

Journal of the Association for Information Systems

JAIS 

Research Article

Rethinking the Meaning of Identifiers in Information Infrastructures*

Owen Eriksson

Uppsala University, Sweden
and Linköping University, Sweden
owen.eriksson@im.uu.se

Pär J. Ågerfalk

Uppsala University, Sweden
par.agerfalk@im.uu.se

Abstract

Identifiers — such as personal identification numbers, student numbers, and license numbers — are used for identifying individual objects and constitute an important part of the information infrastructures of organizations and society. The design, choice, assignment, withdrawal, and replacement of identifiers are significant economic and political issues with more profound consequences than are perhaps commonly perceived. Use of identifiers can result in significant costs because they may include descriptive information, because an inappropriate identifier may be chosen for the object in question, or because there may be a lack of institutional control of the identifier. The objective of this paper is to elaborate on these problems by explaining the identifier construct from a technical, institutional, ontological, and information infrastructural perspective. Based on this understanding, we provide guidelines for how identifiers should be designed, chosen, replaced, and controlled. Accordingly, we address the practical need for improved design principles relating to the increasingly important infrastructural character of computerized information systems that stems from the importance of appropriate identifiers for information infrastructures and society as a whole. In order to understand the role, function, and meaning of identifiers, it is important to acknowledge that the identifier is fundamentally a linguistic construct used when referring to socially constructed institutional objects. Institutional objects are symbolic entities that represent institutional and brute facts, which are the results of human actions.

Keywords: *identifier, institutional fact, brute fact, information infrastructure, ontology, database design, information systems development, speech act, pragmatic meaning.*

Vol. 11 Issue 8 pp. 433-454 August 2010

* Kalle Lyytinen was the accepting senior editor. This was submitted on April 2, 2008 and went through three revisions.

1. Introduction

Our information society, with ever-increasing demands on system integration and information sharing, emphasizes the infrastructural character of information systems more than ever (Hanseth and Lyytinen, 2010; Iannacci, 2010). Unfortunately, the installed base of existing systems, databases, and their interfaces often leads to lock-in situations (Hanseth, 2000; Hanseth and Lyytinen, 2004; 2010). Such a situation is said to occur when necessary changes and maintenance of the infrastructure are too costly, too hard to estimate, and ultimately impossible to pursue. Identifiers, which are used for uniquely referring to, or identifying, objects such as telephone numbers, license numbers, and personal identification (PID) numbers, play a major role in causing lock-in situations (Eriksson, 2003; Milne, 1997; Rood, 2000). The problems associated with identifiers concern:

1. the inclusion of descriptive information in identifiers,
2. the selection of an inappropriate identifier for a particular object, and
3. the lack of institutional control of the identifier.

In the paper we refer to these problems as (1) The Descriptive Identifier Problem, (2) The Identifier Selection Problem, and (3) The Identifier Control Problem. In information systems development (ISD), selection and design of identifiers have traditionally been seen primarily as a technical problem – for example, choosing a compact primary key in a relational database or generating an artificial Object ID in an object-oriented system. The design, selection, or choice is typically considered unproblematic as long as the identifier is unique, stable, and compact (Date, 2004, pp. 269–270). If there is no such suitable property or combination of properties (a natural key) that can be used for identification, we simply create an artificial attribute (a surrogate key)—internal to the information system (IS)—that represents the set of the identifying properties that actually identifies the thing in the real world (Evermann and Wand, 2006). However, this straightforward technical understanding of identifiers is problematic, because it does not reflect the institutional meaning of the identifier construct. Nor does it reflect the social context of its use, its ontological status, or its fundamental role for information infrastructures. The traditional database and ISD literature assumes a green-field situation in which a stand-alone system is developed without any constraints imposed by an already installed base and where the developer is in control of the result. This is rarely the design situation in contemporary systems development projects, which instead require an information infrastructure perspective (Hanseth and Lyytinen, 2010).

The objective of this paper is to elaborate on the problems associated with identifiers outlined above by exploring the meaning of the identifier construct from a technical, institutional, ontological, and information infrastructural perspective. Based on this exploration, we then provide guidelines for how identifiers should be designed, chosen, replaced, and controlled. The research approach adopted is that of conceptual-analytical research (Järvinen, 2000). We provide analytical and logical arguments based on real world examples of typical information infrastructure settings. The paper proceeds as follows: The following section elaborates on the role of identifiers in information infrastructures, and the three problems described above are exemplified by the descriptive Swedish PID number and how it is used as a student identifier. The next section shows how the identifier construct has been explained and discussed previously in the ISD literature. Based on this analysis, which reveals that the identifier is primarily considered only as a technical construct, the following section explores the identifier construct from an institutional and conceptual viewpoint based on Searle's (1969; 1995) speech act theory. This exploration suggests that the identifier construct has a very special meaning in relation to what is identified. We then elaborate on this view using a concrete example—the vehicle example. To illustrate the theoretical argument, we then provide and discuss solutions to the problems associated with the Swedish PID number and suggest guidelines for the design, choice, and institutional control of identifiers from an information infrastructure design perspective. Finally, we conclude by summarizing the main points.

2. The Role of Identifiers in Information Infrastructures

What role do identifiers play in information infrastructures? Certainly, this question can only be answered through understanding the nature of information infrastructures as such. According to Hanseth and Lyytinen (2010, p. 4), an information infrastructure is “recursively composed of other

infrastructures, platforms, and IT capabilities.” Identifiers and registers of identifiers constitute a naming infrastructure, which is an important part of the overall information infrastructure. Hanseth and Lyytinen (2004, p. 234) note that “Registers of such identifiers must be available to all users and they must be common to avoid mistakes and errors. This also requires building an additional information infrastructure to maintain and distribute updated versions of each register and a social and institutional process to define classification semantics.” There are many examples that point to the significance of identifiers for information infrastructures, including the naming infrastructure of the Internet and the European Article Number (EAN). To use the Internet, one must type an address into a computer—a name or a number. This address has to be unique, and the Internet Corporation for Assigned Names and Numbers (ICANN), the American Registry for Internet Numbers (ARIN), and other regional providers coordinate these unique identifiers across the world. The EAN number system is a bar-code standard that is used for identifying products and goods throughout Europe and is compliant with the Global Trade Item Numbers (GTIN) as well as with similar systems in, for example, the USA (UPC) and Japan (JAN). However, the identifiers and registers that constitute the naming infrastructure often prove hard to design and difficult to maintain. The problems associated with the Swedish PID number and student identifier described below are typical examples of problems that may occur in the naming infrastructure (cf. Monteiro, 1998).

2.1. The Problems Associated with the Swedish PID Number

The Swedish PID number is a sequence of ten digits and a punctuation mark – for example, 571129-8337 – where positions one to six represent a person’s day of birth (YYMMDD), seven to nine is a sequence number, and the tenth position is a check number. The ninth position is odd if the number identifies a man and even if it identifies a woman. The punctuation mark is a hyphen (-) when the person is less than 100 years old, and a plus sign (+) when the person is older. Interestingly, eleven of the 14 EU countries that use national PID numbers have descriptive information in their numbers (Otjacques et al., 2007).

The descriptive information contained in the PID number leads to several problems. The General Director of the National Tax Board of Sweden, Mats Sjöstrand (2007), described one of them; namely, that the number can no longer address all people living in Sweden. The number allows for 499 females and 500 males to be born each day, which so far is enough for all native-born Swedes. The problem is that immigrants also need a PID number, and many immigrants’ passports record the first of January or the first of July as their birthdates. This is because many immigrants may not know their exact birthdates. Thus, the custom in many countries is to use these dates as default values. This means that the Swedish PID numbers run out for these dates. It also creates a problem for immigrants who can prove that they were actually born on either of these dates, as they may not get their proper number.

The solution suggested by the National Tax Board in 2007 was to abandon the use of descriptive information in the identifier altogether. However, this would have had huge nationwide consequences, because the PID number is used virtually everywhere (see, e.g., the student identifier discussion below). A subsequent governmental report (SOU, 2008) estimates the cost of modifying PID numbers to the National Tax Board alone at approx. € 46,000,000 (\$ 56,000,000).

2.2. Problems Associated with the Student Identifier in Sweden

The IS used for registering students at Swedish universities uses the Swedish PID number as the identifier of students. This nationwide system is a very important part of the university information infrastructure in Sweden. However, problems occur, because of the large number of foreign students in Sweden who do not have a Swedish PID number. Universities assign temporary Swedish PID numbers to identify foreign students, putting a “T” in the seventh position of the number field. However, this means that there are only 49 female and 50 male PID numbers available for each date and, as a consequence, the numbers run out for certain dates.

Another problem is that there is no nationally managed routine for handing out temporary PID numbers to foreign students. As a consequence, a student can be given several temporary PID numbers should he or she happen to be enrolled in more than one university, and the same

temporary PID number could be given to several foreign students. This is a problem because there is a central system for accepting students into study programs in Sweden, and sometimes universities send the confirmation of acceptance into a program to the wrong student.

A third problem is that if a student eventually gets a proper PID number, the previously registered information about the student may be hard to locate.

