Systems, Signs & Actions
An International Journal on
Information Technology, Action,
Communication and Workpractices
Vol. 6 (2012), No. 1, pp. 46–66
http://www.sysiac.org/

The case for design science utility and quality
- Evaluation of design science artifact within the sustainable ICT capability maturity framework

Markus Helfert [a], Brian Donnellan [b] and Lukasz Ostrowski [a]

[a] Dublin City University, School of Computing, Dublin 9, Ireland

[b] National University of Ireland, Maynooth, Co. Kildare, Ireland

Abstract

In Design Science Research (DSR), evaluation of research outputs in form of design artifacts has been discussed in numerous publications. Many researchers have emphasised the criteria of utility for design artifacts, whereas recent approaches have extended this view to other criteria such as efficiency, consistency, accuracy, performance and reliability to mention a few. In this paper we revisit the evaluation discussion in design science and describe a practical oriented evaluation framework. In order to incorporate the evaluation along on-going design cycles, we propose to build on work related to semiotic and information quality. We argue that design science research is usually complex and requires a more detailed evaluation approach. We review literature and follow a semiotic framework for managing knowledge in complex environments. Together with well established information quality criteria we demonstrate that the proposed framework can help to provide a practical evaluation approach within a complex design environment. The framework was developed within the context of a novel IT Management model. We describe its development during the creation of the maturity model for Sustainable Information and Communication Technology (ICT) called the SICT-Capability Maturity Framework (SICT-CMF). The context was selected as it is particular interesting with design artifacts created within an open and complex innovation community. The research acknowledges the importance of the utility criteria, but also includes other criteria such as artifact quality, consistency and accuracy in form of a more differentiated view of design science evaluation.

Keywords: Design Science, Utility, Evaluation, Quality, Information Quality, Semiotic

Corresponding author: markus.helfert@computing.dcu.ie

This paper is developed from the previous publication: Helfert M & Donnellan B (2012) The case for design science utility - Evaluation of design science artefacts within the IT capability maturity framework, accepted to the International Workshop on IT Artefact Design & Work Practice Intervention, Barcelona

Received: 6 November 2012; Revised: 18 December 2012; Accepted: 30 December 2012
Accepting Editor: Göran Goldkuhl

1 Introduction

The importance of utility and quality of research outputs has been emphasised and reiterated in recent Design Science Research (DSR) discussions (Hevner et al. 2004; Iivari, 2007; Peffers et al. 2007; Hevner & Chatterjee 2010). Evaluation strategies and guidelines have been proposed, but yet practical evaluation of design artifacts is challenging. Many approaches are subjective and designed for a small number of application scenarios or specific projects. Limited contributions have addressed the practicalities of evaluating research outputs that are designed within a complex research environment. Recent attempts have expanded the view of design evaluation and proposed more comprehensive evaluation approaches (Gill & Hevner 2011). However, a particular problem exists when incorporating design evaluation along an on-going design research process within a complex design environment. During the development we are constantly facing the problem of evaluating design artifacts and research outputs. Focusing ex-post on evaluating the end result concerning utility is often not sufficient to ensure high quality research outputs. This calls for revisiting the discussion on DSR evaluation as it becomes increasingly important when research is moving to collaborative and open research environments, in which practitioners and academic researchers contribute to more complex artifact developments.

Revisiting DSR evaluation from a quality aspect, the paper proposes an information quality oriented framework and demonstrates its usage in an effort to develop an information management maturity model. We are arguing for structuring the evaluation along levels of semiotic. Adapting a semiotic based information quality model (Price & Shanks 2005), the framework is structured along three semiotic levels. The result is an information quality oriented framework for evaluating design artifacts during a design process.

The work has been developed in the context of the IT Capability Maturity Framework (IT-CMF), a high-level process capability maturity framework for managing the IT function within an organization (Curley 2004; Curley 2006). The design environment with the IT-CMF is found particular interesting as the design and review processes are based on “open innovation” principles within a collaborative research environment. “Open innovation” as presented by Chesbrough (2003) offers an innovation model where organizations leverage both external and internal resources and expertise. Participants communicate frequently, exchange different views and constantly develop design artifacts by leveraging the collective intelligence of experienced practitioners and researchers in the consortium. Our evaluation framework proposed in this paper is illustrated within a sub-model of the IT-CMF, the maturity model for Sustainable Information and Communication Technology (ICT) called the SICT-Capability Maturity Framework (SICT-CMF).

The remainder of the paper is structured as follows: Section 2 describes the context in which the work is presented; evaluating design artifacts in form of IT management maturity models. Section 3 outlines the research methodology. Section 4 reviews the background and related work of DSR. Subsequently work on information quality and semiotic motivates an information quality oriented evaluation framework for design artifacts. Section 5 described overall design of the IT-CMF design process and the application of evaluation framework with the SICT-Capability Maturity Framework. We conclude our paper in Section 6 by outlining some research directions.

2 Evaluating Information Management Maturity Models

This work is carried out in the area of information management maturity frameworks. Information management (IM) has a wide scope and typical artifact development is extremely complex. It includes areas such as information strategic planning, project management, information systems development, architecture development, resource management, innovation management performance and quality management. In recent years much work has been done on categorizing and describing these areas into processes, describing capabilities and developing models and descriptions. Numerous frameworks and approaches have been documented. One frequent referred example is the Capability Maturity Model (CMM) for software development (Paulk et al. 1993), which has built the foundation for many subsequent developed maturity models and frameworks. Prominent approaches include CMMI, COBIT and ITIL (Johannsen & Goeken 2007), and more recently the IT-CMF. At the same time, maturity models receive increasing attention in practice and literature (Becker et al. 2009; Röglinger et al. 2012). The frameworks include criteria describing distinct maturity levels together with assessment approaches that will assist an organization to identify its specific maturity status (Röglinger et al. 2012). Maturity in this context refers to evolutionary growth in the capability to manage information, processes as well as systems and technology (Humphrey 1989). Together with assessment approaches most maturity models provide some descriptions for improvement as well as assessment guidelines. Classifying capabilities can be useful for maturity assessments and can help companies assessing their own performance in relation to other companies. Furthermore by suggesting how to develop these capabilities, the model can help transforming organizations towards higher maturity levels.

