

INVITED PAPER

## A RECONSIDERATION OF THE FOUNDATIONS OF IDENTITY MANAGEMENT

ROGER CLARKE

Xamax Consultancy Pty Ltd, Canberra, Australia; Australian National University,  
School of Computer Science; UNSW Law, Sydney, Australia.  
E-mail: Roger.Clarke@xamax.com.au

**Abstract** There is widespread recognition that, during the process of digitalisation, much greater care is necessary in relation to the needs of individuals and society. One key area in which tensions exist is identity management. People think that their identities are intrinsic to themselves. Yet organisations represent themselves as 'provisioning' people with their 'identities'. In addition, the model of identity that organisations typically use evidences some important deficiencies. A fresh approach is needed to the model that underpins organisations' management of their relationships with people. This needs to be based on a deeper appreciation by designers of the nature of the phenomena that they seek to document and to exercise control over. A model of those phenomena is needed that is pragmatic, in the sense of fulfilling the needs of information systems (IS) practitioners and organisations, but also of the people whose data the organisation handles. It also needs to reflect metatheoretic insights. This paper presents such a model. It commences by drawing on ontology, epistemology and axiology in order to establish an outline metatheoretic model. The model is articulated, at the conceptual level and at the data modelling level. Initially, a relatively simple model is established, sufficient for inanimate objects and artefacts. The more complex requirements of humans are then addressed. It is contended that the resulting model provides a robust framework for identification and authentication in IS.

**Keywords:**

identity,  
entity,  
metatheory,  
axiology.

## 1 Introduction

To practitioners, an information system (IS) is a set of interacting activities by humans and artefacts that involve the handling of data. New categories of IS have progressively emerged, enabled by increases in the capacity and sophistication of IS and the information technology (IT) used to support IS. Important among them are transaction data processing (DP) systems, information systems, management information systems (MIS), decision support systems (DSS), and autonomous decision-and-action systems. The last of those categories takes advantage of the marriage of computing and communications with robotics, by including actuators that enable direct action by elements of the system on the world, and embodies the delegation of power to artefacts. Many instances are already deployed in, for example, factories, warehouses, mines and water management, and the decades of use of automated teller machines.

The effectiveness of all forms of IS depends on the extent to which the underlying model of the real world appropriately reflects the features of that world that are relevant to that particular system's purpose. Crucial among those features are the entities and identities with which the system needs to associate data. The term 'identity management' is commonly used, particularly in relation to the people whose data organisations handle, but also for inanimate objects such as stock-items and capital equipment. This aspect of IS has been important throughout the phases of eCommerce, eBusiness, eGovernment, social media, and more recently digitalisation and datafication.

This paper re-visits the problem-domains within which 'identity management' is applied. It builds on prior work in philosophy and the information systems (IS) literature. Among the wide variety of possible philosophical assumptions, an approach is selected that reflects the pragmatic world of IS practice. This is directly relevant to that portion of IS research that seeks to deliver information relevant to IS practice. Given the recent, very strong tendency within the IS discipline towards sophistication and intellectualisation, and preference for addressing other researchers rather than IS professionals, the pragmatic metatheoretic model presented here will be relevant to only a moderate proportion of IS research.

The purpose of the model is to reflect the relevant complexities, and hence to guide organisations in devising data architectures and business processes for IS that reflect real-world things and events, with a particular focus on systems in which some of the real-world things are human beings. The scope encompasses all aspects of the handling of data relating to all forms of entities and identities. Wherever possible, the model presented here uses conventional terms in conventional ways. However, many common usages of terms are ambiguous, inconsistent or unhelpful and even harmful to the effective design and operation of information systems. In these cases, terms are used in ways that are materially different from common usage, and in some cases new terms are proposed. For each term, a definition is provided that relates that term to the remainder of the framework. Once defined, all of the key terms are thereafter referred to using an initial capital. A Glossary of the defined terms is provided (Clarke 2010c).

The paper commences with an outline of the philosophical underpinnings of the analysis, comprising metatheoretic assumptions in three areas, relating to existence (ontology), knowledge (epistemology) and value (axiology). A first distinction is drawn between a real-world and an abstract-world. Within the abstract-world, the conventional approach is adopted, with two levels, one conceptual and the other concerned with data. Inanimate entities are addressed first, enabling a relatively mechanistic approach to be adopted. Human entities are then considered, which brings into play interests, rights and values, and necessitates further layers of complexity in the model.

## **2 A Pragmatic Metatheoretical Model**

This section establishes the philosophical foundations underlying the model put forward in the later sections of the paper. The approach developed in the first paper in this series, Clarke (2021) is briefly re-presented and extended. The model is referred to as 'metatheoretic' (Myers 2018, Cuellar 2020), on the basis that it draws on relevant branches of philosophy, in particular ontology (concerned with existence), epistemology (concerned with knowledge) and axiology (concerned with value). These are key areas in which IS theorists and practitioners alike make 'metatheoretic assumptions', often implicitly, and sometimes consciously. Where the assumptions are both conscious and intentional, a more appropriate term for them is 'metatheoretic commitments'.

The model is also 'pragmatic', as that term is used in philosophy, that is to say it is concerned with understanding and action, rather than merely with describing and representing. The author's intention is instrumentalist: To achieve change in the worldviews of IS practitioners and researchers, and hence changes in behaviour and in the management of data. So the model needs to speak to IS practitioners, and to those IS academics who intend the results of their research to do the same. Figure 1 supports the textual explanations with a visual depiction of the key elements of the model.