The Swedish system using PID numbers as student identifiers also creates a demand among foreign students to be able to use their native PID numbers as student identifiers. For example, there are many Finnish students in Sweden, and it would make sense for them to use their Finnish PID numbers as identifiers, however, the Finnish PID number is 11 positions long and is formatted differently from the Swedish one. It has been estimated that changing the system so that Finnish PID numbers also can be used as nationwide student identifiers, and at the same time preparing the system for other PID formats by changing the field to 13 positions, will take approximately 9,700 hours. A working cost of 80 €/hour suggests a total cost of approx. € 776,000 (\$ 950,000).

2.3. The Need for A Deeper Understanding of The Identifier Construct

The examples above point to one conclusion, namely that the design, choice, management, and control of identifiers is not merely a trivial technical matter. On the contrary, it is a complex institutional and business matter, because identifiers have an important meaning and function in society. The examples also show that identifiers are an important part of the installed base (Hanseth and Lyytinen, 2010) of information infrastructures and that poorly designed identifiers cause lock-ins. In order to provide solutions to these problems, it will be useful first to review how the identifier construct has been discussed in the ISD literature to date.

3. The Identifier Construct in Information Systems Development

The role, importance, and robustness of identifiers have occupied centre stage in the relational database literature, where identifiers have been discussed in terms of keys. The key concepts in the relational model are mainly grounded in formal logic and mathematical set theory. The relational database literature embodies a semantic modeling perspective, and the purpose is to represent meaning, i.e., to describe (a part of) the world, in the database. The semantic modeling problem is underpinned by the following assumptions (Date, 2004, p. 411):

1. The world is made up of entities; i.e., distinguishable objects, such as individual vehicles.
2. Entities can be classified into entity types (e.g., vehicles). Entities of a given type have certain properties in common.
3. Every entity has a special property that serves to identify that entity (e.g., the license number)
4. Entities can be related to other entities by means of relationships.

In the relational model the database is represented as a collection of set-theoretical relations, or, more informally, as a number of tables (Elmasri and Navathe, 2004, p. 126). A row is called a tuple and a column header is called an attribute. The special property that serves to identify a tuple is called a key. There are candidate, primary, alternate, foreign, natural, and surrogate keys.

In a table T1, a candidate key is a set of attributes that makes it possible to uniquely identify a row in the table (Date, 2004, p. 269). If a table has more than one candidate key, one of these should be chosen as the primary key, and the other candidate keys are then considered alternate keys. A foreign key is a set of attribute values in T1 that matches the values of the primary key in some other table T2.

One problem is that keys that are commonly used for identifying entities in the world, so-called natural keys (e.g., the PID number of a Swedish resident), tend to be unstable (i.e., prone to change) because they carry attribute values (The Descriptive Identifier Problem). Identifier changes may concern both its value and structure. Another problem is that primary keys constituted by a set of attributes are often cumbersome to use. To resolve these problems, the idea of the “surrogate key” was introduced. Surrogate keys are keys in the usual relational sense, but they have some specific properties (Date, 2004, p. 434):

- They always involve exactly one attribute.

- Their values serve only as surrogates (hence the name) for the entities they stand for. They only represent the fact that the corresponding entities exist—they carry absolutely no additional information or meaning.
- When a new entity is inserted into the database, it is given a surrogate key value, generated internally by the system, that has never been used before and that will never be used again, even if the entity is subsequently deleted.

Two primary definitions of “surrogate key” appear in the literature, which indicates difficulties in interpreting the construct:

- (1) A surrogate key represents an object in the database itself. The surrogate is generated internally by the system and is invisible to the user or application (Codd, 1979; Wieringa and Jonge, 1991).
- (2) A surrogate represents an entity in the outside world. The surrogate is internally generated by the system but is nevertheless visible to the user or application (Hall, Owlett and Todd, 1976; Date, 2004, p. 434).

In both definitions, the surrogate key is considered meaningless, and the solution to the problems associated with natural keys is seen only as technical in nature. There is, for example, no discussion of why natural keys often have embedded attribute information, or why many natural keys are often constituted by a set of attributes, or how natural keys should be designed—only a technical solution is presented: use surrogate keys.

The choice of a primary key among two or more candidate keys (The Identifier Selection Problem) is considered to be important in order to maintain the integrity of the database (Codd, 1988; Date, 2004). However, an acceptable theoretical explanation for how to make the choice is lacking (Date, 2004, p. 272), and the choice is regarded as arbitrary, primarily based on simplicity, and outside the scope of the relational model (Codd, 1988; Date, 2004, p. 272; Elmasri and Navathe, 2004, p. 135). To quote Date (1995, p. 206), “The fact is, however, that the idea of singling out one particular candidate key for special treatment along the lines indicated has always been the source of some slight embarrassment to relational advocates (myself included). One of the strongest arguments in favor of the relational model has always been its claim to a theoretical foundation.”

To understand the problems described above, we have to comprehend the meaning of the identifier construct (the candidate key) and the nature of the entities that the candidate keys represent. The database literature, however, gives little guidance as to what an entity really is. Date (2004, p. 411) even tells us upfront: “we cannot state with any precision exactly what an entity is.” This implies that we have to look elsewhere for answers to the question of the meaning and role of identifiers and the nature of the objects they identify and represent.

Object orientation (e.g., Jacobson et al., 1994; Mathiassen et al., 2000), which has received considerable attention over the last 20 years, can be seen as a natural evolutionary extension of the semantic data models (Hirschheim et al., 1995, p. 193). However, identifiers do not receive the same attention in object-oriented modeling (e.g., Mathiassen et al., 2000, p. 110, 52) as they do in the relational model. All objects in an object-oriented system are assigned a unique object ID upon their creation (Cattell, 1994). Such identifiers are considered implementation specific (e.g., corresponding to memory addresses) and are typically not derived from the way the corresponding “real world” objects are identified, which means that they are surrogate keys according to definition 1. Wieringa and de Jonge (1991) consider object IDs meaningless, although they are still visible to the user, which implies that the identifier is seen as a surrogate key in accordance with definition 2.

Taken together, this means that neither the relational model nor object orientation can provide a satisfactory answer to the question of the meaning and role of identifiers and the nature of the objects they identify and represent. According to Evermann and Wand (2005, p. 147), it is unclear what the modeling constructs of object-oriented design languages mean in terms of the application domain. We agree, and perhaps we can find the answer by following their lead and consulting the so-called Bunge-Wand-Weber (BWW) ontology.

Rule 5 of the seven foundational rules for conceptual modeling (Wand et al., 1999, p. 512) states that

“All attributes and relationships in a class represent properties of things in the class,” which implies that “Identification attributes are not part of a conceptual model.” But, this “does not preclude the use of identifying attributes in systems design. Rather, it indicates that they have no significance in modelling the world” (Wand et al., 1999, p. 512). As Evermann and Wand (2005, p. 149) put it, “No two things have exactly the same set of individual properties. Thus, properties can be used to identify things.” Furthermore, they claim, “In an information system, we may not know or care about all properties and instead use artificial identification attributes to represent the set of identifying properties” (Evermann and Wand, 2005, p. 149). In other words, identifiers have no semantic meaning based on the real world, they are only constructs used for the implementation of the conceptual model in information systems, i.e., technical constructs (surrogates according to definition 1).

It is clear that the mainstream IS development literature (as represented by relational database theory, object orientation, and the BWW ontology) cannot really explain the meaning of the identifier construct and considers it primarily a technical matter. But if the identifier construct is just a technical construct of interest to those who implement computerized information systems, why has the identifier generated such an interest in the philosophy of language over recent centuries?

4. A Speech Act Theory Account of the Identifier Construct

The identifier construct has been thoroughly discussed in the philosophy of language under the label “proper names.” Mill (1872) claimed that an identifier has no meaning above and beyond the object to which it refers (its referent, or reference)—an identifier is an unmeaning mark. This is in line with how the identifier construct is viewed in relational database theory, object oriented modeling and the BWW ontology. However, then an identity statement such as “Mount Everest = Mount Everest” must mean the same as “Mount Everest = Chomolungma.” Yet clearly the latter can convey information in a way that the former cannot. Chomolungma, which is the Tibetan name, means “Mother of the Universe.” “Mount Everest,” the English name, clearly does not convey the same meaning. This example shows the difference between two aspects of meaning, namely sense and reference (Frege, 1892). The identifiers Mount Everest and Chomolungma have the same reference but not the same sense. Consequently, identifiers are not unmeaning marks.

Searle (1969) explains the meaning of the identifier construct (proper name) based on speech act theory. A speech act $F(p)$ consists of a propositional content p associated with an illocutionary force F . For example, the propositional content $\langle \text{car, red} \rangle$ can be used to state a fact (“the car is red”) or to invoke future action (“do paint the car red”). The first is an assertion (stating a fact) and the second is a performative (in this case, a request).

The two components of the speech act show the double structure of language and that the meaning of a speech act must be determined at two levels:

1. Semantic meaning: The function of the propositional content is to assign attributes to identifiable objects, for example, the attribute red to the object car.
2. Pragmatic meaning: The function of the illocutionary force is to convey the action type, that is, the intentions and social obligations that are attached to the speech act.

4.1. Reference Mechanisms

The propositional content of a speech act fulfils two important functions: the referring function and the predicating function. Referring means to pick out or identify an object, while predicating means characterizing or describing it. In order for referring to be successful, the following conditions must be fulfilled (Searle, 1969, p. 82):

- There must exist one and only one object to which the speaker’s utterance of the expression applies.
- The hearer must be given sufficient means to identify the object from the speaker’s utterance.

