulness of the system (Davis *et al.* 1992) (Venkatesh and Davis, 2000). It is noted that perceived usefulness in those studies was translated within the UTAUT model as Performance Expectancy (see sub-section 8.2.1).

When firstly investigating the relevance of Output Quality in influencing adoption of IS, Venkatesh and Davis (2000) defined two different items: (1) "the quality of the output I get from the system is high" (OQ1); and (2) "I have no problem with the quality of the system's output" (OQ2). In this research the task that the system (i.e. open data) is capable of performing is the design and delivery of a commercial service. One external factor and two perceptions emerged from this inductive study across cases in relation to the "Output" which is the focus of this construct. These are *Service Characteristics* (CCA8),

Perceived User-Friendly (CCA14), and *Perceived User Satisfaction* (CCA15). In other words, "Quality" of the open data service (i.e. the Output) emerged in this research as being determined by the ability of: (1) designing a service with an *Unique Content Management System* (CCA8,1) from which secondary data about the specific context is obtained (CCA8,2); (2) designing a lean open data service (CCA14,1) that is *Easy-to Use* (CCA14,2); and (3) designing a personalized experience (CCA15,3) for the end-user that provides them with time (CCA15,1) and resource savings (CCA15,2) when undertaking the specific activities supported by the service itself. Therefore, both of the items defined for this construct (i.e. OQ1 and OQ2) can be related to the three factors considered. In other words, using theoretical language, CCA8, CCA14, and CCA15 were translated into the following statement: Output Quality influences adoption of open data for commercial service innovation.

8.2.8 Voluntariness of Use

The concept of voluntariness is described within the UTAUT model as moderating the effect of Social Influence on adoption (Venkatesh *et al.* 2003). Voluntariness of use was defined in this context as "the degree to which use of an innovation is perceived as being voluntary or of free will" (Moore and Benbasat 1991, p.195). The debate about the role of voluntariness is quite controversial across authors contributing to this academic conversation. For example, Bansler and Havn (2002) conjectured that when people are forced to use a system they frequently use it in ways that do not benefit their organizations, therefore supporting the positive relationship between voluntariness and adoption. On the other hand, Zuiderwijk *et al.* (2015) proposed the role of voluntariness as being "negatively related to the behavioural intention to use and accept open data technologies" (p.4). This controversy is, however, clarified by several other studies by distinguishing adoption in mandatory contexts from voluntary ones. To add clarity to this reflection, it is useful to consider Venkatesh *et al.*'s (2012) extension of the original UTAUT model: "relative to the original conceptualization of UTAUT, we drop voluntariness as a moderating variable. This change is necessary to make UTAUT applicable in the context of voluntary behaviour. While in general, voluntariness can be perceived as a continuum from absolutely mandatory to absolutely voluntary, consumers [in their specific study] have no organizational mandate and thus, most consumers behaviours are completely voluntary, resulting in no variance in the voluntariness construct" (Venkatesh *et al.* 2012, p.159). In this research, the concept of voluntariness emerged as a standalone perception directly connected with adoption. It is again noted that this research focuses on voluntary settings (as opposed to mandatory ones). In other words, adoption open data for commercial service innovation from this research was not observed as being the result of Social Influence (Thompson *et al.* 1991) or Normative Pressures (Venkatesh *et al.* 2003). Therefore, the emergence of *Perceived Voluntariness* was re-interpreted as matching the theory by describing the context in which adoption is actually studied, i.e. voluntary (as opposed to

mandatory). The fact that no links to this factor emerged in any of the cases studied further supports this decision.

8.2.9 Perceived Credibility and Sustainability

The last factor that emerged from this research is *Perceived Credibility and Sustainability* (CCA16). This remains "uncovered" by the UTAUT model (and by all the models combined within it) (Venkatesh *et al.* 2003) as well as by its extension (Venkatesh *et al.* 2012). None of the constructs proposed where found to be linked to the definition of this factor. In this research, *Perceived Credibility and Sustainability* emerged as being critical for several external factors to impact on perceptions and subsequently on adoption. These perceptions were positioned within the Performance Expectancy construct. For example, for *City-Business Collaboration* to have an impact on Performance Expectancy, *Perceived Credibility and Sustainability* of open data needs to be in place. Similar arguments can be made in relation to *EU-Level Factors* and *City Council-Related Factors*. Therefore, *Perceived Credibility and Sustainability* is inserted in the final model achieved by positioning the cross-case findings within the existing IS adoption theories.

At this stage, all factors were positioned within the literature and described with the language and constructs of widely acknowledged IS adoption theories. The following table summarizes this exercise.

Construct	**Cross-Case Factors**
Performance Expectancy	Perceived Usefulness of Open Data
	Perceived Business Opportunity
	Perceived City Opportunity
Facilitating Conditions	Local Ecosystem Factors
	Organizational Skills
	City Council-Related Factors
	City-Related Factors
	City-Business Collaboration
	EU-Level Factors
Output Quality	Service Characteristics
	Perceived User-Friendly
	Perceived User Satisfaction
Innovation Characteristics	Open Data Characteristics
Experience and Attitude	Individual Factors
Perceived Credibility and Sustainability	Perceived Credibility and Sustainability

Table 8.5: Summary of constructs

8.2.10 Positioning Links

In this sub-section the links emerged from this research are located within those proposed in the literature. Table 8.6 gives an overview of all links emerged from the cross-case analysis contextualized

with those constructs defined in the previous sub-sections. Each construct is tackled separately within the following paragraphs.

Performance Expectancy Links: as shown in Table 8.6, from the three factors positioned within Performance Expectancy, three links were established. By looking at the cross-case findings (see Figure 8.9), it can be observed that *Perceived City Opportunity* is linked with *Perceived Business Opportunity* (although mediated by *Perceived Credibility and Sustainability*) which leads to perceive open data useful, which in turn is directly connected with adoption. Therefore, beside two internal links (i.e. between factors included in the Performance Expectancy construct), one link could be established between Performance Expectancy and Adoption. When looking at the theoretical models considered, significant confirmation is found in relation to this connection. For all models combined in (Venkatesh *et al.* 2003) and for the resulting UTAUT (and its extension (Venkatesh *et al.* 2012)), Performance Expectancy was conjectured as linked with adoption.

Facilitating Conditions and Perceived Credibility and Sustainability Links: due to the inclusion of six different CCAs within the Facilitating Conditions construct, a higher number of originally generated links needed to be analysed. However, when considering the integrated constructs, only three links emerged: (1) Facilitating Conditions → Performance Expectancy (derived from: (a) *Local Ecosystem Factors → Perceived Usefulness*; (b) *Organizational Skills → Perceived Usefulness*; (c) *City Council-Related Factors → Perceived City Opportunity*; and (d) *City-Related Factors → Perceived Business Opportunity*); (2) Facilitating Conditions → Perceived Credibility and Sustainability (derived from: (a) *Local Ecosystem Factors → Perceived Credibility and Sustainability*; (b) *City Council-Related Factors → Perceived Credibility and Sustainability*; (c) *City-Business Collaboration → Perceived Credibility and Sustainability*; and (d) *EU-Level Factors → Perceived Credibility and Sustainability*); and (3) Facilitating Conditions → Output Quality (derived from *Organizational Skills → Perceived User-Friendly*).

Therefore, three different links were merged as relevant for this construct. While the factor-factor relationship of point (3) is clear, more reflections are needed for points (1) and (2). Regarding point (1), it represents CCAs related to factors that are directly connected with one of the three perceptions included in Performance Expectancy.

Construct 1 (from)	Cross-case factor	Links		Construct 2 (to)
Performance Expectancy	Perceived Usefulness	Adoption	→	Adoption
	Perceived Business Opportunity	Perceived Usefulness	→	Internal Link
	Perceived City Opportunity	Perceived Credibility and Sustainability	→	Internal Link
Facilitating Conditions	Local Ecosystem Factors	Perceived Credibility and Sustainability	→	Perceived Credibility and Sustainability
		Perceived Usefulness	→	Performance Expectancy
	Organizational Skills	Perceived User-Friendly	→	Output Quality
		Perceived Usefulness	→	Performance Expectancy
	City Council-Related Factors	Perceived Credibility and Sustainability	→	Perceived Credibility and Sustainability
		Perceived City Opportunity	→	Performance Expectancy
	City-Related Factors	Perceived Business Opportunity	→	Performance Expectancy
	City-Business Collaboration	Perceived Credibility and Sustainability	→	Perceived Credibility and Sustainability
	EU-Level Factors	Perceived Credibility and Sustainability	→	Perceived Credibility and Sustainability
Output Quality	Service Characteristics	Perceived Business Opportunity	→	Performance Expectancy
		Perceived City Opportunity	→	Performance Expectancy
		Perceived User Satisfaction	→	Internal Link
	Perceived User-Friendly	Perceived Usefulness	→	Performance Expectancy
	Perceived User Satisfaction	Perceived Usefulness	→	Performance Expectancy
Innovation Characteristics	Open Data Characteristics	Perceived Usefulness	→	Performance Expectancy
Experience and Attitude	Individual Factors	Perceived Business Opportunity	→	Performance Expectancy
Perceived Credibility and Sustainability	Perceived Credibility and Sustainability	Perceived Business Opportunity	→	Performance Expectancy

Table 8.6: Summary of links between constructs

Point (2), on the other hand, shows that some of the factors' impact on Performance Expectancy is mediated by perceptions about credibility and sustainability of open data (*Perceived Credibility and Sustainability* has one unique outbound link to *Perceived Business Opportunity*, considered here within Performance Expectancy). According to these reflections, three different links can be consistently

established from points (1) and (2). Firstly, a direct link is supported from Facilitating Conditions to Performance Expectancy. Secondly, a link is established from Facilitating Conditions to Perceived Credibility and Sustainability. Finally, a link is defined from Perceived Credibility and Sustainability to Performance Expectancy. The two different impacts of Facilitating Conditions on Performance Expectancy reflect the findings of the cross-case analysis. Accordingly, some variables of those factors positioned within Facilitating Conditions were found to be connected directly with one of the perceptions included in Performance Expectancy, whereas some others were linked to *Perceived Credibility and Sustainability*. For example, considering *Local Ecosystem Factors*, both links to *Perceived Credibility and Sustainability* and *Perceived Usefulness of Open Data* were established. This was the result of some variables contributing to the former and some to the latter. For instance, participating in *Hack Competitions* and *Community-Oriented Initiatives* were found to contribute to achieving perceptions about usefulness of open data. On the other hand, the challenging role of the *Intermediary Organization* was found to contribute to *Perceived Credibility and Sustainability*. In summary, considering the thick descriptions achieved for all the links established (section 7.2), Table 8.7 provides a list of variables and cross-case factors integrated within Facilitating Conditions which impact to Performance Expectancy is mediated *by Perceived Credibility and Sustainability*. This table is complemented by Table 8.8 where all the integrated variables directly connected with Performance Expectancy are listed.

Variable	**Cross-case Factor**
Open Data Strategy / Commitment	City Council – Related Factors
IT Department	
Budget and Resources	
Outsourcing of Open Data Services (reverse)	
Intermediary Organization	Local Ecosystem Factors
Win-Win Relationships	City – Business Collaboration
Structure of the Relationship	
EU-Level Commitment and Strategy	EU-Level Factors

Table 8.7: Facilitating Conditions – Perceived Credibility and Sustainability link

Therefore, as shown in Table 8.7, for eight of the twenty-three variables positioned within the Facilitating Conditions construct, the impact on Performance Expectancy is mediated by the *Perceived Credibility and Sustainability* factor. Furthermore, by looking merely at *Perceived Credibility and Sustainability*, it is observed that for *Perceived City Opportunity* to contribute to *Perceived Business Opportunity*, credibility and sustainability should be perceived as well. Therefore, the relationship previously established between these two factors within the Performance Expectancy construct results as mediated by *Perceived Credibility and Sustainability*.

Variable	Cross-case Factor
Size of the City	City-Related Factors
Infrastructural Characteristics	
City Related Opportunity	
Open Data Awareness	City Council –Related Factors
Hack Competitions	Local Ecosystem Factors
Champions	
Community-Oriented Initiatives	
Volunteers Engagement	
Data Analytics and Management Skills	Organizational Skills
Geographic Information Systems Skills	
Communication Skills	
Knowledge of Legal and Policy Context	

Table 8.8: Facilitating Conditions – Performance Expectancy link

Given that *Perceived Credibility and Sustainability* was added to the original UTAUT model as a result of this research, the main reflections in relation to the existing theory are made for the following link established: Facilitating Conditions → Performance Expectancy. This link was not defined in the original UTAUT configuration. Venkatesh *et al.* (2003) reported that if Performance and Effort Expectancy are in place, then facilitating conditions are expected to be non-significant in influencing perceptions (defined in the UTAUT as Usage Intentions), but directly connected with adoption. However, as already mentioned above, "if effort expectancy is not present in the model, then one would expect Facilitating Conditions to become predictive of intention" (p.453). Therefore, considering that this research did not include a specific Effort Expectancy cluster, a link between Facilitating Conditions and Performance Expectancy is consistently positioned within the extant adoption theories.

Output Quality Links: three factors from the cross-case findings were positioned within the Output Quality construct (Venkatesh and Davis, 2000). These are: *Service Characteristics*, *Perceived User-Friendly*, and *Perceived User Satisfaction*. In terms of links within the cross-case adoption model (see Figure 7.9), the first was found to be linked with the third, and for both the second and the third a link was established with *Perceived Usefulness of Open Data*. Therefore, a direct link between Output Quality and Performance Expectancy emerges from this research. Furthermore, an internal link within the Output Quality construct is identified describing how *Service Characteristics* impact the *Perceived User Satisfaction*. Concerning the first link (i.e. Output Quality → Performance Expectancy) confirmation is found from the study where this construct was taken (i.e. TAM2 – extension of the Technology Acceptance Model) (Venkatesh and Davis, 2000). In this model, the original construct was found to be directly linked with Perceived Usefulness (considered here as Performance Expectancy).

Innovation Characteristics Links: following Rogers' (2010) formulation, *Open Data Characteristics* were defined as a factor that influences adoption of open data itself. The only link emerged from the findings is the one with *Perceived Usefulness of Open Data*. Therefore, a link between Innovation Characteristics and Performance Expectancy is here established. It is noted, however, that differently from Rogers' (2010) theory, Relative Advantage (considered in (Rogers, 2010) as part of Innovation Characteristics) was integrated within the *Perceived Usefulness of Open Data*, consistently with the UTAUT's formulation (Venkatesh *et al.* 2003).

Experience and Attitude Links: the last construct defined from this exercise as influencing open data adoption is integrated from several constructs emerged from the combined models within UTAUT (Venkatesh *et al.* 2003) as well as from its extension (Venkatesh *et al.* 2012). The "matched" factor (i.e. *Individual Factors*) was found to have one single outbound link with *Perceived Business Opportunity*. Therefore, a direct link between Experience and Attitude and Performance Expectancy is consistently established. This was found to be consistent with all the integrated constructs (Venkatesh *et al.* 2012) (Ajzen, 1991) (Taylor and Todd, 1995) (Davis *et al.* 1989) (Davis *et al.* 1992) (Thompson *et al.* 1991) (Compeau and Higgins, 1995).

8.2.11 Enriched Model

At this stage, all factors, variables and links emerged from the cross-case analysis, were positioned within the selected theoretical models. As a result, five constructs and seven links were derived. These are represented in Figure 8.2, which constitutes the main contribution of this research to the academic conversation on adoption of IS.

With respect to Figure 8.2, it is noted that for those links identified across two single CCAs, a blue arrow is used to represent this connection. Instead, where a black bold link is depicted, a strong connection supported by all integrated factors is in place. The only exceptions are the links emerged from Facilitating Conditions. In particular, Table 8.7 and Table 8.8 above outline which of the integrated variables was found to be linked with Perceived Credibility and Sustainability, and which instead is directly linked with Performance Expectancy.

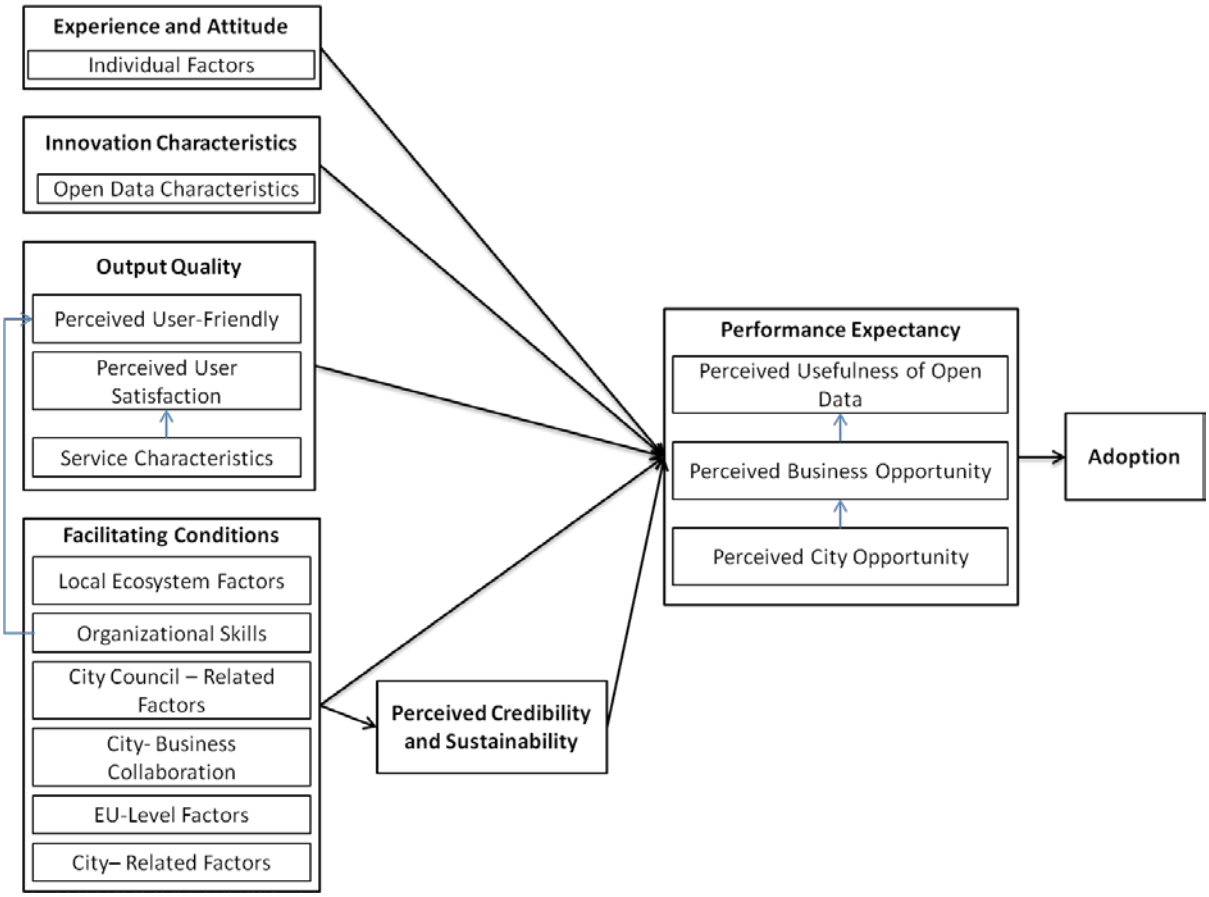

Figure 8.2: Enriched adoption model

9 CONCLUSIONS

According to global open data ranking initiatives such as the Open Data Barometer (http://opendatabarometer.org/) and the Open Data Global Index (http://index.okfn.org/), supply of open data is steadily growing across countries. However, evidence of open data adoption remains scattered and anecdotal. A related research gap was found in the existing literature. Subsequently, the following exploratory research question has been formulated for this study: what are the factors that influence adoption of open government data for commercial service innovation in a city context? This question was addressed through the inductive in-depth study of three cases. The qualitative analysis of the findings across these organizations led to the formulation of a cross-case open data adoption model. These findings were positioned within both the open data services and the IS adoption literature. Through this process, contributions to these fields emerged and are presented in this chapter in sections 9.1.1 and 9.1.2 for Open Data Services and IS Adoption theories respectively.

From the point of view of practice, these findings provide a representation of the dynamics of open data services development. Subsequently, I believe this research will help interpreting and solving the existing inhibitors that are in place for those people or entities that consider to re-use open data for commercial service innovation. As a consequence, I expect that these findings will be leveraged for stimulating and encouraging successful use of open data, i.e. increasing the value of this source as a catalyst for innovation. In addition to contributing to re-users of open data, recommendations for city councils were developed from the findings of this research. These are proposed in the section 9.2. Subsequently, this chapter provides reflections on the limitations of this study (section 9.3). Furthermore, future research avenues in this field are proposed from the contributions of this research to both the IS adoption and the open data services literatures (see section 9.4).

9.1 Contributions to Theory

The previous chapter was dedicated to the positioning of the cross-case findings within the two main bodies of literature tackled by this research: (1) the IS open data services academic debate and (2) the IS adoption literature. Contributions to these fields are summarised in this section (sub-sections 9.1.1 and 9.1.2 for open data services and IS adoption respectively).

9.1.1 Open Data Services Contribution

This research contributes to the IS open data services debate by proposing a model of factors perceived by open data services businesses as the most relevant in explaining adoption of open government data for commercial service innovation in cities. This model includes the definition of sixteen factors, sixty variables, and twenty-two links between factors, emerged across three cases to

be relevant in describing this process. Also, the findings demonstrate that seven stages can be identified within this adoption process. This study and its particular inductive approach are new to this academic conversation. The choice of inductively tackling theory resulted in the generation of thick descriptions extracted, clustered, defined, and validated for each component of the model. This approach was critical to capture the high complexity of the open data services ecosystem perceived by those re-using this data.

The literature acknowledges that value of open data is generated by adopting this source and not by merely releasing it (Kaasenbrood *et al.* 2015) (Susha *et al.* 2015) (Zuiderwijk *et al.* 2015). This research complements the extensive literature to-date which mainly focused on studying socio-technical aspects within the supply-side of this ecosystem (see Chapter 2). Both social and technical insights from the perspective of businesses that achieved sustained adoption of open data for commercial service innovation are proposed from this research. These provide a greater understanding of the adoption process of open data that are expected to be leveraged by businesses, city councils, and other stakeholders in this ecosystem. In other words, the defined factors, variables, and links can be translated into recommendations for city councils on those aspects that need to be in place to foster open data adoption in this context (see section 10.1). Considering the point of view of re-users is consistent with the strongly advocated need for a demand-pulled ecosystem emerging from both this research and the previous literature (Foulonneau *et al.* 2014) (Janssen *et al.* 2012). On the other hand, I expect that those businesses that are considering open data as the source for service innovation will benefit from the identified factors. I believe this model will enable these entities to comprehensively assess their initiatives by proposing richly-defined critical drivers for achieving sustained adoption of open data. These, within the previous literature, were either unknown or deductively derived and presented through thin descriptions. Therefore, I believe that this research will stimulate and encourage sustained adoption of open data for commercial service innovation. Subsequently, I expect that this research will contribute towards meeting those high expectations that were advocated by both academics (Lakomaa and Kallberg, 2013) (Ferro and Osella, 2012, 2013) (Zimmermann and Pucihar, 2015) and institutions (Manyika *et al.* 2013) (European Commission, 2011), and surprisingly not yet achieved in practice (Kuk and Davies, 2011) (Janssen *et al.* 2012) (Barry and Bannister, 2013) (Hjalmarsson *et al.* 2014, 2015).

