http://www.casa.ucl.ac.uk/working_papers/paper82.pdf

ISSN: 1467-1298

82
05 | 04

Codes of Life: Identification Codes and the Machine-Readable World

Martin Dodge
Rob Kitchin

www.casa.ucl.ac.uk

T +44 (0) 20 7679 1782 • T +44 (0) 20 7679 1813 • F +44 (0) 20 7813 2843 • E casa@ucl.ac.uk

Centre for Advanced Spatial Analysis • University College London • 1 - 19 Torrington Place • Gower St • London • WC1E 7HB

[This is a Code 128 format barcode which encodes the paper title, ‘Codes of Life’.]

Abstract

In this paper we present a detailed examine of identification codes, their embeddedness in everyday life, and how recent trends are qualitatively altering their nature and power. Developing a Foucaultian analysis we argue that identification codes are key components of governmentality and capitalism. They provide a means of representing, collating, sorting, categorising, matching, profiling, regulating; of generating information, knowledge and control through processes of abstraction, computation, modelling and classification. Identification codes now provide a means of unique addressing all the entities and processes that make up everyday life – people, material objects, information, transactions and territories. Moreover, they provide a means of linking these entities and processes together in complex ways to form dense rhizomic assemblages of power/knowledge. At present, however, the information that identification codes provide access to are, at best, olgopticon in nature; that is they afford only partial and selective views. In the latter part of the paper we outline four trends – wide scale trawling for data, increased granularity, forever storage, and enhanced processing and analysis – that seek to convert these partial olgopticons into more panoptic arrangements. In turn, we contend that these trends are part of a larger meta-trend – the creation of a machine-readable world in which identification codes can be systematically and automatically ‘read’ and acted on by software independent of human control. This meta-trend is supported and driven by interlocking discourses such

as safety, security, efficiency, anti-fraud, citizenship/empowerment, productivity, reliability, flexibility, economic rationality, and competitive advantage to construct powerful, supportive discursive regimes.

1. Introduction

April 2003 marked the fiftieth anniversary of the discovery of the double helix structure of DNA, unlocking the fundamental genetic code of life (Watson and Crick, 1953). Like the DNA sequences that are unique to each individual, there now exists a plethora of unique codes that correspond to, and represent, individuals. These are the myriad of often taken-for-granted identification codes that we use on an everyday basis (e.g. bank account and credit card numbers, passport number, social security number, computer passwords, and so on). Similarly, material objects are increasingly assigned unique identifiers that allow them to be processed, shipped and traced through complex logistical networks (e.g. manufacturer serial numbers, barcodes, order numbers, shipping numbers, the vehicle license plate, Internet addresses and such like). People, objects, information and transactions can be tied to territorial location through another set of codes (e.g. grid references, latitude and longitude, postal codes, enumeration districts, and geodemographic area types). These identification codes are tangible links into the rhizomatic assemblages of bureaucratic information, logistics, surveillance and security in which we increasing live and interact. While many are banal and are treated ambivalently, identification codes are increasingly vital to everyday interactions, transactions and mobilities. They are essential referents for government agencies and institutions for managing and monitoring populations and businesses for logistics, accounting, and so called ‘customer-relations management’ systems. Their power lies in their utility as devices for representing, collating, sorting, monitoring, matching, profiling; that they allow analytical processes of abstraction, computation, modelling, classification – of generating information and knowledge; to search, sift, differentiate, and regulate complex systems, to provide competitive advantage, to target markets, to reduce real and perceived risks, and enable finer grained “social sorting” (Lyon, 2003a). They provide the practical means for dealing with vast, complex sets of information and to draw sense from these sets.

In this paper, we outline an analytic framework through which to conceptualise identification codes applying it to detail the utility and power of identification codes, to illustrate their embeddedness in everyday life in Western society, and to examine recent trends that qualitatively alter their nature and power. As such, we do not examine either the operation and practice of wider information systems (see Dodge and Kitchin forthcoming; Thrift and French 2002) or broader notions of 'codes' in the form of socio-cultural structures and technical/legalistic protocols of ordering and control, such as national laws, international treaties, etiquette, standards, systems of measurement, institutional customs, professional codes of conduct (see for example Alonso and Starr, 1987; Beniger 1986; Bowker and Star, 1999; Lessig, 1999; Torpey and Caplan, 2001; Wise, 1995 for wider ranging reviews)[1]. Nor do we discuss in detail the implications and consequences of broadening and deepening of surveillance made possible by the increasing use of identification codes, which we detail in another paper.

To structure the discussion the paper is divided into four sections. In the first section we provide an introduction to identification codes and their workings. In the second, we outline their importance employing a Foucaultian framework to illustrate their utility and power. In the next section we use a simple typology to provide an illustrative overview of five sets of identification codes – people, object, information, transaction and territorial. Finally, we examine four recent trends that are transforming the scope, utility and power of identification codes: the drive towards capturing the maximum amount of information as possible, facilitated by the transference from dumb to smart technologies, new dynamic means of monitoring entities, and the mobile tracking of entities; the drive towards finer resolution input and in particular the development of unique identifiers that are infallible; the development of cheap and efficient long-term storage of digital records; and enhanced means of linking datasets and more sophisticated and precise methods of data mining. These trends, we posit, are the result of a larger meta-trend – the move towards a *machine-readable* world. That is, a move to identification codes that are increasingly held on media that can be 'read' and acted on by software independent of human control.

2. Identification codes

Identification is the process of telling like things apart. In practical terms it is the ability to correctly pinpoint a particular individual or item from a group. Once pinpointed and identified, this entity can be worked upon; processed in some fashion that differs from the group as a whole. Consequently, processes of identification are intimately bound to processes of differentiation, wherein the mass is separated into individual or other units of manageable size. Identification in practice is achieved through labels, tokens and names, and increasingly consists of unique identification *codes*, that is strings of numbers and letters. These unique codes are generated automatically by mathematical algorithms using software programs and can be tracked by sophisticated information systems. In many real-time data processing systems (e.g. banking and stock markets, air traffic control) billions of transactions can be logged each day, each transaction encoded with unique identifiers that usually include the time, date and place of the transaction.

The key characteristics of identification code numbers are their form (length and formatting), their granularity, and the extent of recognition/reputation. A code number can be defined by the creator of the system to be any length, particularly when they are only ever handled by computers. Where identification codes are used to authorise access to significant resources (especially money), then code length can be significant in determining security. Many identification code numbers used in business and telecommunications are very long indeed, for example, a standard identification code number, a 'public key', used in secure digital communications (Figure 1; see Kahn, 1967, for the classic historical analysis of cryptographic codes). Codes also vary as a function of the type of valid characters they can contain – numeric, alphabetic. With regards to granularity, the most obvious level for people is that of the individual. For objects and spaces, granularity is more problematic. For example spaces vary as a function of scale (e.g. households, neighbourhoods, districts, towns, regions, and so on). The granularity of codes provide different levels of resolution, with the trend to increasing granularity, thus maximising resolution. Finally, codes vary in their ability to be recognised temporally, geographically and institutionally. Some codes are long lasting and recognised globally (often coordinated through ISO, the International Organization for Standardization), while others only work within a more localised context, for example within a given office or institution.

```
-----BEGIN PGP PUBLIC KEY BLOCK-----
Version: PGP Key Server 0.9.6

mQGiBD84pMERBADr2vRWyNdkhtcrBmqB/mOQ3+Rl2sTuj1Y61sD14Lfru9fJyBfh
4qlOh3Ldn2k0BN/XhpUYxepSpWFU8ChD8+cjnpgSl5V28MuWRXdT/EPByurLk4TN
67hcX41HUNZWcrrBFCiTkLMF8WJNgbyYTXEzJUPh60BOwm+o9oIdKP8ofwCgpNpY
6bp6VTVN2ONrJn+RTfU/iEEEAIVmen9jqqAH6YWidwz3KaI505K1jPbZ3O6jaDdU
pdAnyXOjquJV0FLT0vifDYus+DNovEHndAuQiTcb+Eyeb/B3lc4OEzHQ8o3fDgki
2gPdxDcnL8K+W2LfL7J2+zwOOceg7UP5rYVwSgsmQhIml/KB8jmMJ9KoSvxI2p3N
EFYxBADCiNEjaJfG+g5gRbp4AubgPb11KFq5xU13YEF9uYC+l6jgrNMQ5a7MUsbR
5dIpZCIWlS1XgzxuiWfi1gzntvwyLtQkApKSvtLZe+XchCr1Jfauy/OUbH9PCWyq
p8h5EG3Xx8mb8mRABpExpMPNl9Hc9thgJy8YzUtvTqkgy0hqcrQsU3RlcGhlbiBD
b2FzdCAoU3RldmVDKSA8c3RldmVAZnJhY3RhbHVzLmNvbT6IYQQTEQIAIQUCPzik
wQUJACeNAAYLCQgHAwIDFQIDAxYCAQIeAQIXgAAKCRCoHKYVz0vjXnoUAJ4j4VYQ
+S2KVObgwCgJkYb/FvJTWACgmawBuFy/AOZqfqCuyUPJ8xn708u5AQ0EPzikwRAE
AK/B4Dq6ZCJjNVKls16m3Xx1ReYx8hFFJlk01zy2yGBFNdvk6XgGPl4Jj8ix3bTl
vZuHvybT1A230uvujDT9TMxNFtJHOtbfn5eHNkoK46kEKjAfSmaNoVBUBKn8GwWQ
EkNsxf/mZzG1nwV87u7ETHNOwQJURsEzvxmtyPPdHXh7AAMFA/9zKdGiL2MsLxnM
R+YAm3G2vrktlPUKhR0Pc/qtoH2L3T5Kj9pJiKk6Hyvo7lTgYRJBIS4oA+rB/BZz
WMx4UrYjApnACFTO9DOaGcuhRJSuouSzxXUmskqhV8VuO7FEmgRPO3uRQ/K1qvGL
SIbIcdKjShr1GQtRBKXOR2W+piZpaIhMBBgRAgAMBQI/OKTBBQkAJ40AAAoJEKgc
phXPS+NeMeIAmwaLsjl68Y0x+rK/l+HRN1RrY1S+AJ0R36UzxLuoITpBZoIqxWDa
sDEf/A==
=MLie
-----END PGP PUBLIC KEY BLOCK-----
```


Figure 1. A 1024 bit public key used as an identification code for encrypting messages for secure email communication using the *Pretty-Good Privacy* software. This code provides a digital signature which effectively identifies the sender. (Source: courtesy of Steve Coast, http://www.fractalus.com.)

In combination these identification codes – for people, objects, information, transactions, and territories – are, we contend, key *capta*. Jensen (cited in Becker, 1952) details that capta are units of data that have been selected and harvested from the sum of all potential data. Here, *data* (derived from the Latin *dare*, meaning 'to give') is the total sum of facts that an entity can potentially 'give' to government or business or whomever is constructing a database. Capta (derived from the Latin *capere*, meaning 'to take') are those facts that those constructing the database decide to 'take' given that they cannot record or store everything[2]. They are elements of the continuous material-social reality of everyday life that are most easily intelligible and distinguishable. Capta thus has a specific form, a specific mode of representation, a purpose, is meaningful to the agency that selected it, and has context (Alexander, 2002). Capta is inherently partial and selective, and the distinguishing criteria used thus have consequence.

Identification codes are key capta because they provide the ties that link together separate strands of capta. Capta is important in the contemporary setting of information system design because when placed into a wider context, through say specific queries or forms of analysis, meaning is attributed to it thus converting it into *information* (Checkland and Holwell, 1998; Haklay, 2001). When fields of information are combined to create larger, collective structures they become *knowledge*. Identification codes thus form the basis of a dense rhizomic assemblage of related capta and information within and between systems. This assemblage constitutes what might be termed a 'capta shadow'[3] (capta that uniquely *represents* people, objects, information and transactions) that often has associated 'capta trails' (a record of the *location* of interactions and transactions) across space-time.