According to Searle (1969), we use language to refer in two ways:

1. By using *an identifier*, such as “Mount Everest.” An identifier can also be a number, such as the Swedish traffic vehicle license number “DCA096.”

2. By using *complex noun phrases in the singular*. These expressions are also called “definite descriptions,” for example, “the highest mountain in the world” or “the dented red car.”

This means that there are two types of reference mechanism that can be used: the definite description and the identifier. However, this does not mean that an identifier is the same as a definite description. In fact, trying to present a definite description of an object as the meaning of its identifier would lead to the peculiar consequence that the meaning of the identifier would change if there were any change at all in the object (Searle 1969, p. 166). For example, if the red car with the identifier DCA096 were repaired and repainted blue, the identifier “DCA096” would still refer to the same physical thing, but the definite clause “the dented red car” would not. This means that the two expressions only pick out the same thing in a specific use situation. If they had exactly the same meaning, then the identifier would have a different meaning depending on how the properties in the description change over time; thus, it would not fulfill its referential function, i.e., to represent the existence of the thing over time. It would also imply that we would only be able to refer to a thing by describing it. But this is precisely what the identifier construct avoids and what constitutes the distinction between identifiers and definite descriptions. Searle (1969, p. 172) writes, “The uniqueness and immense pragmatic convenience of proper names in our language lies precisely in the fact that they enable us to refer publicly to objects without being forced to raise issues and come to an agreement as to which descriptive characteristics exactly constitute the identity of the object.”

It is true that the identifier has a semantic meaning, as it has to be connected to attributes that represent properties of the physical thing, since we must be able to substitute the identifier for an identifying description if asked to do so in a certain context of use. For example, if someone asks you which car is “DCA096,” assuming that he or she cannot see the license plate, one could answer “the dented red car.” The important conclusion, however, is that the identifier has a meaning because it is used to refer to the thing without describing it and to represent the existence of a thing over different use contexts (Searle 1969, p. 78), but not to represent a set of identifying properties.

4.2. Institutional Objects

Realizing that identifiers are language constructs used to refer to and represent the existence of objects without having to describe them helps us understand their special meaning and function. However, to fully understand the meaning of the identifier construct, we must also have some idea about how objects exist in the world and the nature of these objects. For this purpose, the distinction between what Searle (1995, p. 27) terms “brute facts” and “institutional facts” is important:

- Brute facts exist independently of human institutions and concern physical (brute) things and their properties. They only rely on the institution of language so that the facts can be asserted; for example, the assertion “The car is red.”
- Institutional facts, on the other hand, require special human institutions for their very existence. For example, performing the declaration, “You are now the owner of that car” presupposes the existence of both language and the human institution of ownership.

In this context, “institution” refers to structures of rules governing cooperative human behavior. Performing speech acts using language implies following and creating such rules (Searle, 1969, pp. 33–34). When language philosophers have discussed performatives (e.g., orders, promises, and declarations), a critical insight is that not only are language constructs employed to describe reality as it is, but also that using language additionally implies constructing social reality (Searle, 1995). Note that the truth of the propositional content (p) of a speech act is only asserted in the special class of speech act called assertions, while performatives bring things about or describe things that ought to be brought about. Performatives such as orders, promises, and declarations create institutional facts, which suggests that speech acts, as well as their propositional contents, are *bona fide* objects (cf. Graham, 1998). This paper focuses on explicit declarative speech acts, such as when an authorized representative of a highways administration authority states, “I declare this red car to be the traffic vehicle with the license number DCA096.” We use the term “institutional object” when referring to this type of object in the paper. An institutional object, as we define it:

1. must be identifiable by some reference mechanism (either by a definite description or by an identifier);

2. is created via a language act at a certain point in time (we prefer the term “language act” to “speech act” because these acts can be performed by different types of media, including IT systems, as exemplified in the next section);
3. is created based on institutional rules that must be followed in order for the object to be valid;
4. is an instantiation of a concept;
5. is something that is referred to in a social context of use;
6. is by itself an institutional fact;
7. represents brute and/or institutional facts;
8. conveys both semantic and pragmatic meaning.

Identifiers are used for representing the existence of and referring to institutional objects. This means that they convey both semantic and pragmatic meaning. Sometimes, but by no means always, the identifier also represents the existence of a physical object (thing). Attributes are used for describing institutional objects and properties of brute facts.

To select the most suitable reference mechanism for an institutional object, we must first analyze whether a definite description or an identifier is used to refer to it. If a definite description is employed, we should consider introducing an identifier to replace it. If a descriptive identifier is used, we should consider redesigning it. If there are candidate keys, we should analyze which key should be chosen as the identifier of the object by analyzing the institutional context. This means analyzing the rules that govern the creation and existence of institutional objects. This may include, but is not restricted to, an analysis of the correspondence relationship between institutional and physical objects (things). Based on such analyses, the semantic and pragmatic meanings of the identifier can be understood and an informed choice be made.

5. The Vehicle Example

In order to exemplify the ontological foundation set out above and its implications for ISD, we will elaborate the vehicle example in this section. We will specifically refer to the Swedish Central Vehicle Registration system (CBR), which maintains information about registered traffic vehicles and their license numbers—an important part of the naming infrastructure of the overall Swedish information infrastructure.

Table 1: Traffic Vehicle table

License number	ISO-VIN	Type of vehicle	Model	Model year	Date of registration	Color	Width	Number of passengers	Engine power(BHP)	Total weight	Tax weight
DCA096	AAABC9LXP1000001	Private light goods vehicle	SAAB 900 2, 3I DX55B	1996	1996-03-14	Red	1.72	4	110	1830	1450

A vehicle object would typically be represented in a relational table as shown in Table 1. Here, we can observe that there are two candidate keys (identifiers that can be used to uniquely identify a physical vehicle): the chassis serial number (the ISO-VIN code), and the license number.

The ISO-VIN is used to identify motor vehicles, trailers, motorcycles, and mopeds and is based on an international standard (ISO 3779:1983). The ISO-VIN consists of 17 positions, divided into three sections. The World Manufacturer Identifier (WMI), in this case “AAA,” occupies the first three positions and uniquely identifies the maker of the vehicle. The Vehicle Description Section (VDS) occupies positions four to nine and may be used by the manufacturer for ascribing to the vehicle attributes such as body style, engine type, etc. In this case (Table 1), “BC99LX” describes the properties body style, gearbox, and engine. The Vehicle Identifier Section (VIS), “P1000001,” is used for the identification of a specific physical vehicle and occupies the last eight positions.

The question, then, is which identifier should be chosen as the primary key. According to the database literature, the choice is arbitrary and should be based on simplicity. As the Swedish license number consists of three letters (A–Z) followed by three digits (0–9) it is clearly simpler than the ISO-VIN. In fact, the ISO-VIN is not a well-designed construct. This is specifically true for the VDS section, which includes descriptive information and makes the ISO-VIN potentially unstable. Notably, the WMI and VIS sections are the only two components that are needed to make it a unique identifier worldwide.

However, a relevant question to ask at this point is: Why are there two identifiers, given that there is only one physical vehicle? As discussed above, relational database theory does not provide an answer to this question. In order to do so, the institutional context, how institutional objects are created, and their relationship to the physical vehicle have to be analyzed.

The physical vehicle (the thing) is typically created in a vehicle factory through material actions such as welding and assembling parts. The manufacturer creates the institutional object identified by the ISO-VIN code using a declarative language act. Apparently, the ISO-VIN is not just an identifier, since it is used for both identifying and describing the vehicle; it is, in fact, a complete institutional object. Although the ISO-VIN is also attached to the vehicle, this imprint is not a property of the physical vehicle; it is a representation of the institutional object. In many countries, the original act is performed by the use of an information system, and the resultant genuine and valid institutional object is stored in the database of the car manufacturer, or in a national VIN database (which is the case in, for example, Australia). A typical representation and instantiation of this institutional object in a relational database is shown in Tables 2 and 3.

Table 2: Manufactured Vehicle Registration Act table

Action ID	Date of action	Action type	VIS
1123989899	1996-02-29	Registration	P1000001
1127989878	2003-06-14	Deregistration	P1000001

Table 3: Manufactured Vehicle table

WMI	VDS	VIS
AAA	BC99LX	P1000001

Here the action ID in the Manufactured Vehicle Registration Act table (Table 2) identifies the declarative act (an institutional object), the action type is the illocutionary force of the object, and the date of the action shows when it was performed. The proposition, thus, consists of the ISO-VIN, which identifies and describes the institutional object (ManufacturedVehicle), which in turn consists of the ISO-VIN parts: WMI, VDS and VIS.

The semantic meaning of the ManufacturedVehicle object is that it is true that there exists a uniquely identifiable physical vehicle having the properties that correspond to the attribute values represented by the VDS column. The pragmatic meaning should be understood as a commitment made by the manufacturer that this is a genuine vehicle, authorized by the manufacturer. This is of great importance for a number of reasons. For example, the information in the Manufactured Vehicle table (Table 3) together with the ISO-VIN stamped onto the vehicle can be used to make it more difficult for criminals to attach an ISO-VIN from a wrecked car onto a stolen car in order to “re-birth” it, as selling a car with an ISO-VIN that has been reported as stolen is almost impossible. It is, therefore, important

to keep a record of another type of declarative registration act, namely, deregistration of the ISO-VIN numbers of wrecked cars, thus invalidating these institutional objects.