The models and framework provide valuable contributions containing best practices and experiences. However, despite the large number of maturity models, many practical relevant approaches and frameworks are limited in providing transparent evaluation approach. Indeed Röglinger et al. (2012) evaluated some prominent maturity models related to business processes and found that rather limited information is available on the empirical validation. This limits the value as well as applications of the models and furthermore underpins the importance of our research presented in this paper. Due to the complexity of the artifacts and the lengthy development process, it is usually a significant effort to evaluate this kind of models (Helgesson et al. 2012). In literature several types of evaluating maturity models can be found (Helgesson et al. 2012)

- Type 1: Evaluation is conducted "offline", only by the authors of the evaluation without involving any outside experts.
- Type 2: Evaluation is conducted by involving practitioners and experts, who have domain expertise. In a type 2 evaluation, no real assessment is carried out; instead, interviews, surveys, or simulated assignments can be carried out.
- Type 3: A type 3 evaluation is conducted through real process improvement activities where the maturity model is used in a practical setting.

Most documented evaluation approaches for maturity models fall into type 1 or type 2, often without having a consistent evaluation approach documented. In addition, as discussed below, the evaluation time (ex-ante versus ex-post) is an important consideration during the design evaluation. Most evaluation approaches using an ex-post strategy with limited attention given to evaluation during the design process. A related more general framework for evaluating tools and models in information system research has been proposed by Ågerfalk (2004). He makes a distinction between internal grounding, external theoretical grounding, and empirical grounding. Vaishnavi et al. (1998) have proposed a validation framework for maturity models, which focuses especially on the requirements for empirical validation of maturity models through application in case studies. Whereas most of the evaluation approaches focus on evaluating the application of the maturity models in an ex-post manner, Mettler (2009) analyses maturity models from a design science perspective and argues that both the design process (i.e. the way the model was constructed) and the design output (i.e. the model, artifact itself) need to be evaluated.

3 Research Methodology

The research presented in this paper introduces an evaluation framework for design science artifacts that is built using semiotic and information quality criteria. The evaluation framework was designed employing a design science research methodology (Peffers et al. 2007; Hevner et al. 2004). In order to developed and validate our evaluation framework we worked together with a group in IVI working on the development of a sustainable ICT maturity model. In addition to the complex design environment, we selected the group and the context of IVI as they employ a design science research methodology to develop the maturity model. In this sense, we apply a design science approach to develop and test an evaluation framework within a design science oriented research environment.

In a first phase of developing our evaluation framework, the problem for building and evaluating design artifacts was scoped and initial concepts developed by reviewing relevant literature. We used chain-referral sampling (Heckathorn, 2002) approach, beginning with the most cited papers gradually reaching towards other relevant publications, and paying particular attention to related journal special issues, prominent books, and conferences related to design science. Through this process a glossary of DSR-related concepts and definitions was compiled. In addition in order to develop our evaluation framework, we followed our earlier work in which we emphasise the importance of complementing the findings by practitioners' input from the field (Ostrowski et al. 2011). Therefore we worked with the group in IVI to discuss and reflect our findings. Using selected quality criteria we examined development practices and discussed these with researchers during the Sustainable ICT development in order to derive suitable questions and quality indicators for the evaluation framework. We held a series of discussion-type focus group meetings with researchers sharing experiences and challenges on building and evaluating research results. The approach attempted to generate discussion and interaction to confirm our evaluation framework. In this environment, the use of design science to develop a maturity model provided us with an environment of considerable degree of richness. In the following we present findings from our research process and conceptual evaluation framework and then in section 5 the application within the SICT-Capability Maturity Framework development.

4 Build and Evaluate

Design science research focuses on creations of artificial systems. It addresses research through the *building* and *evaluation* of artifacts designed to meet stated objectives (Hevner et al. 2004). *Building* refers to the process of constructing an artifact for a specific purpose and *evaluation* assesses how well the artifact meets objectives. Evaluation should be in the centre of DSR focusing on the output of design science research. Thereby we accept the importance of utility for design artifacts as dominant criteria to evaluate satisfying objectives; however at the same time we argue for a more detailed evaluation during the research process and along several levels, namely syntax, semantic and pragmatic. In the following we present findings from literature and describe important characteristics of the build and evaluate activity and output of design science research. The discussion is focused on arguing for the rational of including three levels of semiotic -syntax, semantic and pragmatic. To operationalize the evaluation framework, we follow information quality criteria, which then lead us to the description of the proposed conceptual evaluation framework in section 4.4.

4.1 Build Process

The construction of an artifact is a heuristic search process in which extensive use of theoretical contributions and research methodologies should be made (Schön 1983). Simon's demonstrated that natural science and the science of the artificial are different as the former is about analysis where the latter is about synthesis (Simon 1969). Based on this distinction, researchers distinguished two paradigms: behavioural science research and design science research (Hevner et al. 2004). The former is understood as a "problem understanding paradigm", the latter as a "problem solving paradigm".

A key characteristic of DSR is that it resolves an important, previously unsolved problem, for a class of businesses or environments, while making a contribution to the knowledge-base (Venable 2006). Design researchers investigate the current knowledge and solutions to insure they do not just replicate past work of others. The value of a new solution may come from various activities such as solving a known or expected problem, satisfying needs, or innovating something new. However, the new knowledge comes from "the number of unknowns in the proposed design which when successfully surmounted provides the new information that makes the effort research and assures its value" (Vaishnavi & Kuechler 2004). The research may involve searching the existing knowledge-base, or collecting primary data through empirical work such as case studies, interviews, experiments or surveys. Research should stop if the problem has already been solved, or if it is found to be unimportant for the targeted objectives. Through this research process, the design science researcher satisfies the relevance condition for DSR in IS (Hevner 2007), while also addressing generalizability (Benbasat & Zmud 1999). Characteristic for DSR is that rich phenomena that emerge from the interaction of people, organizations, and technology may need to be *qualitatively* assessed to yield an understanding of the phenomena adequate for theory development or problem solving (Klein & Myers 1999). The process of constructing and exercising innovative IT artifact enable design-science researchers to understand the problem addressed by the artifact and the feasibility of their approach to its solution (Nunamaker et al. 1991).

It is generally agreed, that design science research develops knowledge that can be used by professionals in the field in question to design innovative solutions to their field problems (Van Aken 2005). To obtain knowledge for innovative solutions, Van de Ven (2007) proposed engaged scholarship as a participative form of design science research. It accommodates points of views of key stakeholders to understand complex problems. By exploiting differences between stakeholders, engaged scholarship develops knowledge that is more penetrating and insightful than when researchers work alone. Sein et al. (2011) propose action design research method to interlink the building and evaluation phases and thereby emphasising the organisational context. Illustrating the complexity of developing innovative outputs, Leonard (1995) outlines that working across boundaries between disciplines, specializations, or expertise is a key ingredient for most innovative solutions.

In order to examine communication and knowledge management, Carlile's (2004) proposed a framework based on semiotic that assists to manage knowledge across different parties involved. We illustrate the framework in Figure 1 and discuss it below. It can be used to examine the communication between stakeholders with various domain specific knowledge, identification of their knowledge boundaries, and the barrier to innovation solutions design science.