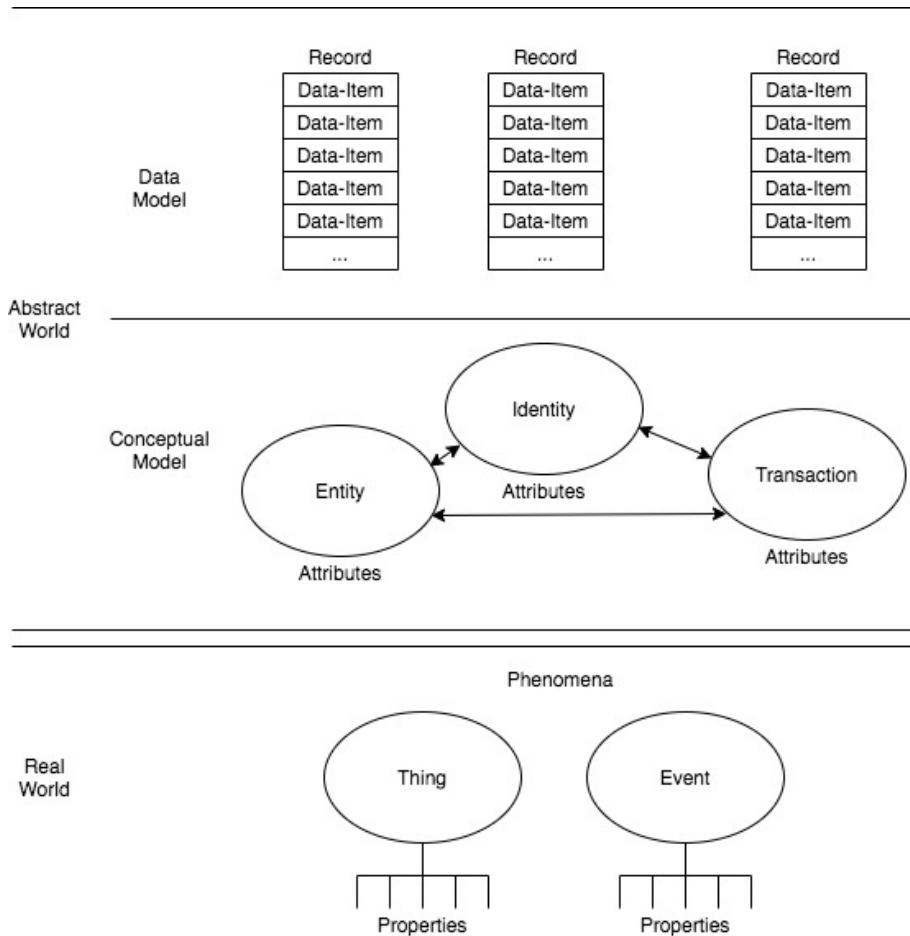


Figure 1: A Pragmatic Metatheoretical Model

## 2.1 Ontology

This section summarises analysis in Clarke (2021). The pragmatic approach adopted is that there is a reality, outside the human mind, where things exist – a position commonly referred to as 'realism'. Humans cannot directly know or capture those things. They can, however, sense and measure those things, create data reflecting them, and construct an internalised model of those things – an assumption related to the ontological assumption referred to as 'idealism'.

The pragmatic model adopted in this paper, and depicted in Figure 1 accordingly distinguishes a Real World from an Abstract World. The Real World comprises Things and Events, collectively Phenomena, which have Properties. These can be sensed by humans and artefacts with varying reliability. Humans create an Abstract World in which Entities are postulated that are intended to correspond to Real-World Things, and Attributes of Entities to represent the Properties of Things. Real-World Events give rise to changes in the Properties of Things, and these are reflected in the Abstract-World as Transactions that give rise to changes in Entities' Attribute-values.

The abstract concept of an Identity, developed further below, caters for the different ways in which Entities present in different circumstances. The various kinds of Entities and Identities have Relationships with one another, represented by arrows in the depiction in Figure 1. The Relationships also have Attributes. Further discussion of these aspects of the model is provided in the following sub-sections.

In the IS field, it is necessary to adopt a flexible conception of what constitutes the Real World. This is because some of the IS that practitioners develop, maintain and operate represent imaginary Things. A design for a new IS is a model of an (as-yet) imaginary Thing. Some IS are just outline representations of an intended future IS, to enable assessment of its likely operational effectiveness, efficiency or security. Other IS create purely formal systems such as games-worlds. Another category of pseudo-Real-Worlds involves past, possible future, and even entirely hypothetical contexts, such as the Earth's atmosphere millions of years ago, or following a large-scale meteorite strike, or 50 years from now, with and without stringent measures to reduce greenhouse gas emissions. The IS profession and discipline need to be able to contribute to and support activities in such areas.

## 2.2 Epistemology

This section summarises analysis in Clarke (2021). Epistemology is the study of knowledge. Two contrasting conceptions of knowledge exist. The proposition of the first, 'empiricism', is that knowledge is derived from sensory experience, and is a body of facts and principles accumulated by humankind over the course of time, that are capable of being stored in the equivalent of a warehouse. This works well in circumstances where the Things represented by Entities are real rather than imaginary, and are inanimate, and their handling is largely mechanical. Examples include aircraft guidance systems and robotic production-lines.

The other, 'apriorist' view is that that knowledge is internal and personal, and the concept is not applicable outside the mind of an individual human. Within this school of thought, knowledge is the matrix of impressions within which an individual situates newly acquired information.

In order to cater for these two extremes, the term 'Knowledge' is best avoided, except when qualified by one of two adjectives:

- 'Tacit Knowledge' exists only in the mind of a particular person, is informal and intangible, and hence is not readily communicated to others. This reflects the apriorist view of epistemology; whereas
- 'Codified Knowledge' refers to Data that has been extracted from individuals' insights, and then structured, expressed and recorded in a more or less formal language (text, formulae, blueprints, procedure descriptions). It is disembodied from individuals, but is communicable among them, and hence capable of delivering a coherent body of Information to individuals in particular contexts. This reflects the empiricist view of epistemology.

A pragmatic metatheoretic approach must support modelling not only in contexts that are simple, stable and uncontroversial, but also where there is no expressible, singular, uncontested 'truth'. The pragmatic assumption adopted here is that both of those categories of philosophical theories are applicable, but in different circumstances.

In Figure 1, the Abstract World is depicted as being modelled at two levels. The Conceptual Model level endeavours to reflect the modeller's perception of the Things, the Events and their Properties, by postulating Entities and Entity-Instances, presentations of Entities called Identities, and Transactions, with Relationships of various kinds among them, all with Attributes.

The notion of an Entity corresponds to a category of Things, and Transaction to a category of Events. In the dialect used by ontologists, the term 'universal' corresponds to a category, and 'particular' refers to an instance. For example, in biology, the notion 'species' (e.g. African Elephant) is a universal, and the notion 'specimen' is a particular. An example that is perhaps more pertinent to IS is the category cargo-containers, which is a universal or Entity, whereas a specific cargo-container is a particular or Entity-Instance. The ideas and terms used in this paper, and articulated further below, are similar to, but not identical with, related ideas in the well-developed and diverse sub-discipline of conceptual modelling.

The other level, referred to here as the Data Model, enables the operationalisation of the relatively abstract ideas in the Conceptual Model level. Central to this level is the notion of 'Data'. The term, used variously as a plural and as a generic noun, refers to any quantity, sign, character or symbol, or collection of them, that is in a form accessible to a person and/or an artefact. The singular term 'datum' has fallen into disuse in recent times, and 'Data-Item' preferred. 'Real-World Data' or 'Empirical Data' is data that represents or purports to represent some Property of a Real-World Phenomenon. That is contrasted with 'Synthetic Data', which is Data that bear no direct relationship to any real-world phenomenon, such as the output from a random-number generator, or data created as a means of testing the performance of software under varying conditions.

Beyond Data, the epistemological aspects of the pragmatic model comprise assumptions made about information, knowledge and wisdom. The term 'information' is used in many ways. Frequently, even in refereed sources, it is used without clarity as to its meaning, and often in a manner interchangeable with Data. The pragmatic model adopted in this paper uses the term 'Information' for a sub-set of Data: that Data that has value. Data has value in only very specific circumstances. Until it is in an appropriate context, Data is not Information, and once it ceases to be in such a context, Data ceases to be Information.

Some people feel very uncomfortable with a definition that embodies such looseness, fuzziness and instability. Rather than a nice, straightforward 'thing', describable in mathematical terms, and analysable using formidable scientific tools, such a definition makes Information rubbery and intangible, a 'will o' the wisp'. I contend that attempts to deny that fuzziness lie at the heart of many problems in IS. By embodying in the IS profession's world-view a too-precise notion that bears little relationship with the Real-World, the modeller pre-destines the resulting IS to be a poor fit with the needs of the IS profession and the people and organisations whose needs they serve.

### 2.3 Axiology

This section summarises analysis in Clarke (2021). The final element of the pragmatic metatheoretic model is concerned with 'Value', in the sense of "the relative worth, usefulness, or importance of a thing" (OED II 6a). The values dominant in many organisations are operational and financial. However, many contexts arise in which there is a pressing need to recognise broader economic interests, and values on other dimensions as well.

Human values are particularly prominent in systems in which people are key players or users, and in systems that materially affect uninvolved people, usefully referred to as 'usees' (Clarke 1992, Fischer-Huebner & Lindskog 2001, Baumer 2015). Examples include people with records in shared industry databases, such as those for police suspects, tenants and insurees; and the conversation-partners of people whose voice and/or electronic communications are subjected to surveillance.

