To Denominate and Characterise in the Context of Information Systems

Owen Eriksson
Dalarna University, Dept. of Computer Science and Informatics (IDI)
SE-781 88 Borlänge, Sweden
Phone: +46 23 778958, Fax: +46 23 778050
oer@du.se

Abstract
The aim of the paper is to discuss the distinction between characterisation (classification) and denomination (naming, identifying) in the context of information systems (IS). The paper has a special focus on the design of identifiers, i.e. terms used for identifying individual phenomenon. The reason for this is that the distinction between denomination and characterisation is especially important when designing identifiers. Identifiers e.g. article numbers, telephone numbers, e-mail addresses and personal numbers, constitute an important part of the information infrastructure of companies and society as a whole, and therefore it is crucial that these terms are well designed. The paper illustrates that who designs, assigns and withdraws electronic identifiers is a significant economic and policy issue both within companies and for society as a whole, with farther reaching consequences than often perceived at first glance

Keywords

Identifier, naming, denominate, proper names, characterise, speech act theory.

1. Introduction

"Nominatur singularia sed universalia significantur" is a sentence, which was formulated by the middle-age philosophers. The sentence means that individuals are denominated (named) and classes (categories) are characterised. With this sentence the middle-age philosophers wanted to stress that it is the name, which gives someone or something (the new-born, a newly founded place, newly discovered mountain peak, or a country) its identity, and the characterisation, which assigns something to a class (Malmberg 1973, p. 44). This important insight, meaning that it is a significant difference between denomination and characterisation, is of great importance when information systems (IS) are designed. If systems developers and users do not recognise this, there is a risk that the systems developed become difficult to maintain, hard to understand and inefficient to use. The paper has a special focus on the design of identifiers, i.e. terms, used for identifying individual phenomenon, e.g. article numbers, telephone numbers, e-mail addresses, and personal numbers. The reason to focus on the design of identifiers is that the distinction between denomination and characterisation is especially important when designing identifiers. A well-designed identifier (Milne 1997, Celko 2000) should meet the requirements listed below:

- It must uniquely identify the individuals in the class.
- It must be possible to address all the individuals in the class; i.e. the value domain of the identifier must be sufficient.

- It should be easy to memorise which implies that it should be short and mnemonic.
- It should be stable, i.e. it should not have to be altered (by accident) neither by content nor to its structure.
- It should be verifiable within itself, e.g. it should be possible to verify with the help of a check number.

Identifiers constitute an important part of the information infrastructure of companies and society as a whole, and therefore it is crucial to design these terms well.

The purpose of this paper is to discuss what it means to denominate and to characterise and practical problems, which can arise if this distinction is not recognised in the context of IS. The structure of the paper is as follows. Section 2 gives the theoretical background for the important distinction between denomination and characterisation. In section 3, the theoretical discussion is further elaborated with the help of a real life example "the vehicle example" with a focus on the design of two vehicle identifiers, the registration number and the chassis number. In section 4, a discussion follows on the danger of using characterising identifiers and the consequences that this may have in the end; this discussion is based on two real life cases. Finally, section 5 concludes the paper.

2. Theoretical BASIS

2.1 The triangle of Ogden

In order to be able to reason about the difference between denomination and characterisation, we need to discuss the relations between words (terms), thoughts and phenomena. This can be done with the help of the triangle of Ogden (Ogden & Richard 1949).

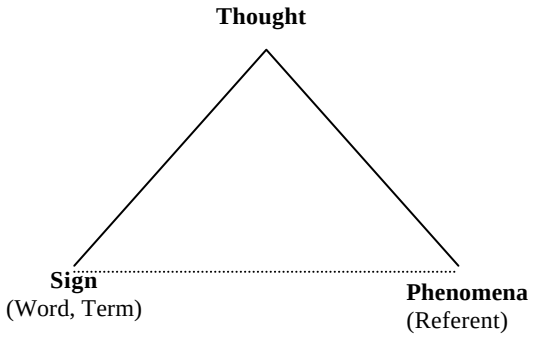

Figure 1. The triangle of Ogden.