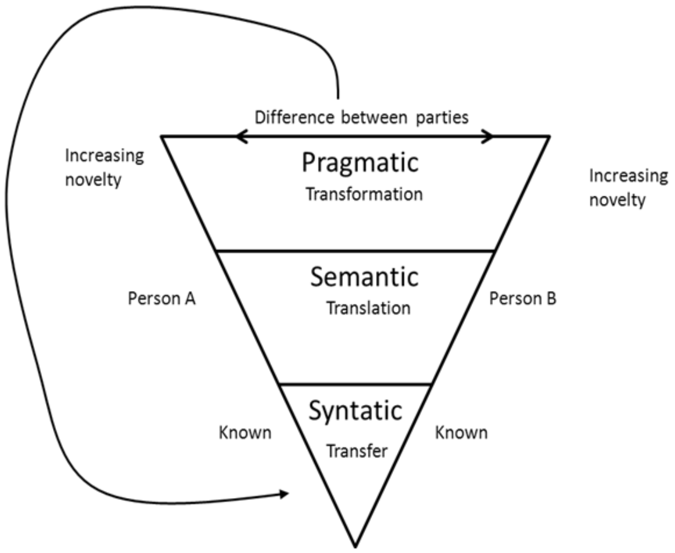

Figure 1 Framework for managing knowledge across boundaries (Carlile, 2004)

Communication between stakeholders requires common knowledge along three semiotic levels; of the *syntax* (structure), *semantics* (meaning), and *pragmatics* (use) of language in order to understand each other's domain-specific knowledge. Stamper (1973) argued for the importance of these different abstraction levels, which have been later extended (Stamper 1994; Liu 2000) and used in several research projects. The most basic level is syntax, which is the grammatical structure of sequence, order and arrangement of words and phrases into sentences of a language. The next interpretive level is semantics, which is expressed with the pattern of words and sentences. Finally, at the most specific and personal level is pragmatics where stakeholders apply their meanings of communication to practical uses in particular research context.

According to Carlile, the difference and dependence of domain-specific knowledge among stakeholders at a boundary determine the complexity of communicating across that boundary. Differences occur in relation to unique amounts of knowledge and types of specialized domain-specific knowledge of stakeholders at the knowledge boundary. Dependence is the degree to which people across boundaries perceive each other's views that meet their goals, such as that of co-authors of a paper, or speakers and listeners. This is a dependence without which different opinions would have no consequences. The triangle illustrated in Figure 1 portrays how increases in the difference and dependence among stakeholders at a boundary create syntactic, semantic, and pragmatic communication difficulties, and hence contribute to progressively more complex forms of communication. If differences and dependences can be specified among stakeholders, then the boundary proves 'unproblematic' and knowledge can be transferred using a conventional information processing view. A common lexicon or glossary can comprise these specifications.

However, this becomes problematic when innovation solution is desired. The current lexicon is no longer sufficient to represent the differences and dependencies faced. Due to the nature of providing novel artifacts, usually these more complex and problematic forms or communication can be observed in DSR projects. Development of an innovation design artifact has its challenges in the limitation of an information-processing because the processing of a common lexicon is assumed to be always a sufficient common knowledge. So while a common lexicon is always necessary, it is not always a sufficient type of common knowledge to share and assess domain-specific knowledge (Carlile 2004). To overcome the problem of insufficient lexicon, researchers can refer to a *form of inquiry* technique that involves leveraging stakeholders' different perspectives to learn about a domain; a *relationship* technique that involves negotiation, mutual respect, and collaboration to produce a learning community; and an *identity* technique of how researchers view their relationships with their communities and stakeholders. In this way we argue not only for more suited forms of inquiry but also for the requirement of more detailed evaluation techniques for design artifacts.

In summary, the central thrust of Design Science attempts to create and evaluate IT artifact intended to solve identified relevant organizational problems. Hevner et al. (2004) propose a set of problem solving guidelines process where the understanding of a design problem and its solution are acquired in the building and application of an artifact. Several researchers have contributed discussions on the nature and process of designing artifacts; however as the discussion of Carlile's model shows, communication within a DSR project along the three semiotic levels are esential to provide novel and problem solving design artifacts.

4.2 Design Output

In the centre of DSR are innovative artifacts (Gericke et al. 2009), that accordingly seek to extend the boundaries of human and organizational capabilities (Hevner et al. 2004). As discussed above, design science creates and evaluates IT artifact intended to solve identified organizational problems. Such artifacts are represented in a structured form that may vary from software, formal logic, and rigorous mathematics to informal natural language descriptions. March & Smith (1995) identify four principle design artifact produced by DSR in IS; namely constructs, models, methods, and instantiations. Constructs are defined as "concepts" and "conceptualizations" (March

& Smith 1995) and "vocabulary and symbols" (Hevner, et al., 2004). These constructs are abstracted concepts aimed for theorizing and trans-situational use. "Conceptualizations are extremely important in both natural and design science. They define the terms used when describing and thinking about tasks" (March & Smith, 1995). Models use the construct to represent a real world situation. They aid "problem and solution understanding and frequently represent the connection between problem and solution components enabling exploration of the effects of design decisions and changes in the real world." (Hevner, et al. 2004). Method then is defined as "a set of steps (e.g. an algorithm) to perform a task" (March & Smith 1995). Finally, an instantiation is a prototype or a specific working system or some kind of tool (Goldkuhl 2004). Artifacts are innovations that provide a degree of novelty into an application context and thus the result of DSR is, by definition, a purposeful artifact created to address an important organizational problem. Some aspect of the artifact must be an original contribution to the existing knowledge base of the application domain. Artifact originality is a defining characteristic of DSR that makes the new artifact an innovation to the field of application.

Hevner et al. (2004) further developed the IS design perspective by developing a holistic framework for IS research – from the perspective of DSR. Hevner determined the necessary functions of IT artifact according to given requirements with the behavioural science perspective that explains and predicts the ways the artifact are used. He tried to reconcile the constraints of prevailing IS research activities extant at that time that either focus on design of IT artifact or their use practices in organizations separately. In this view, the design activity is conducted iteratively, consisting of activities such as elaborating relevant problems in the application domain, building the artifact and evaluating its performance. As a result of design research, the artifact needs to *satisfy the articulated requirements* within a field of application as well as to enlarge the knowledge base of the scientific community. However the process of evaluating the artifacts in relation to requirements is challenging as it involves various stakeholders in a complex design science process.