The pragmatic approach to Value recognises that:

- in some simple contexts, virtue-based evaluation (ethically or morally good/bad) may be applicable;
- in other contexts, deontic approaches are necessary, recognising an obligation, and constraining behaviour to achieve compliance with some externally-imposed norm;
- other contexts are teleologically-driven, by which is meant that the determination of appropriate actions is dependent on the degree of alignment of impacts and outcomes with the designer's purpose. A



common-denominator measure is needed, conventionally referred to as 'utility';

- in simple contexts, the only value-conflicts that arise may be between the system-sponsor's financial objectives and system-users' economic needs; and
- greater complexity arises where the assessment of utility must reflect some or all of:
  - multiple value-sets associated with many different stakeholders;
  - stakeholder interests that extend beyond economic factors to include social, environmental and political objectives;
  - intense conflict among values; and
  - solutions based on tolerance, negotiation and compromise.

### **3 Entities and Identities**

This section defines and discusses the notions of entity and identity, which are the two central features of the pragmatic metatheoretic model adopted in this paper. It draws heavily on an earlier working paper (Clarke 2001a) and published article (Clarke 2010a, b, c, d), but re-casts the model in light of the metatheoretic discussions above. It first considers them within the Conceptual Model level, and then at the Data Model level. The notions are applied in this section to inanimate Real-World Things. The following section addresses additional considerations that arise when the Things are human beings.

#### **3.1 (Id)Entities at the Conceptual Model Level**

The conception of an entity adopted here has a great deal in common with the approach used in a wide range of conceptual modelling techniques. An 'Entity' is an element of a Conceptual Model that corresponds with a Real-World Thing. It is a category or collective notion, or a set of instances. In one sense, recognition of Things and Entities is arbitrary, because a modeller can postulate whatever they want to postulate. Generally, however, a modeller has a purpose in mind, and postulates a category judged likely to be useful in understanding some part of the Real-World, and contributing to its management.

Examples of an Entity are the sets of all cargo-containers and of all mobile-phones assigned by an organisation to its employees. Some objects comprise nested layers of objects. For example, cargo-containers may contain pallet-loads, and within that cartons, and within each carton smaller boxes. Each specific occurrence within the set of objects that makes up an Entity is an 'Entity-Instance'. Hence the Entity cargo-containers comprises many Entity-Instances, one for each particular container, and possibly many nested layers of Entity-Instances.

Each of the many specific conceptual modelling techniques has terms that correspond with those used here. In the case of the original Entity-Relationship Model of Chen (1976), an Entity corresponds with Chen's entity-set ("Entities are classified into different entity sets such as EMPLOYEE, PROJECT, and DEPARTMENT" (p.11)), and Entity-Instance has a degree of correspondence with Chen's entity: "An entity is a 'thing' which can be distinctly identified. A specific person, company, or event is an example of an entity" (Chen 1976, p.10 – but the model presented here does not treat an "event" as an Entity-Instance). An Entity may have 'Entity-Attributes', each of which is an element of a Conceptual Model that represents a Real-World Property. Containers, for example, have a colour, an owner, a type (e.g. refrigerated, or half-height), and various kinds of status (e.g. dirty or clean; and empty or loaded).

Many kinds of Entity are perceived rather differently by the modeller, depending on the context. An 'Identity' is a particular presentation of an Entity, as arises when it performs a particular role. A 'Role' is a pattern of behaviour adopted by an Entity. An Entity may adopt one Identity in respect of each Role, or may use the same Identity when performing multiple Roles.

An 'Identity-Instance' is a particular occurrence of an Identity. For example, any particular motor-vehicle is an Entity-Instance; but a motor-vehicle may at any given time be associated with an Identity-Instance, such as 'the getaway-car', 'the car carrying a person-at-risk' (e.g. the Pope), or 'the lead-vehicle in a convoy'. Another example is a single computing device, which is an Entity-Instance, supporting many processes that interact with one another and with processes running in other devices, each process being an Identity-Instance.

Whereas an Entity commonly has physical form, an Identity may have virtual form. An example of an Identity with physical form is the set of all SIM-cards inserted into mobile phones. Virtual form, on the other hand, is apparent in the case of processes running in consumer computing devices and communicating with other processes running in that or some other device. An Identity is related to the notion of role in Chen's ER Model: "The role of an entity in a relationship is the function that it performs in the relationship" (p.12).

The usage of 'Identity' in the pragmatic model presented here is very different from that attributed to the term during recent decades by most organisations. **What are commonly referred to as 'identity management' services commonly embrace the implicit assumption that Entity and Identity are the same notion or that each Entity is limited to a single Identity. This does not correspond with Real-World phenomena, and this single error in mainstream models has led to a great many difficulties in the use of 'identity management' services.** These difficulties arise with inanimate entities, but are particularly problematic where the entities are human. The term 'identity' has longstanding and widespread use by people to refer to a Real-World phenomenon evidenced by human beings, and it has subtleties that organisations have no use for, and which organisational practices have been ignoring. It is important that IS professionals and researchers, and the organisations that use IS, reflect Real-World phenomena, and respect common usage, rather than trapping themselves into misrepresentation, misunderstanding and mis-design.

An Identity may have 'Identity-Attributes', each of which is an element of a Conceptual Model that represents a Real-World Property. Whereas the colour of a car, and its make and model, are Attributes of the Entity, the dangerousness of its occupants is an Attribute associated with the Identity. Similarly, a SIM-card has different attributes from the mobile handset it is inserted into, and the processes running in a computer have different attributes from the computer that is hosting them.

A 'Transaction' is an element of a Conceptual Model that corresponds with a Real-World Event. It has Transaction-Attributes that reflect Real-World Properties that the modeller considers to be relevant to the purpose. A key function of a

'Transaction-Instance' is to give rise to a change in the state of Attributes for one or more Entity-Instances and/or Identity-Instances.

A 'Relationship' is a linkage between two elements within the Conceptual Model level. Figure 1 depicts a Relationship between an Entity and an Identity with a line ending in an arrow at each end. This applies for example to mobile-handsets and SIM-cards. Entities may also have Relationships with other Entities, and Identities with other Identities. For example, motor vehicles need to be associated with other motor vehicles under joint contracts for roadside assistance, and where they are involved in the same accident. Similarly, containers need to be associated with the organisations that own them. Organisations also own and insure motor-vehicles, and hence the two Entities organisations and motor-vehicles need to have some form of link between them.

A Relationship may have 'Relationship-Attributes'. Cardinality is a particularly important attribute. At each end of the line depicting a Relationship it may be that no Relationship exists in that direction (cardinality 0), or a single linkage (1) may be mandatory, or a range of linkages may be possible (conventionally, 'n' and 'm', or '0-n' or '1-n'). For example, a cargo container must have precisely one linkage with an owner (cardinality 1), whereas the Entity that corresponds to Real-World mobile-phone-handsets may be related to multiple Identity-Instances, associated with different SIM-cards that are inserted into it, successively or even simultaneously. The arrow-head on the other end of that line reflects the fact that a SIM-card may be used in multiple, successive mobile-phone-handsets. Similarly, an Entity for motor-vehicles has a one-to-many relationship with an Identity for 'getaway-cars'. Moreover, escapees may use a succession of vehicles, each of which in turn has the Identity 'getaway-car'; so the arrow depicting this Relationship is also two-headed.

In the remainder of this article, when referring to both Entities and Identities, the abbreviation (Id)Entity is applied, and the same approach is adopted to derivative terms such as (Id)Entity-Instance.