This research responded to the call for inductively and qualitatively investigating this field (Zuiderwijk *et al.* 2014a) (Zuiderwijk *et al.* 2015) (Hjalmarsson *et al.* 2014, 2015). According to Zuiderwijk *et al.* (2015), “there is a need for open data specific theories that address the idiosyncratic nature of open data”, as by leveraging established adoption and diffusion theories deductively “the large part of

variance in the use of open data is not yet explained" (p.9). Although this argument is strongly related to the contribution to the IS adoption academic conversation (section 8.2), this approach to study open data adoption is reflected in the greater understanding of open data services as well.

It is noted that the findings from each single case study (see Chapters 4, 5, and 6) represent contributions specifically tailored for those particular contexts in which open data was re-used, i.e. urban planning (Case A), culture, tourism, and entertainment (Case B), and parking (Case C).

9.1.2 Contribution to Adoption of IS

This research represents the first attempt of open data adoption theory development. By positioning the cross-case findings within the UTAUT model, important contributions to this academic conversation clearly emerge. All constructs and links of the UTAUT were found to be relevant in this field of research (with the exception of Social Influence which was however acknowledged not to be a predictor of adoption in voluntary contexts likewise the one of this research). However, more specific categories within these constructs were defined. In other words, this research proposes the breakdown of those constructs initially established into more specific factors and variables. These are seen as an extension of the original UTAUT theory. The call made in (Zuiderwijk *et al.* 2015) (i.e. "we recommend that adoption theories specifically to open data are developed" (p.9)) was addressed in this research. New constructs were imported from some of the theories combined within the UTAUT formulation as matching the inductive cross-case findings of this research. This was the case of Innovation Characteristics (Rogers, 2010), Output Quality (Venkatesh and Davis, 2000), and Experience and Attitude (Venkatesh *et al.* 2012) (Fishbein and Ajzen, 1975) (Ajzen, 1991) (Venkatesh and Davies, 2000). Therefore, another contribution to the UTAUT is represented by the demonstration of the relevancy of constructs that during the integration process (i.e. the development process of the UTAUT itself) were either "dropped" from the final UTAUT formulation, or simplified and reduced to a smaller number of items. Finally, the findings from this research suggest the inclusion of an additional construct named Perceived Credibility and Sustainability. This newly introduced construct was found to be critical in mediating the impact of (some) Facilitating Conditions to Performance Expectancy, and, according to the cases studied, the cornerstone of the overall model in the context of open data services.

Therefore, by positioning the findings within those theoretical constructs selected as relevant in predicting adoption of IS, a new enriched model has been formulated (see Figure 8.2). These new constructs are defined as predictors of open government data adoption for commercial service innovation in cities. Thus, as part of future research (see section 9.4), I propose this enriched adoption

model to quantitatively investigate the actual importance of each factor, variable and link in explaining adoption of open data.

9.2 Recommendations for city councils

The investigation of the drivers that influence open data adoption from the point of view of those re-using this source enabled the formulation of a number of recommendations for city councils. These should be interpreted as concepts that are expected to generate perceptions among potential re-users that will ultimately lead to increase their adoption of open data for commercial service innovation. In turn, increased adoption will contribute towards the achievement of economic goals stated as part of open data initiatives that so far largely failed to meet these expectations. The recommendations proposed relate to two main topics: (1) addressing users' needs; and (2) ensuring credibility and sustainability of open data programs. For each of these two topics specific recommendations are outlined below.

With respect to addressing users' needs, this research suggests that councils should develop policies with the ultimate goal of:

- **Facilitating access and understandability of open data**: the findings of this research suggest that open datasets are often diverse across domains, stagnant (i.e. not frequently updated), and are not easily understood and navigated. This research recommends that one portal should be developed as a single access point to open data. This portal should be provided with clear information on the datasets' content, the associated licenses, and the standards and formats used. This information should be made available in these portals through structured metadata. Once understandability of the data is enabled, machine readability of the datasets should be ensured through the usage of APIs. Open datasets should have associated licenses that limit restrictions about their commercial re-usage (Creative Commons by Attribution). A point of contention is the prioritization of resources and efforts, in terms of what datasets should be actually published as well as on what standards and formats should be used for releasing these datasets. Acknowledging that the release of all councils' data with these characteristics is impractical (given the current level of maturity of open data programs), this research recommends that city councils should create and engage in an open data services ecosystem.
- **Creating and engaging in an open data services ecosystem**: this research suggests that city councils should build and maintain engagement initiatives with the community of

stakeholders. In particular, two types of initiatives emerged from this research: (1) those led by city councils; and (2) those led by third parties.

In relation to the former, the councils should organize events to enable direct communication with the community of actual and potential re-users. As part of these interactions, re-users should be exposed to the city-related problems and to the data owned by the councils that can be potentially released. In addition, councils should ask the community for help in implementing activities related to the supply of open data. Other forms of initiatives led by the authorities should have the goal of promoting open data through different means including: showcases of successful examples of open data re-usage; periodic informative meetings; and hack competitions. These forms of promotion were found to encourage the re-usage of open data. However, these are connected with early stages of open data adoption and not directly linked with the sustainable development of open data services.

On the other hand, councils should be part in open data initiatives organised by third parties to allow re-users and other organizations to expose their needs. This research suggests that the fact that these initiatives are led by third parties allow re-users to be more vocal in addressing their needs. Based on the experiences of the cases studied, through these forms of engagement constructive contributions towards facilitating re-usage of open data are expected to emerge.

Results of these processes are expected to lead to open data released consistent with the needs of re-users for subsequently develop valuable open data services. However, in order to foster economic activities based on open data, these services need to become sustainable commercial solutions for the businesses in question. This research suggests that open data initiatives need to be perceived by re-users as credible and sustainable. Therefore, the cross-case findings highlight the need for councils to ensure credibility and sustainability of open data. To do so, the following three recommendations are proposed from the findings of this research:

- **Implementing non-fragmented open data strategies**: businesses advocate the need for councils to structure open data strategies across departments. Most of the open data initiatives experienced by the cases were argued to be fragmented across the different functions of the city councils. From their perspective, this approach has several implications. For example, data is provided with different formats and licenses, and updated with different frequencies across departments. This situation was found to inhibit re-usage of open data. Also, the cases stressed the need of a standardized approach to enable integration and mesh-ups of datasets across departments and domains. In addition, having in place structured

strategies helps in spreading trust and perceived sustainability of the open data provision. This was found to be the cornerstone for cases to commit in re-using open data, and also for investors to perceive open data as catalyst for new economic activities. The latter is believed to be a key for fostering the overall ecosystem. To do so, this research suggests that one open data strategy should be developed across the council's functions. People with IT skills should be devoted to champion and implement these initiatives across departments, and budget and resources should be allocated specifically to these programs.

- **Being aligned with EU initiatives**: the three cases studied advocated the need for cities to be part in European open data projects. This is believed to develop EU as "one location for all initiatives". Being part of this broad ecosystem is expected to yield two general benefits. Firstly, it contributes to the development of perceptions about the commitment of city councils towards open data. This aspect was found to increase businesses and investors' expectations on the sustainability of open data which is translated in financial and resource investments in re-using open data. Secondly, having a strategy that is aligned with other cities facilitates businesses to scale their services across these urban environments. This, in turn, augments their perceived business opportunity, which ultimately positively influences adoption of open data.
- **Being open for partnerships and collaboration:** to address the challenges about ensuring long-term provision of valuable open data, the cases have shared the idea of open data published by default and/or "as a service" (e.g. supported by service level agreements). However, participants in this research (from city councils) highlighted how these solutions are not in the thinking of the local authority and that open data initiatives are not at that level of maturity yet. This research suggests that establishing partnerships with open data re-users is the driver for these entities to perceive credibility and sustainability of the availability of open data, and, as a consequence, of the services built on top of open datasets. However, optimal configurations for these relationships need to be investigated and established. In fact, establishing long-term contracts for the procurement of services in a particular context was found to discourage other businesses to re-use open data. Current solutions proposed by this research involve the creation of win-win relationships. Through these processes, open data services businesses commit to deliver outcomes that are beneficial to the city councils in exchange for a long-term commitment in the supply of open data. In particular, four types of benefits emerged to be relevant for cities that engage with re-users: (1) improving existing datasets; (2) improving efficiency and efficacy in managing the specific contexts in which open data is re-used; (3) improving its image as an open and transparent city; and (4) attract new

businesses and investments. Indeed, companies that deliver open data services add value to the datasets re-used through: data cleaning and cleansing, data normalization, and data visualization processes. In relation to point (2), open data is typically used as the enabler of the service rather than being the final service itself. New insights to the contexts of the specific open datasets are typically derived from the open data services delivered. This constitutes additional knowledge about these domains that can be leveraged to augment efficiency and efficacy in how these contexts are managed. Subsequently, having in place sustainable businesses operating on open data can lead to legitimate claims about open and transparent government by the city councils. This, in turn, is expected to improve its image that subsequently contributes to the ability of attracting new economic activities. Outsourcing the creation of open data services (to demonstrate the value behind the initiative) was found to inhibit open data adoption, and leading to the creation of insecure perceptions about potential business opportunities. As highlighted above, in order to facilitate the rise of new start-ups, councils should instead engage with the community of IT developers and entrepreneurs and explore new potential uses of its data.

In summary, releasing open data is not enough to create and foster an open data services ecosystem. This section proposed a set of recommendations for city councils based on the investigation of re-users of open data. In this way, challenges and subsequent actions to be taken were highlighted. These are expected to be crucial for businesses to use and rely on open data.

9.3 Limitations

Before proposing future research avenues in this domain, I acknowledge a number of limitations for this study. It is noted that limitations in relation to the methodology employed were highlighted and addressed in relevant sections of this thesis.

A substantial debate is ongoing in the literature about the appropriateness of conducting cross-case analysis in interpretive studies. Issues are argued to be in place in relation to the concept of generalizability (Khan and Van Wynsberghe, 2008). Section 3.15, provides reflections on how these issues have been conceptually and practically tackled in this research. In a similar vein, some might argue that the overall contribution to the adoption of IS academic conversation might be inconsistent with the ontological assumptions defined in this research, i.e. not recognizing the existence of an objective world. However, also in this way, those reflections on extending the concept of generalization in a probabilistic sense (i.e. typically positivist), towards assuming that findings from case studies (i.e. specific contexts) could be relevant to other contexts (Walsham, 1995) (Flyvbjerg, 2006) (Goetz and Lecompte, 1984) can be also applied. Furthermore, several methodologists have

further proven the ability of interpretivist multiple case studies for theory building (Andrade, 2009) (Lauckner *et al.* 2012). From another angle, the positioning of the findings within IS adoption theories, the subsequent model generated (Figure 8.2), and its ontological potential issue can be seen as strongly related with philosophical challenges of mixed method research (Cameron, 2011). Creswell (2010) identifies "combining philosophical positions, worldviews and paradigms" as one of the potential issues of these methods. In this way, Cameron (2011) reflects on the so called "paradigm war" (p.100) by distinguishing "purists" (who state that paradigms should not be mixed) and "pragmatists" (who advocate the efficient and effective use of multiple approaches). Therefore, although acknowledging the potential limitation (for some school of thoughts) of positioning inductive interpretivist qualitative multiple case study findings within a theoretical model that is proposed to be leveraged quantitatively in future research, the literature recognizes this possibility. Therefore, in order to leverage the proposed model in future studies, I recommend a deep reflection on the paradigmatic stances that will be adopted. To do this, Teddlie and Tashakkori (2010) propose six possible paradigmatic choices that allow combining different worldviews and paradigms.

Another limitation acknowledged for case study research is its poorly defined data analysis process (Yin, 2013). This limitation was firstly addressed by complementing the literature review undertaken for case study research, with a review of specific methods employed in qualitative research and specific procedures for coding the data collected. As shown in section 3.12, an accurate selection and reflection on multiple detailed qualitative data analysis methods has been made prior to the design and implementation of this actual process.

In multiple case studies, the fact that a single researcher undertakes data collection and analysis is seen as a further limitation (Stake, 2006). To cope with this potential issue, it was of fundamental importance for ensuring validity and reliability of the final findings to undertake validation efforts at both the single and the cross-case analysis stages. As carefully explained in this thesis, these additional interactions (among other outcomes) were leveraged to re-visit my interpretations of the data collected, thus ensuring that correct and consistent conclusions were drawn. Ultimately, potential pitfalls in conducting semi-structured interviews were outlined and addressed (see section 3.10.5).

A second class of limitations for this research reflects the nature of the cases that were selected. Stake's (2006) technique for selecting cases was strictly followed. However, I acknowledge that limitations might exist from the fact that the three cases operate in the Irish context. In fact, Case A operates in Ireland and US, Case B in Ireland and Germany, and Case C in Ireland and Spain. In this way, although diversity of locations is partially in place, I acknowledge a potential (minor) limitation.

Finally, further limitations for this study were highlighted during the final validation across cases interactions. The specific scope of this research on adoption of open data for service innovation was argued to contradict other aspects of open data (beyond service innovation). This was mainly the case of two different assertions. Firstly, when asserting that open data should be released in machine readable format for being re-used, one participant argued that in this case specific IT skills would be needed to access, navigate, and understand data (as also asserted in this research). Following the subject's reasoning, consumers of open data that seek information on various aspects of the council's activities (i.e. typical open data usage as a result of open government and transparency initiatives) might be precluded to accessing and consuming this information. The second argument was provided in relation to the advocated need of IT skilled people within councils that are dedicated to open data. In a similar vein, one of the participants argued that the consequence of involving "only IT people" in these initiatives results in the publication of data useful for IT developers and organizations only. In relation to this, S8 stated: "you don't talk about the epistemic community; who is involved in open data in the cities? Who are the champions? They are all with an IT background and the leaders are always the CIO and CTO [as also advocated from all the cases studied]. They are looking at it from an IT perspective. So data are objects and they don't think substantially what's in the datasets". This quote can be interpreted, similarly to the aspect about machine readability of the data, as negatively impacting goals of open data initiatives beyond commercial service innovation. Therefore, in summary, I acknowledge that focusing on open data services might lead to outcomes that constitute potential trade-offs for the achievement of objectives from open data programs such as promoting transparency of governmental agencies.

9.4 Future Research

When reflecting upon future research enabled by this study, the first opportunity is about the possibility of leveraging the extended adoption model (see Figure 8.2). This model includes a number of constructs that were inductively established at a great level of granularity (thick descriptions were generated for each factor, variable, and link in this model) as predicting adoption of open government data for commercial service innovation in cities. As part of future research, these findings entail the opportunity to quantitatively investigate the actual importance of each factor, variable and link in explaining adoption of open data across a wider sample. From the thick descriptions achieved for each factor, variable, and link (see Chapter 7), new items for investigating each construct can be defined and customized for the specific contexts in which open data adoption will be studied. By using this inductively derived model, it is reasonable to expect that the variance explained in the adoption process will significantly increase if compared to the 45% emerged from the previous usage of the UTAUT in the open data domain (Zuiderwijk *et al.* 2015). I recommend these research efforts be

conducted at different stages, thus giving a longitudinal understanding of each component of the model. I believe it would be beneficial to understand where (and how much) each factor has an impact across the different stages of adoption. For example, it is reasonable to expect the impact of *Local Ecosystem Factors,* such as *Hack Competitions,* to be greater in the early stages of adoption, whereas the need of establishing relationships with the councils might emerge during later stages. In addition, the investigation of the potential role of moderating variables, such as gender and age (Venkatesh *et al.* 2003), is proposed as a further scope for these quantitative studies.

The findings from this research suggest that future research is needed in relation to several factors that remain unexplored in the extant literature. These are expected to be critical for fostering the overall open data services ecosystem. Among others (see section 8.1), perceiving the open data ecosystem as credible and sustainable emerged as being of particular importance towards fostering its adoption. According to the findings across cases, this means perceiving commitment and credibility of the various councils in terms of their ability of providing open data (consistent with the characteristics outlined) in the future. The cases studied suggest that a major contribution towards enabling these perceptions is given by three general factors: (1) establishing relationships with the councils; (2) the need for cities to perceive opportunities from both the open data services developed and the fact that businesses are re-using their data; and (3) having in place challenges to the councils; these are perceived as enabling the motivations and those cultural changes that are believed to be critical in fostering this ecosystem. It is noted that enabling these perceptions was found to be crucial not only for the businesses in question, but also for potential investors. These were found to be reluctant in investing financial resources (all the cases shared the idea that investments from both private and public bodies are needed to trigger this ecosystem) for the general lack of trust these bodies have towards public entities in general. All the cases developed these perceptions by establishing various forms of relationships with the councils from which open data is re-used. However, evidence from both the cases and the validation interactions suggest that an optimal configuration for these collaborations has not been found yet. For example, while for Case A establishing contracts with the authorities was found to be the solution, this specific mechanism was found to be a big barrier from the analysis of Case C. As a potential solution to this issue, the possibility of establishing short-term service level agreements for the delivery of open data was tackled during the final validation stage. In this way, councils' representatives clearly highlighted how the overall ecosystem "is not at that level of maturity yet". So far, establishing *Win-Win Relationships* was found as the main option for the cases. Therefore, I propose that future research is needed on how to establish these collaborations. According to the findings of this research, solving these issues would strongly push businesses, investors, and public bodies in recognizing sustainable business

opportunities, i.e. in motivating the investments of financial and other resources, i.e. in fostering open data adoption. In this way, leveraging recent findings in the broader literature on IS Open Innovation is believed to be a viable research option. In particular, collaboration models for inside-out open innovation processes (Chesbrough, 2011) in IS contexts are believed to be a good starting point for exploring new solutions in the open data services domain.

Future research is also recommended in relation to the investigation of additional cases including those for which open data is "only" a partial ingredient of the services offered. It is noted that one of the consequences of choosing "extreme" cases (Pettigrew, 1988) (i.e. where open data is a main requirement for the final service) resulted in the identification of SMEs and start-ups as the candidates for this study. I also believe that big companies and multinationals re-use open data, although its resulting role in the final services is usually marginal. Therefore, including these cases in the overall conversation might help in understanding if different processes and activities are in place for these businesses (e.g. if different sizes of organizations underlie different processes and principles), and, if so, if and how these can be useful in predicting adoption of open data. While recommending this future research, I acknowledge potential difficulties in conducting the analysis and in scoping the research itself. For the cases studied in this research, the unit of analysis was identified in the organizations themselves (as their purpose is to design and deliver open data services). In the instance of a big corporation, defining the unit of analysis for studying adoption of open data might be a difficult and misleading process.

Finally, I propose future research to be conducted to explore the potential trade-offs for open data initiatives between the different scopes that these programs typically have. Initial insights emerged from this research in terms of recommendations, such as the need of machine readable data and of IT people leading these strategies within councils, which were found to be relevant for re-using open data in new services. Notwithstanding this, these factors also emerged as having a negative impact on general principles of transparency and equal access to this data for the overall population, not only developers and IT skilled people. Open data programs are implemented for a variety of reasons (e.g. enabling: service innovation, government transparency, provision for data for more effective research etc.), however frictions seem to be in place between these different streams. Therefore, I propose future research focusing on exploring new solutions for fostering alignment of objectives within unique, non-fragmented, open data strategies by city councils.

In summary, in order to progress towards establishing effective open data services ecosystems, i.e. capturing value from widely implemented and currently scattered open data initiatives, several future

research avenues are recommended from the findings of this research. These were proposed for IS researchers that aim at contributing to this field.

REFERENCES

Adner, R. and Kapoor, R., 2010. Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic management journal*, *31*(3), pp.306-333.

Agarwal, R. and Prasad, J., 1997. The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision sciences*, *28*(3), pp.557-582.

Ajzen, I. and Fishbein, M., 1980. Understanding attitudes and predicting social behaviour. *Prentice-Hall, Englewood Cliffs, NJ*

Ajzen, I., 1991. The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), pp.179-211.

Allport, G.W., 1962. The general and the unique in psychological science. *Journal of personality*, *30*(3), pp.405-422.

Alter, S., 2002. The work system method for understanding information systems and information systems research. *Communications of the Association for Information Systems*, *9*(1), p.6.

Alter, S., 2010. Viewing systems as services: a fresh approach in the IS field. *Communications of the association for information systems*, *26*(1), p.11.

Alvesson, M., 2003. Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research. *Academy of management review*, *28*(1), pp.13-33.

Amichai-Hamburger, Y., McKenna, K.Y. and Tal, S.A., 2008. E-empowerment: Empowerment by the Internet. *Computers in Human Behavior*, *24*(5), pp.1776-1789.

Anaf, S., Drummond, C. and Sheppard, L.A., 2007. Combining case study research and systems theory as a heuristic model. *Qualitative Health Research*, *17*(10), pp.1309-1315.

Andrade, A.D., 2009. Interpretive research aiming at theory building: Adopting and adapting the case study design. *The Qualitative Report*, *14*(1), p.42.

Aoyama, M. and Kojima, H., 2013, June. LISA: linked services architecture based on the linked data and service broker. In *2013 IEEE 20th International Conference on Web Services* (pp. 617-618). IEEE.

Astley, W.G. and Van de Ven, A.H., 1983. Central perspectives and debates in organization theory. *Administrative science quarterly*, pp.245-273.

Ayele, W., Juell-Skielse, G.T., Hjalmarsson, A. and Johanneson, P., 2015. Evaluating Open Data Innovation: A Measurement Model for Digital Innovation Contests. *Proceedings of 19th Pacific Asia Conference on Information Systems, Singapore, 2015.*

Ayres, L., Kavanaugh, K. and Knafl, K.A., 2003. Within-case and across-case approaches to qualitative data analysis. *Qualitative health research*,*13*(6), pp.871-883.

Babüroglu, O.N. and Ravn, I., 1992. Normative action research. *Organization Studies*, *13*(1), pp.019-34.

Baccarini, D., Salm, G. and Love, P.E., 2004. Management of risks in information technology projects. *Industrial Management & Data Systems*,*104*(4), pp.286-295.

Barry, E. and Bannister, F., 2014. Barriers to open data release: A view from the top. *Information Polity*, *19*(1, 2), pp.129-152.

Baskarada, S., 2014. Qualitative case study guidelines. *Baškarada, S.(2014). Qualitative case studies guidelines. The Qualitative Report*, *19*(40), pp.1-25.

Baskerville, R. and Wood-Harper, A.T., 1998. Diversity in information systems action research methods. *European Journal of information systems*, *7*(2), pp.90-107.

Basit, T., 2003. Manual or electronic? The role of coding in qualitative data analysis. *Educational research*, *45*(2), pp.143-154.

Baskarada, S., 2014. Qualitative case study guidelines. Baškarada, S.(2014). Qualitative case studies guidelines. The Qualitative Report, 19(40), pp.1-25.

Becker, H.S., 1990. Generalizing from case studies. *Qualitative inquiry in education: The continuing debate*, pp.233-242.

Bedini, I., Farazi, F., Leoni, D., Pane, J., Tankoyeu, I. and Leucci, S., 2014. Open government data: Fostering innovation. *JeDEM-eJournal of eDemocracy and Open Government*, *6*(1), pp.69-79.

Bellini, P. and Nesi, P., 2013. A Linked Open Data service for performing arts. In *Information Technologies for Performing Arts, Media Access, and Entertainment* (pp. 13-25). Springer Berlin Heidelberg.

Benbasat, I., Goldstein, D.K. and Mead, M., 1987. The case research strategy in studies of information systems. *MIS quarterly*, pp.369-386.