4.3 Design Evaluation

Because design is inherently an iterative and incremental activity, the evaluation phase provides essential feedback to the build phase concerning the quality and utility of the design output under development and its design process. Evaluation delivers evidence that an artifact developed achieves the purpose for which it was designed and consequently provides indications for the design process. Without evaluation, outcomes are unconfirmed declarations that the artifact meet their purpose (i.e. be useful for solving a problem or making some improvement). Design science artifact "are assessed against criteria of value or utility – does it work?" (March & Smith 1995). The essential aim is to rigorously demonstrate the utility of the artifact being evaluated. Rigor in DSR should be approached from two directions. One is to establish if the artifact solves the stated problem and causes an observed improvement, its efficacy. The second direction is to establish if the artifact works in a real situation, its effectiveness (Checkland & Scholes 1990). Gill and Hevner's (2011) recently proposed a Fitness-Utility Model for DSR, in which a fitness-utility function, based on criteria, has been described. Hence, evaluation is quite specific to the artifact, its purpose, and the purpose of the evaluation (Venable, et al. 2012). Consequently utility of artifact is a complex deliverable. It may depend on many different characteristics of the artifact or desired outcomes of the use of the artifact. From the perspective of artifacts, we can distinguish two types, product and process (Gregor & Jones 2007). The former represents tools, models, diagrams or software that people use to solve a problem. The latter is in a form of a method or procedure that guides someone what to do to solve a problem, thus a person must interact to provide utility of the artifact. Following Hevner et al. (2004) many researchers have argued for utility and quality [and efficacy] of DSR artifacts. Quality and other criteria have been mentioned as important in addition to utility. Researchers state that "artifact can be evaluated in terms of functionality, completeness, consistency, accuracy, performance, reliability, usability, fit with the organization, and other relevant quality attributes" (Hevner, et al. 2004), whereas utility is often the defining characteristics of artifact evaluation. These criteria of an artifact in some way contribute to the overall utility. They can act as criteria that are candidates for evaluation in determining the overall utility. In addition, formalising the knowledge of the utility, it can also be perceived through other indicators such as information system design theories (Gregor & Jones 2007; Pries-Heje & Baskerville 2008) or design principles (Sein et al. 2011), which allow to evaluate the artifact independently of its application. As a consequence and following Gamble & Goble (2011), we differentiate between quality and utility in the form of dependent and independent evaluation.

- Quality - a function of the *artifact or process* assessed against a quality specification *independent of the consumer* to provide a specific, objective measure of quality (e.g. accuracy)
- Utility - a function of the *artifact and consumer* to assess whether the output fits the purpose and meet the users subjective needs e.g. relevance.

The discussion above demonstrates the importance of utility and quality in DSR evaluation. Utility as a form of evaluating the fitness of an artifact and degree to meeting users subjective needs. And second, a measure of quality in respect to specifications, standards, guidelines or theories. Schön stressed the importance of constructs in providing the language in which problems and solutions are defined and communicated (Schön 1983). This highlights meaning and interpretation of constructs and its related models and methodologies. Constructs itself are describing models and methods on a syntactical level.

Researchers identified a number of methods that can be used for evaluation of design science artifact. Hevner, et al. (2004) proposed five classes of evaluation methods: (1) Observational methods include case study and field study. (2) Analytical methods include static analysis, architecture analysis, optimization, and dynamic analysis. (3) Experimental methods include controlled experiment and simulation. (4) Testing methods include functional testing and structural testing. (5) Descriptive methods include informed argument and scenarios. Peffers et al. (2007) divide evaluation into two activities, demonstration and evaluation. The former demonstrates that the artifact feasibly works to achieve its purpose in at least one context. The latter considers how well the artifact supports a solution to a problem. Venable (2006) divides evaluation into artificial and naturalistic. Artificial evaluation includes laboratory experiments, field experiments, simulations, criteria-based analysis, theoretical arguments, and mathematical proofs. It evaluates a solution in a contrived and nonrealistic way. Naturalistic evaluation explores the performance of a solution in its real environment. By performing evaluation in a real environment (real people, real systems, and real settings (Sun & Kantor 2006)), naturalistic evaluation embraces all of the complexities of human practice in real organizations. This approach is always empirical, and includes methods such as case studies, field studies, surveys, and action research (Venable, et al., 2012). The dominance of the naturalistic paradigm brings to naturalistic DSR evaluation the benefits of stronger internal validity (Gummesson 1988). However, these authors provide no guidance for choosing between methods, and there is little guidance in the DSR literature about the choice of strategies and methods for evaluation in DSR. The most cited guide selection of evaluation strategies for a DSR project is 2-by-2 framework (Pries-Heje et al. 2008). This state of affairs in DSR constitutes what researchers call an "evaluation gap" (Venable et al. 2012).

In this paper we address the evaluation gap by proposing an evaluation framework based on three levels of semiotic. In order to propose concrete quality criteria we build on common criteria proposed for evaluating the quality of information (see discussion below). The framework was developed in the context of developing the sustainable ICT maturity model. Figure 2 outlines the context as well as the key activities of build and evaluate in a DSR process.