### 3.2 (Id)Entities at the Data Model Level

The previous sub-section had its focus on the Conceptual Model level. The (Id)Entity notions require further articulation at the Data Model level. The terms Data, Real-World Data, Synthetic Data and Information were introduced in s.2.2 above. The pragmatic approach proposed in this paper embodies several further concepts.

In the Abstract World in which IS operate, each Attribute of an (Id)Entity is represented by a 'Data-Item', which is a storage-location in which a discrete 'Data-Item-Value' can be represented. The term 'Value', in this context, is a somewhat generalised form of "a numerical measure of a physical quantity" (OED I 4). For example, Entity-Attributes of cargo-containers may be expressed at the Data Model level as Data-Items and Data-Item-Values of Colour = Orange, Owner = MSK (indicating Danish shipping-line Maersk), Type = Half-Height, Freight-Status = Empty.

A collection of Data-Items all of which relate to a single (Id)Entity-Instance is referred to as a 'Record'. A collection of Records may be referred to as a 'File' or data-set. A Record may relate to a particular Entity-Instance (e.g. a container, or mobile handset) or to an Identity-Instance (e.g. a SIM-card), or to a Transaction-Instance. A File relates to an (Id)Entity or to a Transaction.

The term 'Metadata' refers to data that describes some attribute of other Data. Metadata may be explicitly expressed or captured, by cataloguers; or it may be automatically generated, i.e. inferred by software. It may be stored with the data to which it relates, or stored separately. During the last 2-3 decades, the term has become sufficiently widely-used that hyphenation is no longer common.

A vital question that needs to be addressed is the manner in which each individual (Id)Entity-Instance is distinguished from all of the other instances of the same (Id)Entity. Specific terms are adopted in the pragmatic metatheoretic approach proposed in this paper. The term 'Entifier' refers to any one or more Data-Items held in a Record whose value(s), alone or in combination, are sufficient to distinguish any particular Entity-Instance from all other Entity-Instances of the same Entity. The word 'entifier' is not to be found in the Oxford English Dictionary (OED),

although 'entify' is. Surprisingly (judging by the absence of prior usages found using Google Scholar), 'entifier' appears to be a neologism that I originated, first occurrence in Clarke (2001b), and first published in Clarke (2003), defined at the time as "the signifier for an entity").

Examples of single-item Entifiers include the BIC-code of a cargo-container (BIC being an abbreviation of Bureau International des Containers), the Vehicle Identification Number (VIN) of a motor-vehicle, and the International Mobile Equipment Identity (IMEI) of a mobile-phone. In some circumstances, a proxy-Entifier may be used, e.g. for a computing device, the Network Interface Card Identifier (NICId) of an Ethernet card that is installed in it.

Artefacts are usually distinguished by Entifiers that are purpose-designed, and hence comprise a single Data-Item. However, an example of a multi-data-item Entifier arises in jurisdictions that re-issue motor-vehicle registration-plates previously allocated to a now-defunct vehicle. To achieve the uniqueness that is highly desirable in an Entifier, a date-range needs to be included as part of the Entifier.

An 'Identifier' is any one or more Data-Items held in a Record whose value(s), alone or in combination, are sufficient to distinguish any particular Identity-Instance from all other Identity-Instances of the same Identity. This is a mainstream use of the term, as evidenced by Oxford English Dictionary (OED) definition 1a: "A thing used to identify someone or something".

Examples of single-item Identifiers include a code assigned by a traffic-control authority to a vehicle of interest, for example when monitoring average speed over a section of road, the Integrated Circuit Card Identification (ICCID) of a SIM-card, and a process-id (e.g. for a software agent).

In Figure 2, a visual depiction is provided of the elements of the Conceptual and Data Modelling levels defined so far in this section.

When a Real-World Event occurs, and is reflected in a Conceptual Model-level Transaction, a Record arises, whose function is to cause a change of state in one or more Attributes of one or more (Id)Entities. Means are needed to establish which

(Id)Entity-Instances are affected by the Transaction Record. This is achieved by means of (Id)Entification processes.

The term 'Identification' refers to the process whereby Data is associated with a particular Identity-Instance. This involves acquiring or postulating an Identifier that matches with previously-recorded Data-Item-Values. This application of the term is consistent with dictionary definitions, and has been used in this manner in my works since Clarke (1994c). The term has many other, loose usages, however, particularly as a synonym for 'identifier' (discussed above) or for 'token' or 'authenticator'.

An example of the Identification process in operation is the matching of a SIM-card's ICCID to an existing Record. An example of the use of a multi-Data-Item Identifier is the recognition of a vehicle on the basis of its properties (such as make, model and colour) at each end of a section of roadway over which average speed is being assessed. Another example is the use, as a proxy Identifier for a particular process running in a computing device, of the combination of a port-number and IP-address, together with a date-time range (to allow for IP-addresses being 'dynamic', i.e. subject to being re-assigned).

The term 'Entification' refers to the process whereby Data is associated with a particular Entity-Instance. This involves acquiring or postulating an Entifier that matches with previously-recorded Data-Item-Values. The term exists in some online dictionaries and with a not unrelated meaning, but not in the OED. The term has been used consistently in my work since Clarke (2001b), but to date neither it nor, it seems, any equivalent has become mainstream.

The emergence of some such term is important, because there are material differences between Identification and Entification, variously conceptually, in terms of the Data involved, and in relation to their impacts and implications. **The failure of conventional identity management schemes to differentiate between entities and identities, and between identification and entification processes, has given rise to many IS design, deployment and operational issues.**

Examples of Entification include the matching of a particular cargo-container's BIC-Code, or a motor-vehicle's VIN, to an existing Record. In addition to such purpose-

designed Entifiers, Data-Items of convenience are often relied upon. For example, for computing devices that do not have a reliable, purpose-designed Identifier, the NICId of the Ethernet (or other) card inserted into the computing device, may be used as a proxy. An Ethernet NICId is an example of a multi-data-item Entifier, in that it comprises two Data-Items, an Organizational Unique Identifier (OUI) and a Manufacturer-Serial-No. Dependence on proxies of this nature has varying degrees of reliability.

The acquisition of the Entifier may be by observation followed by either transcription of the Data-Item-Value by a human, or alternatively by technologically-assisted means such as image-recording using a camera followed by application of optical character recognition (OCR) to extract the value. Another approach is to pre-store the Entifier in a machine-readable form, such as a barcode or a chip, and later use an appropriate technology to extract a copy of that pre-stored Data.

From an administrative perspective, (Id)Entification procedures need to be reliable and inexpensive. Achieving that aim can be facilitated by pre-recording an (Id)Entifier on a Token from which it can be conveniently captured. One common form of Token is a card, with the data stored in a physical form such as embossing, or on, or in, a recording-medium such as a magnetic stripe or a silicon chip.

This section has used the simplifying assumption that the Things underlying the (Id)Entities are inanimate, and capable of being treated as mere objects, with minimal concern about the Thing's interests and about clashes among values. The following section relaxes that assumption and considers the additional factors that arise when the underlying Things are people.