The triangle shows the relation between the three factors whenever any statement is made or understood (ibid. p. 10). The triangle illustrates that signs (words, terms) are linguistic expressions of certain ideas (thoughts). What is important about Ogden's triangle is that it shows that the relationship between signs and phenomena is an indirect one (Ogden & Richards 1949). This is important to emphasise because the idea of a direct connection between words and phenomena is the source of many difficulties, which philosophy and thought encounters (ibid.). Saussure (1998, p. 97) claims that "For some people a language, reduced to its essentials, is a nomenclature: a list of terms corresponding to a list of things", which assumes that the ideas already exists independently of words.

To claim that there is no direct relation between symbol and phenomena also implies that the relations illustrated by the sides in the triangle are complicated, and that there is not one answer to the question how, symbols, thoughts and phenomena are connected to each other.

2.2 To characterise (categorize, classify, conceptualise)

To characterise means to categorize and to conceptualise (Malmberg 1973, p. 44-45). Saussure (1998) claims that there are no concepts, which are independent of language. Linguistic concepts are categorizations and abstractions of reality, and when people create concepts it involves that they study characteristics of the phenomena that connect them or separates them from other phenomena. This means that reality is divided with the help of class concepts, and that concepts are used both in order to describe and create reality. For example, phenomena that we categorize as vehicles are alike when it comes to certain characteristics, these characterising attributes connect all phenomena that we look upon as vehicles and separate them from phenomena that do not have these specific characteristics.

2.3 To denominate (name, identify)

To denominate means naming and identifying (Malmberg 1974). For example, when we have created the vehicle class we need a term that we can use in order to talk about and identify unique individuals of this class. The function of this term the identifier (cf. proper name, Searle 1969, p. 75) should be to identify (naming) individuals not characterising and describing them, and this term can be either alphanumeric or numeric. The function of the identifier is to make identifying references to phenomena, and it should help us answer questions like which one, who and what?

The function of the identifier is different from the attributes that are terms used to characterize and describe properties about the phenomena. This does not imply that properties are not used for the identification of individuals, properties can be used for identification and they can be helpful for proving the identity (Searle1969, p. 80). There are also many examples of names, which include properties, a name, as “Harald Blutooth” is obviously an example of a name that is both identifying and characterising. Properties can also be used for assigning identifiers, and properties can be useful in order to create mnemonic identifiers. The personal number used for the identification of persons in Sweden has e.g. the date of birth included in the first 6 positions in the number (571129-8447), positions 7, 8, and 9 is a sequential number, and position 10 is a check number. The reasons why the date-of-birth may be included in the identifier is that it can help people to learn and remember their number and it probably will not change, people are only born once and at a specific time.

However, the examples above do not affect the central function of the identifier i.e. the function of individualisation (Malmberg 1973).

2.4 Linguistic and language theory in the context of information and communication systems

There is a strong connection between IS and linguistic and language theories because IS are used for communication which implies transmission of information where signs and concepts are used. However it is important to recognize that communication does not only mean the transmission of messages; it also implies action. This is also the main idea in speech act theory (Austin 1962, Searle 1969), and several authors have based their definition of IS on

speech-act theory (e.g. Winograd & Flores 1987, Auramäki et. al 1988, Dietz 2001). In these definitions IS are regarded as systems used for performing speech acts (communication acts). For example, Ågerfalk et. al. (2000) claim that IS are action systems used for performing speech acts, and that the result of a single speech is an ae-message which is sent with the help of the IS. An ae-message consists of a *propositional content* and an *action mode*. This can be exemplified with a certificate of registration, which is one type of message that is communicated with the help of the vehicle register system in Sweden (Vägverket 2002) which the Swedish National Road Administration (SNRA) is responsible for.

Certificate of registration

Owner:	Lars Eriksson
Personal number	571129-8447
Street address:	Big street 20
City:	Borlange
Zip Code:	781 88

Registration number:	DCA096
First date of registration:	1996-03-14
Chassis number:	AAABC99LXP1000001
Type of vehicle:	Private light goods vehicle
Model:	SAAB 900 2, 3 I DX55B
Year of Model:	1996
Body:	Station wagon
Length:	4.75 m
Width:	1.72 m
Number of passengers:	4
Motor effect:	110 kW
Total weight:	1830 kg
Tax weight:	1450 kg

Figure 2. The certificate of registration.