Berg, B.L., 2001. *Qualitative research methods for the social sciences*. Boston, MA: Pearson.

Berger, P.L. and Luckmann, T., 1967. *The Social Construction of Reality: A Treatise in the Sociology of Knowledmann*. Anchor books.

Berners-Lee,T., 2006. Linked Data - Design Issues. http://www.w3.org/DesignIssues/LinkedData.html.

Bertot, J.C., Jaeger, P.T. and Grimes, J.M., 2010. Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government information quarterly*, *27*(3), pp.264-271.

Bin, G., 2013. A reasoned action perspective of user innovation: Model and empirical test. *Industrial Marketing Management*, *42*(4), pp.608-619.

Birks, D.F., Fernandez, W., Levina, N. and Nasirin, S., 2013. Grounded theory method in information systems research: its nature, diversity and opportunities. *European Journal of Information Systems*, *22*(1), pp.1-8.

Bizer, C., Heath, T. and Berners-Lee, T., 2009. Linked data-the story so far. *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, pp.205-227.

Blaikie, N., 1993. *Approaches to social enquiry: Advancing knowledge*. Polity.

Bloomfield, B.P. and Vurdubakis, T., 1994. Boundary disputes: Negotiating the boundary between the technical and the social in the development of IT systems. *Information Technology & People*, *7*(1), pp.9-24.

Bonina, C.M., 2013. New business models and the value of open data: definitions, challenges and opportunities. *NEMODE–3K Small Grants Call*.

Brekhus, W.H., Galliher, J.F. and Gubrium, J.F., 2005. The need for thin description. *Qualitative Inquiry*, *11*(6), pp.861-879.

Brown, J. D., 2000. What is Construct Validity?. *JALT Testing & Evaluation SIG Newsletter*, 4 (2) Oct 2000, pp. 8 – 12.

Burr, V., 1995. An introduction to social constructionism. London: Routlegde.

Burrel, G. and Morgan, G., 1979. Sociological paradigms and organizational analysis. *Aldershot, Gower*.

Campbell, D.T., 1975. " Degrees of freedom" and the case study.*Comparative political studies*, *8*(2), p.178.

Cameron, R., 2011. Mixed methods research: The five Ps framework. *The electronic journal of business research methods*, *9*(2), pp.96-108.

Chan, C.M., 2013, January. From open data to open innovation strategies: Creating e-services using open government data. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on* (pp. 1890-1899). IEEE.

Charmaz, K., 2006. Constructing grounded theory: A practical guide through qualitative research. *SagePublications Ltd, London*.

Chase, S.E., 2005. Narrative Inquiry: Multiple Lenses, Approaches, Voices. *Handbook of Qualitative Research*, red. NK Denzin, Y. Lincoln.

Chesbrough, H.W., 2006. *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Chesbrough, H.W., 2011. Bringing open innovation to services. *MIT Sloan Management Review*, *52*(2), p.85.

Cifuentes-Silva, F., Sifaqui, C. and Labra-Gayo, J.E., 2011, September. Towards an architecture and adoption process for linked data technologies in open government contexts: a case study for the library of congress of chile. In *Proceedings of the 7th International Conference on Semantic Systems*(pp. 79-86). ACM.

Coffey, A. and Atkinson, P., 1996. *Making sense of qualitative data: complementary research strategies*. Sage Publications, Inc.

Compeau, D.R. and Higgins, C.A., 1995. Computer self-efficacy: Development of a measure and initial test. *MIS quarterly*, pp.189-211.

Cooperrider, D. and Whitney, D.D., 2005. Appreciative inquiry: A positive revolution in change. Berrett-Koehler Publishers.

Coyne, I.T., 1997. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries?. *Journal of advanced nursing*, *26*(3), pp.623-630.

Creswell, J.W., 2007. Qualitative inquiry and research design: Choosing among five approaches*. (2nd ed., pp. 53-84). Thousand Oaks, CA: Sage Publications, Inc.*

Creswell, J.W., 2010. Mapping the developing landscape of mixed methods research. SAGE handbook of mixed methods in social & behavioral research, 2, pp.45-68.

Creswell, J.W., 2015. Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research, Enhanced Pearson eText with Loose-Leaf Version--Access Card Package. *Pearson Education, Inc.*

Currie, J. L., 2013. The Role of Canadian Municipal Open Data Initiatives: a Multi-City Evaluation. *Department of Geography at Queens University, Kingston, Canada*.

Cyganiak, R., Maali, F. and Peristeras, V., 2010, September. Self-service linked government data with dcat and gridworks. In *Proceedings of the 6th International Conference on Semantic Systems* (p. 37). ACM.

Czarniawska, B., 2004. *Narratives in social science research*. Sage.

Darke, P., Shanks, G. and Broadbent, M., 1998. Successfully completing case study research: combining rigour, relevance and pragmatism. *Information systems journal*, *8*(4), pp.273-289.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pp.319-340.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R., 1989. User acceptance of computer technology: a comparison of two theoretical models. *Management science*, *35*(8), pp.982-1003.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R., 1992. Extrinsic and intrinsic motivation to use computers in the workplace1. *Journal of applied social psychology*, *22*(14), pp.1111-1132.

Davison, R., Martinsons, M.G. and Kock, N., 2004. Principles of canonical action research. *Information systems journal*, *14*(1), pp.65-86.

Denzin, N.K., 1989. Interpretive biography (Vol. 17). Sage.

Devers, K.J. and Frankel, R.M., 2000. Study design in qualitative research--2: Sampling and data collection strategies. *Education for health*, *13*(2), p.263.

Dey, I., 1993. Creating categories. *Qualitative data analysis*, pp.94-112.

Dey, I., 1999. *Grounding grounded theory: Guidelines for qualitative inquiry*. Academic Press.

Eisenhardt, K.M., 1989. Building theories from case study research. *Academy of management review*, *14*(4), pp.532-550.

Eisenhardt, K.M. and Graebner, M.E., 2007. Theory building from cases: Opportunities and challenges. *Academy of management journal*, *50*(1), p.25.

Felicetti, A., d'Andrea, A. and Niccolucci, F., 2012. The Open Data Semantics and the (re) use of Open Information in Cultural Heritage. ECLAP. In *International Conference on Information Technologies for performing Arts, Media Access and Entertainment*.

Feng, Y., Veeramani, A. and Kanagasabai, R., 2012, December. Enabling On-Demand Mashups of Open Data with Semantic Services. In *Parallel and Distributed Systems (ICPADS), 2012 IEEE 18th International Conference on*(pp. 755-759). IEEE.

Ferro, E. and Osella, M., 2012, June. Business models for PSI re-use: A multidimensional framework. In *Using Open data: Policy Modeling, Citizen Empowerment, Data Journalism Workshop, European Comission, Brussels*.

Ferro, E. and Osella, M., 2013, April. Eight business model archetypes for PSI Re-Use. In *Open Data on the Web Workshop, Google Campus, Shoreditch, London*.

Fink, A., 2013. *Conducting Research Literature Reviews: From the Internet to Paper*. Sage Publications.

Fisbein, M. and Ajzen, I., 1975. Belief, attitude, intention and behavior: an introduction to theory and research. *Massachusetts, Addison-Wiley Publishing Company*.

Flowers, P., 2009. Research philosophies–importance and relevance.*Economic Record*, *1*(3), pp.1-5.

Flyvbjerg, B., 2006. Five misunderstandings about case-study research. *Qualitative inquiry, 12(2),* pp.219-245.

Fontana, A. and Frey, J.H., 2000. The interview: From structured questions to negotiated text. *Handbook of qualitative research*, *2*(6), pp.645-672.

Foreman, P.B., 1948. The theory of case studies. *Social Forces*, pp.408-419.

Foulonneau, M., Martin, S. and Turki, S., 2014. How open data are turned into services?. In *Exploring services science* (pp. 31-39). Springer International Publishing.

Fox, M.S., 2013. City data: Big, open and linked. *Municipal Interfaces*, pp.19-25.

Frosterus, M., Tuominen, J., Wahlroos, M. and Hyvönen, E., 2013. The Finnish Law as a Linked Data Service. In *The Semantic Web: ESWC 2013 Satellite Events* (pp. 289-290). Springer Berlin Heidelberg.

Fulk, J., 1993. Social construction of communication technology. *Academy of Management journal*, *36*(5), pp.921-950.

Gautreau, P., Hasenack, H., Lerch, L., Merlinsky, G., Noucher, M. and Severo, M., 2012. A Comparison of the Open Environmental Data Diffusion in Argentina. *Bolivia, and Brazil: some preliminary results from the Baguala Project for methodological discussion, UFRGS, Brazil*.

Geertz, C., 1973. *The interpretation of cultures: Selected essays* (Vol. 5019). Basic books.

Geiger, C.P. and von Lucke, J., 2011. Open Government Data1. In*Conference for E-Democracy and Open Government* (p. 183).

Gibbons, M.T., 1987. Introduction: The politics of interpretation. *Interpreting politics*, pp.1-31.

Given, L.M. ed., 2008. *The Sage encyclopedia of qualitative research methods*. Sage Publications.

Glaser, B.G. and Strauss, A.L., 1967. The discovery of grounded theory: strategies for qualitative research. *Aldin, New York.*

Gläser, J. and Laudel, G., 2004. *The Sociological description of non-social conditions of research*. REPP Discussion Paper 04/2. Canberra: The Australian National University.

Goel, V., and Waechter, R. L., 2004. Inductive and Deductive Reasoning. *Encyclopedia of Human Behavior, 2nd Edition, Amsterdam*, Elsevier.

Goetz, J.P. and LeCompte, M.D., 1984. *Ethnography and qualitative design in educational research* (Vol. 19). Orlando, FL: Academic Press.

Goffman, E., 1959. The presentation of self in everyday life. London: Penguin.

Groen, M., Meys, W. and Veenstra, M., 2013, September. Creating smart information services for tourists by means of dynamic open data. In*Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication* (pp. 1329-1330). ACM.

Guba, E.G. and Lincoln, Y.S., 1994. Competing paradigms in qualitative research. *Handbook of qualitative research*, *2*(163-194), p.105.

Guo, W. and Kraines, S.B., 2010, October. Enriching city entities in the EKOSS failure cases knowledge base with Linked Open Data. In *Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on* (pp. 404-409). IEEE.

Halb, W., Stocker, A., Mayer, H., Mülner, H. and Ademi, I., 2010, September. Towards a commercial adoption of linked open data for online content providers. In *Proceedings of the 6th International Conference on Semantic Systems* (p. 16). ACM.

Hannemann, J. and Kett, J., 2010, August. Linked data for libraries. In *Proc of the world library and information congress of the Int'l Federation of Library Associations and Institutions (IFLA)*.

Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W. and Borriello, G., 2010, December. Open data kit: tools to build information services for developing regions. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development* (p. 18). ACM.

Hartwick, J. and Barki, H., 1994. Explaining the role of user participation in information system use. *Management science*, *40*(4), pp.440-465.

Harvey, L.J. and Myers, M.D., 1995. Scholarship and practice: the contribution of ethnographic research methods to bridging the gap.*Information technology & people*, *8*(3), pp.13-27.

Harvey, L., 1997. A discourse on ethnography. In *Information systems and qualitative research* (pp. 207-224). Springer US.

Haslhofer, B. and Isaac, A., 2011, September. data. europeana. eu: The europeana linked open data pilot. In *International Conference on Dublin Core and Metadata Applications* (pp. 94-104).

Hatch, M.J. and Cunliffe, A.L., 2013. *Organization theory: modern, symbolic and postmodern perspectives*. Oxford university press.

Hausenblas, M., 2009. Exploiting linked data to build web applications. *IEEE Internet Computing*, *13*(4), p.68.

Hermanns, H., 2004. 5.3 Interviewing as an Activity. *A Companion to*, p.209.

Herschel, M. and Manolescu, I., 2012, October. DataBridges: data integration for digital cities. In *Proceedings of the 2012 ACM workshop on City data management workshop* (pp. 1-4). ACM.

Hertog, P.D., 2000. Knowledge-intensive business services as co-producers of innovation. *International Journal of Innovation Management*, *4*(04), pp.491-528.

Hesse-Biber, S.N. and Leavy, P., 2010. *The practice of qualitative research*. Sage.

Hevner, A., March, S.T., Park, J. and Ram, S., 2004. Design science in information systems research. *MIS quarterly*, *28*(1), pp.75-105.

Hielkema, H., and Hongisto, P., 2013. Developing the Helsinki smart city: the role of competitions for open data applications. *Journal of the Knowledge Economy*, *4*(2), 190-204.

Hjalmarsson, A., Johannesson, P., Jüll-Skielse, G., and Rudmark, D., 2014. Beyond innovation contests: A framework of barriers to open innovation of digital services. In *Twenty-Second European Conference on Information Systems (ECIS), Tel Aviv, Israel, 2014.*

Hjalmarsson, A., Juell-Skielse, G., Ayele, W., Rudmark, D., Johannesson, P., 2015. From Contest to Market Entry: a Longitudinal Survey of Innovation Barriers Constraining Open Data Service Development. In *Twenty-Third European Conference on Information Systems (ECIS), Münster, Germany, 2015*.

Holstein, J.A. and Gubrium, J.F., 1995. *The active interview* (Vol. 37). Sage

Hu, P.J., Chau, P.Y., Sheng, O.R.L. and Tam, K.Y., 1999. Examining the technology acceptance model using physician acceptance of telemedicine technology. *Journal of management information systems*, *16*(2), pp.91-112.

Huff, A.S., 2008. Designing research for publication. *Sage*.

Iijima, C., Morita, T., Enomoto, Y. and Yamaguchi, T., 2011. A mobility service based on Japanese linked data. In *Knowledge-Based and Intelligent Information and Engineering Systems* (pp. 246-255). Springer Berlin Heidelberg.

Immonen, A., Palviainen, M. and Ovaska, E., 2014. Towards open data based business: Survey on usage of open data in digital services. *International Journal of Research in Business and Technology*, *4*(1), pp.286-295.

Isaac, A. and Haslhofer, B., 2013. Europeana linked open data–data. europeana. eu. *Semantic Web*, *4*(3), pp.291-297.

Iivari, J. and Venable, J., 2009. Action research and design science research–seemingly similar but decisively dissimilar. In *European Conference on Information Systems* (Vol. 17, pp. 1-13).

Janssen, K., 2011. The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, *28*(4), pp.446-456.

Janssen, M., Charalabidis, Y. and Zuiderwijk, A., 2012. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, *29*(4), pp.258-268.

Juell-Skielse, G., Hjalmarsson, A., Juell-Skielse, E., Johannesson, P. and Rudmark, D., 2014. Contests as innovation intermediaries in open data markets. *Information Polity*, *19*(3, 4), pp.247-262.

Kaasenbrood, M., Zuiderwijk, A., Janssen, M., de Jong, M. and Bharosa, N., 2015. Exploring the factors influencing the adoption of open government data by private organisations. International Journal of Public Administration in the Digital Age (IJPADA), 2(2), pp.75-92.

Kang, W. and Son, S.H., 2008. The design of an open data service architecture for cyber-physical systems. *ACM SIGBED Review*, *5*(1), p.3.

Kaplan, B. and Maxwell, J.A., 2005. Qualitative research methods for evaluating computer information systems. In *Evaluating the organizational impact of healthcare information systems* (pp. 30-55). Springer New York.

Khan, S. and VanWynsberghe, R., 2008, January. Cultivating the under-mined: Cross-case analysis as knowledge mobilization. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (Vol. 9, No. 1).

Kim, S.S. and Malhotra, N.K., 2005. A longitudinal model of continued IS use: An integrative view of four mechanisms underlying postadoption phenomena. *Management science*, *51*(5), pp.741-755.

King, P., 2004. Relativism, subjectivity and the self: a critique of social constructionism in Jacobs. *Kemeny and Manzi (eds) op cit*.

Kitchenham, B.A. and Charters, S., 2007. Procedures for Performing Systematic Literature Reviews in Software Engineering. *Keele University & Durham University, UK*.

Klein, H.K. and Myers, M.D., 1999. A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS quarterly*, pp.67-93.

Kuk, G., and Davies, T., 2011. The Roles of Agency and Artifacts in Assembling Open Data Complementarities. *32nd International Conference on Information Systems*, ICIS, Shanghai.

Kuzel, A.J., 1999. Sampling in qualitative inquiry. " In BF Crabtrree and WL Miller (Eds.) Doing Qualitative Research (second edition). Thousand Oaks, CA: Sage Publlications (pp. 33-45).

Kvale, S., 2008. Doing interviews. Sage.

Kwon, T.H. and Zmud, R.W., 1987, April. Unifying the fragmented models of information systems implementation. In *Critical issues in information systems research* (pp. 227-251). John Wiley & Sons, Inc.

Lakomaa, E. and Kallberg, J., 2013. Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. *Access, IEEE*, *1*, pp.558-563.

Lamb, R. and Kling, R., 2003. Reconceptualizing users as social actors in information systems research. *MIS quarterly*, pp.197-236.

Latif, A., Saeed, A.U., Höfler, P., Stocker, A. and Wagner, C., 2009, September. The Linked Data Value Chain: A Lightweight Model for Business Engineers. In *I-SEMANTICS* (pp. 568-575).

Lauckner, H., Paterson, M. and Krupa, T., 2012. Using constructivist case study methodology to understand community development processes: Proposed methodological questions to guide the research process. *The Qualitative Report*, *17*(13), p.1.

Law, J., 1987. Technology and heterogeneous engineering: The case of Portuguese expansion. *The social construction of technological systems: New directions in the sociology and history of technology*, *1*, pp.1-134.

Lee, T.W., Mitchell, T.R. and Sablynski, C.J., 1999. Qualitative research in organizational and vocational psychology, 1979–1999. *Journal of vocational behavior*, *55*(2), pp.161-187.

Leeds-Hurwitz, W., 2009. Social construction of reality. *Encyclopedia of communication theory*, *2*, pp.891-894.

Lewis, I.M., 2004. *Social and Cultural Anthropology in Perspective*. Transaction Publishers.

Lincoln, Y.S. and Guba, E.G., 1985. *Naturalistic inquiry* (Vol. 75). Sage.

Lindman, J., Rossi, M. and Tuunainen, V.K., 2013a, January. Introduction to Open Data Services Minitrack. In *2013 46th Hawaii International Conference on System Sciences* (p. 1238). IEEE.

Lindman, J., Rossi, M. and Tuunainen, V.K., 2013b, January. Open data services: Research agenda. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on* (pp. 1239-1246). IEEE.

Lindman, J. and Nyman, L., 2014. The businesses of open data and open source: Some key similarities and differences. *Technology Innovation Management Review*, *4*(1), p.12.

Lindman, J., Kinnari, T. and Rossi, M., 2014, January. Industrial open data: case studies of early open data entrepreneurs. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on* (pp. 739-748). IEEE.

Lockyer, S., 2004. Coding qualitative data. *The sage encyclopedia of social science research methods*, *1*(1), pp.137-138.

Lorey, J., 2013. Storing and Provisioning Linked Data as a Service. In *The Semantic Web: Semantics and Big Data* (pp. 666-670). Springer Berlin Heidelberg.

Mainka, A., Hartmann, S., Meschede, C. and Stock, W.G., 2015. Mobile application services based upon open urban government data. *iConference 2015 Proceedings*.

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S. and Doshi, E.A., 2013. Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute. *McKinsey Center for Government and McKinsey Business Technology Office. Available at: http:// www. mckinsey. com/ insights/ business_ technology/ open_ data_ unlocking_ innovation_ and_ performance_ with_ liquid_ information*.

Martin, P.Y. and Turner, B.A., 1986. Grounded theory and organizational research. *The Journal of Applied Behavioral Science*, *22*(2), pp.141-157.

Matheus, R., Ribeiro, M.M. and Vaz, J.C., 2012, October. New perspectives for electronic government in Brazil: the adoption of open government data in national and subnational governments of Brazil. In *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*(pp. 22-29). ACM.

Mayring, P., 2000. Qualitative content analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [On-line Journal], *1*(2), Art. 20. Available at: http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm

Miles, M.B. and Huberman, A.M., 1994. *Qualitative data analysis: An expanded sourcebook*. Sage.

Mir, R. and Watson, A., 2001. Critical realism and constructivism in strategy research: toward a synthesis. *Strategic management journal*, *22*(12), pp.1169-1173.

Mitev, N., 2000, December. Toward social constructivist understandings of IS success and failure: introducing a new computerized reservation system. In *Proceedings of the twenty first international conference on Information systems* (pp. 84-93). Association for Information Systems.

Molenaar, P.C., 2004. A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever.*Measurement*, *2*(4), pp.201-218.

Moore, G.C. and Benbasat, I., 1991. Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information systems research*, *2*(3), pp.192-222.

Morgan, G., 1983. *Beyond method: Strategies for social research*. Sage.

Moustakas, C., 1994. *Phenomenological research methods*. Sage Publications.

Myers, M.D., 1997. Qualitative research in information systems. *Management Information Systems Quarterly*, *21*(2), pp.241-242.

Myers, M., 1999. Investigating information systems with ethnographic research. *Communications of the AIS*, *2*(4es), p.1.

Myers, M.D. and Newman, M., 2007. The qualitative interview in IS research: Examining the craft. *Information and organization*, *17*(1), pp.2-26.

Nakicenovic, N. and Grübler, A. eds., 2013. *Diffusion of technologies and social behavior*. Springer Science & Business Media.

Nandhakumar, J., Rossi, M. and Talvinen, J., 2005. The dynamics of contextual forces of ERP implementation. *The Journal of Strategic Information Systems*, *14*(2), pp.221-242.

Neuman, W.L., 2005. *Social research methods: Quantitative and qualitative approaches* (Vol. 13, pp. 26-28). Boston, MA: Allyn and bacon.

Nicholson, B. and Sahay, S., 2004. Embedded knowledge and offshore software development. *Information and organization*, *14*(4), pp.329-365.

Oh, J., 2013. Literature Review—Open Data in New Zealand. *University of Auckland*.

Okoli, C. and Schabram, K., 2010. A guide to conducting a systematic literature review of information systems research. *Sprouts Work. Pap. Inf. Syst*, *10*, p.26.

Orlikowski, W.J. and Baroudi, J.J., 1991. Studying information technology in organizations: Research approaches and assumptions. *Information systems research*, *2*(1), pp.1-28.

Orlikowski, W.J. and Robey, D., 1991. Information technology and the structuring of organizations. *Information systems research*, *2*(2), pp.143-169.

Orlikowski, W.J., 1991. Integrated information environment or matrix of control? The contradictory implications of information technology.*Accounting, Management and Information Technologies*, *1*(1), pp.9-42.

Orlikowski, W.J., 1993. CASE tools as organizational change: Investigating incremental and radical changes in systems development. *MIS quarterly*, pp.309-340.

Osterwalder, A. and Pigneur, Y., 2013. *Business model generation: a handbook for visionaries, game changers, and challengers*. John Wiley & Sons.

Palka, W., 2013. The Diffusion of IT-enabled Open Government Data. In *10th Workshop on Information Systems and Services Sciences* (p. 158).

Palka, W., Jurisch, M., Leicht, M., Wolf, P. and Krcmar, H., 2013, June. Classification Schemes for Open Government Data Provision. In *European Conference on e-Government* (p. 608). Academic Conferences International Limited.

Palys, T., 2008. Purposive sampling. *The Sage encyclopedia of qualitative research methods*, *2*, pp.697-698.

Patton, M.Q., 1990. *Qualitative evaluation and research methods* . SAGE Publications, inc.