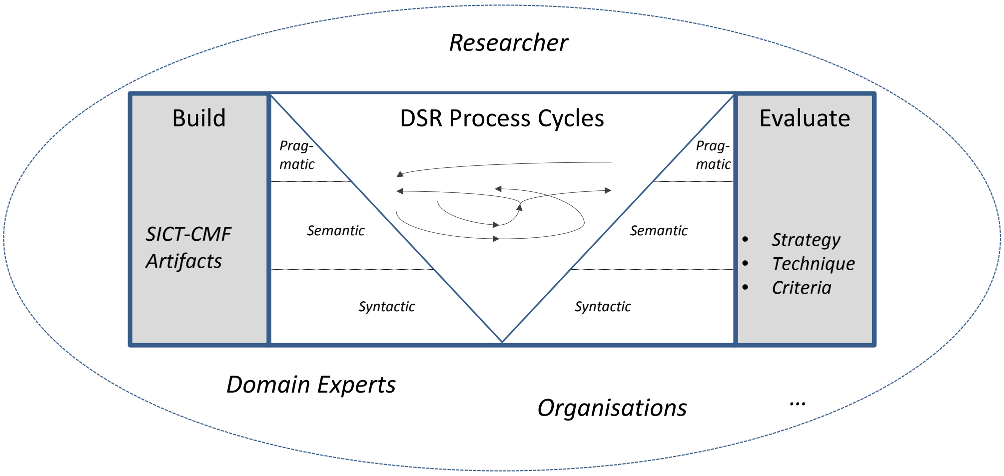

Figure 2 Framework for managing knowledge across boundaries within a DSR process

4.4 An Information Quality Framework for Design Evaluation

We have discussed DSR from a build and evaluate point of view and argued for considering utility as well as quality. We also argued the need for a more detailed evaluation due to the complex nature of design processes and suggested to use the three levels of semiotic in the evaluation. Several researchers have proposed criteria that need to be considered when evaluating design artifacts. In this article we argue that a quality view (Juran 1998) allows, in addition to utility, an objective evaluation of design artifacts, by assessing the conformance to a quality standard or specification. In our study of creating the sustainable ICT model, we build and evaluate information artifacts, and thus the adoption of a prominent information quality framework proposed by Wang & Strong (1996) seems reasonable. Similar to other common information quality frameworks it views information quality (IQ) as a multi-dimensional concept, with dimensions of accuracy, timeliness, reputation etc. (Wang & Strong 1996). Wang & Strong's IQ framework has been proven to be useful in many projects, and established information quality approaches follow similar definitions and classifications. Based on the criteria (Lee et al. 2002) developed the AIMQ methodology providing general evaluation questions for IQ. Questions are the form of 'This information is complete', 'This information is presented consistently', 'This information was objectively collected', 'This information is relevant to our work', 'This information is sufficiently current for our work', 'This information is easy to understand', 'This information is easy to comprehend', 'The meaning of this information is easy to understand', etc. Moody & Shanks (2003) proposed an empirical validation of a quality management framework and later developed a semiotic information quality framework (Price & Shanks 2005). In this paper we employ information quality criteria together with Price & Shanks (2005) semiotic IQ framework as starting point of providing an evaluation framework for DSR. Using selected quality criteria we examined development practices and discussed these with researchers during the Sustainable ICT development in order to derive suitable questions and quality indicators for the evaluation framework. Core element of the framework is a set of relevant quality characteristics (adopted from Wang & Strong 1996) along the 3 semiotic levels as described above. Adapting Price & Shanks (2005) information quality framework, we have listed quality criteria on the three semiotic levels as illustrated in Table 1.

Table 1: Design Evaluation Framework

Semiotic Level	**Example Criteria**
Pragmatic	Relevance, usability, completeness, timeliness, actuality, efficiency,
Semantic	Precise definitions and terminology, easy to understand, interpretability, accuracy (free-of error), consistent content
Syntax	Consistent and adequate syntax, syntactical correctness, consistent representation, accessibility

The rationale behind using the semiotic levels is the correspondence of the pragmatic level to utility, whereas the other two levels can be related to quality standards, specifications or development guidelines. It allows including concepts of quality and utility. The quality criteria at each level are subjective and context depended, however criteria on the level of syntax and semantic allow for application independent evaluation. Moreover, the three semiotic levels correspond directly with the Carlile's framework. This enables design science researchers to apply adequate quality criteria and evaluate each communication level between stakeholders while obtaining their domain-specific knowledge. This is the syntax, semantic, and pragmatic.

5 The IT-Capability Maturity Framework

5.1 Design process and Evaluation within the IT-CMF

This research is being undertaken in the context of the Innovation Value Institute and the development of the IT-CMF. Applying the principles of engaged scholarship (Van de Ven 2007, Mathiassen and Nielsen 2008) and DSR (see above), IT Management is being investigated using a design science process with defined review stages and development activities. During the design process, researchers participate together with practitioners and subject matter experts within research teams to capture the views of key domain experts. The design process and artifact development is divided into four phases separated by stage reviews with key deliverables at each stage. At phase 1 references relating to the artifact are consulted and expanded with input from group of key opinion leaders, subject matter experts, industry and academic literature. At phase 2 comparisons are made with frameworks and best practices in industry. At phase 3 the artifact are reviewed with 3-5 external organisations and key opinion leaders. At phase 4 the artifact are exercised through field experiments in at least three organisations. The output of the design process is components of the IT-CMF maturity model, which contains models and assessment techniques as design artifacts.

The IT-CMF follows design science principles within a rigorous design process that facilitates the engagement of scholars as well as ensures consistency by providing a meta-model for structuring the maturity model, development guidelines and platform (Helfert & Curely 2012). The design science approach used in the IT-CMF is closely aligned with the three DSR cycles proposed by Hevner. Additional detail on the design process development is available from (Carcary 2011; Donnellan and Helfert 2010). In these three closely related cycles of activities can be observed: Relevance Cycle, Rigor Cycle and Design Cycle (Hevner 2007).

The Relevance Cycle inputs requirements from the contextual environment into the research and introduces the research artifacts into environmental field-testing. For example relevance for the sustainable ICT development artifact is driven by the problems of organisations experience in optimizing how they currently manage and measure the business value of their IT investments. Field-testing of the IT-CMF in the application environment helps determine of further development work is required to ensure its relevance in addressing the business problem.

The Rigor Cycle provides grounding theories and methods along with domain experience and expertise from the foundations knowledge base into the research and adds the new knowledge generated by the research to the growing knowledge base. For example Sustainable ICT development is grounded in existing artifacts, methodologies, foundational theories and expertise and draws from an extensive base of industry and academic literature and existing IT standards and frameworks. Contributions to the knowledge base include a detailed framework and a set of practices that help define innovation and change in how organisations manage and use their IT investments to optimise business value.

The Design Cycle supports a tighter loop of research activity for the construction and evaluation of design artifacts and processes. For example, the sustainable ICT development focuses on iterative build and evaluates activities by the working group to address the identified problem, while drawing on existing theoretical foundations and methodologies in the knowledge base. The build process is evolved and refined through evaluation feedback, including committee stage gate reviews to identify further refinements and field-testing of the artifacts within contextually diverse organisations.

The key DSR artifact -the IT-CMF- is developed along a number of Critical Capability (CCs) which are then composed in a number of artifact constructs. For instance, overview and definition section contains definition of critical capability, description of capability building blocks (CBBs) and explanations of key-terms. The maturity profiles describe key characteristics of maturity levels, practices, outcomes, and metrics as well as required transitions to increase maturity. The assessment process, timeframe and questionnaire and interview guidelines are describing the assessment approach. Furthermore, evaluation results together with CC dependencies, comparison to other frameworks and references are documented. In order to facilitate the content development process and ensuring consistency and quality, artifacts are described in detail through templates and development guidelines and are facilitated via a development platform.