The heading “Certificate” indicates the *action mode* of the ae-message and the pragmatic meaning of the message. In this case it means that the car is approved as a vehicle, it also confirms the ownership of the car. This corresponds to what Searle (1969) call the illocutionary act, which is a sub-act of the speech act. The main focus in speech act theory is on the analyses of illocutionary acts. However Searle (1969, pp. 72-127) has also analysed the meaning and function of the propositional content in detail, and it is the two main functions (referring and predicating) of the propositional content, which is in focus in this paper.

The *propositional content* (information content) in the certificate of registration describes a vehicle and the owner of the vehicle. This corresponds to what Searle (1969) call the propositional act that is another sub-act of a speech act. The propositional content is used for both referring and predicating. Referring means to identify a phenomenon, and to predicate means characterising or describing it (ibid. p. 119). From the certificate we can see that the term registration number (the identifier) is used for referring and naming an individual vehicle ’DCA096’, and other terms (the attributes) are used for characterising and describing it, e.g. model, body and length. Searle (1969, p. 174) writes “...we have the institution of proper names to perform the speech act of identifying reference. The existence of these expressions derives from our need to separate the referring and predicating functions of language”.

The example above shows that linguistic terms and concepts are used when IS are used. It also means that the construction of IS implies the design of linguistic terms and concepts (cf Goldkuhl & Lyytinen 1982) where some terms (identifiers) have the main function of identification (denomination) and other terms (attributes) have the main function of characterisation and description. This is important to recognise when IS are designed as will be shown in the next two sections.

3. The vehicle example

In this section, the discussion about denomination and characterisation will develop further with the help of "the vehicle example". In this example, it will show how the class concept vehicle is characterised and identified by the Swedish National Road Administration (SNRA). There will be a focus on the design of two important identifiers, the *registration number* used in Sweden and the *chassis number ISO-VIN*, which is used worldwide.

3.1 The vehicle concept

In the certificate of registration sent to the car owner one can see how the class concept vehicle is characterized and identified from the perspective of the SNRA. Vehicles are characterised and identified with the following terms: registration number, chassis number, type of vehicle, model, year of model, first date of registration, body, length, width, number of passengers, motor effect, total weight and tax weight. The triangle of Ogden illustrates this in the figure below.

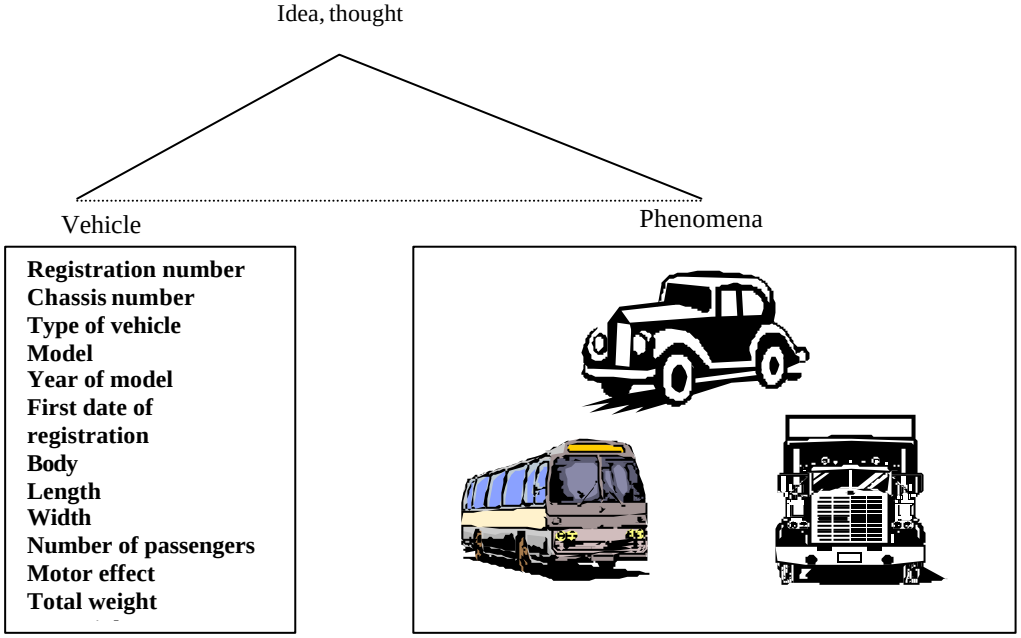

Figure 3. The vehicle concept exemplified with the help of triangle of Ogden

The reason to use this terminology in the registration certificate is that the authorities in Sweden need this information in order to fulfil purposes and aims such tax collection, vehicle control and law enforcement.