Patton, E. and Appelbaum, S.H., 2003. The case for case studies in management research. *Management Research News*, *26*(5), pp.60-71.

Patton, M.Q., 2005. *Qualitative research*. John Wiley & Sons, Ltd.

Peattie, L., 2001. Theorizing planning: Some comments on Flyvbjerg's Rationality and Power. *International Planning Studies*, *6*(3), pp.257-262.

Pettigrew, A.M. ed., 1988. *The management of strategic change*. B. Blackwell.

Piatier, A., 1984. *Barriers to innovation*. London; Dover, NH: F. Pinter.

Pinnegar, S. and Daynes, J.G., 2007. Locating narrative inquiry historically. *Handbook of narrative inquiry: Mapping a methodology*, pp.3-34.

Polkinghorne, D.E., 1995. Narrative configuration in qualitative analysis. *International journal of qualitative studies in education*, *8*(1), pp.5-23.

Polkinghorne, D.E., 2005. Language and meaning: Data collection in qualitative research. Journal of counseling psychology, 52(2), p.137.

Ponterotto, J.G., 2006. Brief note on the origins, Evolution, and meaning of the qualitative research concept Thick Description. *The Qualitative Report*,*11*(3), pp.538-549.

Porter, M.E., 2011. Competitive advantage of nations: creating and sustaining superior performance. Simon and Schuster.

Quaddus, M. and Xu, J., 2005. Adoption and diffusion of knowledge management systems: field studies of factors and variables. *Knowledge-Based Systems*, *18*(2), pp.107-115.

Rapoport, R.N., 1970. Three dilemmas in action research with special reference to the Tavistock experience. *Human relations*, *23*(6), pp.499-513.

Rittenbruch, M., Foth, M., Robinson, R. and Filonik, D., 2012. Program your city: designing an urban integrated open data API. In *Proceedings of Cumulus Helsinki 2012 Conference: Open Helsinki–Embedding Design in Life* (pp. 24-28). Cumulus International Association of Universities and Colleges of Art, Design and Media.

Rogers, E.M., 2010. *Diffusion of innovations*. Simon and Schuster.

Rosen, M., 1991. Coming To Terms with the Field: Understanding And Doing Organizational Ethnography*. *Journal of Management Studies*, *28*(1), pp.1-24.

Rossi, B., Russo, B. and Succi, G., 2007. Open source software and open data standards as a form of technology adoption: a case study. In *Open Source Development, Adoption and Innovation* (pp. 325-330). Springer US.

Saga, V.L. and Zmud, R.W., 1993, October. The nature and determinants of IT acceptance, routinization, and infusion. In *Proceedings of the IFIP TC8 working conference on diffusion, transfer and implementation of information technology* (pp. 67-86). Elsevier Science Inc..

Sandelowski, M., 1995. Sample size in qualitative research. *Research in nursing & health*, *18*(2), pp.179-183.

Sarantakos, S., 2012. *Social research*. Palgrave Macmillan.

Saunders, M.N., 2011. *Research methods for business students, 5/e*. Pearson Education India.

Savelyev, A., Xu, S., Janowicz, K., Mülligann, C., Thatcher, J. and Luo, W., 2011, November. Volunteered geographic services: Developing a linked data driven location-based service. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Spatial Semantics and Ontologies* (pp. 25-31). ACM.

Schultze, U. and Avital, M., 2011. Designing interviews to generate rich data for information systems research. *Information and Organization*, *21*(1), pp.1-16.

Schwandt, T.A., 2001. Dictionary of qualitative inquiry. Sage Publications.

Shim, J.P., Shin, Y.B. and Nottingham, L., 2002. Retailer web site influence on customer shopping: exploratory study on key factors of customer satisfaction. *Journal of the Association for Information systems*, *3*(1), p.3.

Siggelkow, N., 2007. Persuasion with case studies. *Academy of management journal*, *50*(1), pp.20-24.

Smaling, A., 2003. Inductive, analogical, and communicative generalization.*International Journal of Qualitative Methods*, *2*(1), pp.52-67.

Stake, R.E., 2013. Multiple case study analysis. *Guilford Press*.

Stephenson, M., Di Lorenzo, G. and Aonghusa, P.M., 2012, March. Open Innovation Portal: A collaborative platform for open city data sharing. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on* (pp. 522-524). IEEE.

Strang, D. and Meyer, J.W., 1993. Institutional conditions for diffusion. *Theory and society*, *22*(4), pp.487-511.

Susha, I., Grönlund, Å. and Janssen, M., 2015. Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries. *Information Polity*, *20*(1), pp.19-34.

Tammisto, Y. and Lindman, J., 2012. Definition of open data services in software business. In *Software Business* (pp. 297-303). Springer Berlin Heidelberg.

Taylor, S. and Todd, P.A., 1995. Understanding information technology usage: A test of competing models. *Information systems research*, *6*(2), pp.144-176.

Tcholtchev, N., Farid, L., Marienfeld, F., Schieferdecker, I., Dittwald, B. and Lapi, E., 2012, October. On the interplay of open data, cloud services and network providers towards electric mobility in smart cities. In *Local Computer Networks Workshops (LCN Workshops), 2012 IEEE 37th Conference on* (pp. 860-867). IEEE.

Teddlie, C. and Tashakkori, A., 2010. Overview of contemporary issues in mixed methods research. *Handbook of mixed methods in social and behavioral research*, pp.1-41.

Tesch, R., 1990. *Qualitative research: Analysis types and software*. Routledge.

Thompson, R.L., Higgins, C.A. and Howell, J.M., 1991. Personal computing: Toward a conceptual model of utilization. *MIS quarterly*, pp.125-143.

Tinati, R., Carr, L., Halford, S. and Pope, C., 2013. Exploring the Use of# OpenData in UK Open Government Data Community. *Digital Economy 2013*.

Urquhart, C., Lehmann, H. and Myers, M.D., 2010. Putting the 'theory'back into grounded theory: guidelines for grounded theory studies in information systems. *Information systems journal*, *20*(4), pp.357-381.

Van Ark, B., Broersma, L. and den Hertog, P., 2003. Services innovation, performance and policy: a review. *Synthesis Report in the Framework of the Project on Structural Information Provision on Innovation in Services (SIID) for the Ministry of Economic Affairs of the Netherlands. Groningen: University of Groningen and DIALOGIC*.

Van de Ven, A.H., 2007. Engaged scholarship: Creating knowledge for science and practice. *Oxford University Press, Oxford*.

van Manen, M., 1990. Researching lived experience: Human science for an action sensitive pedagogy. New York, NY: state university of New York.

Venkatesh, V., 2000. Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information systems research*, *11*(4), pp.342-365.

Venkatesh, V. and Davis, F.D., 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, *46*(2), pp.186-204.

Venkatesh, V. and Morris, M.G., 2000. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS quarterly*, pp.115-139.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly*, pp.425-478.

Venkatesh, V., Thong, J.Y. and Xu, X., 2012. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, *36*(1), pp.157-178.

Viscusi, G., Castelli, M. and Batini, C., 2014. Assessing social value in open data initiatives: a framework. *Future Internet*, *6*(3), pp.498-517.

Walsham, G., 1993. *Interpreting information systems in organizations* (Vol. 19). Chichester: Wiley.

Walsham, G., 1995. Interpretive case studies in IS research: nature and method. *European Journal of information systems*, *4*(2), pp.74-81.

Wang, H.J. and Lo, J., 2016. Adoption of open government data among government agencies. *Government Information Quarterly*.

Wilde, E., 2010. Linked Data and Service Orientation. In *Service-Oriented Computing* (pp. 61-76). Springer Berlin Heidelberg.

Wolcott, P., Press, L., McHenry, W., Goodman, S. and Foster, W., 2001. A framework for assessing the global diffusion of the Internet. *Journal of the Association for Information Systems*, *2*(1), p.6.

Yin, R.K., 2013. *Case study research: Design and methods*. Sage publications.

Yu, H.Q., Dietze, S., Li, N., Pedrinaci, C., Taibi, D., Dovrolis, N., Stefanut, T., Kaldoudi, E. and Domingue, J., 2011. A Linked Data-driven & Service-oriented Architecture for Sharing Educational Resources; Linked Learning 2011. In *1st International Workshop on eLearning Approaches for the Linked Data Age, Heraklion, Greece*.

Zhang, Y.W. and Wildemuth, B.M., 2010. BM 2009,'Qualitative analysis of content'. *Applications of Social Research Methods to Questions in Information and Library*, pp.1-12.

Zimmermann, H.D. and Pucihar, A., 2015. Open Innovation, Open Data and New Business Models. *Open Data and New Business Models (September 1, 2015)*.

Zorn, T., and Campbell, N., 2006. Improving The Writing Of Literature Reviews Through A Literature Integration Exercise. *Business Communication Quarterly*, *69*(2), 172.

Zuiderwijk, A., Janssen, M. and Davis, C., 2014a. Innovation with open data: Essential elements of open data ecosystems. *Information Polity*, *19*(1, 2), pp.17-33.

Zuiderwijk, A., Helbig, N., Gil-García, J.R. and Janssen, M., 2014b. Special Issue on Innovation through Open Data: Guest Editors' Introduction. *Journal of theoretical and applied electronic commerce research*, *9*(2), pp.i-xiii.

Zuiderwijk, A., Janssen, M. and Dwivedi, Y.K., 2015. Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly*, *32*(4), pp.429-440.

APPENDICES

Appendix 1: Case Studies Documents

Appendix 1.1 Case A Documents

#	Source	Summary of content	Length	Year
D1	CEO	Company and services description; focus on how city's operations are improved by the service; consideration on Open Data and related future strategies.	79 pages	2014
D2	CEO	Open Data Service description and list of cities where it is (in 2013) delivered; strategies to target additional cities.	3 pages	2013
D3	CEO	Feedback to the city council on the hack competition experience; overview of the case's needs to sustainably enter the market and deliver the service; description of the Open Data Services offered.	6 pages	2012
D4	Web	Company's website including news and events update.	N/A	N/A
D5	Web	Overview of the Open Data Services offered with specific focus on US cities.	245 words	2013
D6	Web	Focus on Open Data and how it is being re-used by Case A.	435 words	2013
D7	Web	Reflections and insights on how Open Data can improve the Urban Planning context within cities. Subsequent benefits and opportunities are underlined.	656 words	2014
D8	Web	Open Data Services description; relationship with the councils; focus on challenges faced by the case.	1106 words	2012
D9	Web	Video Interview. Focus on Open Data benefits and how civic engagement initiatives can be implemented as part of the Open Data initiatives.	31 minutes	2013
D10	Web	Councils Open Data Portals (focus on Open datasets used)	Web Page	N/A

Appendix 1.2 Case B Documents

#	Source	Summary of content	Length	Year
D1	CEO	Application to EU H2020 funding. Overview of the company and its future plans. European focus.	20 pages	2014
D2	CEO	Open Data Services description (technical and usefulness); needs and value proposition towards the city council; local focus.	21 pages	2012
D3	CEO	Future trends in the tourism context; Plan for embedding new Open datasets into the service; Structured Open Data Gaming description; Challenges; EU Open Data as the driving solution; Plan for EU application (Horizon 2020).	6 pages	2013
D4	CEO	Overview of the Open Data Services and needs from local council and the EU-driven ecosystem.	3 pages	2012
D5	Web	Workshop follow up; Input for local council; Tourism open data opportunities; Tweets.	7 pages	2014
D6	Web	Open Data and tourism opportunities; Early thinking on the not-yet established Open Data Service.	973 words	2011
D7	Web	Open Data Services characteristics and general information about the current CEO. Local focus.	199 words	2011
D8	Web	Video interview	14 minutes	2012
D9	Web	Open Data Services characteristics and general information about the current CEO. Local focus. City Council perspective (opinion, Open Data strategy and goals)	381 words	2012
D10	Web	Company's website	N/A	N/A
D11	Web	Councils Open Data Portals (focus on Open datasets used)	Web Page	N/A

Appendix 1.3 Case C Documents

#	Source	Summary of content	Length	Year
D1	CEO	Service description, revenue models, market opportunities, and overview of the company's team.	13 pages	2012
D2	CEO	Company's launch document. Service description, idea generation experience.	3 pages	2012
D3	CEO	Focus on technical characteristics and features of the service.	12 pages	N/A
D4	Web	Company's website.	N/A	N/A
D5	Web	Video interview on CEO's experience as an entrepreneur in the parking industry.	Video (35 min.)	2014
D6	Web	Local Council's promoting webpage. Showcase of local Open Data re-usage (including Case C).	Web page	N/A
D7	Web	Local Start-Up consortium's webpage that describes history and offers of the case.	Web page	N/A
D8	Web	Service description, revenue models, technical characteristics, and local market and expansion opportunities.	1277 words	2013
D9	Web	Focus on individual's skills and backgrounds of people working within the case.	580 words	2013
D10	Web	Structured and interactive outline of technical features leveraged by the company in delivering its Open Data Service.	Web page	2014
D11	Web	Councils Open Data Portals (focus on Open datasets used)	Web Page	N/A

Appendix 2: Semi-Structured Interviews Descriptions

Appendix 2.1: Case A Semi-Structured Interviews

Variable	Goals	Description
Stage	Facilitate a productive atmosphere	Places 1 and 2: closed office (no surrounding noise, no interruptions). Equipment: mobile phone for recording, pen, agenda, and SSI protocol paper.
Actors	Identify people involved and their roles	Interviewer: myself; Interviewee: CEO and Founder of Case A.
Audience	--	I and the interviewee were the audience. The Open Data Services academic community is seen as potential audience too.
Script	Cover all aspects that were scripted beforehand and leave time for improvisation	Semi-Structured interview protocol plus enrichment and refinement.
Entry	Maximize impression management variables.	Congratulatory statements were directed to the interviewee to make him feel comfortable. Smart Casual dress.
Exit	Maximize the chances for further interactions, and prepare for the next performances.	Feedback was sent to the interviewee.

Appendix 2.2: Case B Semi-Structured Interviews

Variable	Goals	Description
Stage	Facilitate a productive atmosphere	Places 1 and 2: hotel hall. Internet connection available. Pleasant environment and low surrounding noise. Equipment: pen, agenda, and SSI protocol paper.
Actors	Identify people involved and their roles	Interviewer: myself; Interviewee: CEO and Founder of Case B.
Audience	--	I and the interviewee were the audience. The Open Data Services academic community is seen as potential audience too.
Script	Cover all aspects that were scripted beforehand and leave time for improvisation	Semi-Structured interview protocol plus enrichment and refinement.
Entry	Maximize impression management variables.	Congratulatory statements were directed to the interviewee to make him feel comfortable. Mentioning the collaboration with the Enterprise Partner and the interest of the City Council to stimulate the subject and improve his impressions. Smart Casual dress likewise the interviewee.
Exit	Maximize the chances for further interactions, and prepare for the next performances.	Feedback was sent to the interviewee.

Appendix 2.3: Case C Semi-Structured Interviews

Variable	Goals	Description
Stage	Facilitate a productive atmosphere	Place1: Company Offices. Internet connection available. Pleasant environment and no surrounding noise. Places 2, 3, and 4: spaces isolated from disturbing noise at an Open Data council's event (Place 2) and in a public space (Place 3 and 4). Equipment: mobile phone for recording, pen, agenda, and SSI protocol paper.
Actors	Identify people involved and their roles	Interviewer: myself; Interviewee1: CEO and Founder of Case C. Interviewee2: Co-Founder and Product manager of Case C.
Audience	--	I and the interviewees were the audience. The Open Data Services academic community is seen as potential audience too.
Script	Cover all aspects that were scripted beforehand and leave time for improvisation	Interviews 1 and 2: Semi-Structured interview protocol plus enrichment and refinement. Interviews 3 and 4: new protocol designed from aspects emerged during interview 1 and 2.
Entry	Maximize impression management variables.	Congratulatory statements were directed to the interviewees to make them feel comfortable. Mentioning the collaboration with the Enterprise Partner and the interest of the City Council to stimulate the subject and improve his impressions. Smart Casual dress likewise the interviewees.
Exit	Maximize the chances for further interactions, and prepare for the next performances.	Feedback was sent to the interviewees.

Appendix 3: Case Study A Codes and Evidences

Appendix 3.1 Case A Key Statements and coding

EXTERNAL FACTORS CLUSTERS	
A	**factors related to the Open Data / Open dataset**
B	**factors related to the city**
C	**factors related to the company**
D	**factors related to interaction Business-City**
E	**Others**

PERCEPTIONS CLUSTERS	
F	**Perceived Usefulness of Open Data**
G	**Perceived Business Opportunity**
H	**Perceived City Opportunity**
I	**Other Perceptions**

#	Ref.	KS from SSI	E/P/A	Ext. Fac. Cluster	Perc. Cluster
1	**1,1**	Previous approach to planning not efficient	E	E	--
2	**1,2**	Integration of data from different city authorities	A	--	--
3	**1,3**	Easy to read format of OD and easy to use ODS	P	--	F / I
4	**1,4**	Easy to use format of OD / ODS	P	--	F / I
5	**1,5**	Long time process (BAU)	E	E	--
6	**1,6**	Case enables efficient decision making process;	P	--	G / I
7	**1,7**	Fragmented process across departments (BAU)	E	E	--
8	**2,1**	Open Data is Open Government Data	--	--	--
9	**2,2**	Online availability of content for re-use	E -> P	A	F
10	**2,3**	Machine readability makes it useful and dynamic	E -> P	A	F
11	**2,4**	Content Released online	E -> P	A	F
12	**2,5**	Re-usage licence for being reliable information	E -> P	A	F
13	**2,6**	Standard (built by someone else, before is just raw data stored in city councils);	P -> A	--	F
14	**3,1**	Cities have to come on board (with OD and the business)	E	B / D	--
15	**3,2**	Key for cities: strategy for release	E	B	--
16	**3,3**	It will be a long time process (growth)	A	--	--
17	**3,4**	Moving slowly (scaling across US cities)	A	--	--
18	**3,5**	Cities should imitate leaders	E	B	--

19	**3,6**	Laws - open by default facilitate the process	E	B / E	--
20	**3,7**	Leaders know what OD is and can be	E	B	--
21	**3,8**	We are working with leaders	E	C	--
22	**3,9**	Very early stage (ecosystem adoption)	A	--	--
23	**4,1**	Local authorities on board (strategy) and volunteers engagement	E	B / D / E	--
24	**4,2**	Objective of covering the all country	A	--	--
25	**4,3**	Cities on board	E	B / D	--
26	**4,4**	CC need awareness of reasonable process (release and establish partnerships)	E	B / D	--
27	**5,1**	Leaders have defined organizational positions for OD	E	B	--
28	**5,2**	They are just releasing them	EXCLUDED		
29	**5,3**	CC should not rely on big companies	E	B	--
30	**5,4**	Other cities look at what leaders do	E	B	--
31	**5,5**	Engagement with practitioners/users demonstrates (to the council) actual potential value of OD and convince them	E -> P	B / E	F / G / H
32	**5,6**	Users of OD are partners of the supplier. Should be engaged	E	B / C / D	--
33	**5,7**	Just release and wait not enough; should engage	E	B	--
34	**5,8**	They have to release accurate and valuable content to be useful	E -> P	A / B	F
35	**5,9**	OD needs to be real-time to be useful	E -> P	A	F
36	**5,10**	They have to promote OD as real data	E	B	--
37	**5,11**	Other cities will imitate	E	B	--
38	**5,12**	OD is content for the service	P	--	F
39	**5,13**	Benefits to local economy and image of the CC	P	--	H
40	**5,14**	Improve competitiveness / attracting businesses	P	--	H
41	**6,1**	CEO was working in the field for a long time	E	C / E	--
42	**6,2**	CEO was working with governments	E	C / E	--
43	**6,3**	CEO got the business idea from being involved in the process and from his attitude to work with governments	E -> P	E	G
44	**6,4**	OD started after the idea was conceived	A	--	--
45	**7,1**	Open Data happened	A	--	--
46	**7,2**	Asking for data to city councils	A	--	--
47	**7,3**	Identifying data	A	--	--
48	**7,4**	Difficulties solved to geo-locate UP projects	P	--	I
49	**7,5**	Need of working with local authorities	E	B / C / D	--
50	**7,6**	They need to cooperate because they have the data (with business and community)	E	B	--

51	**7,7**	Initial idea was just to take the data	A	--	--
52	**7,8**	There was a win-win. Thus, partnership	E	B / C / D	--
53	**7,9**	We map solutions in a way that is lean and effective	E -> P	C	I
54	**7,10**	Made it easier for users to see/understand data	P	--	I
55	**7,11**	The data is still owned by the public	A	--	--
56	**7,12**	We share mapping with the public (OD users)	A	--	--
57	**7,13**	Open Data followed the idea	A	--	--
58	**7,14**	Solutions were not available yet, → business opportunity	E -> P	E	G
59	**7,15**	ODS company helps the city (for operations and for OD strategy achievement)	P	--	H
60	**7,16**	"fitted really well" (for users and the city)	P	--	F / G / H
61	**7,17**	Application helps the government;	P	--	H
62	**7,18**	Business opportunity as starting point	P / A	--	G
63	**7,19**	CC sees well new businesses that can run OD commercial activities	P	--	H
64	**8,1**	Key is current data to be useful. Partnership and strategy ensure it	E -> P	A / B / D	F
65	**8,2**	Key is timely data to be useful (strategy CC)	E -> P	A / B	F
66	**8,3**	Standard data not updated is not good enough	E -> P	A / B	F
67	**8,4**	Data today might be different tomorrow. Needs commitment from CC	E -> P	A / B	F
68	**8,5**	Data in near real-time	E -> P	A	F
69	**8,6**	Working with cities to have these characteristics (8,1 – 8,5)	E	D	--
70	**8,7**	Next thing is standard, first thing to be done (by case) to make it useful for the systems	E -> P	A / C	F
71	**8,8**	When building the standard legal aspects were considered in order to structure what can be known and what can/can't be taken, and so what can be useful	E -> P	A / C / E	F
72	**8,9**	Legal aspects are clear at the beginning to be reliable	E -> P	A / C / E	F
73	**8,10**	We thought the city those things (for growth)	A	--	--
74	**8,11**	Standard enables dynamic feed of OD	E -> P	A	F
75	**8,12**	Standards make it easier for cities to come on board (simpler process). It “opens the council’s doors”	E -> P	A	F / H
76	**9,1**	Cities do it to say they are doing it (social value – CC image)	P	--	H
77	**9,2**	Cities are currently partially engaged	A	--	--