Finally, the institutional object Traffic Vehicle is created using the CBR at the Swedish Road Administration (SRA). Since such registration acts are important declarative language acts, we need to complement the information stored in the Traffic Vehicle table (Table 1) with an explicit Traffic Vehicle Registration Act table (Table 4) and an Owner table (Table 5) in order to represent the complete act of registering (and deregistering) a traffic vehicle. The reason is that ownership can only be legally established (in Sweden at least) between a registered vehicle with a license number and an owner with a PID number.

It is important to see that there is an explicit identifier identifying the registration act as such, represented by the Action ID in the Traffic Vehicle Registration Act table (Table 4). The registration act instantiates itself and the objects Owner and Traffic Vehicle. It is the authority that is responsible for the system (in this case the SRA) that creates these institutional objects through the CBR. The physical things outside the system are a physical person and a vehicle, and the act must correspond to this reality. The relationship between the traffic vehicle and owner objects is represented by the license number and owner identity in the Traffic Vehicle Registration Act table (Table 4), because the propositions of the act (as represented by the foreign keys License No and Owner) refer to the Traffic Vehicle and the Owner objects. The Registration Act table is also used for performing and storing information about the deregistration of traffic vehicles (see Table 4).

Table 4: Traffic Vehicle Registration Act table

Action No	Action Date	Action Type	License No	Owner
11234768	1996-02-29	Registration	DCA096	5711298437
12647688	2005-05-19	Deregistration	DCA096	5711298437

Table 5: Owner table

PID Number	Name	Street Address	City	Zip Code
571129-8437	Lars Eriksson	20 Big Street	Stockholm	181 88

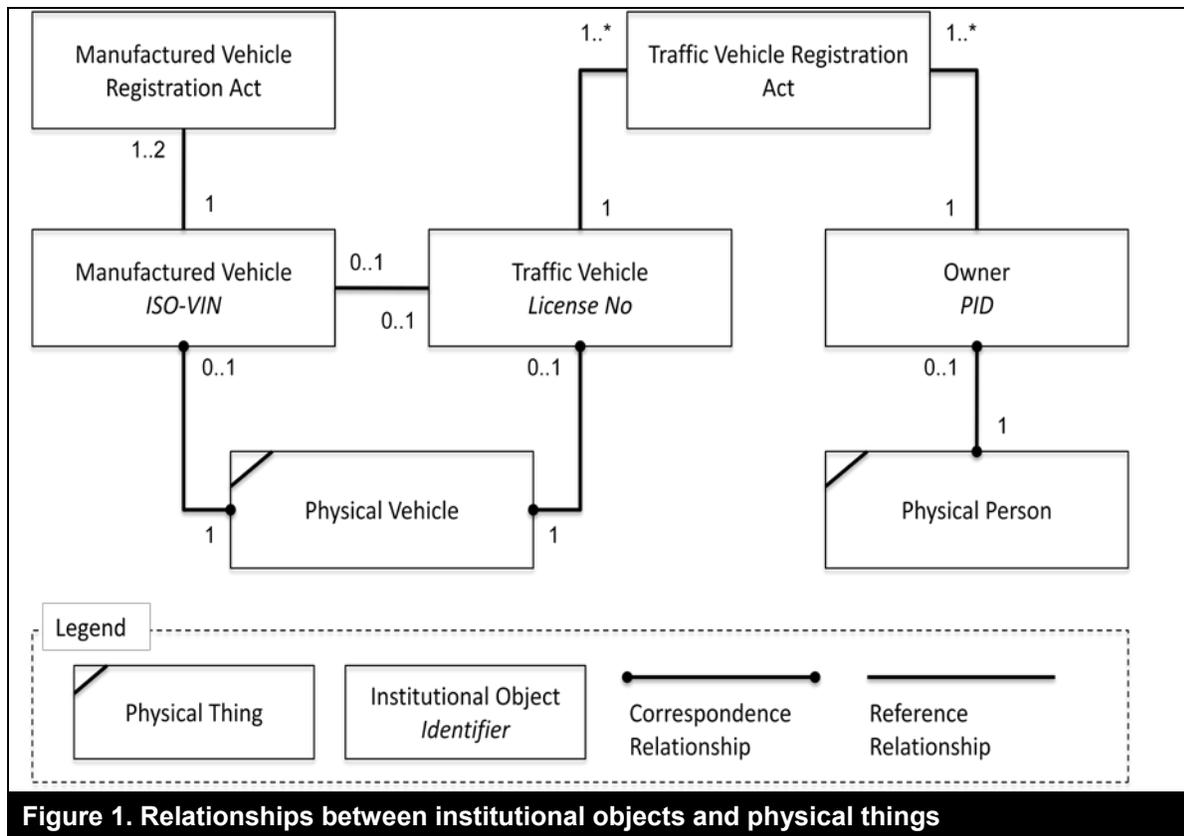
All these institutional objects are created more or less automatically by use of the CBR at the SRA, following strict rules. The registration act has thereby been performed under the authority of the SRA and, therefore, these objects carry not only a semantic meaning but also a pragmatic one related to responsibilities, rights, and obligations. Declaring something to be a traffic vehicle means that the vehicle is allowed for use on public roads, and being an owner of a traffic vehicle carries certain obligations, such as paying road tax.

Following the speech act theory account set out above, the answer to the question of why we need two identifiers of one physical vehicle is that we are actually dealing with two vehicle institutional objects and one physical vehicle, which are interrelated. The identifiers and the institutional objects they represent convey different meanings to the physical thing—similar to the “Chomolungma” and “Mount Everest” example above. It is obvious that this is so, because if someone asks how many Swedish vehicles there are, one must ask that person to be more precise. If the question is how many registered traffic vehicles there are in Sweden right now, then the answer can be given by counting all currently registered traffic vehicles in the CBR. But if the question concerns all vehicles that have been manufactured by a Swedish manufacturer, the answer can be given by counting all currently registered Manufactured Vehicles that have been produced by a Swedish manufacturer. Thus, it is the cardinalities of the two respective types of institutional object that are counted, and one must count them differently based on the differing rules that make them valid, so that neither are the objects the same nor can one of them be seen as a subtype of the other.

The relationships between the institutional objects and the physical things are illustrated in Figure 1 using a UML-like notation that also indicates the distinction between reference and correspondence relationships. The relationships between the institutional objects (Manufactured Vehicle, Traffic Vehicle and Owner) and the corresponding physical things are correspondence relationships. A

correspondence relationship means that the institutional object refers to the physical thing and, at the same time, the attributes of the VDS section of the ISO-VIN, for example, must match the corresponding properties of the physical vehicle. Similarly, the attributes specified for the Traffic Vehicle object must match a selection of properties of the physical vehicle. For example, type of body (estate) and color (red) are attributes typically used for that purpose. However, it is important to understand that these attribute values do not represent all those properties of the physical vehicle that ensure its uniqueness, because this is not the function of attributes. Several physical vehicles may share attribute values in the Traffic Vehicle table. These attributes describe properties of the physical thing and can be used to provide an identifying description in a certain context of use, but they do not represent the unique existence of the individual physical thing itself. It is only the ISO-VIN and the license numbers that represent the unique existence of the individual physical thing, but they are not properties (brute facts) of the physical vehicle. They are institutional constructs and identifiers of institutional objects (representing brute and institutional facts) that have been ascribed to the physical vehicle by declarative speech acts. An identifier of a physical thing can only be an identifier if there is a genuine difference between it and the physical thing identified (Searle, 1969, p. 75).

Figure 1 also shows that for each Physical Vehicle there is zero or one Manufactured Vehicle, because there are physical vehicles that do not have an ISO-VIN—either the physical vehicle has the institutional status of being an authorized manufactured (contemporary) vehicle, or it does not. The same holds for Traffic Vehicle, since there are physical vehicles that do not have a license number. This also means that the functionality of the relationship between the two institutional objects is 0..1 in the direction from Manufactured Vehicle to Traffic Vehicle, and vice versa. However, the relationship between the two institutional objects is not a correspondence relationship; it is a reference relationship. The reason is that two institutional objects do not represent each other, which means that they are not synonyms (cf. the Chomolungma and Mount Everest example above). Specifically, the ISO-VIN code does not represent the existence of the Traffic Vehicle object in the same way as it represents the existence of the physical thing (and vice versa). The relationship between the two institutional objects is an institutional relationship. This is not recognized in relational database theory. The only thing that is recognized is that the two identifiers refer to the same physical thing (the two correspondence relationships are in focus) and, hence, it can be maintained that the ISO-VIN and license number are candidate keys. As a consequence, the choice between them becomes merely a technical matter, since they are assumed to be interchangeable.



However, if we recognize the distinction between physical things (brute facts) and institutional objects (institutional facts), it becomes clear why two identifiers are needed. The simple answer is that they stand for two different institutional objects that have different semantic and pragmatic meanings. If we had the ISO-VIN number and the Manufactured Vehicle object only, we would have to give vehicle manufacturers the right to declare which cars should be allowed for use on public roads and who is to pay the road tax. This is clearly not the responsibility or interest of vehicle manufacturers—it is the responsibility and interest of the highway authorities in each jurisdiction.