Conceptually, it is important to separate the identification code from the material media on which it is stored, moved, represented and retrieved. The lines of the barcode, the magnetic strip on the credit card and the chip in an electronic key are merely a means of holding the code in a convenient form for use. This separation is readily apparent when one considers the power of the credit card number versus the 2x4 inch plastic card that carries it. Although we are primarily concerned with the codes not the carry media in this paper, it is not to say they are unimportant. In certain instances the media is a vital element in the operation of identification, helping to enhance the credentials of the issuer and the legitimacy of the code it carries, such as the multiple design features of passports giving them status and making them hard to forge.

3. The discursive regimes of identification codes

As noted, the power of identification codes comes from their ability to discriminate, to be rapidly and automatically monitored, updated, and processed, and to provide authentication and credentials that dictate various forms of access (e.g. to a bank account, government services, or a location). Identification codes provide information systems their technicity (Mackenzie, 2002; Dodge and Kitchin, forthcoming), that is, the ability to be able to process the capta they hold. In contemporary Western society, consistent and accurate systems of identification are seen as vital, with the proliferation and growing density of identification codes, and the technologies that produce and process them, driven by multiple discourses that primarily relate to governmentality and capital accumulation (see Innes, 2001): codes are useful for both governing people and managing organisations.

Governing People

It is now widely acknowledged that the Enlightenment and the shift into a modern society was accompanied on the one hand by new, more systematic means of managing and governing populations, and on the other by a drive to scale and deliver a uniformity of service across whole populations (e.g. law enforcement, health provision, tax assessment). As a consequence, people became increasingly viewed as components in larger systems as labour commodities, as problems to be solved (e.g. ill-health, criminality, illiteracy), and as citizens. As such, modernity required the 'quantification' of society - cataloguing and classification - much of which necessitated 'objective' means of identification and accurate measurement systems (see Alonso and Starr, 1987; Barnes and Hannah, 2001; Desrosières, 1998; Porter, 1995; Thrift, 2004; Wise, 1995). Identification codes provided the operational means of identifying, verifying and processing people by distant institutions and impersonal management systems as "other tokens of trust were sought, to make up for the lack of visual, body clues and cues, such as handshakes, eye contact, and so on" (Lyon, 2002, page 244). As such, there has been an on-going shift from localised, subjective knowledges to centralised, objective ones. This modernist desire to 'order the world' thus required technical infrastructures of which systems of identification were a key part, including "… that oldest means of ordering space, time and individuals: the list" (Leyshon and Thrift, 1999, page 436). As technologies advanced, first with the use of punch card machinery, then mainframe computer systems, and more recently distributed information systems, the ability to create and process vast quantities of capta in shorter and shorter time periods has progressed significantly. This has led to the development of assemblages of personal identification codes from passports, driver's licences, national insurance or social security numbers, work identification cards, to bank and store loyalty cards, each designed on the one hand to identify and authenticate their owner as 'trustworthy', and on the other to allow them to be 'managed' by the issuer of the card (see Rule et al. 1986; Torpey, 2000; Torpey and Caplan, 2001 for detailed reviews).

In the writings of Foucault, particularly his genealogies (Foucault, 1976, 1978), the development of modern systems of population management – needed to propagate and sustain the emerging the capitalist system and its associated industrialisation, urbanisation and colonisation - provided an apparatus of regulation and 'technologies of the self'. Here, systems of regulation and management were devised to provide a detailed knowledge of a population, "…a way of rendering it into thought, so that it can be examined, evaluated, its ills diagnosed and solutions prescribed … Such *representation* has two major aspects: the articulation of languages to *describe* the object of government and the invention of devices to *inscribe* it" (Rose, 1996, page 70, original emphasis). Here, the translation of individuals into domain of knowledges – 'calculable people' - made it possible to govern according to norms defined and legitimated by the status of science, with professional elites able to ground their authority in an objective knowledge (Rose, 1996). Examples of these capta include censuses, health records, school attendance, criminal records, tax records, registration of births, deaths, marriages, and so on; all of which have grown enormously in volume and scope in recent decades. The systems of regulation that identification codes bound individuals to aimed to both discipline and self-discipline, by producing a particular form of rationality designed to ensure good government through a more efficient and rationalized legal and social field (McNay, 1994). These systems work to instil self-discipline because the mechanism of identification and monitoring threatens observation that is ever present; the potential pervasiveness of observation, combined with its effects, leads to self-regulation, changing the practices of governmentality. As Hannah (1992, page 53) points out, "[j]udgement must seem automatic and impartial, responsibility lies as completely as possible with the perpetrator." Within this context, disciplinary governmentality involves "a complex constellation of group control … (through censuses, bureaucracies, policies, and so forth), an archipelago of disciplinary institutions (schools, workplaces) and an otherwise anonymous territory dotted with millions of staging points (ATM's, surveillance cameras, application forms)" (Hannah, 1997a, page 178).

Throughout modernity, as Hannah (1997b) notes, the panopticon effect of disciplinary forms of governmentality have been open to vertical (within an activity) and horizontal (across activities) fragmentation. That is, within an activity observation, judgement and enforcement has often been undertaken by different agencies who communicated imperfectly; across activities, different organisations for legal or technical reasons have not easily been able to exchange or compare information (Hannah, 1997b). Latour (cited in Amin and Thrift, 2002, page 92) thus refers to the imperfect panopticon of government as an ‘olgoptican’ – vantage points from above, below and within – “a series of partial orders, localised totalities, with their ability to gaze in some directions and not others.” As Hannah (1997b, page 352) states the “modern citizen [becomes] objectified in a life-path comprised of information, as a spatialised dossier. … For the average ‘free’ citizen the life-path … is full of gaps, but retains its unity through the matchability of names, permanent addresses, social security numbers, etc.”

The growing pervasiveness of identification codes and systems to match, trace, monitor and regulate populations, however, seeks to transform the olgoptican into a panopticon. Here, discourse is central as it provides the means through which to articulate, conceptualise and justify governmental goals by making it seem as if common problems are being addressed through shared logic and principles. The justification for the present proliferation of identification codes thus draws on discourses of safety, security, efficiency (the ‘best value’ mantra in public service provision in UK), anti-fraud, citizenship/empowerment (provides citizens with more access to government and governmental services and resources) delivered through a variety of agencies and structures. Innes (2001, no pagination) also notes the importance of key symbolic events in discourses, for example in the case of security surveillance, there are the ‘signal crimes’ “whose mediated coverage is constructed in such a way as to connotatively and denotatively signal to society that there is a ‘public problem’ that needs addressing.” The attacks of 9/11 are the most obvious ‘signal crimes’ of recent decades. Important to note, contra Foucault, is that these discourses work not simply because they act to discipline but because they also seduce. In Althusser’s (1971) terms, these discourses thus interpellate by enticing people to subscribe to and desire their logic and to willingly and voluntarily participate in their ideology and practice (rather than simply disciplining them into docile bodies) (see Dodge and Kitchin, 2004).

Managing Organisations

In tandem with the growth in systems designed to manage populations, there has been a drive to create highly ordered means of operating organisations such as civil services, the military, healthcare providers, local authorities, and businesses. Focusing on the later as illustration, modern business is predicated on systems of practice that work to ensure the maximisation of profits. These systems work to ensure efficiency, productivity (both with regards to workers and production systems), and competitive advantage. Over time the dominance of particular systems, for example Fordist modes of production, have been replaced by new systems that take advantage of new technologies and new forms of transportation, communication and information management. In recent years, the proliferation of distributed forms of communication such as the Internet and the development of complex information systems that can track and manage complex chains of components, commodities and data is now widely acknowledged as having significant effects on the organisation of business globally (Castells, 1996; Dicken, 2003). Here, identification codes and their associated media are essential components. Identification codes, whether they be printed serial numbers, universal product codes (UPC) in grocery barcodes, the Electronic Product Code (EPC) in RFID chips or other kinds of smart label provide a means of unambiguous identification and tracking – the knowing of what, when and where of things at all times is the goal – and the ability to act on this knowledge with regards to sorting, aggregating, directing, linking, processing, and so on. Totally identified processes are seen to reduce the real, and more importantly, the *perceived* risks to business managers.

Like the management of populations, new business systems and practices are driven by powerful discourses designed to discipline, seduce and drive transformation as capital, to use Harvey's (1982) phrase, seeks a new 'spatial fix' that will maximise profit. These discourses include efficiency, productivity, reliability, flexibility, economic rationality, and competitive advantage, each drawing off and reproducing business management and neo-liberal ideology. Here, the rhetoric is that identification codes enable the automation of management, production and logistical systems making them more slim-line and efficient; they allow for the speeding up of processing; they are highly reliable reducing human error; they enable highly complex and flexible systems that can be automatically tracked and processed from a distance. In combination these are used to argue that identification codes provide significant cost savings by reducing staff costs both through automation and by allowing businesses to take advantage of lower wages in different locations without losing quality or control. This in turn allows companies to gain competitive advantage over their rivals. As a consequence, many companies have been seduced by the power of wholesale application identification codes and have transformed their business, re-organising product manufacture, company administration, service delivery, and the ways in which they interact with customers/clients (for example, altering modes of payment through credit cards, introducing loyalty cards, utilising call centres or direct interfacing with information systems via the Internet). Indeed, a noteworthy feature of many interactions between individuals and organisations in contemporary Western society – the telephone query to a customer service centre being the archetypal example – is that they often begin with a 'reference number handshake' (the quoting of an account number or reference code to establish authenticity). Here, identification codes provide a means of authentication and accreditation replacing earlier forms of self-authentication and vouchsafing. For some interactions, especially financial ones, this often involves a lengthy series of identification codes and security phrases. These routine interactions are made personally identifiable and thus provide transaction-generated information (TGI) that enables customer profiling and targeted marketing. Indeed, there seems to be a drive in contemporary business practices in many sectors to remove the potential for anonymous transactions.

The ability to uniquely identify and track individuals has also led to changes in workplace practices by enabling increased security and the ability to monitor work patterns, productivity rates and performance. For example, it is increasingly common for employees to be issued with work identity codes carried on cards; swipe cards, magnetic key-fobs or pin-numbers for keypads to regulate access to rooms and buildings; and usernames and passwords to access networked services. While these systems on the one hand provide security they also increasingly used as a means of workplace surveillance, used to log and track worker whereabouts and performance (particularly in jobs where managers want quantitative measures of employee 'throughput', e.g. working on a checkout or in a call centre).

4. Types of identification codes: people, objects, information, transactions and territorial

In this section we provide illustrative examples of different identification codes used for five different entity types to demonstrate the sophisticated nature, scope and employment of current codes. In each case, the examples we detail are digital machine-readable codes.

Personal codes

> "As every man goes through life he fills in a number of forms for the record, each containing a number of questions. A man's answer to one question on one form becomes a little thread, permanently connecting him to the local centre of personnel records administration. There are thus hundreds of little threads radiating from every man, millions of threads in all. If all these threads were suddenly to become visible, the whole sky would look like a spider's web ….. They are not visible, they are not material, but every man is constantly aware of their existence. …. Each man, permanently aware of his own invisible threads, naturally develops a respect for the people who manipulate the threads …" (Alexander Solzhenitsyn, 1968, 223-224).

> "I am not a number, I am a free man!" (Patrick McGoohan, 1960s television show *The Prisoner*).

Nowadays, pretty much all people in Western societies are shadowed by a voluminous and ever growing number of personalised records in the databases and customer management systems of the businesses, organisations and government agencies that service modern living. On their own, daily, routine computer-mediated economic transactions (e.g. credit card use) entangles the daily life of each person into a denser web of threads, across time and space, than Solzhenitsyn could have imagined when he wrote the above quotation in 1968. This is supplemented by many other systems as more and more daily activities come to require personal identification codes in order to take place.