78	**9,3**	Get cities engaged	E	D	--
79	**9,4**	The key is releasing data more frequently to be useful	E -> P	A / B	F
80	**10,1**	Before starting research on data protection and licenses (for commercial re-use)	E -> P A	A / C	F / G
81	**10,2**	EU data protection law reliable	E -> P	A / E	F
82	**10,3**	No applicants names / we need to know	E	A / C / E	--
83	**10,4**	Lot of companies do not comply with these regulations	E	EXCLUDED	
84	**10,5**	To have the names we should be the owner of the data, but is not worth to do it, we scale where is dynamic and reliable	P	--	F / G
85	**10,6**	The benefits of having names is not worth the costs of owning the data and the related database, CC gives the content	P	--	F
86	**11,1**	We structure the dataset to make it useful and we then adopt it	P -> A	--	F
87	**11,2**	Fix bad data	E -> P -> A	A / C	F
88	**11,3**	Dataset is useful for the data standard and content	E -> P	A	F
89	**11,4**	After of course the accuracy of the data	E -> P	A	F
90	**11,5**	Structuring the data means give to all information the same structure, e.g. Number, street, district code. We do it	E -> P	A / C	F
91	**11,6**	Data is usually stored manually by CC and:	E -> P	A / B	F
92	**11,7**	- inconsistencies → not useful	E -> P	A	F
93	**11,8**	- different formats → not useful	E -> P	A	F
94	**11,9**	The software code is built upon the standard (for Pilot implementation)	A	--	--
95	**11,10**	Software code needs one way of representing the information (for Pilot implementation)	A	--	--
96	**12,1**	Creating a standard is a long time process (for Pilot implementation)	A	--	--
97	**13,1**	Structuring the data is an automated process (for Pilot implementation)	A	--	--
98	**13,2**	We are still actively working on the data	A	--	--
99	**13,3**	As a consequence cities open the doors to us	P	--	F
100	**13,4**	OD provides the content	P	--	F
101	**13,5**	Standard helps in properly structure the data and make it useful	E -> P	A / B / C	F
102	**13,6**	Geo location is added subsequently	A	--	--
103	**13,7**	Visualization is the third step	A	--	--
104	**13,8**	Mash up different datasets is the future	A	--	--
105	**13,9**	For cities to perceive it beneficial, the ODS needs to be aligned with how:	P	--	H

106	**13,10**	- cities work	P	--	H
107	**13,11**	- citizens act	P	--	H
108	**13,12**	We are cooperating with the cities	E	D	--
109	**13,13**	Joined programs to leverage interactions with the cities and to:	E	D / E	--
110	**13,14**	- make them releasing the data and ensure continuity	P	--	G
111	**13,15**	- give the standards back to the related authority	P	--	G / H
112	**14,1**	The standard is specific for the Planning dataset	E	A	--
113	**14,2**	We looked at what other companies are doing with OD	A	--	--
114	**14,3**	Everyone is working with the data	EXCLUDED		
115	**14,4**	Partnership make data reliable and enable short time development	P	--	F / G
116	**14,5**	Cities/users (if possible together) have to create the standards	E A	A / B / C / D	--
117	**14,6**	Unstructured data - differences of data within the same dataset impede people to re-use it	E -> P	A	F
118	**14,7**	Cities subsequently came on board and release data with that standard	A	--	--
119	**14,8**	Get access to as many people as they can is a key	E	B / E	--
120	**14,9**	With right promotion and standard people will adopt it	E -> P -> A	A / B / E	F
121	**14,10**	Open Data was new	A	--	--
122	**14,11**	Cities need to come on board with organizational commitment	E	B / C / D	--
123	**14,12**	Cities need to stimulate re-use	E	B / E	--
124	**14,13**	Cities need to promote standards	E	B / E	--
125	**15,1**	looking at other companies showed the need for cooperatively creating a standard for the Open Dataset	E	A / C / D / E	--
126	**15,2**	(...) and to cooperate with cities	E	D	--
127	**15,3**	(…) and to use the standard and experiences to get more and more cities on board	E A	C / D / E	--
128	**15,4**	But open data was not popular at that time (when case was founded)	A	--	--
129	**16,1**	Six people in the company	E	C	--
130	**16,2**	Team of programmers	E	C	--
131	**16,3**	Programming skills are crucial to create and develop the standard and to access OD	E -> P	C	F
132	**16,4**	Data analysis is conducted (when city targeted)	A	--	--

133	**16,5**	Show how to normalize data is critical	A	--	--
134	**16,6**	From the technical side structured data are generated from unstructured data	A	--	--
135	**16,7**	Normalization process follows	A	--	--
136	**16,8**	Once done it for the first city, the nature and standard of data are shown to other to convince them to come on board	E -> P -> A	C / D / E	F / G / H
137	**16,9**	Then we work with the new city and normalize their data	A	--	--
138	**16,10**	Then the process is consistent with:	A	--	--
139	**16,11**	- Case's standard		--	--
140	**16,12**	- Case 's systems		--	--
141	**17,1**	Half of the people in the company is involved with developing software	E	C	--
142	**18,1**	Graphic design skills are needed	E	C	--
143	**18,2**	Front-end skills are needed	E	C	--
144	**18,3**	Case operates mainly with "the technical side of things"	E	C	--
145	**18,4**	Core skills are:	E	C	--
146	**18,5**	- GIS (mapping)		C	--
147	**18,6**	- Data Analytics		C	--
148	**18,7**	- Data Normalization		C	--
149	**18,8**	- Data visualization		C	--
150	**19,1**	Normal physical resources needed (computers...);	E	C	--
151	**19,2**	No special software licences; Open Source are free, available and easily accessible	E	C	--
152	**19,3**	Open Source for mapping and analysing; that is how it should be	E	B / C / E	--
153	**19,4**	Governments are relying on big contracts with multinationals (but still legacy systems); They should adapt to Open Source	E	B	--
154	**20,1**	Governments needs to be educated about OD:	E -> P A	B / C / D	G / H
155	**20,2**	- what it is		B / C / D	F / G / H
156	**20,3**	- how to use it		B / C / D	F / G / H
157	**20,4**	- promote the standards		B / C / D	F / G / H
158	**20,5**	- legal information		B / C / D	F / G / H
159	**20,6**	- releasing information		B / C / D	F / G / H
160	**20,7**	This was one of the main challenges for making them perceiving opportunity		B / C / D	H
161	**20,8**	Leading cities is good, but this should spread throughout the nations	A	--	--
162	**20,9**	We research and target cities that are leading OD for partnerships	E A	B / C / D	--
163	**20,10**	Big problem to involve new cities (especially small ones)	E	B / C / D	--

164	**20,11**	Communication with local authority was their main barrier	A	--	--
165	**20,12**	Show that other cities are doing it is beneficial	E -> P	B / C / D	H
166	**20,13**	Show what we do is also critical (see opportunity for governments - win-win)	E -> P	B / C / D	G / H
167	**21,1**	Support in releasing the data	E	B / D	--
168	**21,2**	Contracts in the US (commercial)	E -> P	D	F / H
169	**21,3**	Beneficial the hack competition	E	C / E	--
170	**21,4**	OD champions promote (also) their services	E	B / E	--
171	**21,5**	Demonstrations are crucial	E -> P	C / D / E	H
172	**21,6**	Presentations and promotions are critical to show the value behind embracing OD	E -> P	C / D / E	H
173	**21,7**	Local Authority gives support to get known	E -> P	B / C / E	H
174	**21,8**	The price they won is a kind of certificate to be shown to other cities	E -> P	C / D	G / H
175	**21,9**	Price also supported early stages of the business	E	C	--
176	**21,10**	Air space given by posting Case on the city councils websites	E	B / D	--
177	**22,1**	No tax incentives	E	C	--
178	**22,2**	No funds because it is too early; governments have no idea yet	E	B / C	--
179	**22,3**	Governments are unsure	P	EXCLUDED	
180	**22,4**	They provide cash prices	E	B / C	--
181	**22,5**	They say: "if you build the application we will support you"	P A	--	G / H
182	**22,6**	Have a platform with standards and give it to users to develop services	E	B	--
183	**22,7**	ODS stimulates local economy	P	--	H
184	**22,8**	New start ups can be created in the ODS ecosystem	P	--	H
185	**22,9**	If they can't do the platform, then do it with people, engage with the public → cities and businesses will see opportunities	E -> P -> A	B / E	G / H
186	**22,10**	Governments should not build their own apps	E	B	--
187	**22,11**	ODS can be valuable for governments, citizens, and businesses	P	--	F / G / H
188	**22,12**	They should facilitate the creation of new business models	E	B	--
189	**22,13**	/ no big companies outsourcing because they have revenue rules and laws	E	B / E	--
190	**22,14**	SMEs and startups should play the key role in designing and delivering new ODS	P -> A	--	F / G
191	**23,1**	No support from consultants or other forms	E	C	--
192	**23,2**	Case covers all the value network	E	C	--

193	**23,3**	Cities need to be convinced by saying "these is what other councils are doing" and "these are the datasets that they are publishing"	E -> P	D	H
194	**23,4**	If we wait for governments to do it, it will not happen	A	--	--
195	**23,5**	Close to have one universal dataset for all the country (Ireland); nobody else did it so far	A	--	--
196	**24,1**	App competition was a lever	E	E	--
197	**24,2**	It allowed to go and speak with other councils and demonstrate	E -> P	D / E	H
198	**24,3**	It is beneficial to show to cities that others are putting a lot of effort in OD	E -> P	D / E	H
199	**24,4**	If one council is doing it, why shouldn't you do that?	E -> P	D / E	H
200	**25,1**	Size of the city is important variable when targeting one	E -> P	B	G / H
201	**25,2**	Big cities are more likely to embrace OD	E -> P	B	H
202	**25,3**	Small cities are hard to get on board because of lack of resources	E -> P	B	H
203	**25,4**	The money they have to spend are a variable	E -> P	B	H
204	**25,5**	The presence of an established IT team in the council is another key; it allows to "put these things in place	E -> P	B	G / H
205	**25,6**	After creating the standard they "pull information out"	A	--	--
206	**25,7**	The dataset has to go through all the processes to be released	E	B	--
207	**26,1**	CIO is the main contact in councils	E	B / D	--
208	**27,1**	Not having a city CIO is not beneficial; in this case all the work has to be done by the case	E -> P	B / D	G / H
209	**28,1**	Cities are chosen for their digital development (Smart City)	E -> P	B	G / H
210	**28,2**	Cities releasing OD, cities with OD portal	E / A	B	--
211	**28,3**	Active cities in the ODS ecosystem through e.g. Showcases, hack competitions	E / A	B / E	--
212	**28,4**	Cities that "believe" in OD	E / A	B	--
213	**28,5**	Cities that have a CIO	E / A	B	--
214	**28,6**	Cities that are collaborative	E / A	B / D	--
215	**29,1**	Champions	E	E	--
216	**29,2**	Cities aware of needs of users	E	B	--
217	**29,3**	Cities that act concretely, develop and update OD portal	E	B	--

218	**29,4**	Cities worried about costs; but hoping that companies like Case demonstrate that there is value	E -> P	B / C / D	G / H
219	**30,1**	Cities often refuse to release data / don't capture opportunity	P	--	H
220	**30,2**	"We will wait until other councils around us will do it"	P	--	H
221	**30,3**	Even when councils around do it, it is still difficult	P	--	H
222	**30,4**	Breaking down these barriers means opportunity for us	P	--	G
223	**30,5**	CC feel threatened that critics emerge from this data to go against councils (e.g. journalists); scared about potential use	P	--	G / H
224	**30,6**	OD promotes transparency about what is happening	P	--	H
225	**30,7**	Councils refuse other people to teach how to work from data	P	--	H
226	**31,1**	Time to release the data is a barrier	E	B	--
227	**31,2**	Tight budgets is a barrier	E	B	--
228	**31,3**	Lack of resources is a barrier	E	B	--
229	**31,4**	Lack of IT department (well developed) is a barrier	E	B	--
230	**31,5**	Lack of skills to normalize the data is a barrier (rely on community?)	E	B / E	--
231	**31,6**	Case gives them the standard and teach them to bridge these gaps	E A	B / C / D	--
232	**31,7**	Time and resources are excuses; Case provides them standard i.e. data in machine readable format ready to be published	A	--	--
233	**32,1**	In US contract is municipality agreement that covers	E -> P	D	F / G
234	**32,2**	- council commit to provide accurate data	E -> P	D	F / G
235	**32,3**	- data released in daily basis	E -> P	D	F
236	**32,4**	Value of OD is near real-time	E -> P	A	F
237	**32,5**	Standard contracts are formulated	E	D	--
238	**32,6**	Stagnant dataset common negative approach	E	A	--
239	**32,7**	Contracts allow to overcome it	E	D	--
240	**32,8**	The only way it can work is by giving them data in real-time	E -> P	A / B	F
241	**33,1**	Contracts to provide data daily	E -> P	D	F / G
242	**33,2**	Contracts are in cities interests as well; they benefit from Case to provide them a service	E -> P	B / D	G / H
243	**33,3**	ODS companies provide civic engagement platforms for their residents	P	--	H

244	**33,4**	The ODS works because it is a partnership	E	D	--
245	**34,1**	Champions to get data released and cities involved	E	B / E	--
246	**34,2**	Commitment of governments to release data and in near real-time	E	B / D	--
247	**34,3**	No businesses from stagnant dataset;	P	--	F
248	**34,4**	OD pilot project is not enough	A	--	--
249	**34,5**	Businesses have to help and push	E / A	C	--
250	**34,6**	Governments need long-term commitment	E	B / D	--
251	**34,7**	Business landscape on ODS will happen then	A	--	--
252	**34,8**	More apps and services will be developed for them	A	--	--
253	**34,9**	Champions to put in place demonstrations:	E -> P	E	H
254	**34,10**	- to show the value of OD	E -> P	E	F / H
255	**34,11**	- to show what can be done	E -> P	E	H
256	**34,12**	will stimulate all departments to release data	E -> P	B / D	H
257	**34,13**	- to run community events	E	E	--
258	**34,14**	- even at a national scale		E	--
259	**34,15**	Spread awareness is a priority	E	B / E	--
260	**34,16**	Need for more champions	E	E	--
261	**35,1**	Factors preventing:	E	A / B / D	--
262	**35,2**	- stagnant datasets	E	A	--
263	**35,3**	- governments locked to release data	E	B	--
264	**35,4**	- no long-term commitment	E	B / D	--
265	**35,5**	Governments need to be there long time, and ready for the long journey	E	B	--
266	**35,6**	They need to engage new people	E	B / E	--
267	**35,7**	No outsourcing	E	B	--
268	**35,8**	Create rooms where these solutions can be developed	E	E	--
269	**35,9**	Stimulate new companies	E	E	--
270	**35,10**	Outsourcing huge barrier	E	B / E	--
271	**36,1**	Community driven initiatives; Code for Ireland	E	E	--
272	**36,2**	Leaders in the field (Sc domain) helping new services creation	E	E	--
273	**36,3**	Coder dojo	E	E	--
274	**36,4**	Voluntary work	E -> P	E	I
275	**36,5**	Spread awareness among general public as well	E	E	--
276	**36,6**	Governments just publishing OD is not enough (integrated strategy)	E	B	--
277	**36,7**	They need to engage with people	E	B / E	--

278	**36,8**	They need to show and promote what already done (Case is in City Council and SF websites)	E	B / E	--
279	**36,9**	An ecosystem needs to be built by cities to grow (together with existing OD businesses)	E	B / E	--
280	**36,10**	Services need support by cities	E	B / D	--
281	**36,11**	Potential users need to see the value	P	--	F

Appendix 3.2: Assertion 2, City and City Authority Characteristics

ASSERTION 2						
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs**	**Field Notes**	**Validated**
2.1	*Open Data Strategy and Committee*	3,1; 3,2; 4,1; 4,3; 8,1; 8,2; 8,3; 8,4; 9,4; 11,6; 14,8; 14,11; 22,6; 28,2; 28,3; 28,4; 29,3; 32,8; 35,3; 35,5; 36,6; 36,9;	D1;	D5; D6; D7;	When targeting a new city, the presence of an Open Data program is the first aspect considered.	Yes
2.2	*Internal Law Development*	3,6;	--	D6;	--	Yes
2.3	*Organizational Structure*	5,1; 14,11; 25,5; 26,1; 27,1; 28,5; 31,4; 31,5;	D1;	D6;	The cities the case is working with have organizational positions devoted to Open Data.	Yes
2.4	*Open Data Awareness*	4,4; 20,1; 20,2; 20,3; 20,4; 20,5; 20,6; 20,7; 34,14;	D1;	D6; D9;	The need of "educating cities" emerged several times in the notes.	Yes
2.5	*Budget and Resources*	25,3; 25,4; 29,4; 31,1; 31,2; 31,3;	--	--	Main reason provided to the case when cities refused to provide Open Data.	Yes
2.6	*Engagement Initiatives*	5,5; 5,6; 5,7; 7,5; 7,6; 13,5; 14,5; 21,1; 22,9; 28,6; 29,2; 35,6; 36,7;	D1; D4;	D6; D8;	When discussing possible initiatives that cities can implement to foster ODS, workshops and showcases emerged as beneficial programs.	Yes
2.7	*Size of the City*	25,1; 25,2; 25,3;	--	D8;	Bigger cities are targeted for bigger market opportunities.	Yes
2.8	*Support Creation and Sustainability*	21,4; 21,7; 21,10; 22,2; 22,4; 36,8; 36,0;	--	D5; D8;	Both in Europe and in US, Case A has been promoted in the city councils websites.	Yes

	of Open Data Services					
2.9	*Cities should use Open Source Software*	19,3;	--	--	--	No
2.10	*Cities shouldn't outsource ODS creation*	5,3; 19,4; 22,10; 22,12; 22,13; 35,7; 35,10;	--	--	--	Yes

Appendix 3.3: Assertion 3, Organizational Skills Factors

ASSERTION 3						
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Field Notes**	**Validated**
3.1	*Geographic Information Skills*	7,9; 18,5; 19,3;	D1; D4;	D5; D6; D7; D8;	From observing activities, mapping the data emerged as one of the critical step in delivering the ODS.	Yes
3.2	*Programming Skills*	8,7; 8,8; 11,2; 13,5; 14,5; 15,1; 15,3; 16,2; 16,3; 16,8; 17,1; 18,3; 18,7;	D1; D2; D3;	D6; D7; D8; D9;	Developers outlined the importance of their skills for the case.	Yes
3.3	*Knowledge of Legal and Policy Context*	8,8; 8,9; 10,1; 10,3;	D4;	D9;	The CEO gives periodic updates to the overall team of developers about legal and policy context.	Yes
3.4	*Data Analytics Skills*	11,2; 18,6;	D1; D4;	D6;	Part of the business model is enabled by data analytics processes.	Yes
3.5	*Graphic / Web Design and Data Visualization Skills*	18,1; 18,8;	D1; D3; D4;	D5; D6; D7;	Implicitly confirmed by all people approached within the case.	Yes
3.6	*Communication Skills*	20,12; 20,13; 21,5; 21,6; 24,2; 29,4; 34,5;	D2;	D9;	Especially the Business Developer (agreed by all people).	Yes
3.7	*Companies have to cover the entire process*	22,1; 22,2; 23,1;	--	--	--	No

Appendix 3.4: Assertion 9, Perceived Business Opportunity

ASSERTION 9						
Variables		**Key Statements**	**Int. Docs.**	**Ext. Docs.**	**Field Notes**	**Validated**
9.1	*Benefits for End-Users*	1,6; 7,14; 7,16; 7,18; 20,13; 22,11; 29,4; 30,5;	D1; D2; D3; D4;	D5; D6; D7; D8;	The description of services offered and how these improve the process of searching for urban planning information became clear at the time of structuring the contextual data for the case.	Yes
9.2	*Partnership Strength*	13,14; 13,15; 14,4; 16,8; 20,1; 20,2; 20,3; 20,4; 20,5; 20,6; 20,7; 21,8; 22,5; 22,14; 25,1; 25,5; 27,1;28,1;29,4.	--	D9;	Strong emphasis placed by the lead business developer.	Yes
9.3	*Level of Engagement*	5,5; 21,8; 22,9; 30,4;	D1; D2;	D9;	Based on the CEO's experiences within and outside the case.	Yes
9.4	*Scalability*	10,5; 13,14; 13,15; 16,8; 20,4;	D1; D2;	D6;	Mentioned several times by the CEO in relation to business growth across cities.	Yes
9.5	*Demonstrability*	5,5; 8,12; 13,15; 16,8; 20,1; 20,2; 20,3; 20,4; 20,5; 20,6; 20,7; 20,12; 20,13; 21,5; 21,6; 21,8; 23,3; 24,2; 24,4; 29,4; 34,9; 34,10; 34,11; 34,12;	D1; D2;	D6;	Both the CEO and the Business Developer stressed the importance of demonstrating the value to "push cities to come on board", i.e. to grow.	Yes

Appendix 4: Case Study B Codes and Evidences

Appendix 4.1: Case B Key Statements and coding

EXTERNAL FACTORS CLUSTERS	
A	**Factors related to Open Data and Dataset Characteristics**
B	**Factors related to the organization**
C	**Factors related to the city (council and city characteristics)**
D	**Factors related to the Open Data Ecosystem**
E	**Factors related to the service characteristics**
F	**Other Factors**

PERCEPTIONS CLUSTERS	
G	**Factors related to perceptions about Usefulness of Open Data**
H	**Factors related to perceptions about Open Data Service Business Opportunity**
I	**Factors related to perceptions about service users**
J	**Factors related to perceptions about the city**
K	**Other perceptions-related Factors**

#	Ref.	KS from SSI	E / P / A	Ext. Fac. Cluster	Perc. Cluster
1	**1,1**	(step 2) we re-bound the application	A	--	--
2	**1,2**	Passion about the context (Tourism) and IT attitude	E	F	--
3	**1,3**	Passion about Open Data Ecosystem	E	F	--
4	**1,4**	Massive amount of data from EU level wide collaboration	E / A	A / D	--
5	**1,5**	Interactivity of the service, interactive learning experience	E->P	E	I
6	**1,6**	Not his primary job, but passionate about IT (initially voluntary adoption)	E->P	F	K
7	**1,7**	Learning by doing (learning applied - lean startup approach)	E	B	--
8	**1,8**	Sleek Design for making the service easy-to-use	E->P	E	I
9	**1,9**	Design fundamental characteristic	E	E	--
10	**1,10**	Additional Device (Apple Watch) to provide internet to foreigners as Public Wi-Fi not efficient	E	E / C	--
11	**1,11**	Structured Problem formulation (tourism in the area and roaming problems) – but large amount of OD is available	E->P	C	G / J
12	**1,12**	Service includes a challenge and reward system	E->P	E	I
13	**1,13**	Website made content focused	E	E	--
14	**1,14**	Events integrated into the service for dynamic content for users	P	--	G / I

15	**1,15**	Step 2: app development	A	--	--
16	**1,16**	Step 3: fully integrated tourism information services	A	--	--
17	**1,17**	Collect as much information as possible	A	--	--
18	**1,18**	Re-bound information into tiles (focus on technology) – for pilot and when integrating new OD	E /A	E	--
19	**1,19**	Sleeker design and landmarks – GIS Skills	E->P	B / E	I
20	**1,20**	Go and ask information (Open Data) to the city council	A	--	--
21	**1,21**	Make their data (from analytics) openly available to: local authority and other businesses. Participatory approach.	E->P	D / E	J
22	**1,22**	Foster the ecosystem of ODS in EU through more OD published to strengthen sustained adoption	E->P A	D	J
23	**1,23**	Improve local OD strategy by publishing more Open Data and drive ecosystem around – leveraged for growth and for sustaining use	E->P A	C / D / E	J
24	**1,24**	Shared value for the community from data generated	E->P A	D / E	J
25	**2,1**	CEO originally from the area	E	F	--
26	**2,2**	It started as volunteering	P	--	K
27	**2,3**	Commercialization purposes emerged later in the process	A	--	--
28	**2,4**	First goal: build a reasonable user base	A	--	--
29	**2,5**	Expand through more cities	A	--	--
30	**2,6**	Make tours' data openly available – i.e. more OD, more knowledge created	E->P	A / D / E	I / J
31	**2,7**	Open Source as input and as outcome. From access to re-usage and final data sharing	E->P	A / B / E	G / J
32	**2,8**	Despite big potentiality, underexplored area: problem-opportunity	E->P	C	J
33	**2,9**	Tourists usually don't have cars – improving experience	E->P	C	I
34	**2,10**	Goal: incentivise people to visit the area and stimulate local economy	P	--	I / J
35	**2,11**	Other initiatives in Europe are being observed and studied- EU focus	E	D	--
36	**2,12**	Open Data knowledge both sides, i.e. CC and ODS B.	E	A / C / F	--
37	**2,13**	As much data as possible (maps, transport...) consciously from CC to have large datasets analyzable	E->P	A / C	G / I
38	**2,14**	Time saving for us (development) and users (experience)	P	--	G / I
39	**2,15**	Improving user's visit through more content	P	--	G / I
40	**2,16**	Content for enabling new interactive experiences in the context	E->P	E	G / I