This also means that the important decision as to which identifier to choose is not arbitrary and cannot be based on simplicity alone. Such important choices should be based on analyses of the institutional context, or domain, for which the identifier was originally designed. This is required in order to recognize the institutional objects that the identifiers represent and the constitutive rules that govern their existence and validity; that is, to understand the identifiers' institutional meaning.

6. Discussion

We are now in a position where the three problems described in the introduction can be discussed in a new light. To recapitulate, the problems concern:

1. Identifiers that include descriptive information (The Descriptive Identifier Problem).
2. The selection of inappropriate identifiers for objects (The Identifier Selection Problem).
3. The lack of institutional control of identifiers (The Identifier Control Problem).

In what follows, we address each of these in turn and conclude with a discussion of concrete implications for the design of information infrastructures.

6.1. The Descriptive Identifier Problem

We exemplified the Descriptive Identifier Problem with the embedding of descriptive information in the Swedish PID number. According to the traditional ISD literature, an unstable “natural key” is no major problem, since we can instead use a “meaningless” number (a surrogate key) generated by the IS.

But this is not a viable solution to the problem with the descriptive Swedish PID number. Implementing a surrogate key in the installed base in Sweden would be extremely costly; the cost to the National Tax Board alone is estimated at approx. € 46,000,000 (\$ 56,000,000).

The solution suggested by the Swedish government report (SOU, 2008) is to redesign the natural key, taking its use, institutional meaning, and installed base into account by replacing the birthdate (positions five to six) with a number in the range one to 31, or if the month of birth has fewer than 31 days, the number of days in that month. The reason for not allowing numbers higher than 31 is that many systems do not allow larger numbers in positions five to six of the PID number. This solution is much cheaper than introducing a surrogate key, but is still costly. The cost estimate for the National Tax Board alone is approx. € 4,100,000 (\$ 5,000,000). This includes adjusting systems to store birthdates in separate fields, since the correct birthdate can no longer be inferred from the identifier. The major drawback of this solution is that people recognize their date of birth as consisting of year, month and day (in that order), which may make the new number difficult to understand and accept. An alternative solution would be to abandon the check digit, which would increase the available numbers for each day by a factor of 10. However, this solution has been rejected, because the check digit is important for avoiding typing mistakes, as the PID number is often entered manually into various systems. The final decision made by the Swedish government in December 2008 was to allow for an adjacent day to be used in positions five to six if no number is available for the person's actual date of birth. This is in line with the suggestion proposed by the above-mentioned government report.

There are two main problems associated with the surrogate key concept. First, the surrogate key idea assumes there is not an already installed base of information systems and databases using the unstable natural key as the primary key. This assumption gives the false impression that the natural key can easily be substituted. Second, the surrogate key idea is questionable even if we assume a green-field design situation, because it does not consider the identifier's institutional meaning and social context of use.

Two conflicting definitions of the surrogate key concept exist. Both definitions assume that the surrogate key has no meaning, because it includes no descriptive information, but the first definition assumes that the surrogate is invisible to the user and the other assumes that it is not. However, the surrogate key concept is questionable no matter which of the two definitions is adopted.

The first definition is based on Codd (1979, p. 410): "Database users may cause the system to generate or delete a surrogate, but they have no control over its value, nor is its value ever displayed to them." This is the interpretation used by most practitioners who have discussed the pros and cons of surrogate and natural keys (e.g., Ambler, 2010; Berkus, 2010; Celko, 1999; 2007; Richardson, 2007). The advantages of using surrogate keys put forward in this debate are that they are stable and compact and that the database systems are in control of them, since they are generated and maintained internally. The disadvantage is that now two keys have to be maintained within the information system, which means that;

- the uniqueness constraint on the natural key still has to be maintained within the system,
- the facility to perform searches using the natural key must still be provided,
- a strategy for generating the surrogate key internally has to be chosen,
- an additional index has to be maintained within the database.

The idea is that a table in the database stores both the natural and the surrogate key and that the surrogate key substitutes for the natural key as the foreign key in all other tables that reference the object. The advantage is related primarily to internal database considerations; for example, it allows for faster operations on a relational database and allows for easy localization of required database changes due to a change in the natural key. Essentially, only the table that holds the natural key has to be changed rather than every single table where the natural key might appear as a foreign key. Nonetheless, a better approach would be to redesign the natural key if it is an important identifier. The reason for this is that the major costs that occur when important natural keys have to be changed are those that are incurred when changes have to be made to check routines in user interfaces, in programs that search and sort using the natural key, and in formatting the identifier when it is shown to the user. These costs cannot be avoided if this interpretation of the surrogate key idea is assumed.

If instead we consider the second definition as advocated by Date (2004, p. 434), who claims that the surrogate key is a model concept not to be concealed from the user (although it carries absolutely no additional information or meaning), a general design guideline for identifiers should be to always use a completely meaningless number or string of tokens as an identifier. This conclusion, however, is questionable for a number of reasons.

One reason why it is not always appropriate to use a surrogate key according to definition 2 relates to how it is used in business and society. For example, the problem with the Swedish PID number is only partly that it includes date of birth. The date of birth makes the number easier to learn and remember. In fact, most of the national PID numbers used worldwide include date of birth. Nor is the problem that the number has a check digit, because the check digit is important for checking the validity of the number when it is typed in.

Rather, the major problem is that it should have been 11 rather than 10 positions long and that the number includes information about gender, which does not make the number easier to remember. It only makes the PID number unstable, more difficult to change, and is also a source of integrity problems.

Another reason why a surrogate key, using definition 2, cannot always be used as the primary key is that some identifiers have such an important meaning in society that they have to be recognizable. This is why identifiers are typically given a certain pattern (e.g., PID numbers, employee numbers, student numbers, organisation numbers, credit card numbers, and bank account numbers). To give identifiers recognizable patterns, which means that certain positions in the identifier have restrictions on possible digits and tokens, makes them distinguishable, so that, for example, a PID number cannot be confused with an employee number or organization number. Combining the date of birth with a sequential number (as is done in most countries) is one way of creating such a pattern for PID numbers. Combining the number with an identifier of the issuing authority (as in the ISO-VIN) is another way of making the identifier more recognizable. It is also of interest to notice that the identifier does not have to include descriptive information in order to be given a meaningful pattern. The license number in Sweden, where the first three positions consist of letters followed by three digits, is a good example of that.

However, this does not mean that one cannot use a simple number or randomly chosen string of tokens as an identifier. On the contrary, this may be a viable solution in many cases. For example, when the identifier as such is less important to the business and its social context of use, and manual use is limited. There is, for example, a big difference in this respect between a sequential order number generated by an IT system and a PID number. However, it is important to note that even if the identifier is a sequential number, when used in a social context, it has a meaning that depends on the institutional object it represents. This is not recognized in the definition of surrogates where the identifier is considered to have no meaning if it has no descriptive content. There are also technical arguments in favor of using simple numbers as identifiers; for example, they are compact. However, we shall not elaborate further on technical design guidelines, as they are well covered in the existing relational database literature (e.g., Date 2004) and in the discussion among practitioners on the Internet (e.g., Celko, 1999; Richardson, 2007).

Our conclusion is that the surrogate key idea should be abandoned because of its inherent ambiguity, its lack of information infrastructure perspective, and its embodiment of the idea that an identifier is only a meaningless technical construct. This implies that the surrogate key idea oversimplifies the design problem by reducing it to a mere technical problem. Instead, design principles that take technical, usage, institutional, and infrastructural aspects into account should be favored.

First of all, an identifier must fulfill the basic criteria to be able to represent the existence of all the institutional objects it is intended to represent. There are also a number of further criteria (see below) that have to be considered, and these criteria have to be balanced when the identifier is designed:

1. The identifier should be stable, which means that neither its structure nor its value should have to be changed after it has been assigned to the institutional object.
2. If manual use of the identifier is extensive, it should include a check-digit.

3. If the identifier is visible to users either on-screen or through other media, consider giving it a pattern. If a pattern is used, embed minimal descriptive information. Information embedded in the identifier should not be used as attribute values by database administrators and programmers; for example, although date of birth is included in the Swedish PID number, the date of birth should be stored as a separate field in the database, which should be used to determine the exact date of birth.
4. If users have to learn and remember the identifier, consider making it mnemonic. This could be done by using a non-descriptive name, making the identifier as short as possible, or giving it a pattern (as shown above).
5. If the identifier is used extensively to exchange information between different IT systems, departments, and organizations and in society at large, consider giving it a pattern. If it is important, it has to be recognizable. Giving the identifier a pattern is also a way of assigning its uniqueness; for example, include the code of the issuing authority in the identifier.
6. The design of the identifier is also affected by technical concerns. It should be possible to build effective indices on the identifier field, and the identifier should preferably be represented by one column in the database.
7. In a redesign situation, the implications for the installed base have to be considered and a transition strategy should be developed, which will likely affect the new design.