Among the terms listed above there are two terms that could be used in order to denominate (uniquely identify) an individual vehicle; *the chassis number* which is punched into the sheet metal of all vehicles, and the *registration number* which is visible in the back and front of the vehicle.

3.2 Choice of identifier

The SNRA has chosen to use registration numbers to identify vehicles a term constructed by the authority itself. One reason why the SNRA has made this choice is that the authority itself can control the design of the term and assignment of numbers.

Another important reason why the SNRA chooses registration numbers instead of chassis numbers is that chassis numbers identify an individual in the class *manufactured vehicles* whereas the registration number identifies an individual in the class registered *vehicles*, and these classes are not identical. We can see that we are dealing with different concepts by studying the act that creates an individual in each class. In the moment that the vehicle is manufactured, a new individual is created in the class of manufactured vehicles. An individual in the class of registered cars is created when the manufactured vehicle is registered at the SNRA; this is often done after the vehicle has been sold. A physical vehicle can be without registration number in its lifetime, depending on how the vehicle is registered, deregistered and reregistered. In spite of this, the vehicle is still the same physical manufactured object, with the same chassis number punched into the sheet metal.

In the next two sections are the designs of *the registration number* and the international standard for *chassis number the VIN-code* analysed.

3.3 The registration number

The registration number is an example of a well-designed identifier. The registration number is well designed because it is not used for characterising the vehicle; i.e. the term lacks information except for its function to identify a unique vehicle.

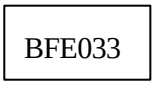

Figure 4. The registration number

There are a number of advantages with this type of design.

1) The term is stable

The registration number becomes stable in the sense that once assigned to a vehicle it will probably not have to be subject to changes by accident to neither its structure nor its content. A problem with the former design of the registration numbers in Sweden (before 1972) was that the County Code of the owner was included in the registration number. Consequently, the identity of the vehicle (accidentally) changed when the owner moved into another county, or if the car was sold between two county boarders. This created problems because it implied administration of the new number and the need for cross-referencing between the old and the new number, in order to keep information about the history of the vehicle. It was also semantically confusing because the identity of the car was dependent on the residence of the owner.

2) The term is short

The term is short only six signs, which is an advantage. A short term is easier to memorize and to feed into a keyboard, compared to a long term. This is a consequence of the fact that the registration number is not used for characterising the vehicle. Identifiers that both identify and characterise the phenomenon often tend to be long because there are several properties included in the term (see the ISO-VIN-number below).

3) The term is easy to memorise

Another advantage with the registration number is that it is easy to memorise because its short and alphanumeric that implies that one can combine numbers and letters and this makes it easier to memorize.

4) The number can identify many vehicles

No position in the number characterises the vehicle. This implies that all six positions can be used for the function of identifying vehicles. This fact, and the possibility to use both numbers and letters in each position, makes it possible to identify a great number of vehicles.

ISO-VIN

The other term that can identify vehicles is chassis number. There is an international standard for chassis numbers ISO-VIN (Vehicle Identification Number) (ISO 3779:1983). The appearance of chassis numbers is guided by demands from the European Union and authorities in the USA. The ISO-VIN is designed to identify motor vehicles, trailers, motorcycles and mopeds and consists of 17 positions (see figure 4). In the ISO-VIN, capital letters A through Z and numbers 1 through 0 may be used, except the letters I, O and Q.