41	**3,1**	Content for solving local inefficiencies (no connections and not promoted)	E->P	C / E	G / J
42	**3,2**	Content for advertising and promoting local places	E->P	C / E	G / J
43	**4,1**	Facilitate visitor's cultural change – benefit for city	P	--	I / J
44	**4,2**	Lot of data (quantity matters)	E->P	A	G
45	**4,3**	Real-time data	E->P	A	G
46	**4,4**	Satisfy a want and need	P	--	I
47	**4,5**	Updated data for dynamic information	E->P	A	G
48	**4,6**	Frequency of data matters for having real-time picture	E->P	A	G
49	**4,7**	Display information-enabled daily decisions to users	E->P	A / E	G / I
50	**4,8**	Sleek fashion (easy to use service)	P	--	I
51	**5,1**	Easy design fundamental – skills needed	E->P	B / E	I
52	**5,2**	Graphic should be "nice" – skills needed	E->P	B / E	I
53	**5,3**	Tourism platform	E	E	--
54	**5,4**	Stagnant dataset is a barrier	E->P	A	G
55	**5,5**	Cities want to state they are embracing Open Data	E->P	C	J
56	**5,6**	Lot of data to be bound into tiles to avoid not having a light solution	E->P	B / E	G / I
57	**5,7**	Technology and design should be "fancy" (attractive)	E->P	B / E	I
58	**5,8**	From rough and ready (step 1) to clean and clear (step 3)	A	--	--
59	**5,9**	Service got cleaner and cleaner	P / A	--	I
60	**5,10**	IOS based	E	B / E	--
61	**5,11**	Google material design integrated	E	B / E	--
62	**6,1**	Previously working in the context	E	F	--
63	**6,2**	Idea generated from work experience	E / A	F	--
64	**6,3**	Idea of the platform for tourists came first but lack of data was main issue	P / A	--	G
65	**6,4**	Data was the key aspect of the idea	P	--	G
66	**6,5**	Hire editors not convenient solution	P	--	G
67	**6,6**	Look for availability of data was the modus operandi	P / A	--	G
68	**6,7**	Developer's knowledge	E	B	--
69	**6,8**	Easy for developer to apply codes on top of a database – also adapting for new city	E->P	B	G / H
70	**6,9**	Open Data useful for:	P	--	G
71	**6,10**	- saving time (for each city)	P	--	G / H
72	**6,11**	- have a database structured	P	--	G
73	**6,12**	- developer pull the database into the app	P	--	G / H
74	**6,13**	- freely available	P	--	G

75	**6,14**	- enabling quicker development of an already existing idea (scalable)	P / A	--	G / H
76	**6,15**	Open Data as the best way of implementing the idea	P	--	G
77	**6,16**	The idea preceded the availability of the data	A	--	--
78	**6,17**	Short time to develop the first app, so decision to use OD	P->A	--	G
79	**6,18**	First develop a MVP	A	--	--
80	**6,19**	Quickest way through OD	P	--	G
81	**6,20**	OD is free	E->P	A	G
82	**6,21**	No effort to search for the data	P	--	G
83	**6,22**	Time saving for not coding maps – but skills needed for cleaning	E->P	A / B	G
84	**6,23**	OD providing localisation for users	E->P	A	G / I
85	**6,24**	Machine readability fundamental for:	E->P	A	G
86	**6,25**	- not coding the map	P	--	G
87	**6,26**	- huge time saving	P	--	G
88	**6,27**	- only pre requisite for success	P	--	G
89	**6,28**	First council to be able to demonstrate others that this is possible and the benefits behind	P / A	--	H / J
90	**6,29**	Advantage for the city for more incomes (valuation)	P	--	J
91	**6,30**	Economic growth of the area	P	--	J
92	**6,31**	Building things quick is perceived as fundamental for business opportunity	P	--	G / H
93	**6,32**	Advantage for hotels and businesses included in the list	P	--	J
94	**6,33**	Businesses are encouraged by the user base to embrace the service	P / A	--	J
95	**7,1**	No idea about the council having the data before. Discovered content from passion on OD	E->P A	F	G / J
96	**7,2**	No idea about OD before hack competition. Discovered content and transparency there	E->P A	D / F	G / J
97	**7,3**	Hack competition fundamental to discover OD	E / A	D	--
98	**8,1**	Hack competition to leverage previous work	E / A	D / F	--
99	**8,2**	Apple Watch game changer	E / A	E	--
100	**8,3**	Aw enables more and more content and service enrichment	E->P A	E	I
101	**8,4**	Directions included	P	--	I
102	**8,5**	Tour guides and so stimulate local economy and improving visitors experience	P	--	I / J
103	**8,6**	Direction without roaming	P	--	I
104	**8,7**	Public Wifi lacks and it is an issue	E		--
105	**8,8**	Providing connection is key	E		--
106	**8,9**	Internet connection not only enables the service but it is a need for visitors – public Wi-Fi	E->P	E / C	I
107	**8,10**	The service wants to own a travel trip – OD is the content	E->P	E	G / I

108	**8,11**	(use case scenarios) of travellers are outlined and covered by the service	E->P	E	I
109	**8,12**	Usefulness of Open Data is great	P	--	G
110	**8,13**	Initial success of Case B	A	--	--
111	**8,14**	Incentivise users is key	E->P	E	I
112	**8,15**	- gamification	E->P	E	I
113	**8,16**	- rewards	E->P	E	I
114	**8,17**	Providing visitors all information	E->P	E	G / I
115	**8,18**	One unique Open Data platform and portal	E	A	--
116	**8,19**	One map with everything (not just travel) available as open data	E	A	--
117	**8,20**	MVP to demonstrate that this is feasible and scale it	P / A	--	H
118	**8,21**	Initial small amount of time dedicated	A	--	--
119	**8,22**	Quickly... x3	P	--	G
120	**8,23**	Usage of APIs critical – but no skills in CC	E	A / B	--
121	**8,24**	Open Source Software critical	E	A / B	--
122	**8,25**	Combination APIs OSS and OD	E	A / B / E	--
123	**8,26**	Context and ecosystem allow effective display of OD benefits	E->P	D	G / J
124	**8,27**	Lots of benefits hidden behind OD	P	--	G
125	**9,1**	OD enables:	E	A	--
126	**9,2**	- user satisfaction	P	--	G / I
127	**9,3**	- travel improvement	P	--	G / I
128	**9,4**	- city efficiency	P	--	G / J
129	**9,5**	- valorisation of tourist attractions	P	--	J
130	**9,6**	- promotions of tourist attractions	P	--	J
131	**9,7**	Alternatively meet up with the locals and write a guide	P	--	G
132	**9,8**	OD enables cheap service development	P	--	G
133	**9,9**	Quickly	P	--	G
134	**9,10**	Real-time is the dream	E	A	--
135	**9,11**	An ecosystem should foster through infrastructural investments	E	A / D	--
136	**9,12**	sensors around the city and real-time open data	E	A /D	--
137	**9,13**	Real-time weather data	E	A	--
138	**10,1**	Timely data	E	A	--
139	**10,2**	Frequent data	E	A	--
140	**10,3**	Real-time data	E	A	--
141	**10,4**	Crime data	E	A	--
142	**10,5**	Data reflecting main daily needs (e.g. transport)	E	A	--
143	**11,1**	Data has to be accurate	E	A	--
144	**11,2**	Data extraction should be simple (machine readability) – Skills in CC needed	E->P	A / B	G
145	**11,3**	Standard development effort should be minimised	E	A / B	--

146	**11,4**	License should be available originally	E	A	--
147	**11,5**	Landmark data	E	A	--
148	**11,6**	Attraction data	E	A	--
149	**11,7**	More data needed	E	A	--
150	**11,8**	Big potential for new companies and ODSs	P	--	H / J
151	**11,9**	OD facilitates app development	P	--	G / J
152	**11,10**	Noise pollution data	E	A	--
153	**11,11**	Traffic data	E	A	--
154	**11,12**	More Open Data from sensors – from ecosystem	E	A / D	--
155	**11,13**	Full information to users	E->P	A / E	G / I
156	**12,1**	The observation of the city should be openly available (data form sensors)	E	A / D	--
157	**12,2**	IT companies with cities should built the infrastructure	E	C / D	--
158	**12,3**	IT companies with cities should build the ecosystem	E	C / D	--
159	**12,4**	I f more OD content new startups would jump in	P	--	G / J
160	**13,1**	Not clear the business model of the OD ecosystem	E	D	--
161	**13,2**	CSR main barrier to open innovation ecosystem	E	D	--
162	**13,3**	Long-term plans needed – from CC and ecosystem	E	C / D	--
163	**13,4**	Ecosystem has to be supervised (CC), built and maintained	E	C / D	--
164	**13,5**	Competition laws threatened?	E->P	D	J
165	**13,6**	CSR has to meet commercial purposes	E	D	--
166	**13,7**	Scalability highly critical for growth	E->P A	E	H
167	**13,8**	Big work is done once only	E->P A	B / E	G / H
168	**13,9**	Open Data is a great opportunity for city and businesses	P	--	G / H / J
169	**13,10**	Need to be the “middle man” to make money	P	--	G / H
170	**13,11**	Software skills key to make money	E	B	--
171	**14,1**	Project management skills	E	B	--
172	**15,1**	Inaccurate data was an issue	E	A	--
173	**15,2**	Data cleaning needed	E	A / B	--
174	**15,3**	Many errors in the Open Datasets	E	A / B	--
175	**15,4**	Huge amount of work for cleaning	E	A / B	--
176	**15,5**	Cleansing the data	E	A / B	--
177	**15,6**	Double checking before loading time consuming process	E->P	A /B	G

178	**15,7**	Latitude and longitude were main errors	E	A	--
179	**15,8**	Data fix and then load	E	A / B	--
180	**15,9**	Data integration within the system	E	A / B	--
181	**15,10**	Data visualisation skills	E	B	--
182	**15,11**	Data cleansing skills	E	B	--
183	**15,12**	Data integration skills	E	B	--
184	**15,13**	Ensure the right APIs is in place	E	A / B	--
185	**15,14**	Visualisation is key	E	B	--
186	**15,15**	Transform excel table into something easy to see and detect	E->P	A / B	G / I
187	**16,1**	No language issues	E	EXCLUDED	
188	**17,1**	Lack of trust in the data was an issue – needs to be promoted	E->P	A / C	G
189	**17,2**	Takes time to trust the data and the supplier	E->P	A / C	G
190	**17,3**	Initial difficulties for many errors in the data	E->P	A	G
191	**17,4**	Lack of data is an issue (much more data could be made available)	E	A	--
192	**17,5**	Lack of awareness from the city (they have more data)	E	C	--
193	**17,6**	Lack of free wifi	E	C	--
194	**17,7**	Errors in latitude and longitude	E	A	--
195	**18,1**	Three people	E	B	--
196	**18,2**	Two volunteers – and I started as volunteer	E->P	B	K
197	**19,1**	I am Project manager	E	B	--
198	**19,2**	I am designer	E	B	--
199	**19,3**	I am developer	E	B	--
200	**19,4**	Small lean team	E	B	--
201	**20,1**	Commercialization:	P / A	--	H
202	**20,2**	- charging hotels (businesses in the lists)	P	--	H
203	**20,3**	- charging events to be hired up	P	--	H
204	**20,4**	Commercialize deals offers	P / A	--	H
205	**20,5**	User base fundamental to commercialize	P / A	--	H
206	**20,6**	Funds needed for commercialization	P / A	--	H
207	**20,7**	Improve local OD website is a goal (through ODS)	E->P A	C / E	G / J
208	**21,1**	Expansion will come after funding	A	--	--
209	**21,2**	Perceived voluntariness of use is great	P	--	K
210	**21,3**	Availability of data is the key	E->P	A	G
211	**22,1**	Community initiatives	E	D	--
212	**22,2**	- should be accessible (free and at appropriate time)	E	D	--
213	**22,3**	Coder Dojos	E	D	--

214	**22,4**	Hackathons	E	D	--
215	**22,5**	Data from EU next goal	E	D	--
216	**22,6**	Funding seek from EU (H2020)	E / A	D	--
217	**22,7**	Content Management System is scalable, Generalizable, and interoperable	E->P A	E	H / I
218	**22,8**	Funding seek then wider adoption and commercialization and expansion across cities	P / A	--	H
219	**22,9**	Funds fundamental	P / A	--	H
220	**23,1**	H2020 funding scheme	E->P	D	H
221	**23,2**	Awareness about EU funding	E->P	D / F	H
222	**23,3**	Partners needed from ecosystem	E	D	--
223	**24,1**	Game theory critical for incentivise users	E->P	B / E	I
224	**24,2**	Individual creativity for combining the "buzzwords"	E	B / F	--
225	**24,3**	Cloud based platform	E	B / E	--
226	**24,4**	Analytics	E	B / E	--
227	**24,5**	Social Local and Global	E	E	--
228	**25,1**	Open Data discovered as a way of developing solutions quickly – and scale (only if machine readable)	E->P	A / B / E	G / H
229	**26,1**	Champions local (within and outside CC)	E	C / D	--
230	**26,2**	National level champions	E	D	--
231	**26,3**	Innovation Network	E	C / D	--
232	**26,4**	UK example	E	D	--
233	**26,5**	Working at European level for:	E	D	--
234	**26,6**	- get funding	E	D	--
235	**26,7**	- develop best practices	E	D	--
236	**26,8**	- knowledge sharing at the wide level	E	D	--
237	**26,9**	- wide collaboration (but management critical)	E	B / D	--
238	**26,10**	- content for promoting the local area	E->P	C / D	G / J
239	**26,11**	- great culture in EU: tourism huge opportunity to be exploited through OD	E->P	D	G / J
240	**27,1**	Cities will be targeted once OD will be made available	E / A	A / C	--
241	**27,2**	We develop commercial solutions on top of the dataset	P	--	G / H
242	**27,3**	Councils need to be educated	E	C	--
243	**27,4**	Scalability crucial for growth	E->P A	E	H
244	**27,5**	- enables generalization	E->P A	E	H
245	**27,6**	- applicable to other cities (if data is made available)	E->P A	C / E	H
246	**28,1**	Cities need however to be pushed and educated of the benefits	E->P	C	J

247	**28,2**	Demonstrability crucial	P / A	--	H / J
248	**29,1**	Service fully scalable	E->P	E	H
249	**29,2**	Get the data key focus – it provides the content	E->P	A / C	G
250	**29,3**	Other cities can be designed quickly and easily	P / A	--	G / H
251	**29,4**	Demonstration to convince other councils	P / A	--	H / J
252	**29,5**	Engagement with innovation network – right communication and approach	E / A	B / D	--
253	**29,6**	Promotion within the OD ecosystem (presentations) to showcase benefits for cities	E->P A	B / D	J
254	**29,7**	Mega trends in the context (tourism): more informed decisions → local economic growth	P	--	J
255	**29,8**	- mobile (skills and opportunity)	E->P	B	H
256	**29,9**	- social (skills and opportunity)	E->P	B	H
257	**29,10**	- clouds (skills and opportunity)	E->P	B	H
258	**29,11**	- analytics (skills and opportunity)	E->P	B	H
259	**29,12**	Business network critical	E	B / D	--
260	**29,13**	Building chain for OD initiatives	E	D	--
261	**29,14**	Innovation network	E	D	--
262	**29,15**	Industry cluster	E	D	--
263	**29,16**	Industry led responses	E	D	--
264	**29,17**	This ecosystem to facilitate Open Data – cities should engage more	E	C / D	--
265	**29,18**	Social means engagement (e.g. reviews) that will be crowd sourcing for the city council	E->P	E	J
266	**29,19**	Analytics enable: track, pattern discovery, enjoyment level as new knowledge created	E->P	B / E	G / J
267	**29,20**	Analytics conducted at an OD level with API support	E->P	B / E	J
268	**29,21**	Much more information and knowledge for the CC	P	--	J
269	**29,22**	More informed decisions from CC	P	--	J
270	**29,23**	HTML5 key aspect for robustness and interoperability	E->P	B / E	I
271	**29,24**	Enabling live experiments and real-time OD (from analytics)	E->P	A / B / E	G / J
272	**29,25**	OD games because it is dynamic	E->P	E	G / I
273	**29,26**	Scale it in short time – but depends on data	E->P	A / B	H
274	**29,27**	Lack of data main issue – CC need commitment	E	A / C	--

275	**30,1**	Economic growth potential	P	--	H / J
276	**30,2**	Europe should play a key role	E	D	--
277	**30,3**	H2020 funding fundamental	E / A	D	--
278	**31,1**	Social value:	P	--	J
279	**31,2**	- local to meet visitors	P	--	J
280	**31,3**	- opportunity for locals (in underexploited areas)	E->P	C / E	J
281	**31,4**	Engage with local to give them touring opportunities	E->P	C / E	J
282	**31,5**	Information rich service	E->P	E	G / I
283	**31,6**	More Open Data to promote and maintain local traditional culture	E->P	A / C / E	G / J
284	**32,1**	Key is to incentivize people (locals) to:	E->P	E	I / J
285	**32,2**	- sell tour guides	P	--	J
286	**32,3**	- more money for them	P	--	J
287	**32,4**	- voluntary engagement of locals	P	--	K
288	**32,5**	Involving social is the key to commercialize	P / A	--	H / J
289	**33,1**	Barriers:	E	A / C	--
290	**33,2**	- lack of data	E	A / C	--
291	**33,3**	- lack of Wifi	E	C	--
292	**33,4**	- lack of funding from city (EU needed)	E	C / D	--
293	**34,1**	Funding at EU level	E / A	D	--
294	**34,2**	EU level programs needed	E	D	--
295	**34,3**	Create spaces for people to come together to facilitate re-usage	E->P	D	
296	**34,4**	City councils need to be encouraged	E / A	C	--
297	**34,5**	EU provides money	E / A	D	--
298	**34,6**	Volunteers are fundamental in the Case and local area can provide it	E->P	B / D	J
299	**34,7**	Funding needed for commercialization	P / A	--	K
300	**34,8**	One location for all initiatives	E	D	--
301	**34,9**	Submission to H2020 funding	E / A	D	--
302	**34,10**	Partners are crucial for ODS businesses	E	B / D	--
303	**34,11**	- academic	E	B / D	--
304	**34,12**	- tourist associations	E	B / D	--
305	**34,13**	- SMEs	E	B / D	--
306	**34,14**	- Local Innovation Network / Business Network	E	B / D	--
307	**34,15**	Massive scale is the goal for the ecosystem	E / A	A / D	--
308	**34,16**	Once developed is cheap to re adapt	P / A	--	G / H
309	**34,17**	With the data it can be done everywhere easy and quickly	P / A	--	G / H

310	**34,18**	From local to European funded one	E	C / D	--
311	**35,1**	ODS need to be useful as consumers are the key for success	P / A	--	H / I
312	**35,2**	OD ambassadors to:	E	C / D	--
313	**35,3**	- spread awareness from government side and consumer side	E	C / D	--
314	**35,4**	- stimulate the ecosystem	E	D	--
315	**35,5**	Each council should have its ambassador	E	C	--
316	**35,6**	City –led ecosystem building (to regularly provide data)	E	C / D	--
317	**35,7**	Community focus	E	D	--
318	**35,8**	Sharing at EU level	E	D	--
319	**35,9**	Knowledge sharing across Europe	E	D	--
320	**35,10**	- promote best practices	E	D	--
321	**35,11**	- joint efforts of similar contexts/services	E	D	--
322	**35,12**	Collective learning – city / ecosystem / users	E->P	C / D	I / J
323	**35,13**	Repository of best practices	E	C / D	--
324	**35,14**	- from which develop a repository of MVPs	E	C / D	--
325	**35,15**	Lean Start Up approach	E	B	--
326	**35,16**	- (build measure and learn) suitable for OD	E	B	--
327	**36,1**	One Open Data initiative(for city and EU)	E	C / D	--
328	**36,2**	- no fragmented projects	E	C / D	--
329	**36,3**	More Open Datasets	E	A	--
330	**36,4**	Focused strategy needed	E	C	--
331	**36,5**	Areas should be tackled in depth, not just the superficial data	E	C	--
332	**36,6**	All data about a context should be published from all departments of CC	E	A / C	--
333	**36,7**	Then new apps will be developed and used as demonstration for other councils	E->P A	D	H / J
334	**36,8**	Lack of funding (local and EU)	E	C / D	--
335	**36,9**	Lack of data	E	A / C	--
336	**37,1**	EU level Open Data incubators	E	D	--
337	**37,2**	EU to establish best practices for businesses and making profit	E->P	D	H
338	**37,3**	Cooperation across countries	E	D	--
339	**37,4**	EU to foster the ecosystem	E	D	--
340	**37,5**	OD should be EU driven (probably)	E	D	--
341	**37,6**	If they state the big potential they should invest (EU)	E	D	--

342	**37,7**	The innovation ecosystem needs to be fostered	E	D	--
343	**37,8**	EU ambassador/champions to go around cities and educate them how to do it	E	C / D	--
344	**37,9**	Innovation awards for cities embracing OD	E->P	C / D	J
345	**37,10**	Cities compete for being Pioneer of Open Data	E->P	C	J
346	**37,11**	From EU travel across cities to educate the councils	E->P	C / D	J
347	**37,12**	- to promote best practices	E	C / D	--
348	**37,13**	- to export toolkits	E	C / D	--
349	**37,14**	- to provide examples	E	C / D	--
350	**38,1**	Interaction with academia	E	D	--
351	**38,2**	Funding for SMEs	E	D	--
352	**38,3**	EU finding → business opportunity → OD useful	P	--	G
353	**38,4**	Cooperation with academia	E	D	--
354	**38,5**	Success rate is low	EXCLUDED		
355	**38,6**	Barrier to ecosystem: responsibility share/allocation	E	D	--
356	**38,7**	If the all community and CC gain awareness of the benefits it facilitates the process	E->P	C / D	G
357	**38,8**	Triple Helix approaches	E	D	--
358	**38,9**	- but needs to be improved	E	D	--

Appendix 5: Case Study C Codes and Evidences

Appendix 5.1: Case C Key Statements and Coding

EXTERNAL FACTORS CLUSTERS	
A	**Open Data and datasets characteristics**
B	**City Council related factors**
C	**Intermediary organization factors**
D	**Service Characteristics**
E	**Skills and Resources**
F	**Events / Community Factors (outside the city council)**
G	**City Factors**
H	**Individual Factors**
I	**City Council - Business Relationships**
L	**Other External Factors**

PERCEPTIONS CLUSTERS	
M	**Perceived Usefulness of Open Data**
N	**Perceived Business Opportunity**
O	**Perceived Ecosystem Awareness**
P	**Perceived User Satisfaction**
Q	**Perceived Commitment/Trust from City Council**
S	**Perceived City Opportunity**
T	**other Perceptions**