The proliferation of codes means that individuals are effectively accompanied by a parallel capta shadow. A capta shadow is the digital version of Solzhenitsyn's threads – the personal records and transaction-generated information held by institutions and businesses about an individual. While an individual's capta shadow is probably best characterised as multiple - capta shadows (to paraphrase McGoohan, individuals are not just a number, but many numbers) – with the records held by one organisation not necessarily visible to another, it is true to say that the capta shadow for each individual is sizable and growing. Indeed, a capta shadow is inevitable and necessary to be able to function in contemporary society, for example to work legitimately and pay taxes, to access government services, to travel, or to buy commodities (Clarke, 1994a; Lyon, 2002). That is not to say that capta shadows do not have their consequences with regards to privacy, confidentiality and civil rights (see for example Gandy's 1993 analysis). For example, individual capta shadows are now routinely used to 'socially sort' individuals, being used to evaluate perceived worth and risk through activities such as customer profiling, crime profiling, and constructing credit ratings (Lyon, 2003a; Graham 2004). Consequently, Stalder (2002, page 120) notes,

> "… this shadow body does more than follow us. It has also begun to precede us. Before we arrive somewhere, we have already been measured and classified. Thus, upon arrival, we're treated according to whatever criteria has been connected to the profile that represents us."

While individuals actively produce their capta shadow through their actions and transactions, they do not have full control over its form, extent, or how it is used. Moreover, capta shadows have become valuable, tradable commodities as evidenced by the growth of credit reference agencies, lifestyle profiles and geodemographics systems (Goss, 1995; Larson, 1992). Personal identification codes are important here as they are the *means* by which an individual and their capta shadow are continually and consistently bound together; they can no more be separated from it than be separated from their physical shadow cast on a sunny day. In their digital form they make people machine-readable. Clarke (1994b) has classified personal identification coding systems into three general types, (1) 'something you have' (a physical token like a passport), (2) 'something you know' (e.g. a password or PIN), (3) 'something you are' (e.g. physical characteristics, particularly the face).

It is possible to illustrate these codes and an individual's capta shadow by deconstructing a wallet, identifying all the codes contained within on any given day. As Table 1 reveals, the wallet contains identification codes relating to personal identity and authentication, financial accounts, travel, consumption, security. Other such individual codes can be found by sorting through bills and statements such as customer loyalty cards, gas and electricity utility account numbers, insurance policies, other financial accounts, and health insurance policies. Other codes are imposed without our knowledge, starting from birth. These codes bind individuals and their capta shadows together, with Table 1 illustrating the function, usage and processes of particular codes. In most cases the identification codes work to authorise payment or access.

The personal identification codes in Table 1, with the exception of the photo on the work id card, are forms of Clarke's first type of identification coding system. However, there is increasing interest in the third type of personal identifiers, 'something you are', and later in the paper we discuss in detail the development and use of new biometric identification systems.

Object Codes

> 'This little footprint … has built a gigantic structure of improvements – of size and speed, of service, of less waste, of increased efficiency. This lousy little footprint is like the tip of an inverted pyramid, and everything spreads out from it.' (Alan Haberman, chairman of the barcode selection committee, quoted in Leibowitz, 1999, page 132)

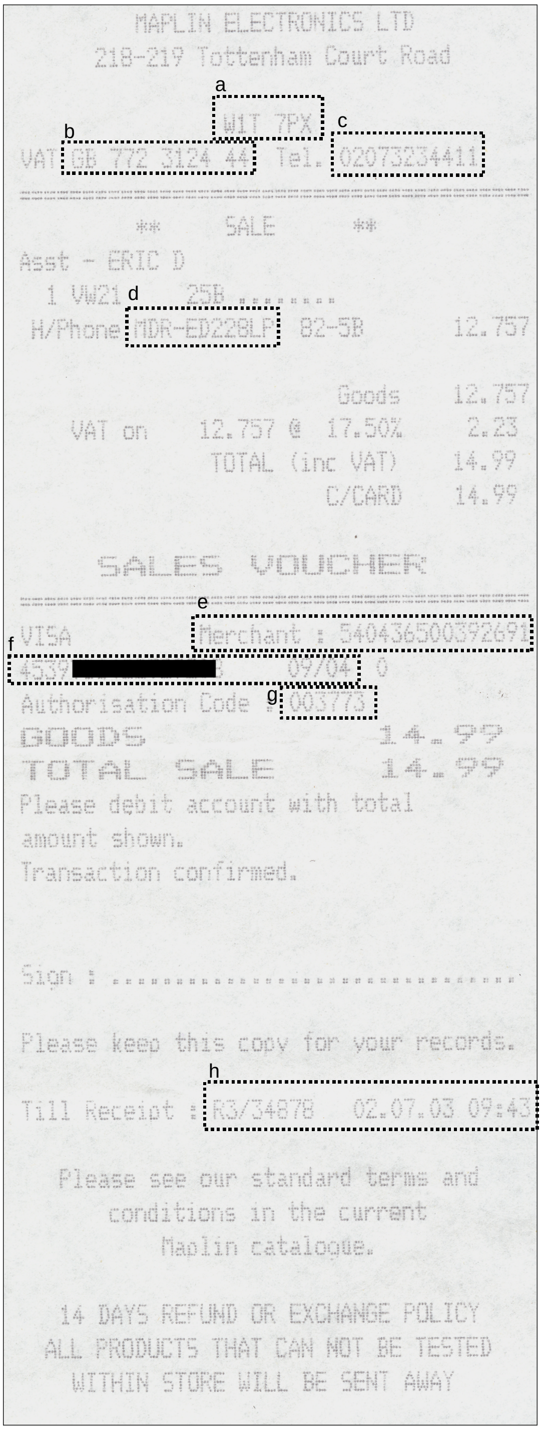
Capta shadows are not confined to people, they also almost universally accompany objects. Like people, objects are assigned unique identification codes including individual product code numbers (e.g. car chassis number[4]) and barcodes that facilitate manufacture, distribution, processing sales, and account management. In this section we examine UPC/EAN codes – the common barcodes on groceries - and a till receipt to illustrate the ubiquity of these codes and their utility.

The UPC (Universal Product Code) is undoubtedly the epitome of both the banality and power of identification codes in daily life. Little notice is taken of barcodes, yet they are present on virtually all retail products, manufactured goods, address labels on parcels, and an increasing number of official forms, and it is estimated that some 5 billion barcodes are scanned daily (Thrift, 2004, page 184). As a society we largely ignore barcodes primarily because they are not meant for humans, instead designed to be read by scanners and processed by information systems, to make physical objects machine-readable[5]. Each barcode has two distinct components. First an agreed allocation of unique numbers and, second, parallel black and white bars that can be read automatically using a laser scanner. The impact of the barcode system has been a reorganisation of retail businesses, transforming inventory control, logistic systems needed for 'just-in-time' operations, and the point of purchase. The first of UPC barcode was used on the 26th of June 1974 in a supermarket in Troy, Ohio to scan a pack of chewing gum (Morton, 1994). The original concept has been traced back to 1940s with current UPC standard being developed in the early 1970s by a group of U.S. retailers and food manufacturers, based on a design by IBM (see Brown, 1997; Savir and Laurer, 1975). Their successful development and diffusion was driven, in part, by the replacement of manual tills by new computerised point-of-sales technologies, so that by April 1976, some seventy-five percent of goods in U.S. supermarkets had a UPC barcode (Dunlop and Rivkin, 1997). The most common barcode code system is the UPC/EAN codes printed on consumer goods in the retail sector, but there are in fact a wide range of different formats of barcodes, with a diversity of printed symbol styles deployed for different application areas and industries, including the Code 128 format barcode used in the title of this paper (see Harmon, 1994; Palmer, 1995). Many large organisations and industrial sectors have developed their own particular form of barcode representation and protocols for allocating numbers. Each particular barcode has a defined visual structure to represent its code number and these can be relatively easily decoded (Figure 2).

Figure 2. Decoding the barcode. Shown is a 12 digit UPC (universal product code) barcode used in the U.S. and managed by the Uniform Code Council. Each digit is encoded as a binary number, represented by a particular sequence of seven black bars (1) and white spaces (0). The number 3 for example is represented by the sequence: space-bar-bar-bar-bar-space-bar (0111101). The binary representation in the left and right hand sides are inverses of each other, so the barcode can be scanned from either direction.

The barcode only provides a single identification code number, the rest of the details describing the object (e.g. product type, date of manufacture, price, etc) have to be held in a information system. In this way the true importance of barcodes is that they provide a very cheap, simple, unambiguous and reliable 'digital thumbprint' that can be placed on almost an object and consistently read back by a computer. Barcodes link material things to their virtual representation, making them machine-readable. Barcodes provide an easily distributable means of identification vital to maintaining distanctiated, global supply chains. Hosoya and Schaefer (2001, page 157) thus describe them as the 'bit structures' that organise and synchronise flows, acting as "the mechanism by which the virtual establishes its logic in the real." Further they provide a means, when accompanied by loyalty cards, of tracking consumer tastes and undertaking target marketing. As we detail below, the strengths of the barcode (simplicity and lack of granularity) are also its major weakness and, as a result, they are being replaced by new coding systems held by smart labels and RFID chips (that have a finer resolution and can hold more capta).

Returning to our wallet, the capta shadow associated with the retail sector can be illustrated by examining a till receipt (Figure 3). Receipts are almost all now computer generated by 'point of sales' terminals (tills in shops, barcode readers, ATMs). At one level they are a material manifestation of the banal nexus between daily life and routine, large-scale data processes that make retail possible. And yet, the receipt is also vital to the management of retail businesses, providing a record that links together several capta shadows at a particular time and place. This allows individual purchasers, store managers, retail chains, suppliers, and business regulators and administrators (e.g. tax offices) to track and monitor transactions.


MAPLIN ELECTRONICS LTD
218-219 Tottenham Court Road
W1T 7PX
VAT GB 772 3124 44 Tel. 02073234411

** SALE **
Asst - ERIC D
1 VW21 25B ....
H/Phone MDR-ED228LP 82-5B 12.757

Goods 12.757
VAT on 12.757 @ 17.50% 2.23
TOTAL (inc VAT) 14.99
C/CARD 14.99

SALES VOUCHER

VISA Merchant : 540436500392691
4539 [redacted] 09/04 0
Authorisation Code : 003773
GOODS 14.99
TOTAL SALE 14.99
Please debit account with total
amount shown.
Transaction confirmed.

Sign : ....

Please keep this copy for your records.

Till Receipt : R3/34878 02.07.03 09:43

Please see our standard terms and
conditions in the current
Maplin catalogue.

14 DAYS REFUND OR EXCHANGE POLICY
ALL PRODUCTS THAT CAN NOT BE TESTED
WITHIN STORE WILL BE SENT AWAY

a. `W1T 7PX`: Postal code, territorial identification system for mail delivery. Code is much more machine-readable than the full postal address.

b. `GB 772 3124 44`: VAT number, unique registration code for this business assigned by Customs and Excise for taxation. (VAT - 'value added tax', i.e. sales tax.)

c. `02073234411`: Telephone number of the store.

d. `MDR-ED228LP`: Product code, a manufacturer's object identification code. In this case, Sony super-light in-the-ear headphones.

e. `540436500392691`: Merchant number identifies the store to their bank.

f. `4539 xxxx xxxx xxxx`: VISA card number used for payment. Identifies the purchaser and provides link between the store's bank and the purchaser's bank.

g. `003773`: Authorisation code from the VISA payment system, generated in real-time in verification from the card issue to confirm for the retailer that payment is approved.

h. `R3/34878`: Receipt number generated automatically by the till. Date and time of transaction are also recorded.

Figure 3. The range of unique identification code on a typical consumer retail receipt. This example is for purchase of headphones in a high street electrical store in Britain.

Information codes

Moving beyond physical objects, there has been a move to generating capta related to information (text, sound, image or video). Conventionally, industry specific identification systems have developed focused around the particular media format that information is usually stored and distributed on, such as ISBNs for books, ISSNs for periodicals (*Society and Space's* being 02637758, which is represented as a barcode on the back cover of the printed issue) and ISWC (International Standard Musical Works Code) for music[6]. These codes are often inscribed on the physical media through barcodes. Many of these identification code numbers are developed and operated as necessary supports for the intellectual property regimes of publishers.