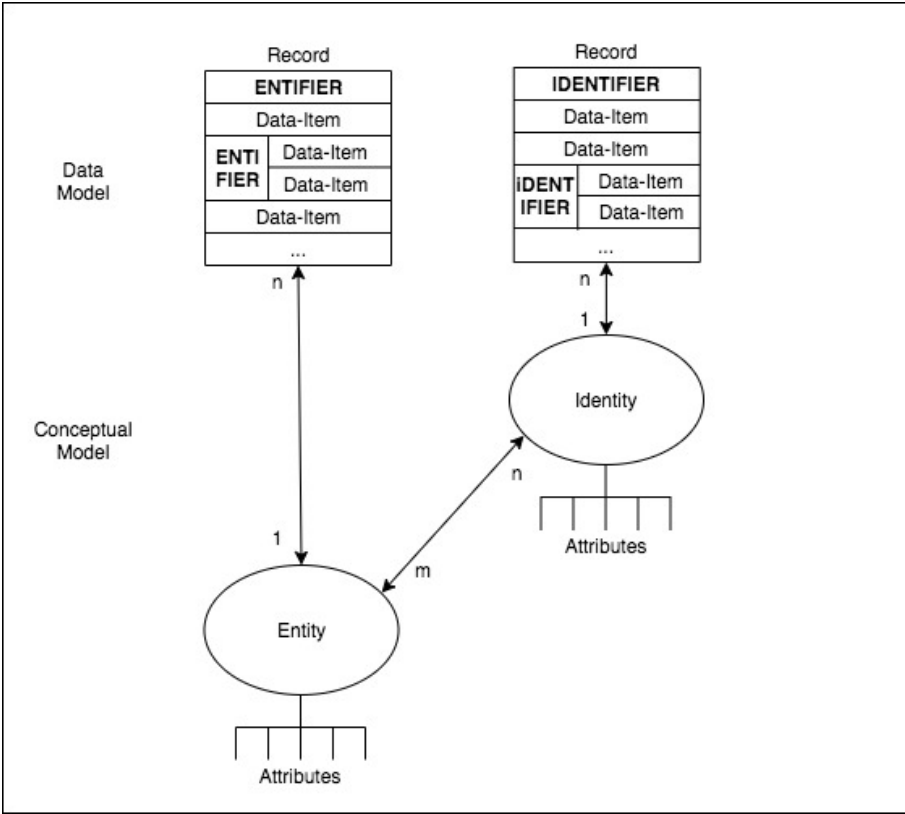


Figure 2: (Id)Entities and (Id) Entifiers

#### 4 The Model Applied to Humans

Limiting the model's focus to inanimate objects and their representations enabled a straightforward, mechanistic approach to be adopted, and the values (axiological) aspects left in the background. In many circumstances, animals are also treated as objects. Flies and mice are variously poisoned and injected, and the impacts are rendered as Data. Cattle are entified using brands and ear-tags, and pets have chips injected. On the other hand, animal welfare constraints are placed on the handling of vertebrate animals during life and in relation to the manner of death. In some circumstances, Data is required by law to be gathered and stored, such as stocking densities for caged chickens and inoculation records, and some forms of animal slaughter are subject to monitoring and Data-recording.

Where the Entities being modelled are human beings, however, further factors come into play, and hence both the Conceptual and Data Modelling levels need to be adapted in order to reflect those factors. One consideration is the 'free will' or volitional aspect of human beings: inanimate objects do not act of their own accord, and do not have interests that influence their behaviour. In addition, values and rights loom far larger when the Entities involved are human beings. The terms 'objectification', in its sense of "the demotion or degrading of a person or class of people ... to the status of a mere object" (OED 2), and the recent terms 'digitalisation' (Brennan & Kreiss 2016), and 'datafication' (Lycett 2014) or 'datification' (Newell & Marabelli 2015), all carry a pejorative tone when used in respect of people. This is because **the mechanistic application of data-handling notions to humans involves a clash of values between administrative efficiency on the one hand and humanism on the other**. This section considers the impact on the modelling approach firstly at the conceptual and then at the data level.

#### 4.1 The Conceptual Model Level

In section 3.1, a series of concepts was discussed and defined. The application of these concept to humans requires care. The notion of human Entity is (at least to date) uncontroversial, with Entity-Instances confined solely to specimens of the species homo sapiens. A great many Entity-Attributes are applicable specifically to human Entities. Some are physiological in nature, such as the person's hair-colour, gender, and date-of-birth or age-range. Others arise from the person's behaviour, such as their gait, how they write a signature or type a password, where they live, and their capacity to act as an agent for another Entity-Instance.

Each human Entity-Instance may present many Identity-Instances, to different people and organisations, and in different contexts. **The notion of Identity is especially important to humans, because each Entity-Instance (person) plays many roles in many contexts, and these in many cases give rise to separate Identity-Instances. Examples in economic contexts alone include seller, buyer, supplier, receiver, debtor, creditor, payer, payee, principal, agent, franchisor, franchisee, lessor, lessee, copyright licensor, copyright licensee, employer, employee, contractor, contractee, trustee, beneficiary, tax-assessor, tax-assessee, business licensor, business licensee, plaintiff,**

**respondent, investigator, investigatee, and defendant. A similar richness exists in social contexts.**

In many circumstances, an Identity-Instance is a presentation or role of a single, specific underlying Entity-Instance, e.g. 'I' (an Entity-Instance) am the sole 'author of this paper' (an Identity-Instance). On the other hand, some roles are filled by different people, in some cases only serially and in other cases in parallel as well. Examples of serial ambiguity include club treasurer and journal editor-in-chief, and examples of parallel ambiguity include club committee-member, journal senior editor and fire warden.

Human Identity-Attributes are related to a presentation or role, rather than reliably to a particular Entity-Instance. For example, an eConsumer has a profile comprising such features as demographics, interests, user-interface preferences and prior purchases. These Attributes may be common across some or all of the Identity-Instances a human Entity-Instance adopts; but very commonly many are not.

People performing roles in organisations inherit authorisations, permissions or privileges. While acting in their manager's absence, a person may be able to sign sick leave forms for their peers, and during an emergency, as fire warden, they can give orders to the CEO's secretary, and even the CEO. A major issue in data security and in fraud is the phenomenon of individuals abusing powers that they have by virtue of one role that they play, by applying them for extraneous purposes unrelated to that role. The Identity-Attribute commonly referred to as authorisation is accordingly very significant in many IS, and is further examined in a later paper in this series.

Transactions represent Real-World Events that give rise to changes in (Id)Entity-Attributes. Events involving humans can be both significant and sensitive, and hence considerable care is needed in the design and processing of such Transactions.

In Figure 2 above, Entities and Identities are shown as having a Relationship. The complexities of this Relationship are particularly significant where the Real World Things are humans. Relationship has a Relationship-Attribute of cardinality. Any particular Relationship-Instance may be:

- prohibited (usefully indicated by the absence of an arrow-head, or a '0' at that end of the arrow), e.g. a person under 18 is not permitted to have a driver's licence;
- mandatory and mandatorily singular ('1'), e.g. in many jurisdictions, a person may have only one spouse;
- mandatory but of any value ('1-n'), e.g. in some IS, each human Entity-Instance reflected in a Record may be required to have at least one Entifier;  
or
- optional and of any value ('0-n' or just 'n'), e.g. each human Entity-Instance may have any number of associated children.

Each human Entity-Instance may relate to multiple Identity-Instances (hence 'n'). Further, because many Identity-Instances can be adopted by multiple Entities (multiple fire wardens at once, multiple journal-editors in succession), the other end of the arrow is marked with an 'm' – equivalent to 'n', but implying that it is a variable independent from the 'n' at the other end of the arrow.