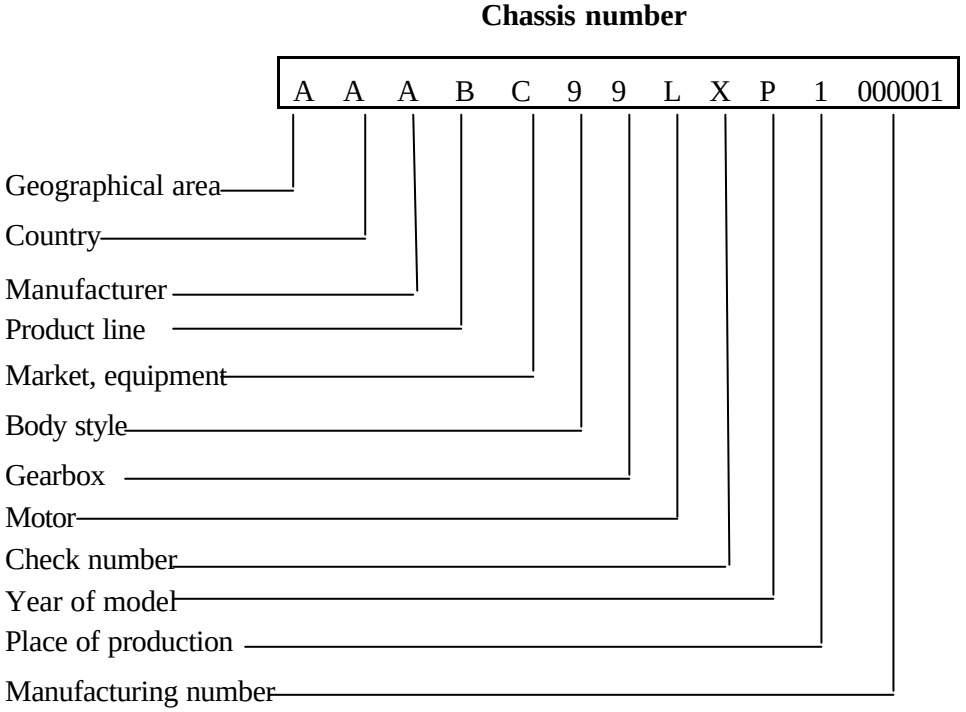

Figure 5. VIN-code the standardized chassis number

The ISO-VIN is not a well-designed identifier. The problem with the term is that is has double functions, i.e. it is used both for identifying and characterising the vehicle. The EU and US authorities want a lot of information punched into the sheet metal of the cars, nothing to say about that, but the ISO-VIN is not a semantically and from a IS point-of view well-designed identifier. The problem is that too many properties of the vehicle have been included in the identifier. This type of identifier, where the functions of identifying (denomination) and characterising has not been separated leads to a number of problems:

1) The term is difficult to use and mistakes occur

Mistakes occur because this very long code string is not easy to feed in from a keyboard and the number is not easy to memorize. One mistake that easily occurs is that two signs in the long code string are switched. This gives way to a lot of unnecessary mistakes and a lot of extra work to correct these mistakes, because the chassis number is used in many activities and IS at manufacturing companies, whole sellers and retailers in the car industry. If the content of the identifier has to be altered, i.e. updated, the problem will be to update the identity everywhere where it exists, and updating an identifier is always questionable (Builder Com 2002). From the standard, we can also see that the term is unnecessarily long for the function of identifying vehicles; only the last 8 characters are actually needed in order to identify a vehicle uniquely.

2) The identifier becomes instable

The more information that is put into the identifier, the greater the probability that it will have to change (by accident); this is due for both the content and structure of the identifier. There is e.g. a significant risk that a term as VIN-Code, which contains so much one-position dependent information, cannot represent the information that it is intended to represent. Each position can only represent 32 different values. A big problem is if the structure of the identifier has to be changed, for example extend the amount of positions; this leads to unacceptable costs (see the examples in section 4 below).

3) It gives rise to “code mysticism”

One aim with the system development process is to create concepts and terms that are comprehensible. The problem with a term like ISO-VIN is that the information that is communicated with the help of the term becomes fuzzy and hard to understand. You have to be familiar with “code mysticism” in order to understand the information. For example, you must know that in position number 2 you find information about the manufacturer, and that the digit 4 in this position means Cadillac. A number of special rules also apply to different positions of the ISO-VIN. According to the standard, position 10 contains information about either the year of model, or the building-year of the car. This is fuzzy and may lead to misinterpretations because the year of model and the year the car was built are really two different properties. Another example is the rule that specifies that when a manufacturer builds less than 500 vehicles a year, should the third digit always be a 9 and that the 12th, 13th and 14th position of the VIN should then identify the manufacturer. This implies that position 3, 13, 14, and 15 have a different meaning depending on how many cars the manufacturer is making a year.