#	REF.	KEY CONCEPTS FROM SSI	E / P / A	Ext. Fac. Cluster	Perc. Cluster
1	**1,1**	Engagement businesses developers city council and researchers for ensuring commitment and spreading awareness	E->P	F / I / L	O / Q
2	**1,2**	Need of an unique initiative both CC and Int.	E	B / C	--
3	**1,3**	Workshops and public events also outside CC	E	B / C / F	--
4	**1,4**	Council and intermediary org. strategy	E	B / C	--
5	**2,1**	City council datasets	E	A / B	--
6	**2,2**	Dataset about parking was the most usable	E->P	A / D	M
7	**2,3**	Some data provided by the council but not Open Data yet (engaged with CC for release)	E	A / B / I	--
8	**2,4**	Lot of work on the data	E->P	A / E	M
9	**2,5**	Time consuming to re-use the data, skills needed	E->P	A / E	M
10	**2,6**	Council working at new updated datasets	E->P	A / B / I	M / Q
11	**2,7**	Company pushes council to publish data	E->P	B / I	M / Q

12	**2,8**	Multi-storey feed	E	A	--
13	**2,9**	Identified the useful source and now included into the service	E->P	A	M
14	**2,10**	Access was guaranteed before from relationships	E->P	A / B / I	M / Q
15	**2,11**	Revenue data of the meters (initially not public)	E	A / D / I	--
16	**2,12**	Lot of information can be structured from that for the city and the business	P	--	M / N / S
17	**2,13**	Trip information (initially not public)	E	A / D / I	--
18	**2,14**	Initially a sample of the dataset useful for:	E->P A	A	M
19	**2,15**	Modelling knowledge in the city, best business opportunity	E->P A	D	M / N
20	**2,16**	Test the datasets	E->P A	A	M
21	**2,17**	Discover new patterns and sell to the market	P	--	M / N
22	**2,18**	Traffic data also included in the service	E	A / D	--
23	**3,1**	Intermediary organization provides Open Data	E	A / C	--
24	**3,2**	The rest comes from the council	E	A / B / I	--
25	**4,1**	Open Data in machine readable format enables quick time development	E->P	A	M
26	**4,2**	csv files format	E	A / E	--
27	**4,3**	Not all are machine readable (but should be)	E	A	--
28	**5,1**	Promotion of existing businesses	E	D	--
29	**5,2**	Personalised service (for parking operators)	E->P	D	P
30	**5,3**	Basic service Free, Premium Option	E->P	D	N
31	**5,4**	Tier based revenue model	E->P	D	N
32	**5,5**	Flexible revenue model	E->P	D	N
33	**6,1**	Payment support system as a service for the city	E->P	D	S
34	**6,2**	Tier based revenue model in relation to client dimension (for satisfying users)	E->P	D	N / P
35	**7,1**	Privacy in the data and in the service	E->P	A / D	P
36	**7,2**	Choose between monthly fee and transaction fee	E->P	D	N / P
37	**7,3**	Revenues based on the amount of client supplied	E->P	D	N / P
38	**8,1**	Scalable central platform and personalisation	E->P	D / E	N / P
39	**8,2**	Personalization of the service	E->P	D	N / P
40	**8,3**	Business model conceptualization	P / A	--	N
41	**8,4**	Central platform to protect the brand (logo and structure of the app) and establish business	E->P	D	N
42	**8,5**	Central Platform to develop solutions easily and quickly	E->P	D	M
43	**8,6**	Customised app as additional cost	E->P	D / E	N / P
44	**9,1**	Three full time	E	E	--

45	**9,2**	Two interns	E	E	--
46	**9,3**	Volunteers engagement	E	E / L	--
47	**9,4**	Eight people in 2012	E	E	--
48	**9,5**	Software development skills	E	E	--
49	**9,6**	Business development skills	E	E	--
50	**9,7**	Individual entrepreneurial spirit	E	H	--
51	**9,8**	Attitude towards Software-oriented business	E	H	--
52	**9,9**	Product Design skills	E	E	--
53	**9,10**	Communication and Social Media Skills	E	E	--
54	**9,11**	Service ecology and usability support	E	E	--
55	**9,12**	Data scientist	E	E	--
56	**9,13**	Data scaling	E	E	--
57	**9,14**	Information protection, need of internal knowledge	E	A / E	--
58	**9,15**	Database management skills	E	E	--
59	**9,16**	Huge amount of information	E->P	E	M
60	**9,17**	Lean team	E	E	--
61	**9,18**	Data Analytics	E	E	--
62	**10,1**	Idea generation in 2011	A	--	--
63	**10,2**	Launch end of 2012	A	--	--
64	**10,3**	Private Pilot (private beta) in end of 2012	A	--	--
65	**10,4**	(List of milestones from website: generate adoption statements)	A	--	--
66	**10,5**	May 2012: Started Phase 1 of [program]	A	--	--
67	**10,6**	July 2012: Started Phase 2 of [program]	A	--	--
68	**10,7**	Nov 2012: Launched Registration for Private Beta of Case C	A	--	--
69	**10,8**	Dec 2012: Launch of private beta	A	--	--
70	**10,9**	Feb 2013: xxx Launch pad Finalists	A	--	--
71	**10,10**	May 2013: Winners City Beta – People's choice	A	--	--
72	**10,11**	June 2013: Launch xxx as public beta	A	--	--
73	**11,1**	Initially small team (two co-founders)	E	E / H	--
74	**11,2**	Open beta released in 2013	A	--	--
75	**11,3**	In 2013 partnership with newspaper	E / A	L	--
76	**11,4**	Later partnership with major news agency	E / A	L	--
77	**11,5**	The goal was to use data to reach city drivers where parking problems are in place	E->P	G	M / N
78	**11,6**	Free advertisement for one month (worth Euro38,000)	E	L	--
79	**12,1**	Hack competition fundamental for:	E->P A	B / F	M / O / Q
80	**12,2**	Knowing open data and educate the councils	E->P A	B / F	M
81	**12,3**	Leverage previously existing idea and think how to commercialize	E->P A	F	M/ N

82	**12,4**	Idea generated from the data available	E->P A	A / F	M
83	**12,5**	Idea was on taxies before then adapted to parking because this was the data available	E->P A	A / F	M
84	**12,6**	Open Data provides the content	E->P	A	M
85	**12,7**	Open Data resource saving for business and the city	P	--	M / S
86	**12,8**	Open Data time saving for business and the city	P	--	M / S
87	**13,1**	Previously aware of Open Data but not interested (because no commercial opportunity)	P / A	--	N
88	**13,2**	Interest from council initiative (saw commitment and awareness spread in society)	P / A	--	O / Q
89	**13,3**	Councils to show the value of Open Data	E->P	B	Q
90	**13,4**	Cash prices for incentivizing developers	E->P	B	Q
91	**14,1**	Multiple ideas before hack competition / awareness from community and commitment was a push	P / A	--	O / Q
92	**14,2**	Idea generated from Open Data available	P / A	--	M
93	**14,3**	Open Data selected because of the amount of content within it	E->P	A / F	M
94	**14,4**	Biggest dataset was chosen	E->P	A / F	M
95	**14,5**	Majority of data was about this context	E->P	A / F	M
96	**14,6**	Perception that it was a push from the city council to develop solutions in the context	P	--	M / N / Q
97	**14,7**	A previously designed solution was leveraged with Open Data for parking	E / A	A / F	--
98	**15,1**	Open Data provides content	E->P	A	M
99	**15,2**	Location fundamental	E	A	--
100	**15,3**	Standard important	E	A	--
101	**15,4**	Unified location standard	E	A	--
102	**15,5**	Need for meta data to describe and navigate the dataset	E->P	A	M
103	**15,6**	Longitude and Latitude should be used	E	A	--
104	**15,7**	GIS skills	E	A / E	--
105	**15,8**	Data cleansing skills	E	E	--
106	**16,1**	Open Data accuracy	E	A	--
107	**16,2**	Long time data cleaning (should be faster)	E->P	A / E	M
108	**16,3**	Datasets should have meta-data	E->P	A	M
109	**16,4**	Important to display all the standards for all the characteristics of the dataset (format, location)	E->P	A	M
110	**16,5**	Council should be responsible of the accuracy of data	E->P	A / B	M / Q
111	**17,1**	Since councils not responsible, it was put on the back of the drivers	E	B / D	--
112	**17,2**	Regulation compliance (not use the app while driving)	E	D	--
113	**17,3**	Council should update the data in real-time (clearways)	E->P	A / B / D	M / Q
114	**17,4**	Real-time Open Data	E	A	--
115	**17,5**	The city payment service is integrated in the app	E	D	--

116	**17,6**	Partnerships with competitors	E	L	--
117	**17,7**	MVP development	A	--	--
118	**17,8**	Payment integrated to derive additional information and charge transactions	E->P	D	N / Q / S
119	**17,9**	Drivers patterns in the city can be detected and sold or shared with CC	P	--	M / N / S
120	**17,10**	Will to structure behaviours of drivers in the city from using analytics on the service	E->P	D / E	M / N / S
121	**17,11**	Value for city council for parking infrastructure management	P	--	S
122	**17,12**	From meters transactions knowledge generated for them because of partnership	E->P	D / I	M / S
123	**17,13**	This was not Open Data initially	E / A	A / B / I	--
124	**18,1**	Councils not responsible	P	--	Q
125	**18,2**	Instead of the enabler Open Data is seen as a commitment for helping developers	P	--	M / Q
126	**18,3**	City related factors:	E	G	--
127	**18,4**	Payment for parking	E	G	--
128	**18,5**	Parking congestion	E	G	--
129	**18,6**	Long time to find a parking space	E	G	--
130	**18,7**	Users satisfaction achieved	P	--	P
131	**19,1**	No contracts with the city council	E	I	--
132	**20,1**	Partnership would be beneficial for ensuring commitment	E->P	I	Q
133	**20,2**	Commercial relationship was the goal	E->P	I	M / N / Q
134	**20,3**	Commercial relationship negative for risk of lock in	E->P	I	N / Q
135	**20,4**	Thinking studio with entrepreneurs as brainstorming sessions from council	E	B / H	--
136	**20,5**	Invited the company for looking at parking solutions	E->P	B / I	N / Q
137	**20,6**	Commercial relationships existing made insecure perceptions	E->P	B / I	N / Q
138	**20,7**	Procurement rules are limitations for businesses	P	--	N / Q
139	**20,8**	This impeded the supply of commercial services to the council	E->P	B / I	N / Q
140	**21,1**	Real-time data	E	A	--
141	**21,2**	More datasets from information that can be gathered from existing systems (meters) but lack	E->P	A / B	M / Q
142	**21,3**	Commercially sensitive (barrier)	P	--	N / Q
143	**21,4**	IT skills in the city council	E	B	--
144	**21,5**	Up to date software for city councils	E	B	--
145	**21,6**	Sensors data from council initiative	E	A / B	--
146	**21,7**	Infrastructure investments	E	B	--
147	**21,8**	Who pays? Who maintains them? (barrier)	P	--	Q
148	**21,9**	Real-time feeds	E	A	--
149	**21,10**	Weather and climate data	E	A	--

150	**21,11**	Opportunity for city councils to avoid damages	P	--	M / S
151	**21,12**	Increase efficiency in the city councils	P	--	M / S
152	**21,13**	Predict potential problems (flooding and parking) is another opportunity	P	--	M / N / S
153	**21,14**	More informed decisions enabled for us and the council	P	--	M / N / S
154	**21,15**	Data integration for structuring knowledge in the context	E->P	A / D	M / N / S
155	**22,1**	Investment needed	E	B / L	--
156	**22,2**	Organizations need to connect the governments and re-users	E->P	B / C	O / Q
157	**22,3**	Should be no profit	E->P	C	O
158	**22,4**	Other countries showed it works	E->P	C	O
159	**22,5**	Should be independent	E	C	--
160	**22,6**	Independent	E	C	--
161	**22,7**	Experts dedicated to Open Data	E	C	--
162	**22,8**	Counter part of the city councils	P	--	O / T
163	**23,1**	If too close to the council barrier	E->P	C	O / T
164	**23,2**	Open Data portals need to be effectively managed	E	A / C	--
165	**24,1**	Legal Openness of the data	E	A	--
166	**24,2**	Academic say "the license is holding us back"	P	--	M
167	**24,3**	Restrictions should be limited to privacy concerns	E->P	A	M
168	**24,4**	PSI directive was in place at the beginning	E	A	--
169	**24,5**	Commercial openness critical to sell data on the back of Open Data Services	E->P	A / D	M / N
170	**24,6**	Tokens supply is a "way around that"	E->P	A / D	M / N
171	**24,7**	Open License like creative commons for commercial	E->P	A	M / N
172	**24,8**	Re-usage open license	E	A	--
173	**24,9**	Not enough commitment from city council for new policy making	E->P	B	M / Q
174	**24,10**	CEO employed outside the case to promote open licensing	E	F / H	--
175	**24,11**	Councils care if it causes additional workloads or costs	P	--	Q / S
176	**24,12**	Legal licensing critical in planning future activities	E->P	A	M /N
177	**24,13**	Partnership with national parking association	E	L	--
178	**25,1**	Opportunity to deliver data back to the council	P	--	Q / S
179	**25,2**	Not undertaken at the moment because of previously established commercial relationships	E->P	B / I	Q / S
180	**26,1**	Win-win relationship should be established	E	I	--
181	**26,2**	But commercial relationships big barrier	E->P	B / I	Q
182	**26,3**	Lot of value potentially provided to the council	P	--	Q / S
183	**26,4**	But: presentation at the council about an opportunity they handled it back to the partner	P	--	Q /T

184	**26,5**	Relationships saw idea translated in suppliers profits	P	--	Q / T
185	**26,6**	No exclusivity of the data barrier	P	--	Q / T
186	**26,7**	Risk of copying solutions at councils events	P	--	T
187	**26,8**	Keep key ideas internally	E->P	I / L	T
188	**26,9**	More incentives for collaboration	E	B / I	--
189	**27,1**	Open Data main requirement for the service	E->P	A / D	M
190	**27,2**	Open Data hidden from the website	E->P	D	T
191	**27,3**	Open Data leveraged to talk to the council	P	--	M / T
192	**27,4**	Open Data leveraged to enter the context	P	--	M / T
193	**28,1**	Local promotion from the council	E	B	--
194	**28,2**	Council awareness on who and how is using the data	E->P	B	O / Q
195	**29,1**	Data cleaning	E	E	--
196	**29,2**	Data integration	E	E	--
197	**29,3**	Graphic Design	E	E	--
198	**29,4**	Mobile programming skills	E	E	--
199	**29,5**	Product design	E	E	--
200	**29,6**	Data Analytics	E	E	--
201	**29,7**	Data mining	E	E	--
202	**29,8**	Database management	E	E	--
203	**29,9**	Data Visualisation	E	E	--
204	**30,1**	Service easy to use	P	--	P
205	**30,2**	Easy to detect (also OD)	P	--	M / P
206	**30,3**	Collect information that users seed (feedback loop)	E->P	D	M / N / S
207	**30,4**	A lot of information can be captured	P	--	M / N / S
208	**30,5**	Push notifications	E	D	--
209	**30,6**	Understand how people use cities	P	--	M / N / S
210	**30,7**	Help the utility and what we can get on the back of this	P	--	M / S
211	**31,1**	The main business and city opportunity is from the data you can extract	P	--	N / S
212	**31,2**	Detail knowledge creation to be sold to retailers	E->P	D	M / N
213	**31,3**	Confirmed by extensive market analysis	E->P	G	N
214	**31,4**	Opportunity for retailers to push notifications in relation to the users' habits	P	--	N
215	**31,5**	User base is critical	P / A	--	N
216	**31,6**	Existing system was small and inefficient	E	G	--
217	**32,1**	Personal investments	E	E	--
218	**32,2**	3 people	E	E	--
219	**32,3**	150,000 Euros for product development, trade shows, designers etc.	E	E	--
220	**32,4**	Local funding provision (hack competition)	E	F	--
221	**32,5**	Local Funding provision (Enterprise associations)	E	F / L	--

222	**32,6**	Lean Start up approach	E	L	--
223	**33,1**	Expansion in different cities	P / A	--	N
224	**33,2**	Competitiveness increase for cities	P	--	S
225	**33,3**	Opportunities based on the problem occurrence in the cities	E->P	D / G	M / N / S
226	**33,4**	Important having a scalable solution	E->P	D	N
227	**33,5**	Language can't be scalable	E->P	D / L	N
228	**33,6**	Currencies can't be scalable	E->P	D / L	N
229	**33,7**	Localisation across Europe achieved with partners	E / A	A / D / L	--
230	**33,8**	5,000 car parks in Europe will have a white label solution	E / A	D / L	--
231	**33,9**	Most European languages are covered	E	D / E / L	--
232	**33,10**	The partnership brings the opportunity of integrating these localisation in the Open Data Service	E->P	D / G / L	N
233	**33,11**	Also public parking such as in train stations, airports, ports etc. are covered.	E	A / D	--
234	**34,1**	Testing in new cities always first step	A	--	--
235	**34,2**	Ask for the data to the councils	A	--	--
236	**34,3**	Show as demonstration what they did originally	P / A	--	Q
237	**34,4**	Need for the standard is highlighted	E	A	--
238	**34,5**	Cities that have open data programs are targeted	E->P	B / G	M / Q
239	**34,6**	Cities that have sensed data too	E->P	B / G	M / Q
240	**34,7**	Cities can be met in international workshops ("trade shows")	E	B / F	--
241	**34,8**	First thing to look if it is possible to work with the councils	P / A	--	N / Q
242	**34,9**	Language barrier	E	A / L	--
243	**34,10**	Partners for localisation (language)	E->P	L	N / T
244	**34,11**	Strategy to provide the ODS and up sell the white label solution	P / A	--	M / N / P
245	**35,1**	Open Datasets should be structured	E	A	--
246	**35,2**	Online portals	E	A / B	--
247	**35,3**	Value to transparency and Open Data from councils	E->P	B	M / S
248	**35,4**	City councils need to be educated	E->P	B / F	Q
249	**35,5**	From Open Data potential (Open government)	E->P	B / F	Q / T
250	**35,6**	From the IT perspective	E->P	B / F	Q
251	**36,1**	Open Data as a badge of honour	P	--	M
252	**36,2**	European level strategy	E	L	--
253	**36,3**	Community awareness (investors)	E->P	B / F / L	O
254	**36,4**	Trust in the city council (investors)	E->P	L	T
255	**36,5**	Investment at EU level	E	L	--

256	**36,6**	Guarantee that information will be available in the future	E->P	A / B / C	Q / T
257	**36,7**	Commitment by city councils	E->P	B	Q
258	**36,8**	Open Data is ready (risk)	E->P	A / D / E	T
259	**36,9**	Value is on top of Open Data (collect the data and structure it manually would be more worth)	P	--	M / T
260	**37,1**	Risk in communicating the solutions	P	--	Q / T
261	**37,2**	Open Data is integrated with internal data collection processes	E	D / E	--
262	**37,3**	Lack of trust in sustainability of Open Data Provision	E->P	B / C / L	T
263	**38,1**	Inaccuracy of data but can be fixed	E->P	A / E	M
264	**38,2**	Relationship with the council main obstacle	E->P	I	Q
265	**38,3**	Educate councils about:	E->P	B / F / I	Q
266	**38,4**	Open Data needed (need right language)	E	A / B / F / I	--
267	**38,5**	What and how the company is doing the service	E->P	B / F / I	Q
268	**38,6**	First mover advantage is not in place	P / A	--	T
269	**38,7**	Relationship is needed if you want the data	E->P		Q
270	**38,8**	Sell cycle to the council is too long	P	--	N / Q
271	**39,1**	Commitment from city councils	E->P	B / I	Q
272	**39,2**	Champions in city councils	E	B / I	--
273	**39,3**	External champions as well as internal have limited roles for ensuring fair competition	E->P	B	O / Q
274	**40,1**	Procurement issues	P	--	T
275	**40,2**	Competition laws an issue	P	--	T
276	**40,3**	Workshops as a mean to stimulate councils to evolve in this way	E->P	F	O
277	**40,4**	Small contracts for piloting a solution from councils	E / A	B / I	--
278	**40,5**	Cash incentives from councils	E	B / I	--
279	**40,6**	Councils states the problems	E	B / I	--
280	**40,7**	Shared investments between ODS businesses and councils (for pilot stage)	E / A	I	--
281	**41,1**	Availability of Open Data	E->P	A / B / C	M
282	**41,2**	Scalable solutions	E->P	D	N
283	**41,3**	Hack competitions	E	F	--
284	**41,4**	Sustainable long-term plans	E->P	B / C / L	Q
285	**41,5**	Integrated and singular Open Data program	E	B	--
286	**41,6**	Focus on quality and not on quantity (see CSO discussion)	P	--	M / Q
287	**41,7**	Just competition to state that the council is working on Open Data is unsustainable	E->P	B / F	Q

288	**41,8**	Focus on tangible value	E	A / B / C / L	--
289	**41,9**	Too much focus on job creation	P	--	Q
290	**42,1**	Public Private Partnership needs a sustainable business model	E->P	B	Q
291	**42,2**	Government shouldn't do the service development work	E	B	--
292	**42,3**	There would be not balance with the public sector:	E	B	--
293	**42,4**	They are paid to do it	E	B	--
294	**42,5**	Private people lack resources	E	B	--
295	**42,6**	Same opportunity should be given to everyone	P	--	T
296	**42,7**	Councils should incentivize ODS businesses to educate them	E->P	B	Q
297	**42,8**	Councils should incentivize ODS businesses to take part in community initiatives	E->P	B	O / Q
298	**42,9**	Councils should incentivize ODS businesses to take part in workshops and public meetings	E->P	B	O / Q
299	**42,10**	Private businesses have limited time and resources but no big pushes to local authorities	E->P	I	Q
300	**43,1**	Opportunity cost if participating to events and workshops	P	--	O
301	**43,2**	Volunteers in the company	E	E	--
302	**44,1**	Cities are targeted if:	E->P A	B / G	N
303	**44,2**	There is an open data program	E-A	B / G	--
304	**44,3**	There is opportunity to engage with the council	E->P A	B / I	N
305	**44,4**	Opportunity for partnerships	E->P A	G / L	N / O / Q
306	**44,5**	Presence of investors	E->P A	G / L	N / O
307	**44,6**	Based on traditional market research	E->P A	G	N
308	**44,7**	Market size	E->P A	G	N
309	**44,8**	Smart-phones penetration	E->P A	G	N
310	**44,9**	Socio-economic status of the city	E->P A	G	N
311	**44,10**	Traffic congestions	E->P A	G	N
312	**44,11**	Number of cars driven in the city	E->P A	G	N
313	**44,12**	Relative number of parking spaces	E->P A	G	N
314	**44,13**	Open Data follows the market research	P / A	--	N / Q
315	**45,1**	It was asked because it wasn't available for some regions	P / A	--	T