Subtleties in the Relationships between human Entities and Identities need to be well-understood by the designers and users of IS, and reflected in data models and business processes. A particular human Entity-Instance may strongly desire to be the only user of a particular Identity-Instance (e.g. people are very particular about who exercises the capacity to operate on their various bank accounts). Similarly, an organisation may be very concerned that a particular Identity-Instance is used only by one or more specific Entity-Instances (e.g. for the signing of contracts that bind the organisation, and for making statements to the media). **There are circumstances in which it is strongly desirable to prevent use of particular Identities by inappropriate parties. Undesirable activities of these kinds are described by such terms as impersonation, masquerade, spoofing, identity fraud and identity theft. The pragmatic model presented in this paper enables representation of these concepts. The (often implicit) models underlying some identity management schemes fail to do so.**

## 4.2 The Data Model Level

In section 4.2, further (Id)Entity notions were defined at the Data Model level. These too require further articulation where humans are involved. Because of the high valuation placed on human-ness, many aspects of the manner in which Data relating to inanimate objects is handled is inadequate where the Data relates to human Entities. Since c.1970, the collection and management of Data about humans has exploded under the pressures of increased organisational scale, increased social distance, and increased IT capabilities. So a great deal of public concern has arisen about the use and abuse of this Personal Data by organisations. Personal-Data-Items vary enormously in their degree of sensitivity. However, no simple formula exists for assessing sensitivity. It is dependent on individuals, their personal histories and concerns, and the contexts that they find themselves in from time to time.

The problems have been compounded by widespread, casual re-use of data for additional purposes. The meanings of data-items and their content, and the choices made in relation to data-quality, are seldom clear to the recipients. Yet more problems arise where data is drawn from multiple sources. Incompatibilities among the quality-levels and the meanings inevitably lead to inappropriate inferences.

To address the risk that the activities of government and business might be negatively affected by these public concerns, laws relating to 'personal data/information' and 'data protection' emerged. The early, largely nominal protections have proven inadequate to placate an increasingly concerned public. **Data protection laws now place considerable constraints on organisations' data-handling activities, and make considerable demands on identity management schemes. The model of (Id)Entification presented in this paper is intended to enable those challenges to be met.**

Organisations are confronted with challenges in relation to the collection, storage and use of particular Personal-Data-Items (e.g. religion, marital status, ethnicity, disability), and particular Personal Data-Item-Values (e.g. non-binary gender-choices, and gender-preferences other than hetero-sexuality). Depending on the jurisdiction, overt discrimination based on such information, even if demonstrably relevant, may be precluded by law.

Common examples of Identifiers used by or for humans include the particular name or name-variant that a person commonly uses in a particular context, such as with family, with a particular group of friends, or when working in a customer-facing role such as a prison officer, psychiatric nurse, counsellor or telephone help-desk. Names are highly variable and error-prone. They do not represent convenient Identifiers for operators of information systems, and are often supplemented by synonym-breakers, such as date-of-birth or some component of address. More effective and efficient business processes can be achieved by means of an organisation-imposed alphanumeric code, such as a customer-code or a username (Clarke 1994c). Each human Identity-Instance may themselves use many Identifiers including variants of names, and may be assigned many more Identifiers by organisations.

As discussed above, some Identifiers comprise more than one Data-Item. In rich datasets, however, a large number of multi-data-item Candidate Identifiers may be available. Examples are particularly prevalent in the kinds of data-collections about which most people feel the greatest sensitivity: health data and financial data. For example, **uniqueness can readily arise from unusual medical conditions and postcode of residence; or even place, gender and date of birth (Sweeney 2000). See also Ohm (2010) and Slee (2011). Yet it is precisely these kinds of rich data-collections that are being expropriated by governments obsessed by the 'big data' mantra, and blind to the issues of incommensurable data definitions, a-contextual applications of data, and low data-quality.**

Camouflage techniques in the form of Personal Data De-identification have been attempted, but rich data-sets are inadequately resistant to Personal Data Re-identification techniques. Personal Data Falsification is necessary if balance is to be achieved between personal values and collectivist values (Clarke 2019b).

For human entities, the primary form of Identifier is a biometric. This is a measure of some aspect of the physical person that is unique (or is claimed, or assumed, to be so). Examples include a thumbprint, fingerprints, an iris-pattern and DNA-segments. The uniqueness is not guaranteed. In theoretical terms, some biometric measures are capable of providing a very high probability of uniqueness. On the other hand, the practice of biometrics is far less reliable than theory suggests it could be, because a very substantial set of challenges have to be overcome. The literature

on biometric challenges and resulting quality is somewhat sparse, but see Mansfield & Wayman (2002) and Clarke (2002b). In some circumstances, occasional errors may matter very little and/or be easily discovered and corrected. On the other hand, some errors remain concealed, and serious consequences can arise from them, varying from psychological, social and economic harm to cases of conviction, imprisonment and even execution of the wrong person.

Another category of human Entifier is usefully referred to as an 'imposed biometric'. Examples include a brand imposed by tattooing or other techniques on a person's skin, and a unique code pre-programmed into an RFID tag that is closely associated with the person, or implanted in them (Clarke 1994b, 1997, 2001a, 2002a).

The term Nymity refers to circumstances in which the relationship between Entity and Identity is unclear. The term Anonymity refers to a characteristic of an Identity-Instance, whereby it cannot be associated with any particular Entity-Instance, whether from the data itself, or by combining it with other data. In the case of Pseudonymity, on the other hand, association of an Identity with a particular Entity may be achieved, but only if legal, organisational and technical constraints are overcome (Clarke 1999). Where either form of Nymity applies, it is inappropriate to use the term 'Identifier'. The term Pseudonym refers to a circumstance in which the association between the Identifier and the underlying Entity is not known, but in principle at least could be known. For example, a carefully-protected index may be used to sustain a link between a client-code and the name and address of the AIDS-sufferer to whom the record relates. If an Identifier cannot be linked to an Entity at all, then it is appropriately described as an Anonym. The term Nym usefully encompasses both Pseudonyms and Anonyms.

The term Pseudonym is widely used, and has a large number of synonyms (including aka, 'also-known-as', alias, avatar, character, handle, nickname, nick, nom de guerre, nom de plume, manifestation, moniker, persona, personality, profile, pseudonym, pseudo-identifier, sobriquet and stage-name). In contrast, only a small number of authors have used the term Nym, although it is readily traceable back prior to 1997. Even fewer have used the term Anonym, but it is far from unknown and I have used it consistently in my work since Clarke (2002c).

There are many circumstances in which an Identifier is unnecessary and a Nym is entirely adequate. A common example is enquiries in which a set of circumstances is described by the enquirer, and a response is provided explaining the applicability of the law, or of an organisation's policies, to those circumstances. Enquiries are in many cases conducted as a single contiguous conversation. However, it is also possible for multiple, successive interactions to be connected with one another by means of a Persistent Nym, such as <meaningless-string>@gmail.com.

Identification refers to the process whereby Data is associated with a particular Identity-Instance, in this case an Identity-Instance used by a human. It involves the acquisition of an Identifier, such as a person's commonly-used name or one of their nyms, or a customer number, or a ticket-number for a particular queue. This may be provided by the person concerned, by voice or in textual form, by displaying a Token such as a membership card or a ticket, or by making a Token available that contains a pre-stored Identifier capable of being read by a device operated by an organisation.

Entification refers to the process whereby Data is associated with a particular Human Entity-Instance. This depends on the acquisition of an Entifier such as a biometric, or an imposed biometric such as an implanted chip. All forms of biometric acquisition are highly personal and threatening, and many are demeaning. For example, high-quality recording of a thumbprint or fingerprints involves a skilled operator grasping the person's wrist and controlling the hand's movement, and iris-scans and retinal-scans involve submission of the body to whatever device the measuring organisation imposes on the individual. The moderate quality of those biometric measures results in a material degree of error and hence mistaken identity. The negative impacts of those errors commonly fall on the individual. Entifiers pre-stored on a Token such as a chip-card can be captured in a technologically-assisted or -performed manner. However, that greatly increases the risk of the Entifier being associated with the wrong person.