5.2 Example: Sustainable ICT within the IT-CMF

The IT-CMF content development and review process is implemented by members of the IVI community where consortiums members are invited and encouraged to participate in the research and development activities through working group contributions. A work group (WG) exists for each of the CCs and include a mix of Subject Matter Experts (SMEs) and Key Opinion Leaders (KOLs), including academic researchers, industry-based practitioners and consultants. The working group on sustainable ICT was established in 2008 with participation from over 10 organizations. As a result of the design process, the collective learning’s and experience of the group were captured within a maturity model for Sustainable ICT called the SICT-Capability Maturity Framework (SICT-CMF). The sustainable ICT working group consisted of 12 members (4 academic, 6 practitioners, & 2 consultants) with experiences from domains including ICT Hardware & Software, IT Services, Power Utilities, Insurance, and Financial Services. Members of the working group had experience with sustainable ICT operations at multiple levels including: site, national, regional, and global.

The development of the SICT-CMF followed the general design science oriented process within IVI, in which we adapted an engaged scholarship design process (Becker et al. 2009, Donnellan and Helfert 2010) to cater for constraints often faced when working in collaboration with practitioners. Working group development output evolves through a series of stages and is reviewed at the end of each stage by a technical committee.

As development of the work progresses through the various stages the material is subject to review and validation assessments. The final design process is divided into four main phases separated by stage reviews with key deliverables at each stage.

- *Phase 1:* The objective of the phase is to establish the need for the maturity model and to define the scope of the model. Initial background research is performed by reviewing relevant industry and academic literature. These are expanded with input from group of key opinion leaders and subject matter experts.
- *Phase 2:* The work group is established with participation from both academia and industry. The core objective of this phase is to develop definition of the model and the capability building blocks within it. The working group also makes

comparisons with artifacts in industry frameworks and industry best practices. Subject matter experts and key opinion leaders are interviewed directly, and the model is created for example using a Delphi Study approach.

- *Phase 3:* The objective of this phase is to complete the definition of the model by developing the maturity profiles for the model. The artifacts are reviewed with external organisations and key opinion leaders.
- *Phase 4:* The full model artifacts are exercised through field experiments. The model is adjusted based on experiences. The model is submitted for final review and approval.

In addition to reviewing literature and publicly available material, personal interviews with Subject Matter Experts (SMEs) and Key Opinion Leaders (KOLs) are the preferred technique for the model development, because they allow expansive discussions which illuminate factors of importance (Oppenheim 1992; Yin 2009). The questions for the content development are largely open-ended, allowing respondents freedom to convey their experiences and views, and expression of the contexts (Yin 2009; Oppenheim 1992). Interviews are undertaken by at least two members of the working group, both taking notes and reconciling into a single report. The interviews are conducted in a responsive (Wengraf 2001; Rubin & Rubin 1995), or reflexive (Trauth & O'Connor 1991) manner, allowing the researchers to follow up on insights uncovered mid-interview, and adjust the content and schedule of the interview accordingly. In any cases of ambiguity, clarification is sought from the corresponding interviewee, either via telephone or e-mail. Findings were continuously presented and discussed. Venting is used, whereby results and interpretations are discussed with the working group to avoid the problem of what Kaplan and Duchon (Kaplan & Duchon 1988) call multiple realities.

5.3 Design Evaluation within the SICT-CMF

The initial version of the maturity model was defined using a mode of argument called induction by simple enumeration by that rationalizes that "if a generalization has been positively instantiated in a large number of cases, and negatively instantiated in none, this makes it more likely that the generalization has been positively instantiated in a wide variety of circumstances" (Cohen 1970), and hence validity of inferences is increased. In order to provide indication of the quality and utility of the provided artifacts, we employed a simple questionnaire. The questionnaire was discussed with some researchers within the work group and is provided in the appendix.

Syntactical elements are verified during the entire development process. This could be verified internally by the researcher or editors. Thus no further involvement of domain experts is required. As mentioned above a large number of templates, guidelines and editorial support are available within IVI, so that the syntactical quality of the results is expected to be high. In phases 2 and 3 the core content is developed and *semantic* can be used to evaluate the quality of the development involving interviews with domain experts. *Pragmatic*: Evaluation on phase 4 is on-going and involved the application of the SICT-CMF to organisations. Here traditional evaluation approaches used in DSR can be applied to evaluate relevance, usability, timeliness and actuality, as well as efficiency. The core data collection was conducted using the SICT-CMF assessment processes and by capturing a detailed case study protocol that ensures consistency between case study questions and data gathering.

An organisation assessment had two phases: An online survey, and in-depth interviews. The online survey confirmed the relevance of the material whereas the assessment process provided insight in the usability and efficiency. Comparing a number of cases timeliness and actuality of the SICT-CMF could be confirmed. Several IVI consortium members are currently using the SICT-CMF to support their sustainable IT efforts, to help understand their current maturity and to set future directions. Since the initial development, the model has been applied in a number of organisations that underpins its relevance and quality further.

6 Concluding Remarks and further Research

The work presented in this article revisited the evaluation of artifacts within DSR. We discussed the build cycle, design outputs and evaluation approaches in DSR. Many researchers have described the challenges of evaluating design artifacts often referring to utility as the dominant evaluation criteria. Recent approaches have extended the utility centred argumentation to other criteria and included quality, completeness, consistency, accuracy, reliability and usability. Discussions have argued for evaluation of research outputs as well as research process. Furthermore due to the innovative character of artifact design, research processes are usually complex involving a number of stakeholders. Motivated by Carlile's (2004) framework, we have argued for reflecting the complexity and called for extending the utility focused discussion on artifact evaluation. Guided by semiotic and information quality in this paper we have presented an evaluation framework. We listed common information quality criteria on three semiotic levels and presented an example questionnaire. We described how the framework was applied during the development of the maturity model for the SICT-Capability Maturity Framework (SICT-CMF). Syntactical evaluation is based on templates, standards and development guidelines. Semantic can be evaluated by examining the content, whereas pragmatic evaluation is based on results from interviews with domain experts and the application in organisations. In this way we employed the tree types of evaluating maturity models discussed in section 2.

As the maturity model is developed within an open innovation community we found the work particular interesting. Our work describes the importance of incorporate design evaluation along the on-going design process. By using well established information quality criteria this was achievable and researchers found the questions and simple criteria useful to use. During the artifact development and expert interviews, a brief evaluation of the created design artifacts was feasible with reasonable effort.

The research shows that not only utility of the final design artifact is important, but also quality evaluation during the build (design) phase. However, although we believe our research provides a valuable contribution to DSR, the research in this article is limited. We could only revisit selected DSR contributions, evaluation approaches and outline an overview of the evaluation framework. We illustrate its application within the work at the Innovation Value Institute by which we discussed the general design process and the evaluation approach. Further research would be required to apply the information quality oriented framework to other similar DSR developments (e.g. IT maturity models). Further research to validate the criteria list and framework is suggested, and indeed in our further research we aim to detail the design steps and evaluation techniques along the proposed information quality oriented evaluation criteria. We aim to incorporate further evaluation techniques and relate these to quality criteria. Another route for further research within the Innovation Value Institute is the continuing development and improvement of the SICT-CMF.