To create identifiers like ISO-VIN with coded and position-dependent information buried in the term is a common way of designing identifiers, and it is sometimes believed to be a smart design. However, what seems as a smart move at the first sight may create problems in many cases in the end. The danger of using characterising identifiers and the consequences that this might have will be presented in this section with the help of two examples based on real life cases.

3.4 The problem with the article number

Within a company of three divisions in Sweden, division A, B and C a common article number was used to identify articles (spare parts). The article number was a numeric term of 7 positions. One day division C faced a problem; they would run out of article numbers within a year. Division C worked intensely on this problem because without available article numbers the division would not survive.

The reasons why the article numbers were running out was explained this way:

1) The article number was characterising

The company had decided that the first position of the number should determine to which division the article belonged. The numbers 1, 2 or 3 in the first position of the number showed that the article belonged to division A. The numbers 4, 5 or 6 showed the belonging to division B and the numbers 7, 8 or 9 that the article belonged to division C.

2) The article number contained a check number

The very last number in the article number was a check number which meant that only every ten numbers could be used in the seven-position number. This implied that the total domain of article numbers were restricted to the range of 700000-999999 for division C.

3) Routines for re-usage of article numbers was missing

The following suggestions to solve the problem were discussed:

A. To change the format of the article number, extend it from 7 positions to 8 positions.

The consequences would be that approximately 4000 programs would have to be changed. Approximately 50 000 work hours would be needed which would cost 20 million SEK (€2, 2 mil.), as the estimated cost of a work hour was 400 SEK. In addition to that, the staff of system developers would largely be tied up with this work and this would imply that other important system changes would not be performed.

B. Not care about the check number and be able to access all seven positions in the article number

The consequences would be more mistakes in the ordering and delivery process of articles. This would not be in line with the main business goals of division C regarding quality, reliability and customer service.

C)Abandon the use of characterising information in the identifier

The advantage would be that division C could use the 4000000 series of numbers that was not used by division B. This would make 100 000 article numbers available for division C, which would be sufficient for another 15 years approximately. The only disadvantages with this solution would be that the ownership of the article could no longer be easily distinguished by the first position of the article number. The assignment of article numbers had also to be co-ordinated in another way, because the former rule of using the first position to determine the ownership of the article number had also been an easy way of assigning article numbers to different divisions.

THE DECISION MADE WAS OF COURSE TO CHOOSE ALTERNATIVE C. IT WAS ALSO DECIDED TO DEVELOP A STANDARD FOR A NEW ARTICLE NUMBER. THE NEW ARTICLE NUMBER IS A NUMERICAL TERM OF 9 POSITIONS INCLUDING A CHECK NUMBER. THE NEW ARTICLE NUMBER IS NOT CHARACTERIZING. IT IS USED ONLY TO UNIQUELY IDENTIFY ARTICLES. THIS SOLUTION ALLOWS THE IS TO BE GRADUALLY ADJUSTED, OVER A PERIOD OF 10 YEARS.

3.5 The problem with the telephone number

Another example of characterising identifiers is the area codes in telephone numbers which has become a big problem in a number of countries because of the huge growth in telecommunications. In the 1990s a collective change in telephone numbers have imposed substantial costs in many countries. In England the telephone numbers have been changed

three times over the last five years. These changes force people to inform friends, relatives and business relations on the changed numbers, and it also leads to a major raise of unsuccessful calls. This implies that the change of telephone numbers raise huge costs and create problems for people and society (Rood 2000).
In electronic networks identifiers perform three different functions; naming, addressing and routing, and the origin of the problem with the telephone numbers is that a single identifier sometimes combines more than one of these functions. Telephone numbers of the fixed telephone network are well-known examples of this. Phone-numbers with properties included, i.e. area information, restrict the domain of available numbers, creating a big problem when the demand for available numbers increases. It also leads to a number of other problems e.g. that people have to change their telephone number when they move.
The trend in order to solve the problems is that all telephone numbers shift towards being identifiers with one function i.e. the naming function. Until recently only free phone, premium rate and short codes have been used more or less as a name; the emergency number 112 in Europe is a well-known example of this. New modern communications networks deploy databases, which return addresses or routes when queried with a name identifier. This solution works as long as there is at least one database in the world that is connected to the network and that performs the translation of the name sent to it in a query into an address.