The rapid shift of information production to digital formats, most particularly in terms of Internet content in the 1990s has spurred the development of many new schemes of identification and metadata. The rapid take-off of the Web in the mid 1990s can be put down, in part, to its simple but effective identification scheme of URLs (uniform resource locators). This is a distributed scheme, that enables anyone to specify a globally unique identifier and network location of their homepage. However, a significant limitation with URLs is their lack of persistence – giving broken links and 404 error messages – due the failure to separate location from identity in the scheme. One attempt to solve this, which has found particular favour with libraries and academic publishers that require a high degree of permanence, is the DOI (digital object identifier) scheme. This allows information publishers to assign a unique identifier (e.g. Pion has assigned this article a DOI listed above the title) that never changes and lets the actual location of the information be stored separately. The DOI system was developed in late 1990s and is operated by the International DOI Foundation (www.doi.org) (Paskin, 2003). It is a distributed system (so publishers can use their own id scheme for items they 'own') and also compatible with existing Web software and URLs.

The growth in sharing of digital information (especially MP3 format for music) is forcing media corporations to develop new identification codes systems for enforcing control of the distribution and use of content. These codes will identify each individual instance of a piece of music and be tied to individual customers. As such, codes are at the heart of ‘digital rights management’ and will be key in future intellectual property ‘wars’ (Lessig, 1999).

Transaction codes

Many transactions in everyday life, particularly those involving communication and money of some kind, are also made machine-readable through the application of identification codes. These are usually generated automatically, often comprising some combination of exact time, location and nature of the transactions. Transaction identification codes serve especially as a means of authentication and are readily apparent on banal objects like receipts (Figure 3) and tickets. Indeed, in some circumstances, like air travel, the physical ticket is being replaced by e-tickets, which are simply transaction identification code sans media (see Dodge and Kitchin, 2004). Importantly, the assignment of unique identification codes for the many transactions which are short-lived processes and leave no physical trace of their happening (e.g. a phone call) provides a lasting record that is used for billing, statistical monitoring or in the event of future queries. (The ‘reference number handshake’ often involves the quoting of identification codes of transactions gone awry.) The identification code is usually required to link a transaction to particular person and/or place and time. As noted, the recording and long-term storage of the identification codes of ephemeral electronic transactions which involve people creates TDI, a vital component of individual’s capta shadow (see Table 1 and Figure 3 for examples of a number of transactional identification codes).

To a large extent money has become an electronic transaction, with vast flows at the global scale in terms of trading on financial markets, and more locally through retail banking. The growth in business-to-business transactions on EDI networks, flows on EFT (electronic funds transfer) networks and more recently online via the Internet have spurred on the development of many important forms of transaction identification. EFT networks now interconnect millions of ATMs and point-of-sales terminals at retailers. It is impossible to reliably quantify the range of identifications codes, and total transaction volumes, used in these movements of money across the world but it is in the order of billions a day. For example, in the U.S. in 2002 there were an estimated 19,028 million PIN-based payment transactions (13,968m at ATMs and 5,060m at point-of-sales terminals), up from 7,495 million transactions in 1992 (ATM&Debit News, 2002, page 2). While in Britain there were 4.7 billion payment transactions made with debit and credit cards in 2002, amounting to some £211 billion, an average of £86 per adult per week (APACS, 2003).

Information transactions over telecommunications networks also generate huge daily volumes of capta, as each message is uniquely coded to insure successful transmission. Much of this identification coding is automated and unseen by sender or receiver; for example, all Internet email is given a unique identification upon despatch which is contained in the header information (this is not visible by default in most email reading software). A typical example, is

`5.1.0.14.0.20040202131004.036c2378@Mail.may.ie`

taken from an email between the authors, created by the Maynooth mail server. To further illustrate the range of transaction identification codes in telecommunications Table 2 provides an audit of codes used in a typical online interaction. To enact this transaction at least 15 different unique identification codes are used, some of which are person identifiers and others which are assigned automatically by software in various parts of the assemblage for the length of the transaction. As a consequence accessing a bank account online generates much more capta than a face-to-face transaction in a branch.

Territorial codes

Spatial identification codes are systems that locate people, places and objects across the terrain of the globe. Such spatial identification systems, in a UK context, are detailed in Figure 4 and include latitude and longitude, grid references, and administrative units such as postcodes. These systems vary in their scale and resolution (granularity) but they all standardise continuous, ongoing spatial relations into fixed, knowable territorial units by generating unique locational capta. Particularly apposite historical examples of this spatial standardisation are the large-scale topographic surveys of the Enlightenment and Victorian eras, especially in colonial contexts (e.g. India survey (Edney, 1997), U.S. (Johnson, 1976)), focussed on capturing *all* geographic knowledge deemed important to the government. As such, territorial identification codes have long been used to manage both populations and businesses, with governments in particular creating systems with regards to collecting capta, identifying and classifying land-use, and providing services (e.g. registration of land/property ownership, collection of taxes, the election of officials, delivery of mail). These systems have been adopted by other organisations as a means to deliver services.

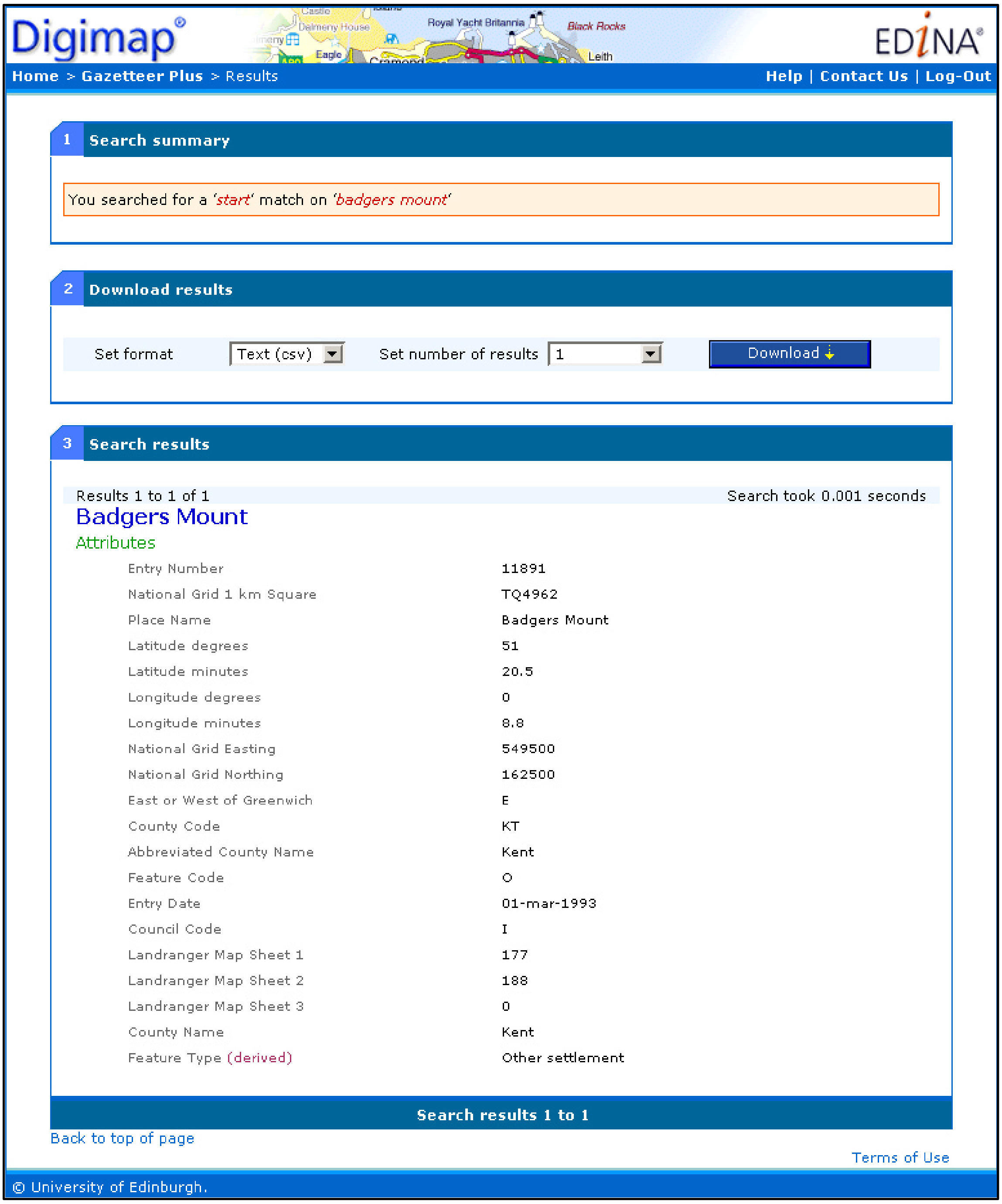


Figure 4. A range of territorial identifications codes generated for places in the UK from the Ordnance Survey's 1:50,000 gazetteer. (Source: Digimap Gazetteer Plus, part of the EDINA digital library service, University of Edinburgh, http://www.edina.ac.uk.)

While traditionally the power of territorial identification codes has been that they tie people and objects to particular addressable units, allowing them to consistently located, tracked and managed, more recently that power has been magnified by four developments that we examine in more detail in the following section. First, more and more capta is being tied to spatial referents, with this being greatly enhanced by the use of relational databases that share common fields and can be matched to other databases. Second, the granularity of much spatial capta has been increased, with data being increasingly tied to finer spatial resolutions (e.g. individual households and land parcels as opposed to neighbourhoods or routine use of sub-metre GPS co-ordinates). Third, static snapshot capta trails of TGI (e.g. the location of credit card purchases) are gradually being replaced by dynamic data trails gathered by technologies that actively and continuously track spatial location using GPS and radio location. Fourth, there is a marked increase in the ability to spatially analyse, model and simulate territorial capta, as evidenced in the adoption of GIS in many organisation.

5. Towards the fully machine-readable world

At present, personal, object, information, transaction and territorial codes provide a series of olgopticons, that is segments of much larger capta shadows. By piecing together segments more of a person's or organisation's capta shadow becomes revealed. The growing pervasiveness of identification codes and informational systems that match, trace, monitor and regulate populations, however, seeks to transform the various olgopticans into one universal panopticon: a dense rhizomic assemblage that will enable its users to know simultaneously and in real-time the what, when and where of people and things; to make all of Solzhenitsyn's thread visible *and* usable. The undeclared logic of the machine-readable world is 'all data, all the time, on all people, at all places', where risk is eliminated through perfect knowledge. Here, we use the phrase rhizomic assemblage rather than say grid because we want to acknowledge that that the capta fields are complexly interlinked and are not uniform, varying as a function of granularity, scale and context.

The nature of this spatialised rhizomic assemblage is, we contend, in a process of transformation due to four recent trends: (a) the drive towards generating capta wherever and whenever possible, which is being facilitated by the transference from ‘dumb’ to ‘smart’ technologies, new dynamic means of monitoring, and the mobile tracking of entities; (b) the drive towards far finer capta granularity and in particular the development of unique identifiers that are infallible; (c) the development of cheap and efficient means of storing digital capta shadows, meaning some are likely to be held forever; (d) enhanced means of processing and analysing capta including the increased convergence of separate capta fields and more sophisticated and precise methods of data mining. These trends, each of which we discuss below, coalesce as a meta-trend, that of the creation of a machine-readable world.

(a) Voracious capta generation

There is presently a trend towards massively increasing capta generation which seems to be driven partially by a ‘collect it if you can mentality’ and partially by a believe that increased volumes of key capta will provide enhanced marketing, competitive advantage, safety, security, and so on. These discourses have been accompanied by the process of ‘control creep’, “whereby the social control apparatus progressively expands and penetrates (or ‘creeps’) into different social arenas, in response to a set of inchoate fears about a sense of security in late-modernity” (Innes, 2001, no pagination). ‘Control creep’ can perhaps be best seen in action surrounding intensification of air travel security with rise of aggressive passenger profiling and biometric authentication (Lyon, 2003b). As we explore elsewhere, this clearly has implications for issues of privacy and confidentiality, hence the proliferation of data protection and freedom of information laws. The transference from ‘dumb-to-smart’ technologies, dynamic and mobile tracking are all aspects of the desire to generate as much capta as possible, taking advantage of the development of new technologies.