Pastoralists have had no qualms about clipping RFID-cards onto the ears of entire herds of stock-animals. Pet-owners have accepted the injection of chips into their beloved animals because they perceive it to increase the chances of a lost pet being returned to its owner. The same approaches, however, have historically excited revulsion when applied to humans. Early applications to humans have included



chips in 'anklets' for convicts, and even remandees, in 'prisons without walls', in military 'dog-tags' to assist the identification of combat casualties, and chips injected into the bodies of staff in research facilities; but also in a few consensual contexts, such as patient-tags to assist in ensuring that operations are performed on the right person and to the right body-part; and chips injected into the bodies of customers of fashionable bars, who want a fashion-statement and/or doors to open automatically for them. On the other hand, the insertion of chips into the tooth-enamel of children to identify victims of kidnapping and abduction was not attractive to parents, and considerable concern has been expressed about entification tokens imposed on the aged. It remains to be seen whether and to what extent human values will be overridden and/or voluntarily sacrificed through this form of objectification of individuals.

IS are designed to assist organisations in administering their interactions with humans by recording Data-Item-Values for relevant (Id)Entity-Instances. The Data-Item-Values for each particular (Id)Entity-Instance are stored in a Record that contains one or more of their Entifiers or Identifiers. Data-Item-Values contained in each new Transaction can be used to locate the appropriate Record on the basis of an (Id)Entifier that the Transaction contains. Hence decisions can be made, actions taken, and amendments made to the Record for that person.

During the early decades of IS, the primary source of organisations' Data about individuals was Transactions between that organisation and the individual concerned. However, **since the late 20th century, organisations have increasingly drawn Data from multiple, additional sources, and consolidated it into individuals' Records. The reliability of the association, the potential conflicts among meanings of apparently similar Data-Items, and the nature of the original collection and subsequent handling, result in a degree of doubt about data-quality standards in such circumstances. The degree of expropriation of Personal Data has intensified enormously since the emergence of the Digital Surveillance Economy c.2005.** This commenced with the inversion of the originally user-driven World Wide Web by means of Web 2.0 technologies, and the explosion of social media and technology platforms more generally (Zuboff 2015, Clarke 2019a, Clarke 2022).

The term 'Digital Persona' refers to a model of an individual's public personality based on Data and maintained by Transactions, and intended for use as a proxy for the individual. The term is my own coinage, first presented at the Computers, Freedom & Privacy Conference in San Francisco (Clarke 1993), and published in Clarke (1994a). But it is in any case an intuitive term and has gained some degree of currency, as documented in Clarke (2014). It is quite common to see the term 'identity' used to refer to what is called here a Digital Persona; but 'identity' has many meanings, and to avoid ambiguity it is far preferable that some other term be used. Another candidate term is e-persona.

A 'Projected Digital Persona' is under the control of the individual, and is fundamental to an individual's sense of self and self-esteem. An 'Active Projected Digital Persona' is capable of taking actions as an agent for the individual. An agent may be as simple as an auto-responder to emails, or as complex as a bot that conducts social interactions intended to create and maintain a Digital Persona radically different from the individual it nominally represents. An active agent may have varying degrees of autonomy.

A Projected Digital Persona may have an 'Avatar' associated with it. This is a visual representation or embodiment of a digital persona., static or moving, which enhances, or substitutes for, an image of the (or an) underlying Entity.

An 'Imposed Digital Persona' is one controlled by someone other than the individual it is associated with. It is a model, based on data held in records and extended by transactions, which an organisation treats as an adequate impression of the Entity or Identity that it assumes the persona to represent. Under the approach adopted until the mid-twentieth century, an organisation's decisions about each individual were made by an employee local to that individual, based on available information, much of it provided at the time by the individual themselves. This exposed the organisation to the risk of uneven application of policies, and required human resources, widely dispersed. By making decisions instead on the basis of an Imposed Digital Persona, automation was facilitated, and considerable staff-savings are achieved. For individuals, **organisations' use of Imposed Digital Personae give rise to enormous social distance from the organisations they deal with, errors and inequities that are very difficult to even challenge let alone get fixed, and a great deal of dissatisfaction and frustration.**

## 5 Conclusions and Further Work

This paper has presented a model of the important area of entities and identities and their representation in information systems. The model is pragmatic, in that it reflects the needs of IS practitioners and researchers performing practice-relevant research, but also of those affected by IS. The model also reflects metatheoretic insights arising from the relevant branches of philosophy, reported in a predecessor article, Clarke (2021).

The notions have been introduced progressively, with definitions provided for each term, and extracted into a Glossary (Clarke 2010c). In a first pass across the territory, the simpler kinds of entities, inanimate objects and artefacts have been addressed. Further concepts have then been introduced that are necessary for IS to deal appropriately with human entities and identities.

The adoption of the model, or the re-working of existing models to take into account these key points of difference, will enable a number of weaknesses to be overcome in existing IS that assist in the management of digital personae for entities and identities. The first of these key points of difference is clear differentiation between an entity (representing a real-world physical thing), and an identity (representing a real-world virtual thing). Another is the allowance for an m:n relationship between entities and identities. These ideas lead to an appreciation of the challenges involved in appropriately associating entity-instances with identity-instances, and in appropriately associating data with an entity-instance or identity-instance. Another insight is the need to accept the existence of nymity, where an association between an identity-instance and the underlying entity-instance cannot be reliably established.

The work has laid the foundations for extensions to the basic model presented in this paper. These include the authentication of assertions of (id)entity, authorisations and access control, data and information quality, bias in organisational decision-making, and 'fake news'. In addition to the inanimate objects, artefacts and people addressed in this paper, other categories of (id)entity are also important in IS. So-called 'incorporated' organisations are key to economic and social activities, because they enable scale and longevity. Another looming challenge for humankind is delegation to active-artefact entities including decision-and-action systems and

robots. This harbours the risk of autonomous artefacts making decisions and taking actions with direct and material effects on human beings, without any real-world entity being accountable for the decision, the action and the outcomes. Further model articulation for organisations and decision-and-action systems will be presented in subsequent papers in this series.

### Acknowledgements

The author gratefully acknowledges feedback provided by participants in workshops on aspects of this topic over the last two decades, and the insightful comments of a reviewer.