6.2. The Identifier Selection Problem

The Identifier Selection Problem concerns the choice of primary key from amongst a number of candidate keys and was exemplified by the choice of the Swedish PID number as the identifier of students. The example shows that choosing the right identifier is crucial. Arguably, such choices should be based on the identifiers' institutional meaning. It is important to recognize that the chosen student identifier, the PID number, was originally designed for identifying the existence of an instance of the institutional object "Resident of Sweden," not "Student of a Swedish university." In order to make an informed choice in this type of situation it is important to recognize that we are dealing with two distinct and important institutional objects that sometimes, but not always, correspond to the same physical thing (analogous to the vehicle example). By analyzing the institutional context, we learn that the rules that govern the enrollment of students at Swedish universities have never required that they be Swedish residents. A special student identifier should have been designed and used when the system was originally developed (back in the 1980s).

However, the special student identifier is not the same as a surrogate key (assuming definition 2). The student identifier should not be considered a surrogate even if it were, for example, a random number. One problem with the distinction between natural and surrogate keys is that it gives the false impression that there is a special set of identifiers (natural keys) just because they include descriptive information. The only thing that makes, for example, the Swedish PID number seem natural to native Swedes is that it is given to the person very soon after birth and that Swedes are familiar with using it. What is considered a natural key is clearly as "artificial" as any other identifier, and its meaning does not depend completely on whether it includes descriptive information or not. Rather, it depends on which institutional object it represents, how it is used and acknowledged, and the rules that govern its validity, which may or may not prescribe that descriptive information should be included in the identifier.

However, one major problem in the current situation is that the cost of replacing the Swedish PID number with a special student identifier will be high, approx. € 776,000 (\$950,000). The major costs of changing the number are related to work that has to be done in order to change how the identifier is presented on 900 screens (76 percent of the cost), change programs that use the PID number for sorting (16 percent of the cost), make changes to the database (7 percent of the cost) and make other changes (2 percent of the cost). It is of interest to note that only 7 percent of this cost is related to internal database changes, which is the only cost that could have been avoided if the surrogate key idea (assuming definition 1) had been used when the system was originally designed.

In order to replace the Swedish PID number with a special student identifier and lower the associated costs, a transition strategy should be developed (cf. Hanseth and Lyytinen, 2010). Such a strategy should involve the idea that students already registered can keep their identifiers while new students

receive the new identifier, and the pattern of the new identifiers should be adjusted to the old one. This could be done by formatting the new student identifier similarly to the Swedish PID number; i.e., YYMMDD-XXX9. The difference compared to the Swedish PID number is that positions (seven through nine) could be any alphanumeric character. The tenth position would still be a check digit. A valid number could then look like this: 571129-BA59.

Adjusting the student identifier to the pattern of the Swedish PID number would reduce the cost of reformatting all the fields that show the PID number on screen-viewed documents as well as the costs incurred by changing the computer programs that sort on the Swedish PID number, which constitute the major costs. This design of the new identifier would also ensure that there are enough numbers for foreign students, and the number would also be quite easy to remember and recognize.

A service should also be provided to help students find their student identifier by using another identifier that they are familiar with, for example, their native PID number. This identifier should then be stored together with the student identifier in the student table, see Tables 6–8.

Table 6: Student Registration Act table

Action ID	Date of action	Action type	Student ID
1123989899	2008-08-29	Registration	731129-BA59
1123989900	2008-08-29	Registration	750330-KL08
1123989901	2008-08-29	Registration	770227-LL09
1127989878	2009-06-14	Deregistration	731129-BA59

Table 7: Student table

Student ID	Name	Surname	Date of birth	Nationality	Native PID
731129-BA59	Sven	Eklund	19731129	Swedish	731129-8637
750330-KL08	Kalle	Koivo	19750330	Finnish	300375-8916
770227-LL09	Susanne	Raaby	19770227	Danish	270277-4538

The proposed solution of designing a special student identifier also eliminates the privileged position of the Swedish PID number in the Swedish university information infrastructure. Indeed, it is the foreign students that suffer the most from the mistake of making the Swedish PID number the student identifier.

Table 8: Student Enrolment table

Student ID	University
770227-LL09	Uppsala
750330-KL08	Lund
731129-BA59	Stockholm

An important lesson to be learned from this case is that the PID number should not be used as a primary key excessively, as is the case at Swedish universities today (as well as in other areas of Swedish society).

6.3. The Identifier Control Problem

The Identifier Control Problem concerns the lack of institutional control of identifiers, something that is rarely touched upon in the ISD literature (Hanseth and Lyytinen (2004; 2010) being notable exceptions). Identifiers have to be established at an institutional level and often involve a great deal of political complexity. To describe the identifier as a surrogate of interest only in the implementation of an information system conceals this complexity. This means that a standardized and authorized system for the assignment and control of identifiers is a basic condition for the efficient exchange of information between information systems, organizations, and society at large, for example, between the Swedish universities and the Swedish_National Agency for Services to Universities and University Colleges (VHS).

In the case of the Swedish student identifier, there is a most urgent need for a standardized routine

that ensures that the same physical person cannot be assigned two different temporal PID numbers and that two physical persons cannot be assigned the same temporal PID number. The first step would be to introduce a control in the Swedish admission system that generates temporal PID numbers, an issue that has to be dealt with at the right institutional level, that is, at inter-university level. So far, system developers and database administrators who do not have the institutional authority to implement the required routine have primarily driven the system design. One of the most important challenges is to know when something requires an institutional and organizational solution and not just a technical fix (Edwards et al., 2007). In order for a standardized routine to be implemented, the administrative management at the universities and VHS must be made aware that this is primarily an institutional issue. It would be a first step toward developing an authorized institutional control system that can be used for the creation, maintenance, and verification of student identifiers.

The second step would be to make a decision at the inter-university level not to use the Swedish PID number or the temporal PID numbers as student identifiers. Instead, a separate student identifier should be introduced. As there is no standard that can be used, a specification and rules for the control system have to be developed. This second step would give the educational authorities full institutional control of the identifier. This is important because there has to be a trusted source that creates, maintains, and verifies the identifier.

6.4. Design Principles for Information Infrastructures

The examples provided above show the fundamental role identifiers play in information infrastructures and how important it is to design them appropriately in order to avoid lock-ins. The solutions proposed with regard to the problems above are all examples of how to avoid technology lock-ins in information infrastructure design (Hanseth, 2000; Hanseth and Lyytinen, 2010). Based on these solutions, the following design principles can be identified:

- 1) The first key principle is to design identifiers based on technical, usage, institutional, and information infrastructural aspects. For redesign and replacement of poorly designed identifiers, a transition strategy has to be developed that may affect the design of the identifier.
- 2) The second key principle is to choose the appropriate identifier or identifiers for important institutional objects and not to overuse identifiers that seem interchangeable. This is a way of dividing infrastructures into sub-infrastructures (cf. Hanseth and Lyytinen, 2010), which makes them easier to change and maintain.
- 3) The third key principle is to develop a standardized specification and an authorized institutional control system that can be used for the creation, maintenance, and verification of identifiers. This can be accomplished by developing standards that contain rules and specifications that can be used as the basis for the control system. However, it is important not to overuse already established standards and identifiers (see item 2 above). If there is no standard that can be used, specifications have to be developed. A decision must also be made about who should create and maintain the registers of identifiers and how they should be distributed. Establishing the institutional control system must involve people with both technical and business competence as well as sufficient authority.

In order to use these design principles it is important:

- to understand the function and meaning of the identifier construct as well as the vital difference between identifiers and definite descriptions;
- to understand the idea of the institutional object and how institutional objects are related to physical objects;
- to analyze the rules that govern how institutional objects are created and exist.

7. Conclusion

In this paper we have provided principles for the design of identifiers in order to avoid lock-in situations, inefficiency, and quality problems in information infrastructures. The proposed set of guidelines is an important step toward addressing “the difficulty of translating vivid empirical descriptions of IIs [information infrastructures] evolution into effective *socio-technical* design principles that promote their evolution, growth and complexity coordination” (Hanseth and Lyytinen 2010, p. 2,

emphasis as in original). The information infrastructure perspective is quite different from the traditional perspective touted in the database design and ISD literature, in which a green-field situation is assumed, where the IT artifact is designed for a specific task, for a limited set of users, and with no requirements imposed by an already installed base.

The guidelines are based on an alternative ontological commitment compared to what is commonly proposed in the database design and ISD literature. One basic assumption within the fields of entity-relationship modeling, object orientation, and BWW ontology is that information systems store representations of the state of affairs in the real world outside the system (e.g., Date, 2004; Jacobson et al., 1994; Evermann and Wand, 2005). Wand and Wang (1996, p. 88) capture this view quite succinctly in their “representation assumption”: “An information system is a representation of a real-world system as perceived by users.” Basically, it is assumed that IT systems represent brute facts and that a sharp distinction can be made between the IT system and its internal representations on the one hand and the outside “real world” on the other. An ontological outlook, as presented in this paper, that recognizes the idea of institutional objects that represent brute as well as institutional facts would provide another view. This ontological position implies that we are not restricted to what exists in the physical world, such as physical cars and people. In addition to such physical things, it is important to take into account the existence of institutional objects (such as traffic vehicles and owners), which are actually instantiated by the use of IT systems, as well as the relationships between physical things and such institutional objects. This leads to the important conclusion that the demarcation between IT systems on the one hand and the business domain (the real world) on the other should be abandoned.