‘Dumb-to-smart’

> ‘The telescreen received and transmitted simultaneously … There was of course no way of knowing whether you were being watched at any given moment.’
> (George Orwell, *Nineteen Eighty-Four*, 1987, page 4)

Many of the material objects and environments used in everyday life are shifting from ‘dumb’ to ‘smart’ and in the process becoming capable of creating significant new streams of personally identifiable capta. In this context, ‘smart’ means programmed *awareness* of use, rather than intelligence. So household objects (e.g. televisions, VCRs, cameras), cars and spaces are, through the application of sensors and software, being made aware of how they are being used (e.g. time, location) and, crucially, which people use them. This awareness of usage is easily captured as logs and transmittable to third parties (especially through wireless technologies such as Bluetooth). In the last decade or so, there has been burgeoning interest in so called ubiquitous computing, reactive environments and 'smart' objects, but it is only relatively recently that widespread deployment has been occurring.

Importantly, in the shift from 'dumb' to 'smart' the overt nature of these objects and environments does not change – they still look the same, operate in much the same way and give the same results (although often with enhanced features). The capta enabling software is invisible to users. The 'smartening' of the things is widely trumpeted in terms of increased consumer convenience, but more insidiously there is no doubt that they are opening new conduits for surveillance of individual activity patterns. Home entertainment appliances, particularly televisions and VCRs provide good exemplars of 'smart' appliances. The newest interactive televisions and digital VCRs look similar to their 'dumb' forebears, but the addition of software and network feedback means they are able to 'watch' the viewers watching them. By 'watching' we mean they are aware of the behaviour of users and can transmit this back to the service provider[7]. The mundane nature of television, easy to use and freely available in almost every household, belies the degree to which viewing preferences reveal inner thoughts and the value to which media corporations place on obtaining a direct feed on this kind of capta. 'Smart' entertainment appliances are also likely to become cornerstones, along with personalised identification codes (digital certificates) on media objects, in new regimes of so-called 'digital rights management' (DRM). In the DRM future envisaged by large media corporations, interlocking of 'hard' identification systems on both the entertainment content and playback equipment, creating a secure 'content control system', is seen as vital to continued growth in revenue. A potential scenario of future is that a customer will buy a license to watch a given film, playable only on their television for a fixed length of time (and will quite likely include targeted advertising that can not be skipped), all made possible by specific codes for identification and authorisation.

The car is another prime example of an environment that is rapidly shifting from 'dumb' to 'smart'. As Thrift (2002, page 10) notes, "[a]lmost every element of the modern automobile is either shadowed by software or software has become …. the pivotal component." While the car still appears the same and drives in the same fashion, its systems are increasingly aware of their performance through on-board diagnostics, its location through GPS-tracking, and the actions of the driver. The result is cars are now rich capta generators, logged in event data recorders hidden from view (analogous to the 'black box' flight data recorders mandatory in all aircraft for many years) but recallable. The implications of 'smart' cars are only just being felt, in terms of capta being used in accident court cases and 'spy' scare stories in the press (e.g. Austen, 2003); as Elliot (2004) reports the GPS navigation systems such as Hertz's NeverLost were initially intended to help motorists find their way, however, more recently they have been used to catch renters who drive out of state or break speed laws. Longer term, the implications of 'smart' cars are likely to be felt much more deeply on driving behaviour, as the combination of telematic streets and in-car capta on drivers will enable new regimes of variable road pricing and 'flexible' insurance rates calculated on a journey-by-journey basis.

Always on

Conventionally, the capta shadow consists of a limited snapshot revealing person or object actions at particular known times. In most cases capta generation is usually triggered by a transaction of some kind (e.g. purchase by a credit card, entry into a country by passport). Increasingly, however, the ability to *continuously* generate capta about people and objects in real time is being realised. For example, the performance of a worker in a supermarket checkout is continuously recorded in order to monitor worker performance. Similarly, a browser, 'web bugs' and other software such transaction identification codes set on a computer by websites, known as cookies, continuously records Internet usage in the form of 'clickstream' surveillance. As Bennett (2001, page 206) notes, "[t]he metaphor of the 'data trail' is still valid, but that trail is extended and deepened by "mouse-dropping" left by unwary browser of the world wide web." Turning off 'cookies' (and their surveillance) serious degrades the functionality of a large part of the Web. In these cases, for the entire period that a worker remains on the till or a person is online capta generation is always on; individuals or objects continuously remain 'in view' of the information systems. Here, the technology of unique identification that enables both parties to perform their task is also the means of surveillance. In effect this is form of 'casual surveillance' that largely goes unnoticed and unchallenged because it seems to be inherently part of the technology.

Related to issues of privacy is the level of capta generation with the architectures of information systems designed to process and monitor personal transactions often set to gather excessive amounts personal data (Lessig, 1999), with default settings rigged so that people have to consciously choose to 'opt-out'. Moreover, systems are promoted in ways that encourage people to adopt, with penalties such as degrading of service, or additional costs (in time or money), or service denial. This promotion becomes more compulsory in nature where the controlling organisation has significantly more power in the transaction than the individual, for example in a monopoly situation like welfare payments (see Gilliom, 2001).

Mobile

In conventional databases only a rudimentary capta trail can be constructed for a person or object by mapping the time and date of each transaction. However, the movements in-between the point of direct identifiable transactions remains unknown. These spatial gaps are though being filled in through the employment of embedded, cheap and accurate, always on GPS and radio tracking systems. In the case of wireless ICT, particularly mobile phones, the location of all new devices can be continuously monitored by network companies, even if the device itself is switched off from the users' point of view. In effect, the network operating companies know to within a few metres everywhere a person's phone has been, and by implication the person. In other words, a complete spatial capta trail is revealed. Such continuous spatial tracking will become commonplace in next few years, encouraged in large part by the development of novel location-based services (LBS). In essence, LBS is about providing individualised information based on people's geographic location and is highly likely that such data will be automatically logged and become part of the capta shadow. Clearly, being able to spatially track people and objects in real-time has significant implications with regards individual privacy as the spatial gaps in the olgoptic 'capta shadow' disappear. Consequently, Dobson and Fisher (2003) have argued this trend may lead to geoslavery for some.

(b) Increased granularity

Many of the discourses surrounding identification codes interlock in contemporary Western society in technological drives to eliminate the fallibilities of many existing analogue forms of identification. Here, we consider two illustrative examples of this in terms of biometrics and smart labels.

Personal identification, in the passport for example, has traditionally consisted of printed matter, whether that be text or photographs. The problem with some means of documentary identification is that they are open to forgery or doctoring. With technical advancements, these forms of fallible identification are in the process of being replaced by unique biometric forms that are extremely difficult to fake. They promise a technical fix to what are often more complex social issues (e.g. the pressure of underage drinking driving the market for fake id). Biometric identification systems seek to render people machine-readable and by "fusing together flesh and machine" (Davies, 1994) they enforce the infallible linkages of individuals to their records in information systems[8]. At present, the most well known biometrics utilise distinctive physiological patterns on fingers-tips and the eyeball, while the 'ultimate' biometric is 'hidden' internally in the patterning of DNA. Here, the fingers, eyes and genes are the physical media carrying unique biological codes (e.g. the loops, arches and whorls of skin on the fingertip) that is scanned in some form to produce digitised capta suitable for storage and processing by software. The computer does not store the finger print per se, but a unique numerical code representation of it. This code is a digital biometric which can then be automatically verified against other (documentary) identification, usually for purposes of authentication (e.g. to enter a building or logon to a computer). In some instances several different types of biometric are taken to increase the effectiveness of identification.

These biometric forms of identification seek to fulfil Clarke's (1994b) list of desirable characteristics for effective human identification codes:

(1) universality of coverage (everyone should have one);

(2) uniqueness (each person should have only one, and no two people should have the same one);

(3) permanence (should not change, nor be changeable);

(4) indispensability (a natural characteristic that can not be removed);

(5) collectibility;

(6) storability;

(7) exclusivity;

(8) precision;

(9) simplicity;

(10) cost and convenience;

(11) acceptability.

Until now, biometrics perform well on criteria one to four, because they are naturally available, globally unique, cannot be lost or easily faked. However, biometrics have not be widely deployed for routine identification thus far because they have traditionally performed poorly in terms of criteria five through eleven. For example, ink finger printing has been costly, inconvenient, hard to store and process, and deemed socially unacceptable because of its stigmatisation through criminal associations. Consequently, biometric use could only be justified in special circumstances, such as controlled access to high-security facilities. In the last five years or so, technological changes in terms of non-invasive, cost-effective capture of biometrics (e.g. hand geometry scanning) and ease of digitally storing and processing have made it much easier to deploy; consequently use has been growing, for example in workplace time and attendance systems to prevent 'buddy-punching' fraud. The dramatic upsurge in securitisation post-9/11 has focused extensively on the 'need' (read desire) to deploy biometrics, with many states moving to include biometrics on passports and national id cards due to the concern that visual face-to-passport verification by immigration officials at borders is insufficient. The International Civil Aviation Organisation, the UN agency which co-ordinates passport standards world-wide, nominated facial recognition as the primary biometric for travel documents with iris pattern and fingerprint scans as secondary but not mandatory in May 2003[9]. In January 2004 U.S. immigration began mass finger scanning of all visitors to the America with visas. In the UK, a facial image biometric is scheduled to be introduced on all new passports from mid-2005.

A second example of the trend of finer granularity is 'smart labels'. As noted earlier, barcodes are limited because of their level of granularity. All objects of the same product possess the same barcode, meaning that they cannot be discriminated. As a result, in recent years, considerable effort has been placed on developing 'smart labels' with increased granularity which enable individual object recognition. This increased granularity is on the one hand to increase the ability to track individual items (to reduce and track theft), and on the other to provide greater personal surveillance. One obvious extension are 2D barcodes which encode data both horizontally and vertically in a speckled of dots, so are able to store unique capta (an example is the encrypted PDF417 barcodes used on majority of U.S. drivers' licenses). The U.S. Postal Service plan for I-Mail (intelligent mail) and Royal Mail's introduction of 'SmartStamps', both use 2D barcodes as a form of personalised stamps (Figure 5) that identifies the sender in terms of the person and their location (plus other information such as billing details). This it is hoped will improve the efficiency of tracking mail through the sorting and delivery system, while simultaneously improving surveillance of the sendee[10].

At their simplest, 'smart labels' such as 2D barcodes are really extensions of existing systems of object identification except that they hold more data, but more sophisticated technologies such as RFID (Radio Frequency identification) tags represent and communicate the code quite differently. The goal is to provide objects some level of 'awareness' and an ability to communicate over short distance using wireless networking technologies (Want et al, 1999).

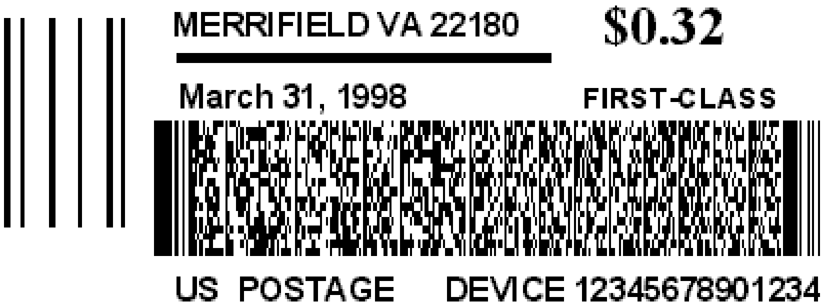


Figure 5. An example of a high capacity 2D barcode, known as PDF417. It is used here in a design for the U.S. Postal Service's proposed new personalised, secure postage stamp. (Source: http://www.symbol.com.)

RFID tags aim to provide a means to ‘identify any object anywhere automatically’. They are small chips composed of a simple digital circuit and an antenna that are embedded with a unique identification code and can be read at a distance by a radio transponder, rather than having to be passed in line of sight a laser scanner. They are presently most widely used in vehicle dashboard tags for automatic toll payment (the main system in the U.S. being EZPASS) and in livestock to facilitate “ ‘farm-to-fork’ traceability” with growing concerns over food safety (Wired News, 2003). Their main application though is likely to be in retail where they are seen as a major advance in inventory management and the fight against shoplifting (Ferguson, 2002).

In terms of the identification codes stored by the RFID chip the leading standard is known as the Electronic Product Code (EPC) developed by the Auto-ID Center, an industry-sponsored R&D lab at MIT, and now being actively commercially implemented by EPCglobal Inc. (a joint venture of the Uniform Code Council and EAN International, the main players in UPC barcode management). The definition of the EPCs creates a truly huge number space, enough to uniquely identify every object on the planet (Figure 6).