References

Ågerfalk P J (2004) Grounding through operationalization – constructing tangible theory in IS research. In *Proceedings of the European Conference on Information Systems 2004*

Becker J, Knackstedt R, Pöppelbuß J (2009) Developing Maturity Models for IT Management – A Procedure Model and its Application, in *Business & Information Systems Engineering*, 1(3), pp. 213-222

Benbasat I, Zmud R (1999) Empirical Research in Information Systems- The Practice of Relevance, *MIS Quarterly*, 23(1), pp. 3-36

Carcary M (2011) Design Science Research: The Case of the IT Capability Maturity Framework (IT CMF), *The Electronic Journal of Business Research Methods,* 9 (2), pp. 109-118

Carlile P (2004) Integrative Framework for Managing Knowledge Across Boundaries, *Organization Science*, 15(5), pp. 555-68

Checkland P, Scholes J (1990) *Soft Systems Methodology in Practice*. J.Wiley, Chichester

Chesbrough H W (2003) *Open Innovation: The new imperative for creating and profiting from technology*, Harvard Business School Press, Boston

Cohen L J (1970) *The implications of induction*, Methuen, London, England

Curley M (2004) *Managing Information Technology for Business Value*, Intel Press, USA

Curley M (2006) The IT transformation at Intel, *MIS Quarterly Executive*, 5(4), December, pp. 109 – 122

Donnellan B, Helfert M (2010) The IT-CMF: A Practical Application of Design Science. In: *DESRIST 2010* (Winter R, Zhao JL, Aier S, Eds.), LNCS Vol. 6105, Springer, Heidelberg. pp. 550–553

Gamble M, Goble C (2011) Quality, Trust, and Utility of Scientific Data on the Web: Towards a Joint Model, In *Proceedings of the ACM WebSci'11*, June 14-17 2011, Koblenz, Germany. pp. 1-8

Gericke A, Fill H, Karagiannis D, Winter R (2009) Situational Method Engineering for Governance, Risk and Compliance Information Systems, In *Proceeding DESRIST '09 Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology,* Article No. 24, May 7-8, 2009, Malvern, PA, USA.

Gill G T, Hevner A R (2011) A Fitness-Utility Model for Design Science Research. In *Proceeding DESRIST '11 Proceedings of the 6th International Conference on Design Science Research in Information Systems and Technology*, pp. 237-252

Goldkuhl G (2004) Design Theories in Information Systems – A Need for Multi-Grounding. *Journal of Information Technology and Application*, 6(2), pp. 59-72

Gregor S, Jones D (2007) The Anatomy of a Design Theory, *Journal of Assoc. Information Systems*, 8(5), pp. 312-335

Gummesson E (1988) *Qualitative Methods in Management Research*. Chart-well-Bratt. Lund, Sweden

Heckathorn D (2002) Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations, *Social Problems*, 49(2), pp. 11-34

Helfert M, Curley M (2012) Design Science in Action: Researching and Developing the IT-CMF, In *European Design Science Symposium 2011. Communications in Computer and Information Science Series* (Helfert M, Donellan B. Eds.), Springer, Heidelberg, pp. 50-59

Helgesson Y Y L, Höst M, Weyns K (2012) A Review of Methods for Evaluation of Maturity Models for Process Improvement, *Journal of Software Maintenance and Evaluation: Research and Practice*, 24(4), pp. 436 -454

Hevner A R (2007) A Three Cycle View of Design Science, *Scandinavian Journal of Information Systems*, 19(2), pp. 87-92

Hevner A R, March S T, Park J, Ram S (2004) Design Science in Information Systems Research, *MIS Quarterly*, 28(1), pp. 75-105

Hevner A R, Chatterjee S (2010) *Design Science Research in Information Systems*, Springer Science and Business Media, Heidelberg, Germany

Humphrey W (1989) Managing the Software Process. Addison Wesley, Reading, MA

Iivari J (2007) A paradigmatic analysis of Information Systems as a design science. *Scandinavian Journal of Information Systems*, 19(2), pp. 39-64

Johannsen W, Goeken M (2007) *Referenzmodelle für ITGovernance*, dpunkt verlag GmbH, Heidelberg

Juran J M (1998) How to think about Quality, in *Juran's quality handbook* (Juran J M, Godfrey A B, Eds.), 5th ed., McGraw-Hill, New York, pp. 2.1-2.18

Kaplan B, Duchon D (1988) Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study, *MIS Quarterly*, 12(4), pp. 571–586

Klein H K, Myers M D (1999) A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems, *MIS Quarterly*, 23(1), pp. 67-93

Lee Y W, Strong D M, Kahn B K, Wang R W (2002) AIMQ: a methodology for information quality assessment, *Information & Management*, 40(2), pp.133-146

Leonard-Barton D (1995) *Well Springs of Knowledge: Building and Sustaining the Sources of Innovation*, Harvard Business School Press, Boston, MA.

Liu K (2000) *Semiotics in information systems engineering*, Cambridge University Press, Cambridge, England

March S T, Smith G G (1995).Design and natural science research on information technology, *Decision Support Systems*, 15(4), pp. 251-266

Mathiassen L, Nielsen P A (2008) Engaged Scholarship in IS Research, *Scandinavian Journal of Information Systems*, 20(2), pp. 3-20

Mettler T (2009) A Design Science Research Perspective on Maturity Models in Information Systems, Institute of Information Management, Universtiy of St. Gallen, St. Gallen

Moody D L, Shanks G (2003) Improving the quality of data models: empirical validation of a quality management framework, *Information Systems*, 28(6), pp. 619–650

Nunamaker J F, Dennis A R, Valacich J S, Vogel D, George J F (1991) Electronic Meeting Systems to Support Group Work, *Communications of the ACM*, 34 (7), pp 40–61

Oppenheim A N (1992) *Questionnaire Design, Interviewing and Attitude Measurement*, Pinter, London, England.

Ostrowski L, Helfert M, Hossain F (2011) A Conceptual Framework for Design Science Research, In *Perspectives in Business Informatics Research, Lecture Notes in Business Information Processing* (Grabis J, Kirikova M), Volume 90, pp. 345-354

Paulk M C, Weber C V, Curtis B, Chrissis M B (1993) Capability Maturity Model for Software, Version 1.1. Technical Report, Carnegie Mellon University / Software Engineering Institute, Pittsburgh, http://www.sei.cmu.edu/library/abstracts/reports/93tr024.cfm

Peffers K, Tuunanen T, Rothenberger M A, Chatterjee S (2007). A Design Science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), pp. 45-77

Price R, Shanks G (2005) A semiotic information quality framework: development and comparative analysis, *Journal of Information Technology,* 20 (2), pp. 88–102

Pries-Heje J, Baskerville R (2008) The Design Theory Nexus, *MIS Quarterly*, 32(4), pp. 731-755.