### References

- Baumer E.P.S. (2015) 'Uses' Proc. 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15), April 2015
- Brennen S. & Kreiss D. (2016) 'Digitalization and Digitization' International Encyclopedia of Communication Theory and Philosophy, October 2016, PrePrint at <http://culturedigitally.org/2014/09/digitalization-and-digitization/>
- Chen P.P.S. (1976) 'The Entity-Relationship Model - Toward a Unified View of Data' ACM Transactions on Database Systems 1 (March 1976) 9-36
- Clarke R. (1992) 'Extra-Organisational Systems: A Challenge to the Software Engineering Paradigm' Proc. IFIP World Congress, Madrid, September 1992, PrePrint at <http://www.rogerclarke.com/SOS/PaperExtraOrgSys.html>
- Clarke R. (1993) 'Computer Matching and Digital Identity' Proc. Computers Freedom & Privacy, Burlingame CA, March 1993, at <http://cpsr.org/prevsite/conferences/cfp93/clarke.html/>, PrePrint at <http://rogerclarke.com/DV/CFP93.html>
- Clarke R. (1994a) 'The Digital Persona and its Application to Data Surveillance', The Information Society 10, 2 (June 1994), PrePrint at <http://rogerclarke.com/DV/DigPersona.html>
- Clarke R. (1994b) 'Information Technology: Weapon of Authoritarianism or Tool of Democracy?' Proc. IFIP World Congress, Hamburg, September 1994, at <http://www.rogerclarke.com/DV/PaperAuthism.html>
- Clarke R. (1994c) 'Human Identification in Information Systems: Management Challenges and Public Policy Issues' Information Technology & People 7, 4 (December 1994) 6-37, PrePrint at <http://rogerclarke.com/DV/HumanID.html>
- Clarke R. (1997) 'Chip-Based ID: Promise and Peril' Proc. Int'l Conf. on Privacy, Montreal September 1997, PrePrint at <http://www.rogerclarke.com/DV/IDCards97.html>
- Clarke R. (1999) 'Anonymous, Pseudonymous and Identified Transactions: The Spectrum of Choice' Proc. IFIP User Identification & Privacy Protection Conference, Stockholm, June 1999, PrePrint at <http://www.rogerclarke.com/DV/UIPP99.html>
- Clarke R. (2001a) 'Authentication: A Sufficiently Rich Model to Enable e-Business' Xamax Consultancy Pty Ltd, December 2001, at <http://rogerclarke.com/EC/AuthModel.html>
- Clarke R. (2001b) 'The Re-Invention of Public Key Infrastructure' Working Paper, Xamax Consultancy Pty Ltd, 22 December 2001, at <http://rogerclarke.com/EC/PKIREinv.html>
- Clarke R. (2002a) 'Biometrics in Airports: How To, and How Not To, Stop Mahommed Atta and Friends' Xamax Consultancy Pty Ltd, February 2002, at <http://www.rogerclarke.com/DV/BioAirports.html>

- Clarke R. (2002b) 'Biometrics' Inadequacies and Threats, and the Need for Regulation' Xamax Consultancy Pty Ltd, April 2002, at <http://www.rogerclarke.com/DV/BiomThreats.html>
- Clarke R. (2002c) 'The Mythology of Consumer Identity Authentication', Statement for a Panel Session on 'Understanding e-Business: Can we remain anonymous in the marketplace?' Proc. 24th Int'l Conf. of Data Protection & Privacy Commissioners, Cardiff UK, 9-11 September 2002, PrePrint at <http://www.rogerclarke.com/DV/AnonDPPC02.html>
- Clarke R. (2003) 'Authentication Re-visited: How Public Key Infrastructure Could Yet Prosper' Proc. 16th Int'l eCommerce Conf., Bled, Slovenia, 9-11 June 2003, PrePrint at <http://www.rogerclarke.com/EC/Bled03.html>
- Clarke R. (2010a) 'A Sufficiently Rich Model of (Id)entity, Authentication and Authorisation' Proc. IDIS 2009 - The 2nd Multidisciplinary Workshop on Identity in the Information Society, LSE, London, June 2009, rev. February 2010, PrePrint at <http://rogerclarke.com/ID/IdModel-1002.html>
- Clarke R. (2010b) 'A Sufficiently Rich Model of (Id)entity, Authentication and Authorisation: Supplementary Materials' Xamax Consultancy Pty Ltd, February 2010, at <http://rogerclarke.com/ID/IdModel-Supp-1002.html>
- Clarke R. (2010c) 'A Sufficiently Rich Model of (Id)entity, Authentication and Authorisation: Glossary of Terms' Xamax Consultancy Pty Ltd, February 2010, at <http://rogerclarke.com/ID/IdModel-Gloss-1002.html>
- Clarke R. (2010d) 'A Sufficiently Rich Model of (Id)entity, Authentication and Authorisation: Application of the Model' Xamax Consultancy Pty Ltd, February 2010, at <http://rogerclarke.com/ID/IdModel-App-1002.html>
- Clarke R. (2014) 'Promise Unfulfilled: The Digital Persona Concept, Two Decades Later' *Information Technology & People* 27, 2 (Jun 2014) 182 - 207, PrePrint at <http://www.rogerclarke.com/ID/DP12.html>
- Clarke R. (2019a) 'Risks Inherent in the Digital Surveillance Economy: A Research Agenda' *Journal of Information Technology* 34, 1 (March 2019) 59-80, PrePrint at <http://www.rogerclarke.com/EC/DSE.html>
- Clarke R. (2019b) 'Beyond De-Identification: Record Falsification to Disarm Expropriated Data-Sets' Proc. 32nd Bled eConference, June 2019, PrePrint at <http://www.rogerclarke.com/DV/RFED.html>
- Clarke R. (2021) 'A Platform for a Pragmatic Metatheoretic Model for Information Systems Practice and Research' Proc. Australasian Con. Infor. Syst., December 2021, PrePrint at <http://rogerclarke.com/ID/PMM.html>
- Clarke R. (2022) 'Research Opportunities in the Regulatory Aspects of Electronic Markets' *Electronic Markets* 32, 1 (Jan-Mar 2022) 179-200, PrePrint at <http://rogerclarke.com/EC/RAEM.html>
- Cuellar M.J. (2020) 'The Philosopher's Corner: Beyond Epistemology and Methodology - A Plea for a Disciplined Metatheoretical Pluralism' *The DATABASE for Advances in Information Systems* 51, 2 (May 2020) 101-112
- Fischer-Hübner S. & Lindskog H. (2001) 'Teaching Privacy-Enhancing Technologies' Proc. IFIP WG 11.8 2nd World Conf. on Information Security Education, Perth, Australia
- Lycett M. (2014) 'Datafication: Making Sense of (Big) Data in a Complex World' *European Journal of Information Systems* 22, 4 (December 2014) 381-386, at <https://core.ac.uk/download/pdf/19784773.pdf>
- Mansfield A.J. & Wayman J.L. (2002) 'Best Practices in Testing and Reporting Performance of Biometric Devices: Version 2.01' National Physical Laboratory Report, CMSC 14/02, United Kingdom, August 2002
- Myers M.D. (2018) 'The philosopher's corner: The value of philosophical debate: Paul Feyerabend and his relevance for IS research' *The DATA BASE for Advances in Information Systems* 49, 4 (November 2018) 11-14
- Newell S. & Marabelli M. (2015) 'Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of "Datification"' *The Journal of*

- Strategic Information Systems 24, 1 (2015) 3-14, at <http://marcomarabelli.com/Newell-Marabelli-JSIS-2015.pdf>
- Ohm P. (2010) 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' 57 UCLA Law Review 1701 (2010) 1701-1711, at <http://www.patents.gov.il/NR/rdonlyres/E1685C34-19FF-47F0-B460-9D3DC9D89103/26389/UCLAOhmFailureofAnonymity5763.pdf>
- Slee T. (2011) 'Data Anonymization and Re-identification: Some Basics Of Data Privacy: Why Personally Identifiable Information is irrelevant' Whimsley, September 2011, at <http://tomslee.net/2011/09/data-anonymization-and-re-identification-some-basics-of-data-privacy.html>
- Sweeney L. (2000) 'Simple Demographics Often Identify People Uniquely' Data Privacy Working Paper 3, Carnegie Mellon University, 2000, at <https://dataprivacylab.org/projects/identifiability/paper1.pdf>
- Zuboff S. (2015) 'Big other: Surveillance capitalism and the prospects of an information civilization' Journal of Information Technology 30, 1 (March 2015) 75-89, at <https://cryptome.org/2015/07/big-other.pdf>