RFIDs chips and their EPCs are part of a much larger global information network which provides the means to automatically ‘look-up’ details on any tagged object from any location (with an Internet connection). Borrowing the domain name schema used on the Internet, the EPC network will use a distributed Object Naming Service (ONS) to link each EPC number to appropriate naming authority database that provides the detailed information[11]. Importantly, the querying of the ONS as RFID tagged products move through supply chains will automatically create a richly-detailed audit trail of capta, identifying everyone who handles the goods. Further, this capta trail will be geocoded as the location of fixed RFID readers will be known. The result will be much greater degree of routine ‘machine-to-machine’ generated knowledge on the positioning of many millions of physical objects in time and space. Kevin Ashton, a Proctor & Gamble executive who headed the Auto-ID Center, conceptualised the outcome in the memorable phrase, “…an internet for things, a standardized way for computers to understand the real world” (Schoenberger, 2002, no pagination). Aggregated patterns of these RFID-generated capta trails will give an unprecedentedly detailed view of the material flows of products across space, a newly revealed micro economic geography of the machine-readable world.

Evangelists for RFID tagging and the EPC network envisage a wide range of innovations in the handling of physical objects arising from these networked capta trails (see Ferguson, 2002, for typical speculation). Inside stores the ‘holy grail’ is so-called ‘smart-shelving’ that is aware of its own stock-levels and can reorder as required, while in the home pundits are talking of microwave ovens checking the best cooking settings for ready-meals, washing machines choosing the most appropriate cycle for clothing and medicine cabinets able to ‘spot’ out of date or recalled pharmaceuticals. There could also be potential for tracking goods at the end of the life-cycle, alerting bin-men to items containing toxic substances for example. There are also many privacy concerns, but discussion is beyond the scope of this paper.

Figure 6. The structure of a 96 bit EPC (Electronic Product Code), deliberately partitioned into 4 distinct parts to allow flexibility in object identification using RFID tags. The header defines the type of EPC. The 28 bit manager element allows for 268,435,455 different users (typically product manufacturers), who are able to uniquely define 16,777,215 different types of object. Each of these is then assigned a 36 bit serial number, giving a total of 68,719,476,735 potential unique identification numbers per object. In total the EPC-96 number space can uniquely identify 3^{26} (309484990217176000000000000) different entities.

This increased granularity for identifying material objects is being mirrored with respect to territorial units. The drive towards totalising spatial knowledge at finer scales is clearly evident in the development of many national digital geospatial infrastructures across different countries in the last decade (Rhind, 1997). The development in the United Kingdom of the National Land & Property Gazetteer (NLPG) and the Ordnance Survey’s (OS) Digital National Framework, known commercially as MasterMap, have given rise to new systems of highly detailed territorial identification coding. The NLPG is a digital cadastre which assigns 8 digit 'unique property reference number' (UPRNs) to all property entities[12] and the OS MasterMap goes further to assign 16 digit numeric 'topographic identifiers' (TOIDs) to all significant mapped natural and manmade features (as defined by the formal rules and everyday practices of the OS surveyors who produce the underlying spatial capta). The promotional rhetoric for these two examples of intensified territorial identification coding speaks to the increased versatility, functionality and efficiency, becoming 'common hooks held on any number of databases within and between different organisations' (Triglav, 2002, page 12), while an OS marketing piece 'Join-up geography' notes, '[l]ike barcodes on products, TOIDs given to landscape features provide more intelligence for the relevant user' (PCSA, 2002, page 182). As of summer 2003 there were around 448 million TOIDs assigned (Hodgson, pers. comm. 2003), a raw metric of the extent of the standardised British landscape capta, the available record of geographic reality that can be uniquely addressed and 'worked' directly by software.

(c) Forever storage

The capability to log, process and *permanently* store capta shadows has become feasible for most businesses and organisations as the real cost of computer storage has plummeted in the last decade. Developments in hard disks technologies, in particular, have relentlessly pushed up the capacity and thus driven down the cost per megabyte of data storage, making it considerably cheaper than paper or film (Grochowski and Halem, 2003). Gilheany (2000) has estimated that since the introduction of the first commercial magnetic disk storage by IBM in 1956, the cost per gigabyte has fallen by a factor of a million to one. The growth in storage density, as measured in bits per inch on magnetic disks, has even outpaced the curve of Moore's Law and shows little signs of slowing down in the near future. The actual physical space required for data storage has also shrunk dramatically, as hard drives have got smaller and the density of packing increased. This growth in storage capabilities enables radically different strategies of information management with the deleting old files virtually unnecessary. The technical and cost barriers to storing large parts of people's capta shadows over an entire life course have thus been erased. As a result, transactions undertaken today (e.g. post a message to an email list, paying in a store with a credit card, speaking to a friend on the phone) may well be logged and kept beyond the death of the person, with the potential to be recalled and analysed at any point in the future[13]. In addition, many more dimensions of daily life are being logged and stored simply because the cost of doing so is marginal. Such long term retention is driven on the governmental side for security and surveillance purposes and on the business side for its present commercial value, or likely future potential value. The machine-readable world is likely to be a place that does not forget.

(d) Enhanced capta processing and analysis

The preceding trends primarily focused on changes in the *production* of key capta. In this section we detail two related trends that alter capta processing and use.

Convergence

The enormous growth in relational databases means that databases with shared fields can be easily matched and compared[14]. At present, there is a drive towards marrying relational databases as the convergence allows more of an individual's or object's capta shadow to be revealed. As such, this convergence means that previously unknown relationships between capta become visible. This is particularly the case given the vast amounts of capta generated for each individual, entity and transaction. Moreover, this convergence facilitates the widening and deepening of geographical coverage as more and more capta is tied to spatial referents, and territorial capta itself becoming more widespread with increased granularity. A good example of the exploitation of convergence are the various data resellers such as US Search and Equifax that allow consumers to 'pull together' vast amounts of capta from related databases to profile individuals (including social security, health, education, court, criminal, property records; and to screen contractors, nannies, neighbours and so on), or the Galileo system used in air and holiday bookings that similarly links several databases to construct a security profile.

Precise, predictive profiling

Huge volumes of capta are presently accumulating, particularly from TGI, in information systems that are then being 'data mined' to exploit its latent value. Data mining or 'knowledge discovery from databases' (KDD) in the technical community, consists of a set of exploratory analytical techniques (see Table 3) premised on the "belief that information is hidden in very large databases in the form of interesting patterns. These are non-random properties and relationships that are valid, novel, useful and ultimately understandable" (Miller and Han, 2001, page 4). KDD is at heart a modus operandi for differentiating, it a modern means of discrimination by software algorithm in the machine-readable world. In terms of commercial and public services delivery these rankings, in the name of efficiency, will be used as a filter to channel particular goods, services, information, opportunities and life chances to some people and not others. Business see this as a way to focus onto the their 'high lifetime value customers' and exclude the other segments of customers (these might not even be unprofitable, simply not profitable enough) (Danna and Gandy, 2002).

The goal of KDD is a better under people's behaviours through the analysis of what they do, resulting in evermore precise profiles. As such, KDD analysis is used by commercial organisations, particularly retail and fraud detection in financial services, and increasing in the realm of policing. The use of predictive profiling particularly affects life chances of millions in terms of credit-scoring (Leyshon and Thrift, 1999). This has led Miller and Han (2001, page 5) to state that, "KDD is to information space as microscopes, remote sensing and telescopes are to atomic, geographic and astronomical spaces respectively. KDD is a tool for exploring domains that are too difficult to perceive with unaided human abilities."

Closely allied to KDD is geodemographic profiling and geoprocessing founded on GIS. The GIS industry that is largely centred on the aggressive acquisition and exploitation of territorial capta, creating systems able to enact evermore realistic models of 'reality', particularly the profiling of people. The combination of GIS and aligned technologies and much greater amounts of social capta on large numbers of people (in some cases whole populations) is proving irresistible for risk management in two dominate discourses in contemporary (Western) societies – consumption and crime. The panoply of agencies tasked with 'keeping us safe', from police, immigration, military to corporate security firms and consultancies, are making use of crime mapping and geographic profiling with the goal of identifying 'suspect' individuals and interdicting them based on their socio-spatial behaviour patterns (Crampton, 2003). This is most readily apparent at airports (Lyon, 2003b). In the arena of consumption, the last two decades have seen the development of a geodemographics industry which draws in evermore detailed social and territorial capta to produce segmentations of space in terms of ‘purchasing power’ and fetishised lifestyles. While often portrayed in a positive light, using the simple premise ‘that birds of a feather flock together’ and the marketing mantra of ‘getting inside the consumer’s head’, there are negative consequences to geodemographics (Curry, 1997; Goss, 1995). For example, the ability to determine the kinds of people who live in particular places, and their propensity to consume, implies an inherent ranking based on ‘worth’ and has led to discriminatory practices, such as the ‘redlining’ of communities deemed unprofitable or high risk by insurers and banks.

Conclusion

In this paper we have provided an analytic framework for thinking through the role of identification codes in Western society, an easily overlooked but vital part of the “technological unconscious” (Thrift, 2004). Utilising a Foucaultian approach, we have illustrated the salience of identification codes in the governance of society and the managing of organisations and businesses through a variety of examples. Identification codes are vital elements in so many small, seemingly benign rituals of social control, operating at the intersections of power, knowledge and the person. Our thesis is that the application and processing of *digital* identification codes are key to the evolving forms of contemporary governmentality and capitalism. They provide a means of representing, sorting, collating, matching, profiling, regulating; of generating information, knowledge and control through processes of abstraction, computation, modelling and classification. Identification codes take on special significance because they are a key form of capta providing the vital links between separate strands of capta. Being able to monitor and process identification codes thus provides a certain power to owners and users by enabling access and analysis of valuable capta and then to act on this knowledge (to vet, to segment, to authorise, to regulate, and so on). As our simple typology demonstrated, identification codes now provide a means of addressing all the entities and processes that make up everyday life – people, material objects, information, transactions and territorial units. What’s more they provide a means of linking these entities and processes together in complex ways to form a dense rhizomic assemblage, creating information and knowledge that not so long ago was the preserve of science fiction. At present, however, the capta identification codes provide access to are, at best, olgopticons; that is they afford only partial and selective views.

The four trends we outlined – voracious capta generation, increased granularity, forever storage, and enhanced capta processing and analysis - seek to convert these partial olgopticons into more panoptic arrangements; that is to offer a more detailed and yet wider ranging and continuously updating view composed of several, interlinked capta shadows, that can be data mined using more sophisticated techniques. In turn, we contend, that these trends are part of a larger meta-trend – the creation of a machine-readable world in which identification codes can be 'read' and acted on by software independent of human control. They work to eliminate the messy, subjective realities of everyday life, ordering and objectifying people, striving to make them *consistently* addressable by software. Again drawing on our Foucaultian framework, this meta-trend is supported and driven by interlocking discourses such as safety, security, efficiency, anti-fraud, citizenship/empowerment, productivity, reliability, flexibility, economic rationality, and competitive advantage to construct powerful discursive regimes. Contra Foucault, these discourses and regimes work not only by disciplining people to their logic, but because they are seductive. Following Althusser (1971) they induce a process of interpellation wherein people willingly and voluntarily subscribe to and desire their logic.

Clearly the trends we have detailed, in combination with the discursive regimes which provide a driving logic to their implementation, have significant implications with regard to issues of privacy, confidentiality, community equity and the organisation and functioning of society. However, in spite of the intensity and sophistication of contemporary identification coding systems, they remain open to resistance and subversion, as the lucrative market in fake ids for underage drinkers and the growth in identity-theft crimes clearly testifies. The extent, individual motivations and implications of resistance to machine-readability is debatable. A full analysis of these consequences and implications is, however, beyond the scope of this paper and we address them more fully elsewhere. Suffice to say, given that identification codes are becoming more prevalent, embedded in everyday life, refined in nature, spatially referenced, interlinked and open to sophisticated analysis, their importance as means of governing populations and managing organisations is becoming more pervasive. As such, these codes of life and their supporting discursive regimes are worthy of further analysis.