Certainly, from a semiotic point of view, institutional objects are signs and as such exist also physically in some form (Stamper, 2001). However, this should not be understood as that an institutional object relates to a physical thing in the problem domain in the same sense that a spoken word relates to the sound waves that carry it. The physical manifestation of institutional objects is typically in the form of stored data in an information system. The data represents institutional objects, which, in turn, represent socially constructed facts and obligations related to some physical or social reality. Sometimes, however, also physical imprints outside the system represent institutional objects, for example, the ISO-VIN attached to a vehicle or BAR codes on goods. This implies that the definition of information infrastructure must include identifiers regardless of whether or not they are used within or outside IT systems. Attaching identifiers to physical things will also increase with the development of the so-called Internet of things (ITU, 2005), where RFID technology and sensor networks will be used to connect things to the Internet. This suggests that proper design of identifiers and understanding the distinction between physical things and institutional objects will be even more important in the future.

In this paper, we have provided guidelines for how identifiers should be designed. Our position is that the concept of the surrogate key should be abandoned and replaced by the conscious design of identifiers based on technical, usage, institutional, and information infrastructural aspects.

We have also presented a theoretical explanation of why there are candidate keys, along with guidelines for how to select from amongst them. This problem is considered to be of great importance in relational database theory but has, to date, not been given a satisfactory theoretical explanation. We have also showed how important this choice is for the design and subdivision of information infrastructures.

Finally, we have explained why institutional control of important identifiers is fundamental to the quality and efficiency of information infrastructures and that problems with identifiers in many cases have to be solved at the institutional level. This is important, since the design of information infrastructures is a highly political endeavor, where “[c]onflict and multiplicity are often buried beneath layers of obscure representation” (Bowker and Leigh Star, 1999, p. 47), and since identifiers and institutional objects form critical boundary objects that mediate between different communities of practice and enable shared social worlds.

Similar to Iannacci (2010, p. 2), our approach to information infrastructures and institutional facts

relies on a Searlian approach that “stresses the importance of constitutive rules in the process of institutionalization in order to dissect the very background where information infrastructures are cast.” It can be seen to rest on the cultural-cognitive pillar of institutional theory (Scott, 2001; 2003) in which institutions are conceived of as systems of constitutive rules (Searle, 1995) that govern the existence of social facts and validity of institutional objects. “Attention is directed to the shared conceptions that constitute the nature of social reality and provide the symbolic frames that support social sense-making” (Scott, 2003).

The concepts discussed in this paper open up an exciting new research agenda—in line with, yet challenging, Wand and Weber’s (2002) suggestions for research opportunities on conceptual modeling (cf. Ågerfalk, 2010). Views similar to those presented in this paper have been suggested in the past by, for example, Goldkuhl and Lyytinen (1982) and Winograd and Flores (1986), and even as “a new foundation for design” (cf. Weigand, 2006). However, when speech act theory has been used in IS research it has focused so heavily on the illocutionary aspect of language that the propositional aspect has been almost completely neglected (Ågerfalk and Eriksson, 2004). This unfortunate state of affairs is, not surprisingly, reflected also by Wand and Weber’s (2002, p. 369) account of the possible role of speech act theory in conceptual modeling, which suggests it only be used “as a means of identifying interactions.” In this paper we have shown that speech act theory has much more to offer. We suggest, based on the ontological foundation presented in this paper, that the focus of conceptual modeling should be the institutional objects—how they exist and how they represent both brute and institutional facts. Future research could build on these insights in order to explore how ontology and different language systems can help us solve important theoretical and practical problems—a contemporary “linguistic turn” for the evolving digitally interactive society.

Acknowledgements

We are grateful to Kalle Lyytinen, Jeffrey Parsons, Rikard Lindgren and the anonymous reviewers for the many insightful and helpful comments on various draft versions of this paper.

References

- Ågerfalk, P.J. (2010) “Getting Pragmatic,” *European Journal of Information Systems*, 2010(19), pp. 251–256.
- Ågerfalk, P.J., & Eriksson, O. (2004) “Action-Oriented Conceptual Modelling,” *European Journal of Information Systems*, 13 (1), pp. 80–92.
- Ambler, S. W. (2010) *Choosing a Primary Key: Natural or Surrogate?*, [WWW document] <http://www.agiledata.org/essays/keys.html>, (accessed 2 June 2010).
- Berkus, J. (2010), *Database Soup: Primary Keyvil, Part I.*, [WWW document] <http://blogs.ittoolbox.com/database/soup/archives/primary-keyvil-part-i-7327>, (accessed 2 June 2010).
- Bowker, G.C., and Star, S.L. (1999) *Sorting Things Out: Classification and Its consequences*, Cambridge, MA: MIT Press.
- Cattell, R.G.G. (1994) *Object Data Management: Object-oriented and Extended Relational Database Systems*, Reading, MA: Addison-Wesley.
- Celko, J. (1999) *Joe Celko’s Data and Databases: Concepts in Practice*, San Francisco, CA: Morgan Kaufmann Publishers.
- Celko, J. (2007) *Celko on SQL: Natural, Artificial and Surrogate Keys Explained*, [WWW document] <http://www.intelligententerprise.com/showArticle.jhtml?articleID=201806814> (accessed 2 June 2010).
- Codd E.F. (1979) “Extending the database relational model to capture more meaning,” *ACM Transactions on Database Systems*, 4 (4), pp. 397–434.
- Codd, E.F. (1988) “Domains, keys and referential integrity in relational databases,” *InfoDB*, 31 (1), pp. 53–78.
- Date, C.J. (1995) *Relational Database Writings, 1991-1994*, Boston, MA: Addison-Wesley.
- Date C. J. (2004) *An Introduction to Database Systems*, 8. ed. International ed., Boston, MA: Addison-Wesley.
- Edwards, P.N., Jackson, S.J., Bowker, G.C., and Knobel, C.P. (2007) *Understanding Infrastructure: Dynamics, Tensions, and Design*, Ann Arbor: Deep Blue.

- Elmasri, R., and Navathe, S.B. (2004) *Fundamentals of Database Systems*, 4. ed. International ed., Boston, Mass : Pearson/Addison-Wesley.
- Eriksson, O. (2003) "To Denominate and Characterise in the Context of Information Systems," in *Proceedings of the 11th European Conference on Information Systems ECIS 2003*, Naples, Italy, June 19–21, 2003.
- Evermann, J., and Wand, Y. (2005) "Ontology Based Object-oriented Domain Modelling: Fundamental Concepts," *Requirements Engineering*, 10 (2), pp. 146–160.
- Evermann, J., and Wand, Y. (2006) "Ontological Modelling Rules for UML: An Empirical Assessment," *The Journal of Computer Information Systems*, 46 (5), pp. 14–29.
- Frege, G. (1892) "Über Sinn und Bedeutung," in *Zeitschrift für Philosophie und philosophische Kritik*, 100: 25-50, Translated as "On Sense and Reference" by M. Black in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds. and trans.), Oxford: Blackwell, third edition, 1980.
- Goldkuhl, G., and Lyytinen, K. (1982) "A language Action View on Information systems," in Ginzberg, M. and Ross, C. (eds.), *Proceedings of the 3rd International Conference on Information Systems*, pp.13–29, Ann Arbor Mi.
- Graham, I. (1998) *Requirements Engineering and Rapid Development: An Object-Oriented Approach*, Harlow, UK: Addison-Wesley.
- Hanseth, O. (2000) "The Economics of Standards," in Ciborra, C., et al. (eds.), *From Control to Drift. The Dynamics of Corporate Information Infrastructures*, Oxford: Oxford University Press, pp. 56–70.
- Hanseth, O., and Lyytinen, K. (2004) "Theorizing about the Design of Information Infrastructures: Design Kernel Theories and Principles," *Sprouts: Working Papers on Information Environments, Systems and Organizations*, 4 (4), pp. 207–241.
- Hanseth, O., and Lyytinen, K. (2010) "Design theory for dynamic complexity in information infrastructures: the case of building Internet," *Journal of Information Technology*, 25 (1), pp. 1–19.
- Hall P., Owlett J., Todd S. (1976) Relations and entities, in Nijssen G. M. (ed.), *Modelling in Database Systems*, pp. 201-220, North Holland.
- Hirschheim, R., Klein, H., and Lyytinen, K. (1995) *Information Systems Development and Data Modeling: Conceptual Foundations and Philosophical Foundations*, Cambridge : Cambridge University Press.
- Iannacci, F. (2010) "When is an information infrastructure? Investigating the emergence of public sector information infrastructures," *European Journal of Information Systems*, (2010) 19, pp. 35–48.
- ISO (3779:1983) *Road vehicles – Vehicle identification number (VIN) – Content and structure*, Geneva: International Organization for Standardization.
- ITU (2005) ITU Internet Reports 2005: The Internet of Things, International Telecommunication Union, [WWW document] <http://www.itu.int/internetofthings/> (accessed 2 June 2010)
- Jacobson, I., Ericsson, M., and Jacobson, A. (1994) *The Object Advantage: Business Process Reengineering with Object Technology*, Wokingham: Addison-Wesley.
- Järvinen, P. (2000) Research Questions Guiding Selection of an Appropriate Research Method, in Hansen, Bichler and Mahrer (eds.), *Proceedings of the 8th European Conference on Information Systems*, 3–5 July, 2000, Vienna, pp. 124–131.
- Mathiassen, L., Munk-Madsen, A., Nielsen, P. A., and Stage, J. (2000) *Object-Oriented Analysis and Design*, Aalborg: Marko.
- Mill, J. S. (1872) *A System of Logic*, definitive 8th edition. 1949 reprint, London: Longmans, Green and Company.
- Milne, C (1997) "The design and management of Numbering Systems" in *Telecom reform: Principles, policies and regulatory practices*, Den Private Ingenirfond, Lyngby: Technical University of Denmark.
- Monteiro, E (1998) "Scaling information infrastructure: the case of the next generation IP in Internet," *The Information Society*, 14 (3), pp. 229–245.
- Otjacques, B., Hitzelberger, P. and Feltz F. (2007) Interoperability of E-Government Information Systems: Issues of Identification and Data Sharing, *Journal of Management Information Systems*, 23 (4), pp. 29-51.