3.6 Lessons learnt

The lessons learnt from the cases above are that using characterising identifiers is not recommendable. The problem is that properties included in identifiers restricts the domain of numbers which can be used and create demands on expanding the structure of the identifier. We can also see that a change of the structure is not an easy task, because the identifiers are frequently used in many IS and human activities. The examples also show that who designs, assigns and withdraws electronic identifiers is a significant economic and policy issue both within companies and for society as a whole, with farther reaching consequences than often perceived at first glance.

4. Conclusions

Information systems design implies the construction of linguistic concepts and terms (Goldkuhl & Lyytinen 1982). This also implies that we have to create well-designed linguistic entities and adapt them to the context of where the systems are used. While creating class concepts the characteristics (attributes) that unite and divide phenomena are studied. We also need a way of identifying individual members of a class and we do that by naming the individuals, i.e. individuals are denominated. This also means that some of the terms designed and used (the identifiers) have the main function of identification (denomination), and other terms (the attributes) have the main function of characterisation and description. This is important to recognise when IS are designed.

In this paper, the focus has been to analyse how to design identifiers. It is important that we have well-designed identifiers because they serve important purposes in society and in human activities, and they are everywhere. Terms like article numbers, register numbers, chassis numbers, social security numbers (personal numbers), telephone numbers, e-mail-addresses, course codes, etc., are all examples of identifiers that we have to use in order to make things work. These identifiers are an important part of the infrastructure of society and organisations. This implies that we need well-designed identifiers but unfortunately there are far too many identifiers which are not, and it creates significant problems. It is also important to realize that technical solutions are not the ones that really can help us to solve the problem,

although technical solutions might help to reduce the problems. The thing that really can help us to solve the problem is to realize what the middle-age philosophers meant with the sentence "Nominatur singularia sed universalia significantur", i.e. individuals are denominated (named) and classes are characterised, and to apply that wisdom when we develop IS.

Acknowledgements

I wish to thank Björn Lundell, Pär J. Ågerfalk and four anonymous reviewers for constructive comments on the paper, and Susanne Corrigaux for helping me with the English.

References

Ågerfalk, P, Goldkuhl, G, & Cronholm S (2001), 'Actability Design – Developing IT-systems for Business Action', Version 1.0.

Lehtinen, E & Lyytinen, K (1988) 'A Speech Act Based Office Modeling Approach', *ACM Transactions on Office Information Systems*, 6(2), pp. 126–152.

Austin, JL (1962) 'How to Do Things with Words', Oxford University Press, Cambridge.

Builder COM (2002), 'The great primary-key debate', [www dokument], http://builder.com.com/article.jhtml;jsessionid=MX3CMTMYWJALJTQQACQCFFQ?id=u00320020327ssh01.htm&page=2

Celko, J (2000) 'Playing the NameGame', [www dokument], http://www.intelligententerprise.com/000908/celko.shtml

Dietz, JLG (2001) 'DEMO: Towards a Discipline of Organisation Engineering', *European Journal of Operational Research,* 128(2), pp. 351–363.

ISO 3779:1983, 'Road vehicles -- Vehicle identification number (VIN) -- Content and structure'.

Goldkuhl, G & Lyytinen, K (1982) 'A language Action View on Information systems', in *Proceedings of 3rd International Conference on Information Systems*, Ann Arbor

Malmberg, B (1973) 'Teckenlära' In Swedish. Aldus/Bonniers, Stockholm.

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.

Ogden, CK & Richard, IA (1949) 'The meaning of meaning', Routledge & Kegan Paul. London.

Rood, H (2000) 'What's in a name, what's in a number: some characteristics of identifiers on *Telecommunications Policy*, vol: 24 issue: 6-7. pp. 533-552 Pergamon.

Saussure, F (1998) 'Course in general linguistics', Pergamon, Oxford

Searle JR (1969) 'Speech Acts: An Essay in the Philosophy of Language', Cambridge University Press, Cambridge.

Vägverket (2002), 'Bilregistret', [www dokument], http://www.vv.se/i_bilregistret.shtml

Winograd, T & Flores, F (1987) ‘Understanding Computers and Cognition: A New Foundation for Design’, Reading, MA: Addison-Wesley