References

Alexander P M, 2002 *Towards Reconstructing Meaning When Text is Communicated Electronically* Unpublished PhD, University of Pretoria, http://upetd.up.ac.za/thesis/available/etd-08192002-155431/unrestricted/00front.pdf

Alonso W, Starr P, 1987 *The Politics of Numbers* (Russell Sage Foundation, New York)

Althusser L, 1971 *Lenin and Philosophy and Other Essays*. Trans. Ben Brewster. (NLB, London)

Amin A, Thrift N, 2002 *Cities* (Polity Press, Cambridge)

APACS, 2003 *UK Payment Markets: Trends and Forecasts In Brief 2003* (Association for Payment Clearing Services, London), http://www.apacs.org.uk

ATM&Debit News, 2002 *EFT Data Book: The Complete Guide to the ATM and POS Debit Markets*, 2003 edition (Thomson Corporation and ATM&Debit News, New York)

Austen I, 2003, "Your brake pads may have something to say (by email)" *New York Times*, March 27, http://www.nytimes.com/2003/03/27/technology/circuits/27next.html

Barnes T J, Hannah M, 2001, "The place of numbers: Histories, geographies, and theories of quantification" *Environment and Planning D: Society and Space* **19** 379-383

Becker H, 1952, "Science, culture, and society" *Philosophy of Science* **19**(4) 273-287

Beniger J R, 1986 *The Control Revolution: Technological and Economic Origins of the Information Society* (Harvard University Press, Cambridge, MA)

Bennett C J, 2001, "Cookies, web bugs, webcams and cue cats: Patterns of surveillance on the world wide web" *Ethics and Information Technology* **3** 197-210

Bogard W, 1996 *The Simulation of Surveillance: Hypercontrol in Telematic Societies* (MIT Press, Cambridge, MA)

Bowker G, Star S L, 1999 *Sorting Things Out: Classification and its Consequences* (MIT Press, Cambridge MA)

Brown S A, 1997 *Revolution at the Checkout Counter: The Explosion of the Bar Code* (Harvard University Press, Cambridge, MA)

Castells M, 1996 *Rise of the Network Society* (Blackwell, Oxford)

Checkland P, Holwell S, 1998 *Information, Systems and Information Systems: Making Sense of the Field* (John Wiley & Sons, Chichester, England)

Clarke R, 1994a, "The digital persona and its application to data surveillance" *The Information Society* **10(2)**, http://www.anu.edu.au/people/Roger.Clarke/DV/DigPersona.html

Clarke R, 1994b, "Human identification in information systems: Management challenges and public policy issues" *Information Technology and People* **7(4)** 6-37, http://www.anu.edu.au/people/Roger.Clarke/DV/HumanID.html

Crampton J, 2003, "Cartographic rationality and the politics of geosurveillance and security" *Cartography and Geographic Information Science* **30(2)** 135-148

Curry M R, 1997, "The digital individual and the private realm" *Annals of the Association of American Geographers* **87** 681-699

Danna A, Gandy O H, 2002, "All that glitters is not gold: Digging beneath the surface of data mining" *Journal of Business Ethics* **40** 373-386

Davies S G, 1994, "Touching big brother: how biometrics technology will fuse flesh and machine" *Information Technology and People* **7(4)**, http://www.privacyinternational.org/reports/biometric.html

Deleuze G 1992 "Postscript on the societies of control" *October* **59** 3-7

Desrosières A, 1998 *The Politics of Large Numbers: A History of Statistical Reasoning* trans. Naish C. (Harvard University Press, Cambridge, MA)

Dicken, P, 2003 *The Global Shift* 4th edition (Sage, London)

Dobson J E, Fisher P F, 2003, "Geoslavery" *IEEE Technology and Society Magazine* spring 47-52

Dodge M, Kitchin R, 2004, "Flying through code/space: The real virtuality of air travel" *Environment and Planning A* **36(2)** 195-211

Dodge M, Kitchin R, forthcoming, "Code and the transduction of space", *Annals of the Association of American Geographers*

Dunlop J T, Rivkin J W, 1997 "Introduction" in *Revolution at the Checkout Counter: The Explosion of the Bar Code* Ed Brown S A (Harvard University Press, Cambridge, MA) pp 1-38

Edney M H, 1997 *Mapping an Empire: The Geographical Construction of British India, 1765-1843* (University of Chicago Press, Chicago)

Elliot C, 2004, "Some rental cars are keeping tabs on the drivers" *New York Times*, January 13, http://www.nytimes.com/2004/01/13/business/13gps.html

Ferguson T, 2002, "Have your objects call my object" *Harvard Business Review* June 1-7

Foucault M, 1976 *Discipline and Punish* (Allen Lane, London)

Foucault M, 1978 *The History of Sexuality, volume one* (Random House, New York)

Gandy O H, 1993 *The Panoptic Sort: A Political Economy of Personal Information* (Westview Press, Boulder, CO)

Gilheany S, 2000, "Projecting the cost of magnetic disk storage over the next 10 years" *White paper 22011v039* (Archivebuilders.com, Manhattan Beach, CA), www.archivebuilders.com/whitepapers/22011v039h.html

Gilliom J, 2001 *Overseers of the Poor: Surveillance Resistance and the Limits of Privacy* (University of Chicago Press, Chicago)

Goss J, 1995, " 'We know who you are and we know where you live': The instrumental rationality of geodemographics systems" *Economic Geography* **71** 171-198

Graham S, 2004, "The software-sorted city: Rethinking the 'digital divide' " in *The Cybercities Reader* Ed S Graham (Routlege, London) pp 324-331

Grochowski E, Halem R D, 2003, "Technological impact of magnetic hard disk drives on storage systems" *IBM Systems Journal* **42** 338-346

Haklay M, 2001 *Public Environmental Information Systems: Challenges and Perspectives* Unpublished PhD, University College London, http://www.casa.ucl.ac.uk/muki/thesis.htm

Hannah M, 1992 *Foucault Deinstitutionalised: Spatial Prerequisites for Modern Social Control* Unpublished PhD thesis, Pennsylvania State University

Hannah M, 1997a, "Space and the structuring of disciplinary power: an interpretative review" *Geografiska Annaler Series B, Human Geography* **79** 171-180

Hannah M, 1997b, "Imperfect panopticism: Envisioning the construction of normal lives" in *Space and Social Theory* Eds G Benko, U Strohmayer (Blackwell, Oxford) pp 344-360

Harmon C K, 1994 *Lines of Communication: Bar Code and Data Collection Technology for the 90s* (Helmers Publishing, Inc., Peterborough, NH)

Harvey D, 1982 *The Limits to Capital* (Blackwell, Oxford)

Hodgson R, 2003, "Information on TOIDs", personal communication, 28 August to Marketing Manager, OS MasterMap Strategy and Delivery

Hosoya H, Schaefer M, 2001, “Bit structures” in *Harvard Design School Guide to Shopping* Eds C J Chung, J Inaba, R Koolhaas, S T Leong (Taschen, Köln) pp 156-162

Innes M, 2001, “Control creep” *Sociological Research Online* **6(3)**, http://www.socresonline.org.uk/6/3/innes.html

Johnson H B, 1976 *Order Upon the Land: The U.S. Rectangular Land Survey and the Upper Mississippi Country* (Oxford University Press, New York)

Kahn D, 1967 *The Codebreakers: The Story of Secret Writing* (Macmillan, New York)

Lander E, 1992, “DNA fingerprinting: Science, law and the ultimate identifier” in *The Code of Codes: Scientific and Social Issues in the Human Genome Project* Eds D J Kevles, Hood L (Harvard University Press, Cambridge, MA.)

Larson E, 1992 *The Naked Consumer: How Our Private Lives Become Public Commodities* (Henry Holt and Company, New York)

Leibowitz E, 1999, “Bar codes: Reading between the lines” *Smithsonian* February, **29(11)** 130-146

Lessig L, 1999 *Code and Other Laws of Cyberspace* (Basic Books, New York)

Leyshon A, Thrift N, 1999, “Lists come alive: Electronic systems of knowledge and the rise of credit-scoring in retail banking” *Economy and Society* **28** 434-466

Lyon D, 2002, “Everyday surveillance: Personal data and social classifications” *Information, Communication & Society* **5** 242–257

Lyon D, 2003a, “Surveillance as social sorting: Computer codes and mobile bodies” in *Surveillance as Social Sorting: Privacy, Risk and Digital Discrimination* Ed D Lyon (Routledge, London) pp 13-30

Lyon D, 2003b, “Airports as data filters: Converging surveillance systems after September 11th” *Information, Communication and Ethics in Society* **1** 13-20

Mackenzie A, 2002 *Transductions: Bodies and Machines at Speed* (Continuum Press, London)

McNay L, 1994 *Foucault: A Critical Introduction* (Polity Press, Oxford)

Miller H J, Han J, 2001, “Geographic data mining and knowledge discovery: An overview” in *Geographic Data Mining and Knowledge Discovery* Eds H J Miller, J Han (Taylor & Francis, London) pp 3-32

Morton A Q, 1994, “Packaging history: The emergence of the uniform product code (UPC) in the United States, 1970-75” *History and Technology* **11** 101-111

Orwell G, 1987 [1949] *Nineteen Eighty-Four* (Penguin, London)

Palmer R C, 1995 *The Bar Code Book: Reading, Printing, Specification, and Application of Bar Codes and Other Machine Readable Symbols* 3rd edition (Helmers Publishing Inc. Peterborough, NH)

Paskin N, 2003, "Identification and metadata" in Eds E Becker, W Buhse, D Günnewig, N Rump *Digital Rights Management: Technological, Economic, Legal and Political Aspects* (Springer, Berlin)

Porter T M, 1995 *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton University Press, Princeton, NJ)

Rhind D, 1997 *Framework for the World* (John Wiley & Sons Inc)

Rose N, 1996 *Inventing Our Selves: Psychology, Power and Personhood* (Cambridge University Press, Cambridge)

Rule J B, McAdam D, Stearns L, Uglow D, 1983, "Documentary identification and mass surveillance in the United States" *Social Problems* **31(2)** 222-234

Savir D, Laurer G J, 1975, "The characteristics of and decodability of the universal product code" *IBM Systems Journal* **14** 16-35

Schoenberger C R, 2002, "The internet of things" *Forbes Magazine* March 18, http://www.forbes.com/technology/forbes/2002/0318/155.html

Solzhenitsyn A, 1968 *Cancer Ward*, part I, trans. Bethell N, Burg D (Bodley Head, London)

Stalder F, 2002, "Privacy is not the antidote to surveillance" *Surveillance & Society* **1(1)** 120-124, http://www.surveillance-and-society.org/articles1/opinion.pdf

Torpey J, 2000 *The Invention of the Passport: Surveillance, Citizenship, and the State* (Cambridge University Press, Cambridge)

Torpey J, Caplan J, 2001 *Documenting Individual Identity: The Development of State Practices in the Modern World* (Princeton University Press, Princeton, NJ)

Thrift N, 2002, "Driving in the city" mimeo, http://www.its.leeds.ac.uk/projects/mobilenetwork/downloads/driving_in_the_city.doc

Thrift N, 2004, "Remembering the technological unconscious by foregrounding knowledges of position" *Environment and Planning D: Society and Space* **22** 175-190

Thrift N, French S, 2002, "The automatic production of space" *Transactions of the Institute of British Geographers* **27** 309-335

Triglav J, 2002, “OS MasterMap, heralding a new era in GIS applications” *GeoInformatics Magazine* January/February 12-15

Want R, Fishkin K, Gujar A, Harrison B, 1999, “Bridging physical and virtual worlds with electronic tags” *Proceedings of CHI’99* 370-77, http://doi.acm.org/10.1145/302979.303111

Watson J, Crick F, 1953, “A structure for Deoxyribose Nucleic Acid” *Nature* 171 2 April, 753

Wired News, 2003, “Mad cow spurs livestock tracking” *Wired News* December 31 http://www.wired.com/news/medtech/0,1286,61770,00.html

Wise M N, 1995 *The Values of Precision* (Princeton University Press, Princeton, NJ)

[1] Although we acknowledge that identification code numbers are often the outcome and requirement of software systems or legal processes (e.g. license plates on vehicles required by law and now allocated and processed by centralised computer systems).