Pries-Heje J, Baskerville R, Venable J (2008) Strategies for Design Science Research Evaluation, In *Proceedings of the 16th European Conference on Information Systems (ECIS 2008)* (W. Golden, T. Acton, K. Conboy, H. van der Heijden & V. K. Tuunainen, Eds.) Galway, Ireland: National University of Ireland., pp. 255-266.

Röglinger M, Pöppelbuß J, Becker J (2012) Maturity models in business process management, *Business Process Management Journal*, 18(2), pp. 328 - 346

Rubin H J, Rubin I S (1995). *Qualitative Interviewing: The Art of Hearing Data*, Sage, Thousand Oaks, CA

Schön D A (1983) *The Reflective Practitioner: How Professionals Think in Action*, Basic Books, New York

Sein M, Henfridsson O, Purao S, Rossi M, Lindgren R (2011) Action Design Research. *MIS Quartely*, 35(1), pp. 37-56.

Simon H A (1969) *The Sciences of the Artificial. Cambridge, Massachusetts*, MIT Press, Sociology of Knowledge, Doubleday

Stamper R K (1973) *Information in Business and Administrative Systems*, John Wiley and Sons, New York

Sun Y, Kantor P B (2006) Cross-Evaluation: A New Model for Information System Evaluation, *Journal of the American Society for Information Science and Technology*, 57(5), pp. 614-62.

Trauth E M, O'Connor B (1991) A study of the interaction between information technology and society: An illustration of combined qualitative research methods, In "Information Systems Research: Contemporary Approaches & Emergent Traditions (Nissen H E, Klein H K, Hirschheim R, Eds.), Amsterdam: North-Holland, pp. 131-144

Vaishnavi V K, Fraser M D (1998). Validation framework for a maturity measurement model for safety-critical software systems, In *Proceedings of the Annual Southeast Conference*, Marietta, GA, USA, pp. 314–322

Vaishnavi V K, Kuechler W (2004) Design Science Research in Information Systems, January 20, 2004, last updated September 30, 2011. URL: http://www.desrist.org/desrist

Van Aken, J E (2005) Management Research as a Design Science: Articulating the Research Products of Mode 2 Knowledge Production in Management, *British Journal of Management*, 16(1), pp. 19-36

Van de Ven A H (2007) *Engaged Scholarship: A Guide for Organizational and Social Research*, Oxford Univ. Press, Oxford, England

Venable J R (2006) A Framework for Design Science Research Activities, In *Proceedings of the Information Resource Management Association Conference*, , 21-24 May, Washington, DC, USA

Venable J R, Pries-Heje J, Baskerville R (2012) A Comprehensive Framework for Evaluation in Design Science Research, In *Design Science Research in Information Systems. Advances in Theory and Practice*, *Lecture Notes in Computer Science* (Peffers K, Rothenberger M Marcus, Kuechler B, Eds.) Volume 7286, Springer, Berlin, Heidelberg, pp. 423-438

Wang R Y, Strong, D M (1996) Beyond accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), pp. 5 – 33.

Wengraf T (2001) *Qualitative Research Interviewing: Semi-Structured, Biographical and Narrative Methods*, Sage Publications, Thousand Oaks, CA

Yin R K (2009). *Case Study Research: Design and Methods*, *Essential guide to qualitative methods in organizational research*, 5th Ed., Sage Publications, Thousand Oaks, CA.

Appendix: Questionnaire

Syntax

Syntactical completeness	Templates are used in accordance to development instructions All required material is completed
Consistent representation	Material is consistently presented in the same format and according to the development instructions (e.g. Spelling, Names, Colour Schema)
Accessibility	All material is available on the development platform All workgroup members have access to the material on the development platform

Semantic

Precise definitions and terminology	Definitions and terms are described within the global lexicon
Interpretability	All key terms and definitions are available on the global lexicon Expert terminology which is not defined in the global lexicon is avoided Abbreviations are defined
Easy to understand	The material is easy to understandable for information managers The material is easy to comprehend The CC is positioned in the overall framework and relationships to other CCs are described
Accuracy and Objectivity	The content review with key opinion leaders and peer reviews is documented The material is evaluated in respect to other relevant external frameworks Reference list of sources is included
Consistent content	Consistent use of terms in accordance to the global lexicon (Synonyms and Homonym are avoided) Relations between CCs are described consistently and reciprocal
Concise	Short and clear sentences with no repetition

Pragmatic

Relevance	Material is relevant and important for organisations
Usability	The maturity assessment can be operated in an organisation Usable to derive conclusions and recommendations for an organisation Usable to compare organisations (benchmark)
Timeliness and Actuality	This material reflects current practises of organisations with a typical distribution in maturity profiles
Efficiency	The assessment approach is in accordance to comparable assessment approaches within the IT-CMF (timeframe, effort, resources, personnel)

About the Authors

Markus Helfert is the Director of the Business Informatics Group and a Lecturer (Tenured) in Information Systems at Dublin City University, Ireland. He is programme chair of the European M.Sc. in Business Informatics at Dublin City University. His research interests include Design Science, Information Quality, Cloud Computing, Enterprise Analytics and Digital Learning. He holds a Doctor in business administration from the University of St. Gallen (Switzerland), a Master-Degree in business informatics from the University Mannheim (Germany) and a Bachelor of Science from Napier University, Edinburgh (UK-Scotland).

Brian Donnellan holds the Chair of Information Systems Innovation at the National University of Ireland Maynooth (www.nuim.ie) and is Co-Director of the Innovation Value Institute (www.ivi.ie). Prior to joining NUI Maynooth, Prof. Donnellan was a faculty member in the National University of Ireland, Galway. He has spent 20 years working in the ICT industry where he was responsible for the provision of IS to support Product Development. He is an expert evaluator for the European Commission and has been guest and associate editor of several leading IS journals including Journal of IT, Journal of Strategic Information Systems and MIS Quarterly.

Lukasz Ostrowski is a PhD student at Dublin City University, Ireland. He holds a First Class Honours Bachelor Degree in Information Systems Management and an Ordinary Bachelor Degree in Business from Galway-Mayo Institute of Technology, Ireland. His research interest includes business process and data modelling, design science, application of information quality, and ontology engineering. He is currently working on a research process to design IT-Management processes at Dublin City University. Since 2010, he has been serving as a member of Business Informatics Group, which takes an active role in aligning strategy, processes and information technology.