- Richardson L., VA, (2007) *Surrogate vs Natural Primary Keys – Data Modeling Mistake 2 of 10*, [WWW document] <http://rapidapplicationdevelopment.blogspot.com/2007/08/in-case-youre-new-to-series-ive.html>, (accessed 2 June 2010).
- Rood, H (2000) "What's in a name, what's in a number: some characteristics of identifiers on electronic networks," *Telecommunications Policy*, 24 (6-7), pp. 533-552.
- Scott, W.R. (2001) *Institutions and Organizations*, 2 ed., Thousand Oaks, Calif. : Sage.
- Scott, W.R. (2003) "Institutional Carriers: Reviewing Modes of Transporting Ideas over Time and Space and Considering their Consequences," *Industrial and Corporate Change*, 12 (4), pp. 879–894.
- Searle, J. R. (1969) *Speech Acts: An Essay in the Philosophy of Language*, Cambridge, UK: Cambridge University Press.
- Searle, J. R. (1995) *The Construction of Social Reality*, New York: The Free Press.
- Sjöstrand (2007) "Personnumren tar snart slut och måste därför ersättas," In Swedish, *Dagens Nyheter*, 2007-02-23, [WWW document] <http://www.dn.se/opinion/debatt/personnumren-tar-snart-slut-och-maste-darfor-ersattas-1.451552>, (accessed 2 June 2010).
- SOU (2008) *Personnummer och samordningsnummer*, Swedish official government report (SOU) series 2008:60. Stockholm.
- Stamper, R. K. (2001) "Organisational Semiotics: Informatics without the Computer?" in Liu K. et al. (eds.) *Information, Organisation and Technology: Studies in Organisational Semiotics*, Boston: Kluwer Academic Publishers, pp. 115–171.
- Wand and Wang (1996) Anchoring data quality dimensions in ontological foundations, *Communications of the ACM*, 39 (11), pp. 86 – 95.
- Wand, Y., and Weber, R. (2002) "Research Commentary: Information Systems and Conceptual Modeling—A Research Agenda," *Information Systems Research*, 13 (4), pp. 363–376.
- Wand, Y., Storey, V. C., and Weber., R (1999) "An Ontological Analysis of the Relationship Construct in Conceptual Modeling," *ACM Transactions on Database Systems*, 24 (4), pp. 494–528.
- Weigand, H. (2006) "Two Decades of the Language-Action Perspective," *Communications of the ACM*, 49 (5), 44– 46.
- Wieringa R., De Jonge W. (1991) *The identification of objects and roles - Object identifiers revisited*, Technical Report IR-267, Faculty of Mathematics and Computer Science, Vrije Universiteit.
- Winograd T., and Flores F. (1986) *Understanding Computers and Cognition: A New Foundation for Design*, Reading: Addison-Wesley.

About the Authors

Owen Eriksson is a Senior Lecturer of Information Systems in the Department of Informatics and Media at Uppsala University and an Associate Professor of Informatics in the Department of Management and Engineering at Linköping University. He received his Ph.D. from Linköping University and has held positions at Dalarna University, University of Borås and the Viktoria Institute. His main research fields are conceptual modeling and database design based on language/action theories, e-infrastructures, and Intelligent Transport Systems (ITS). His research is action-oriented and design-oriented and he has been the research leader of a number of externally funded research projects in close co-operation with authorities and industry. His work has appeared in journals such as *European Journal of Information Systems* and *Journal of Information Technology*.

Pär J. Ågerfalk is a Professor in the Department of Informatics and Media at Uppsala University, where he holds the Chair in Computer Science in Intersection with Social Sciences. He received his Ph.D. from Linköping University and has held positions at Örebro University, Lero – The Irish Software Engineering Research Centre, Jönköping International Business School, and University of Limerick, where he is also currently an Adjunct Professor. His work on open source software, global software development and information systems development methodology has appeared in a number of leading information systems journals, including *MIS Quarterly*, *Journal of Database Management*, and *Communications of the ACM*. He is a Senior Associate Editor of the *European Journal of Information Systems*, a Secretary of IFIP WG 2.13 on Open Source Software, the founding Chair of the AIS Special Interest Group on Pragmatist Information Systems Research, and Dean of the Swedish Research School on Management and IT.

Copyright © 2010, by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers for commercial use, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints, or via e-mail from ais@gsu.edu.

Editor
Kalle Lyytinen
Case Western Reserve University

Senior Editors			
Michael Barrett	University of Cambridge	Robert Fichman	Boston College
Dennis Galletta	University of Pittsburgh	Varun Grover	Clemson University
Jeffrey Parsons	Memorial University of Newfoundland	Suzanne Rivard	Ecole des Hautes Etudes Commerciales
Carol Saunders	University of Central Florida	Avi Seidmann,	University of Rochester
Ananth Srinivasan	University of Auckland	Bernard Tan	National University of Singapore
Michael Wade	York University	Ping Zhang	Syracuse University
Editorial Board			
Steve Alter	University of San Francisco	Kemal Altinkemer	Purdue University
Michel Avital	University of Amsterdam	Cynthia Beath	University of Texas at Austin
Michel Benaroch	University of Syracuse	Avi Bernstein	University of Zurich,
Anandhi S. Bharadwaj	Emory University	Marie-Claude Boudreau	University of Georgia
Susan A. Brown	University of Arizona	Andrew Burton-Jones	University of British Columbia
Traci Cart	University of Oklahoma	Dubravka Cecez-Kecmanovic	University of New South Wales
Patrick Y.K. Chau	University of Hong Kong	Mike Chiasson	Lancaster University
Mary J. Culnan	Bentley College	Jan Damsgaard	Copenhagen Business School
Elizabeth Davidson	University of Hawaii	Jason Derdrick	University of California, Irvine
Samer Faraj	McGill university	Chris Forman	Carnegie Mellon University
Peter Gray	University of Virginia	Ola Henfridsson	Viktoria Institute & Halmstad University
Traci Hess	Washington State University	Qing Hu	Iowa State University
Jimmy Huang	University of Warwick	Kai Lung Hui	National University of Singapore, Singapore
Bala Iyer	Babson College	Hemant Jain	University of Wisconsin-Milwaukee
Zhenhui (Jack) Jiang	National University of Singapore	Bill Kettinger	University of Memphis
Gary Klein	University of Colorado, Colorado Springs	Ken Kraemer	University of California, Irvine
Mary Lacity	University of Missouri-St. Louis	Liette Lapointe	McGill University
T.P. Liang	National Sun Yat-Sen University	Kai H. Lim	City University of Hong Kong, Hong Kong
Lihui Lin	Boston University	Ji-Ye Mao	Renmin University
Anne Massey	Indiana University	Ramiro Montealegre	University of Colorado at Boulder
Michael Myers	University of Auckland, New Zealand	Fiona Fui-Hoon Nah	University of Nebraska-Lincoln
Fred Niederman	St. Louis University	Mike Newman	University of Manchester
Brian Pentland	Michigan State University	Geert Poels	Katholieke Universiteit Leuven
Jaana Porra	University of Houston	Sandeep Puroo	Penn State University
T. S. Raghuram	Arizona State University	Dewan Rajiv	University of Rochester
Neil Ramiller	Portland State University	Matti Rossi	Helsinki School of Economics
Suprateek Sarker	Washington State University	Susan Scott	The London School of Economics and Political Science
Ben Shao	Arizona State University	Olivia Sheng	University of Utah
Choon-ling Sia	City University of Hong Kong	Carsten Sorensen	The London School of Economics and Political Science
Katherine Stewart	University of Maryland	Mani Subramani	University of Minnesota
Burt Swanson	University of California at Los Angeles	Jason Thatcher	Clemson University
Ron Thompson	Wake Forest University	Christian Wagner	City University of Hong Kong
Dave Wainwright	Northumbria University	Eric Walden	Texas Tech University
Eric Wang	National Central University	Jonathan Wareham	ESADE
Stephanie Watts	Boston University	Tim Weitzel	Bamberg University, Germany
George Westerman	Massachusetts Institute of Technology	Kevin Zhu	University of California at Irvine
Administrator			
Eph McLean	AIS, Executive Director		Georgia State University
J. Peter Tinsley	Deputy Executive Director		Association for Information Systems