[2] Jensen (cited in Becker, 1952, page 278) states that “it is an unfortunate accident of history that the term datum … rather than captum … should have come to symbolize the unit-phenomenon in science. For science deals, not with ‘that which has been given’ by nature to the scientist, but with ‘that which has been taken’ or selected from nature by the scientist in accordance with his purpose …”.

[3] This idea of an individual having a ‘shadow’ inside information systems has been conceptualised by a number of scholars – ‘shadow order’ (Bogard, 1996), ‘digital persona’ (Clarke, 1994a) ‘dividuals’ (Deleuze, 1986), and ‘data doubles’ (Lyon, 2003a). We choose the term ‘capta shadow’ as the metaphor of the shadow usefully conveys the idea that it is something you can never be get rid of.

[4] This is the 17 digit Vehicle Identification Number (VIN), a global standard defined by ISO 3779:1983. (http://www.vehicleidentificationnumber.com/)

[5] The actual numbers written underneath barcodes on consumer products are for the convenience of humans and are not read by the scanner.

[6] The ISBN (International Standard Book Number) system began in 1966 and was adopted as international standard ISO 2108 in 1970 (http://www.isbn-international.org/). ISSN (International Standard Serial Number) was designated in 1974 and is managed by a world-wide network of 77 National Centres and an International Centre based in Paris (http://www.issn.org). ISWC, ISO 15707:2001, is managed by the International Confederation of Societies of Authors and Composers (http://www.iswc.org/).

[7] A telling example being the tracking of viewing habits during the 2004 Super Bowl by TiVo digital VCRs which revealed the popularity of Janet Jackson’s ‘wardrobe malfunction’. See for example, “TiVo: Jackson stunt most replayed moment ever”, *CNN*, February 3, 2004, http://www.cnn.com/2004/TECH/ptech/02/03/television.tivo.reut/index.html

[8] In fact biometrics are not as infallible as is as widely claimed, see for example Lander’s (1992) detailed analysis of the weaknesses in DNA evidence.

[9] International Civil Aviation Organisation's proposed standards for machine readable travel documents. (http://www.icao.int/mrtd/Home/Index.cfm)

[10] The I-Mail system is still in the consulting stage, but under the proposed system, mail that lacks a personalised 2D barcode would be subject to much more rigorous screening for hazardous contents.

[11] The contract to operate the ONS look-up element of the EPC Network was awarded to VeriSign in January 2004. VeriSign is the controversial U.S. corporation that operates the Internet DNS.

[12] The UPRN itself is just part of a much larger British Standard BS7666 (Spatial datasets for geographical referencing).

[13] Note, there are, of course, still major technical challenges in the intelligent summarisation and analysis of such huge and detailed terabyte databases.

[14] Indeed, the potential of matching records in disparate 'data banks' has been an enduring fear surrounding computerisation since the 1960s.

Table 1. 'Codes in your wallet'. An audit of identification codes and their media carried by one person. The use of codes and linkage to information systems are specified.

Physical object	Type	Codes	Media	Usage	Code reading	Information system linkages
Season ticket for train	Transaction	29279o5099s01; AMS3950	Paper card with magnetic strip on back; printed on the front	Travel	Magnetic reader in turnstile; visual inspection	Fare structure and travel zones for the railway; potential linkage to person via photocard number
British Rail photo card	Person	AMS3950	Paper card with printed code, name and photograph	Travel	Visual inspection	Unclear, may be a database of code assignments which link to full personal details
British Library reader card	Person	402257DOD	Plastic card with magnetic strip on back; visibly printed on the front. Plus name, photograph and expiration date on the front	Physical and Network Access	Visual inspection for access; code manually entered into library OPAC to search and order books	Library patron database with direct linkage to personal details and borrowing history
Scrap of yellow paper with PIN number	Transaction	3447	Handwritten on paper	Network Access	Authorise transactions at ATMs	Bank account database with full personal details and account transaction records
UCL library photocopy card	Transaction	23116	Plastic card with magnetic storage of value	Payment	Photocopier card reader; recharging machine	No linkage; No personally identifiable capta
British Library photocopy card	Transaction	92859	Plastic card with magnetic strip storing value	Payment	Photocopier card reader; recharging machine	No linkage; No personally identifiable capta
University of London paper photocopy card	Transaction	None visible; plus contact telephone number of card manufacturer	Paper card with magnetic strip storing value	Payment	Photocoper card reader. Disposable, one-time use	No linkage; No personally identifiable capta
Book 6 first class Royal Mail stamps.	Object	5014721112268; plus telephone, fax and url	Barcode in EAN13 format; printed on plastic paper	Retail	Barcode scanner at till; visual inspection	EAN manufacturers database; no personally identifiable data or linkage possible
Promotional business cards for *Atlas of Cyberspace* book.	Object, Territorial	ISBN 0201745755; plus publisher contact details (postal codes, telephone/fax, urls)	Paper card with codes visibly printed	Information	Help interested people to obtain the book. ISBN uniquely identifies the book	Publisher industry books in print databases; contact details linked to postal address systems, telecommunications databases
Business card for university department	Person, Territorial	telephone number (+44 (0) 20 7679 1255), fax (+44 (0) 20 7813 2843), email (m.dodge@ucl.ac.uk), url (www.casa.ucl.ac.uk/)	Paper card with codes visibly printed	Information	Help interested people to contact the person	Contact details linked to postal address systems, telecommunications databases.
Bank ATM/debit card	Person	Card number (4539 7486 2721 9238), sort code (20-55-10), account number (30168352); contact telephone number for lost/stolen cards	Plastic card with magnetic strip and gold security chip on front; codes printed on the front	Payment	Swiped in EFTPOS machine; ATMs; visual inspectation to make payment	Bank account database with full personal details and account transaction records; electronic payment networks such as VISA
Credit card	Person	card number (4533 9520 0731 7604), security code (7604 371)	Plastic card with magnetic strip and gold security chip on front; codes visibly printed on the front and back	Payment	Swiped in EFTPOS machine; ATMs; visual inspectation to make payment	Bank account database with full personal details and account transaction records; electronic payment networks such as VISA
Donor card	Person	telephone number	Paper card with codes visibly printed	Information	Visual inspection of contact number of next of kin in the event of death	Contact telephone number links into telecommunications databases
Scrap of yellow paper with old password for university central computer system.	Person	ulcc24!	Handwritten on paper	Network Access	Used in login procedure to network computer system.	Authenication database of computer system; direct linkage to personal details and records of prior usage
£5 note	Object, Transaction	HH10907701; possibly other hidden security codes	Printed on high security paper	Payment	Visual inspection; machine-readable by OCR by banks	Bank of England databases for money in circulation. No personal linkage possible
Bromley council library card	Person	A20182005784058A; telephone numbers for different branches	Plastic card with barcode; visibly printed	Physical and Network Access	Visual inspection; scanned to authorise book borrowing	Library patron database with direct linkage to personal details and borrowing history
Door access swipe card for university.	Person	3110; plus telephone number / fax of card manufacturer	Plastic card with magnetic strip; visibly printed on the front	Physical Access	Swipe card reader on doors	University card allocation database linked to person; provides log of card usage showing all times of entry/exits from buildings when card used
Staff photo ID card	Person	0282263462; name, photograph	Plastic card with code for library on reverse as a barcode and visibly printed	Physical and Network Access	Visual inspection; barcode scanner for access to library; barcode reader to borrow books	Library patron database with direct linkage to personal details and borrowing history
Scrap of yellow paper giving various computer and website passwords	Person	8 digit alphanumeric codes	Handwritten on paper	Network Access	Used in login procedures to computer, networks systems and websites.	Authenication database of computer system; direct linkage to personal details and records of prior usage

University of London Library borrower card	Person	01 9433594 4; postal code, telephone numbers, and url for library	Plastic card with barcode; visibly printed	Physical and Network Access	Visual inspection; barcode scanner for access to library; barcode reader to borrow books	Library patron database with direct linkage to personal details and borrowing history
Scrap of paper with friend's mobile phone numbers	Person	0781-7811111; 0793-9111111; 0785-5411111	Handwritten on paper	Communication	Make telephone calls (Note, does not own a mobile so has to write numbers down)	Contact telephone number links into telecommunications databases; potential linkage to person
London buses ticket (route 24)	Transaction	42715 34809 17272 (unclear if it is unique id)	Slip of paper, codes visibly printed	Travel	Visual inspection to proof right to travel with details of date, time and fare paid	Fare structure and routing of bus company; no linkage to personally identifiable data possible
Visa debit receipt for Malborough Arms, local pub	Transaction, Person	Card number (4539 7486 2721 9238); Merchant codes (M1001004249) (T17239874); Authorisation code (046256).	Slip of paper, codes visibly printed	Payment	Provide material record of electronic transaction, giving date, time and amount; visual inspection	Bank account database with full personal details and account transaction records; electronic payment networks such as VISA
Receipt from Boots, chemist shop (very faded)	Transaction	Not legible	Slip of paper, codes visibly printed	Payment	Provide material record of electronic transaction, giving date, time and amount; visual inspection	Bank account database with full personal details and account transaction records; electronic payment networks such as VISA
Visa debit receipt from Maplin Electronics. (Reproduced as Figure 3)	Transaction, Person	Card number (4539 7486 2721 9238); Merchant code (540436500392691); Authorisation code (0033773) till receipt number (R3/34878); plus VAT number, store postcode, telephone number	Slip of paper, codes visibly printed	Payment	Provide material record of electronic transaction, giving date, time and amount; visual inspection	Bank account database with full personal details and account transaction records; electronic payment networks such as VISA
ATM receipt	Transaction, Person	LC3CU51 (ATM location code) and 004687 (transaction reference code)	Slip of paper, codes visibly printed	Payment	Provide material record of electronic transaction, giving date, time, place and amount; visual inspection	Bank account database with full personal details and account transaction records; electronic payment networks such as VISA

** Full personal details meaning name, date of birth, sex, address and contact*

(Source: The wallet of one of the authors, 14th July 2003. Note, some code numbers have been randomly altered, but their format remain the same.)

Table 2. The identification codes involved in a typical online transaction. This example is the use an Internet banking service in the UK by a home user with dial-up access.

1. Login to computer. Person identification codes: username, password
2. Establish dial-up Internet connection via telephone provider. Person identification codes: telephone number, username, password
3. Caller line identification (CLI) number identifies the dialling location is automatically logged in telephone network system
4. Authenticated by ISP using RADIUS protocol. Transaction codes: unique session identifier, assigned IP address
5. Open web browser, enter domain name and load bank homepage. Transaction codes: assigned permanent and session cookies by the bank web server

```
email.barclays.co.uk FALSE / FALSE 1753985730 ssuserid v2.1060167458.639.62.7.132.29.595.www11
www.personal.barclays.co.uk FALSE / FALSE 2006273884 brcvisitor 1060167612227269483371107091
```

6. Click on hyperlink to secure Internet banking webpage. Transaction codes: another cookie set, 128 bit Secure Sockets Layer (SSL) certificate to encrypt communication
7. First part of the login/authentication process for Internet banking. Person identification: surname, 12 digit internet banking membership number
8. Second part of the process. Person identification: 5 digit passcode, 2 letters randomly chosen from your memorable word

Table 3. Key data-mining techniques. (Source: Miller and Han 2001, page 8.)

Data mining task	Description	Techniques
Segmsentation	*Clustering*: Determining a finite set of implicit classes that describes the data *Classification*: Mapping data items into pre-defined classes	• Cluster analysis • Bayesian classification • Decision or classification trees • Artificial neural networks
Dependency analysis	Finding rules to predict the value of some attribute based on the value of other attributes	• Bayesian networks • Association rules
Deviation and outlier analysis	Finding data items that exhibit unusual deviations from expectations	• Cluster and other data mining methods • Outlier detection
Trend detection	Lines and curves summarizing the database, often over time	• Regression • Sequential pattern extraction
Generalization and characterization	Compact descriptions of the data	• Summary rules • Attribute-oriented induction