316	**45,2**	They knew the council has this data	E->P	A / B	O
317	**45,3**	The council was approached	A	--	--
318	**45,4**	Different departments had different datasets in terms of:	E	A / B	--
319	**45,5**	Content	E	A / B	--
320	**45,6**	Format	E	A / B	--
321	**45,7**	Structure	E	A / B	--
322	**45,8**	Frequency of update	E	A / B	--
323	**45,9**	Real-time (live) data was not available	E	A	--
324	**45,10**	But it was a webpage	E	A	--
325	**45,11**	At an event from the intermediary org. the dataset was asked and given	E / A	A / B / C	--
326	**45,12**	No questions were asked	P	--	Q
327	**46,1**	The head of IT (no CIO in place) was approached	E	B	--
328	**46,2**	The data was coming from another department (urban planning)	E	A / B	--
329	**46,3**	For the departments that were not at the event a phone call was made	E	B	--
330	**46,4**	They gave the data (transport department)	E / A	A / B	--
331	**47,1**	Lot of work to be done because it was a KML file	E->P	A / E	M
332	**47,2**	Too much text	E->P	A	M
333	**47,3**	Map file	E->P	A	M
334	**47,4**	Not detailed information but general	E->P	A	M
335	**47,5**	Only parking areas not parking slots	E	A	--
336	**47,6**	Slots were derived by the case	E	A / E	--
337	**47,7**	Geo-codes were reversed and lot of work to see how many spaces	E->P	A / E	M
338	**47,8**	Field work was done to improve the dataset	E / A	A / D / E	--
339	**48,1**	The dataset was then created by the case	A	--	--
340	**48,2**	Only raw data was available	E	A / B	--
341	**49,1**	Results and improved dataset not shared with the council	P	--	Q / S
342	**49,2**	Because they never asked for it	P	--	Q /S
343	**50,1**	In addition to the data provided by the council, the one from the intermediary organization was:	E->P	A / C	M / O
344	**50,2**	Legally open	E->P	A / C	M
345	**50,3**	Meta-data provided on how to use it	E->P	A / C	M
346	**50,4**	Justification on why they needed the city council non-open data was asked and provided	P	--	Q
347	**51,1**	The one from I.O. was more structured	E->P	A / C	M
348	**51,2**	Re-usability was facilitated	E->P	A / C	M
349	**51,3**	Datasets were understandable	E->P	A / C	M
350	**51,4**	The rest was raw data	E	A / B	--

351	**52,1**	I.O. data had only to be cleaned	E->P	A / C / E	M
352	**52,2**	The other had to be also interpreted	E->P	A / B / E	M
353	**53,1**	The I.O. one was more thoughtful	P	--	M
354	**53,2**	More work put into it to refine it	E->P	A /C	M
355	**54,1**	Dataset characteristics:	E	A	--
356	**54,2**	Machine readability	E	A	--
357	**54,3**	Format / standard	E	A	--
358	**54,4**	No pdfs	E	A	--
359	**54,5**	Frequency of update (real-time not fundamental but once a year the data has to be cleaned)	E	A	--
360	**54,6**	Supported by API	E	A	--
361	**54,7**	Linked data	E	A	--
362	**54,8**	Usefulness in terms of providing a picture of the city	P	--	M
363	**54,9**	Effectiveness of reporting	P	--	M
364	**54,10**	Structured	E	A	--
365	**54,11**	Updatable makes it useful	E->P	A	M
366	**54,12**	Updated quickly and easily	P	--	M
367	**55,1**	Relationship with the council helpful	E->P	I	Q
368	**55,2**	Lack of certainty about the future	P	--	T
369	**55,3**	Lack of long-term commitment (lack of guarantees)	E->P	B / I	Q / T
370	**55,4**	Investors worried about scaling the solution	P	--	N
371	**55,5**	Investors worried about the availability of the data in the right forms	P	--	T
372	**55,6**	Investors lack trust in the council	P	--	T
373	**55,7**	Concerns about other small cities	P	--	T
374	**55,8**	Big risk if existing relationship is in place	P	--	Q / T
375	**55,9**	Partners need to be opened for collaborations	E->P	I / L	O / T
376	**55,10**	Company relied on the existing partners	E->P	B / I / L	T
377	**56,1**	Commercial relationship:	E->P	I	Q
378	**56,2**	Companies are paid to manage the context	E->P	I	M / Q
379	**56,3**	Companies are given the data needed	E->P	I	Q
380	**56,4**	Integrating more open data into the existing service was the strategy to convince CC	E / A	D / I	--
381	**56,5**	But big risk of being replaced by the original company	P	--	N / Q / T
382	**56,6**	The partner has not blocked the case but they could	P	--	T
383	**56,7**	Investors worried about how the case controls this market	P	--	T
384	**56,8**	There is no will to control the market jut to "exist" and be able to provide this additional service.	P	--	T
385	**57,1**	I.O. should be connected with academia	E	C	--

386	**57,2**	But the council runs it so:	E->P	B / C	Q / T
387	**57,3**	No credibility	P	--	T
388	**57,4**	No challenge for city council	P	--	T
389	**57,5**	I.O. should be not for profit organizations (ODI)	E->P	C	O
390	**57,6**	Independent body from the city council to be:	E->P	B / C	O
391	**57,7**	More connected to the user community	E->P	C	O
392	**57,8**	Flexible in supporting businesses	E->P	C	O
393	**57,9**	More successful	E->P	C	O
394	**58,1**	If too closely related to the council:	P	--	O
395	**58,2**	Complacency	P	--	O / T
396	**58,3**	Relationships are not sustainable for competition	P	--	O / T
397	**59,1**	I.O. should be able to challenge the council	P	--	O / T
398	**59,2**	To be more vocal	E->P	C	O
399	**59,3**	The businesses are but the I.O. should be too	E->P	C	O
400	**60,1**	City councils need to be educated at events	E->P	B / F	O / Q
401	**60,2**	Events are the mean to do that	E	B / F	--
402	**60,3**	Events are crucial for:	E->P	B / F	O / Q
403	**60,4**	Educating the councils	E->P	B / F	O / Q
404	**60,5**	Ask for effort in directions of the services	E->P	B / F	O / Q
405	**60,6**	Engage with the councils	E->P	B / I / F	O / Q
406	**60,7**	Demonstrate value of Open Data and what can be done if...	E->P	B / F / L	O / Q
407	**60,8**	Stimulate emergence of new ideas for the city	E->P	B / F	O / S
408	**60,9**	Ask help to the councils	E->P	B /F	O / Q
409	**60,10**	The company presented the idea at council's events	E	B	--
410	**60,11**	Presented what they were doing and what needs	E	B	--
411	**60,12**	But meetings outside the councils are more effective:	E->P	F	O
412	**60,13**	The council is invited	E->P	F	O /Q
413	**60,14**	A comprehensive community is present	E->P	F	O
414	**60,15**	Co-collaboration enabled	E->P	F	O / Q
415	**60,16**	Users and potential users can create events	E->P	F	O
416	**60,17**	Face to face interactions is important	E->P	F	O / Q
417	**60,18**	Workshops	E	F	--
418	**61,1**	Attribution not a big issue for the company and the service	E->P	A / D	M
419	**61,2**	Commercial contracts are real problem	P	--	Q / T
420	**61,3**	Legal openness should allow everyone the possibility of developing a service	E->P	A	M / Q / T

421	**62,1**	Lock in situation to be avoided: they retain potential innovations	P	--	Q / T
422	**63,1**	Incentives should be given to businesses contributing to the ecosystem (events)	E	B / F / L	--
423	**63,2**	Structure within the council	E	B	--
424	**63,3**	Unique strategy	E	B	--
425	**63,4**	No fragmented projects	E	B	--
426	**63,5**	Unique effort across departments	E	B	--
427	**63,6**	Open Data established committee	E	B	--
428	**64,1**	Real-time data	E	A	--
429	**65,1**	Trade off of apps developed by councils and public bodies:	E->P	B / L	Q
430	**65,2**	PLUS they gain credibility	P	--	Q / T
431	**65,3**	MINUS they usually own it and are closed to collaborations	P	--	Q
432	**65,4**	They should entail partners engagement	E	B / L	--
433	**65,5**	Government should not be the only leaders	E	B / C / L	--
434	**65,6**	Collaboration and co-creation are fundamental	E	B / L	--
435	**66,1**	In councils:	E->P	B	Q
436	**66,2**	Waste of money	E	B	--
437	**66,3**	Fragmented structures	E	B	--
438	**66,4**	Not individual budgets but council's budget	E	B	--
439	**66,5**	Central fund body	E	B	--
440	**66,6**	Central Open Data committee	E	B	--
441	**66,7**	With its own budget	E	B	--
442	**66,8**	Unique effort across departments	E	B	--
443	**66,9**	With IT skills	E	B	--
444	**66,10**	Aware of the Open Data users' needs	E->P	B	O / Q

Appendix 5.2: Adoption Stages

ADOPTION STAGES					
Stage		**Key Statements**	**Internal Documents**	**External Documents**	**Validated**
1	*Idea Generation*	10,1; 12,4; 12,5; 13,1; 14,1;	--	D5; D8;	Yes
2	*Open Data Adoption*	10,5; 10,6; 10,9; 12,1; 12,2; 12,3; 12,4; 12,5; 13,1; 13,2; 14,2; 14,7; 17,13;	D4;	D8;	Yes
3	*Market Analysis*	44,6; 44,7; 44,8; 44,9; 44,10; 44,11; 44,12;	D2;	D5; D8; D9; D10;	Yes
4	*Private Beta Release*	2,14; 2,15; 2,16; 10,3; 10,7; 10,8; 10,10; 17,7; 40,4; 40,7; 47,8;	D2; D4;	D5; D8;	Yes
5	*Public Beta Release*	10,2; 10,11; 11,2; 38,6; 45,11; 46,4; 48,1;	D4;	D8;	Yes
6	*Commercialization*	8,3; 31,5; 34,11;	D4;	D5; D8; D9; D10;	Yes
7	*Growth*	11,3; 11,4; 33,1; 33,7; 33,8; 34,3; 34,8; 44,1; 44,2; 44,3; 44,4; 44,5; 45,1; 45,3;	--	D8;	Yes
8	*Sustained Use*	2,1; 2,2;	D4;	D6; D7; D8; D10;	Yes

Appendix 6: Cross-Case Assertions

Appendix 6.1: Cross-case Assertion 2, Individual Factors

CROSS-CASE ASSERTION2					
Variables	**CASE A**	**CASE B**	**CASE C**	**Retained**	**Validated**
Experiences in the Context	x	x		No	Yes
Attitude to Work with Governments	x			Yes	No
Attitude towards Open Data and IT-oriented Business	x	x	x	No	Yes
Entrepreneurial Spirit	x		x	No	Yes
Individual Creativity		x		Yes	No

Appendix 6.2: Cross-Case Assertion 3, City-Related Factors

CROSS-CASE ASSERTION 3					
Variables	**CASE A**	**CASE B**	**CASE C**	**Retained**	**Validated**
Size of the city / Market Size	x		x	No	Yes
Public Free Wi-Fi		x		Yes	**No**
Infrastructural Characteristics		x	x	No	Yes
*City Related Opportunity**		x	x	No	Yes

*Integrated after Option 3 of External Factors Cross-Case Analysis.

Appendix 6.3: Cross-Case Assertion 4, City Council-Related Factors

CROSS-CASE ASSERTION 4					
Variables	**CASE A**	**CASE B**	**CASE C**	**Retained**	**Validated**
Open Data Strategy / Commitment	x	x	x	No	Yes
Open Data Awareness	x	x	x	No	Yes
Support Creation and Sustainability of Open Data Services	x			Yes	No
Organizational Structure / IT skills and Department / City CIO	x		x	No	Yes
Budget and Resources	x	x	x	No	Yes
Engagement Initiatives	x	x	x	No	Yes
Internal Laws	x			Yes	No
Outsourcing of Open Data Services	x			Yes	Yes

Appendix 6.4: Cross-Case Assertion 5, Local Ecosystem Factors

CROSS-CASE ASSERTION 5						
Variables	**CCTA**	**CASE A**	**CASE B**	**CASE C**	**Initially Retained**	**Validated**
Hack Competitions	*Local Community Factors*	x	x	x	No	Yes
Champions		x	x	x	No	Yes
Community-Oriented Initiatives		x	x	x	No	Yes
Volunteers			x		Yes	Integrated
Intermediary Organization	*Open Data Ecosystem*		x	x	No	Yes
Partners Engagement			x	x	Yes	**No**
Volunteers Engagement			x	x	No	Yes

Appendix 6.5: Cross-Case Assertion 6, Organizational Skills

CROSS-CASE ASSERTION 6					
Variables	**CASE A**	**CASE B**	**CASE C**	**Retained**	**Validated**
Programming Skills	x	x	x	No	Yes
Data Analytics and Data Management Skills	x	x	x	No	Yes
Geographic Information Systems Skills	x		x	No	Yes
Graphic Design and Data Visualisation Skills	x	x	x	No	Yes
Project Management / Communication Skills	x	x	x	No	Yes
Knowledge of Legal and Policy Context	x		x	No	Yes

Appendix 6.6: Cross-Case Assertion 7, City Council – Business Relationship

CROSS-CASE ASSERTION 7				
VARIABLES	**CASE A**	**CASE C**	**Retained**	**Validated**
Win-win relationship	x	x	No	Yes
Contracts / Structure of the Relationship	x	x	No	Yes
Collective Learning		x	Yes	No

Appendix 6.7: Cross-Case Assertion 8, Service Characteristics

CROSS-CASE ASSERTION 8				
VARIABLES	**CASE B**	**CASE C**	**Retained**	**Validated**
Unique content Management System	x	x	No	Yes
Secondary Data Generation	x	x	No	Yes
Incentivizing / Challenging	x		Yes	**No**
Enable Interactive Experience	x		Yes	**No**
Engagement with Local Community	x		Yes	**No**
Attractive Technology	x		Yes	**No**
Addressing Users' Needs	x		Yes	**CCA11,1**
Open Source Software	x		Yes	**No**

Cloud-Based	x		Yes	**No**
Interoperability	x		Yes	**No**
Valuable Service in the Context		x	Yes	**CCA3,3**
White Label Solution		x	Yes	**No**
Flexible Revenue Model		x	Yes	**No**
Complexity		x	Yes	**No**

Appendix 6.8: Cross-Case Assertion 10, Perceived Usefulness of Open Data

CROSS-CASE ASSERTION 10					
Variables	**Case A**	**Case B**	**Case C**	**Retained**	**Valid.**
Provision of Content	x	x	x	No	Yes
Time Saving / Short Time Development	x	x	x	No	Yes
Resource Saving	x		x	No	Yes
Enabling Dynamic Reliable Information	x	x		No	Yes
"Opens the Councils' Doors"/"Badge of Honour"	x		x	No	Yes
Free of Charge		x		Yes	No
Large Datasets Analysable and Shareable		x		Yes	No
Easier Process		x		Yes	No
Discoverability and Understandability			x	Yes	Yes

Appendix 6.9: Cross-Case Assertion 11, Perceived Business Opportunity

CROSS-CASE ASSERTION 11					
Variables	**Case A**	**Case B**	**Case C**	**Retained**	**Validated**
Benefits for End-Users / Market Opportunity	x		x	No	Yes
Partnership Strength	x			Yes	No
Level of Engagement	x			Yes	No
Scalability Across Cities	x	x	x	No	Yes
Demonstrability	x	x		No	Yes
Funding		x		Yes	Yes
User Base		x		No	Yes
Knowledge potentially derived			x	No	Yes
Revenue Model Structure			x	Yes	No

Appendix 6.10: Cross-Case Assertion 11 Links

CCA11 Links	**Case A**	**Case B**	**Case C**	**Validated**
Service Characteristics → Perceived Business Opportunity		x	x	Yes
Individual Factors → Perceived Business Opportunity	x	x	x	Yes
City-Related Factors → Perceived Business Opportunity	x	x	x	Yes
Perceived Business Opportunity → Perceived Usefulness of Open Data	x	x	x	Yes

Appendix 6.11: Cross-Case Assertion 12, Perceived City Opportunity

CROSS-CASE ASSERTION 12					
Variables	**Case A**	**Case B**	**Case C**	**Retained**	**Validated**
Improve city Image / transparency	x	x	x	No	Yes
City Competitiveness	x	x	x	No	Yes
Internal Efficiency and Efficacy Gain	x	x	x	No	Yes
Participation / Collaboration		x		Yes	**CCA7**

Appendix 6.12: Cross-Case Assertion 12 Links

CCA12 Links	**Case A**	**Case B**	**Case C**	**Validated**
Local Ecosystem Factors → Perceived City Opportunity	x	x	x	Yes
City Council-Related Factors → Perceived City Opportunity	x	x	x	Yes
Service Characteristics → Perceived City Opportunity		x	x	Yes

Appendix 6.13: Cross-Case Assertion 13, Perceived Voluntariness

CROSS-CASE ASSERTION 13				
Variable	**Case A**	**Case B**	**Retained**	**Validated**
Voluntary Use of Open Data	x	x	No	Yes

Appendix 6.14: Cross-Case Assertion 14, Perceived User-Friendly

CROSS-CASE ASSERTION 14				
Variables	**Case A**	**Case B**	**Retained**	**Validated**
Lean Service / Sleek Design	x	x	No	Yes
Being able to get information	x		Yes	**CCA15**
Easy to Use	x	x	No	Yes
Interoperability across browsers		x	Yes	No

Appendix 6.15: Cross-Case Assertion 14 Links

CCA14 Links	**Case A**	**Case B**	**Case C**	**Val.**
Organizational Skills → Perceived User Friendly	x	x	(x)	Yes
Perceived User Friendly → Perceived Usefulness of Open Data		x	(x)	Yes

Appendix 6.16: Cross-Case Assertion 15, Perceived User Satisfaction

CROSS-CASE ASSERTION 15				
Variables	**Case B**	**Case C**	**Retained**	**Validated**
Gain Experience Efficiency / Time Saving	x	x	No	Yes
Gain Experience Efficacy / Resource Saving	x	x	No	Yes
Learning Opportunity	x		Yes	No
Personalised Service		x	No	Yes
Easy to Use		x	Yes	**CCA14**

Appendix 6.17: Cross-Case Assertion 15 Links

CCA15 Links	Case A	Case B	Case C	Validated
Service Characteristics → Perceived User Satisfaction		x	x	Yes
Perceived User Satisfaction → Perceived Usefulness of Open Data	(x)		x	Yes

Appendix 6.18: Cross-Case Assertion 16, Perceived Credibility and Sustainability

CROSS-CASE ASSERTION 16				
CCTA16 Variables	**Variables Integrated**	**Case**	**Retained**	**Validated**
Level of Commitment	*Partnership Strength*	A	No	Yes
	Level of Engagement	A		
	Participation / Collaboration	B		
	Account for Responsibilities on Open Data Released	C		
Challenges to the Council	*Challenges to the Council*	C	Yes	Yes

Appendix 6.19: Cross-Case Assertion 16 Links

CCA16 Links	Case A	Case B	Case C	Validated
City Council-Related Factors → PC&S	(x)	(x)	X	Yes
Cities – Business Collaboration → PC&S	(x)		x	Yes
Perceived City Opportunity → PC&S	(x)	(x)	x	Yes
PC&S → Perceived Business Opportunity	(x)	(x)	x	Yes
EU-Level Factors → PC&S	Result of Validation Stage			
Local Ecosystem Factors → PC&S	Result of Validation Stage			

Appendix 6.20: Cross-Case Adoption Stages

#	ADOPTION STAGES	CASE A	CASE B	CASE C	Retained	Validated
1	*Idea Generation*	x	x	x	No	Yes
2	*Open Data Adoption*	x	x	x	No	Yes
3	*Market Analysis*			x	Yes	**No**
4	*Pilot Implementation*	x	x	x	No	Yes
5	*Acquire Funding*		x		Yes	**Yes**
6	*Wider Adoption*		x		Yes	**No**
7	*Commercialization*	x	x	x	No	Yes
8	*Growth*	x	x	x	No	Yes
9	*Sustained Use*	x	x	x	No	Yes

Appendix 7: Cross-Case Analysis – Finally Retained Variables

#	Variables	Original Factors	Case
1	*Lean Start Up Approach*	*Organizational Factors*	B
2	*Partners Engagement*	*Organizational Factors*	B
3	*Attitude to Work with Government*	*Individual Factors*	A
4	*Support Creation and Sustainability of ODS*	*City-Related Factors*	A
5	*Internal Laws*	*City-Related Factors*	A
6	*Creativity*	*Individual Factors*	B
7	*Public Free Wi-Fi*	*City-Related Factors*	B
8	*Incentivizing / Challenging*	*Service Characteristics*	B
9	*Enable Interactive Experience*	*Service Characteristics*	B
10	*Engagement with Local Community*	*Service Characteristics*	B
11	*Attractive Technology*	*Service Characteristics*	B
12	*Open Source Software*	*Service Characteristics*	B
13	*Cloud-Based*	*Service Characteristics*	B
14	*Interoperability*	*Service Characteristics*	B
15	*White Label Solution*	*Service Characteristics*	C
16	*Flexible Revenue Model*	*Service Characteristics*	C
17	*Complexity*	*Service Characteristics*	C
18	*Lean Start Up Approach / Lean Team*	*Skills and Resources*	C
19	*Collective Learning*	*Council-Business Relationship Factor*	C
20	*Rewards*	*Perceived User Interest*	B
21	*Amusement*	*Perceived User Interest*	B
22	*Interactivity*	*Perceived User Interest*	B
23	*Promote Existing Businesses*	*Perceived Potential Economic Growth*	B
24	*Job Creation*	*Perceived Potential Economic Growth*	B
25	*Community Awareness*	*Perceived Ecosystem Awareness*	C
26	*City Council Awareness*	*Perceived Ecosystem Awareness*	C
27	*Investors Awareness*	*Perceived Ecosystem Awareness*	C
28	*Replicability of the Open Data Service*	*Perceived Credibility and Sustainability*	C
29	*Large Datasets Analysable and Shareable*	*Perceived Usefulness of Open Data*	B
30	*Free of Charge*	*Perceived Usefulness of Open Data*	B
31	*Easier Process*	*Perceived Usefulness of Open Data*	B
32	*Revenue Model Structure*	*Perceived Business Opportunity*	C
33	*Interoperability Across Browsers*	*Perceived User-Friendly*	B
34	*Learning Opportunity*	*Perceived User-Satisfaction*	B
35	*Level of Engagement*	*Perceived Business Opportunity*	A
36	*Partnership Strength*	*Perceived Business Opportunity*	A
37	*Market Analysis*	*Adoption Sub-Factor*	C
38	*Wider Adoption*	*Adoption Sub-Factor*	B

Appendix 8: Effort Expectancy - Root Constructs

Construct	Reference	Definition
Perceived Ease of Use	(Davis, 1989) (Davis *et al.* 1989)	The degree to which a person believes that using a system would be free of effort.
Complexity	(Thompson *et al.* 1991)	The degree to which a system is perceived as relatively difficult to understand and use.
Ease of Use	(Rogers, 2010)	The degree to which using an innovation is perceived as being difficult to use.

Appendix 8.1: Effort Expectancy Items; adapted from (Venkatesh *et al.* 2003)

Construct	Items	UTAUT
Perceived Ease of Use (Davis, 1989)	-Learning to operate the system would be easy for me. -I would find it easy to get the system to do what I want to do. -My interaction with the system would be clear and understandable. -I would find the system to be flexible to interact with. -It would be easy for me to become skilful at using the system. -I would find the system easy to use.	EE1: My interaction with the system would be clear and understandable. EE2: It would be easy for me to become skilful at using the system. EE3: I would find the system easy to use. EE4: Learning to operate the system is easy for me.
Complexity (Thompson *et al.* 1991)	-Using the system takes too much time from my normal duties. -Working with the system is so complicated; it is difficult to understand what is going on. -Using the system involves too much time doing mechanical operation (e.g. data input). -It takes too long to learn how to use the system to make it worth the effort.	
Ease of Use (Rogers, 2010)	-My interaction with the system is clear and understandable. -I believe it is easy to get the system to do what I want to do. -Overall, I believe that the system is easy to use. -Learning to operate the system is easy for me.	