K + K = 120: Papers dedicated to L. Kálmán & A. Kornai on the occasion of their 60th birthdays, pp. 529–540

Interpretation of the concept of namespace

ISTVÁN SZAKADÁT
Budapest University of Technology and Economics
i@syi.hu

KEYWORDS

interoperability
name space
proper name space
Knowledge Organization System
semantic structure

ABSTRACT

The significance of namespaces is becoming more and more recognized with the spread of the world of networks and digital culture. Global, national and local namespaces are being built around the world. Before clarifying what the concept of namespace means (and what it does not), I will present some special namespaces in order to show how they work and what their function is. After that, I will interpret the concept of namespace and then briefly examine the practical implications of the theoretical findings.

1. File names and the namespace

Perhaps the simplest but surely the most commonly used namespace is the computer file system, whose only task is to ensure that each file has a unique name on the computer. It depends on the operating system what rules can be used to create names, but this is a secondary issue as far as the functioning of the namespace is concerned. However, the creation of unique names is facilitated by the possibility to build names from components. Files can be collected into groups (folders, directories), folders can be embedded into each other, and the complete (unique) file name can be obtained by concatenating the (short) name of the file and the names of the folders containing the file. This is a hierarchical method of name-identification, with all the advantages and disadvantages of hierarchical control. From a linguistic point of view, hierarchical organization means that names can be interpreted as compositions of distinct name elements, so a name can be used more than once as a name element when composing a specific file name. This is made possible by the fact that file names are compound names, i.e. the full and complete (thus unique) file names contain at least as many name components as the number of folders the file

itself is embedded in, starting from the root of the file system. However, these name elements do not have to be unique in themselves, they can be duplicated.

With the requirement that each file should have a name and this name should be unique, it can be ensured that each file is identified unambiguously, and always found on the computer. This expectation is compatible with a file having multiple names in the file system, since if the uniqueness of names is guaranteed and all names are linked to files, then all files can be found on the basis of the names. (It would not be true the other way round, but it seems to be pointless to find a name starting from a file.) In linguistic terms, we can say that if two names belong to the same thing (here: file), synonymy is allowed in the namespace, but if a name can only belong to one thing (file), homonymy is not allowed.

It is also important to note that the file system as a namespace is local in the sense that there is at least one file system on each computer, so there are at least as many namespaces as computers. Of course, these local namespaces can be combined (by connecting them to a network) and therefore larger namespaces can be created. Their uniqueness is easy to maintain if the machines connected have unique identifiers. With respect to a file system, it is easy to answer the following question (which is important for each namespace): Who can manage (create, delete, change) the names? To simplify things a little, the answer is that it is the owner (user) of the computer.

2. Domain names and the namespace

Similarly to the case of the file system, the expectation for the uniqueness of names and for the accessibility of the computer resources denoted by names arises with respect to the domain names of the World Wide Web, too. The system of domain names is a namespace as well. Both the domain names on the web and the hierarchically structured set of web page addresses can be regarded namespaces (the latter containing the former). What is the main function of domain names? To mark the common roots of web page addresses (facebook.com, microsoft.com, mit.edu) with character series that can be unambiguously identified and easily memorized by humans. There are rules for creating domain names (concerning what characters can and cannot be used as name elements, or stating the fact that separators have to be used among the name elements, etc.) and there is a basic expectation for the whole set of domain names: each name must be unique. This can be achieved by distributing the task of registering

domain names among specific institutions, which are expected to comply with the principle of unique name distribution in the namespace of their responsibility. This naming system is based on hierarchical organization, too.

Maintaining the domain namespace is necessary in the first place so that, in the case of a new naming request, it would be possible to know whether the required domain name exists in the namespace. If it is a name already present in the namespace, it cannot be registered. If the required name is not yet in the namespace, it can be registered (provided it complies with the naming rules). What does this registration mean? It entails establishing a connection between domain names and the owners of domain names. The other important function of the domain namespace is to create a clear link between domain names and IP addresses that belong to computers.

Domain namespace is thus a hierarchical registry of unique names and the dynamic connections between names and owners on the one hand and names and physical addresses on the other. The uniqueness of the domain names ensures the unambiguous identification of their IP addresses and the registration provides the right of use of the domain name for its owner. As the main function of the domain namespace is the clear identification of all communication nodes on the web, it is obvious that this is a global namespace. This namespace allows synonymy but forbids homonymy, and uses the technique of creating composite names, since domain names can be divided into components, and the same name elements may be reused in other domain names as well.

3. Person names and the namespace

After the two examples from the world of technology, let us look at why and how person names are used and how to build namespaces for person names. In a family, the use of different first names obviously has the same meaning as in the two examples above: by using unique names, we create the possibility of referring unambiguously to the entities denoted by the names, the children within the family. For a family, it is enough to choose names from a collection of a few thousand first names that have become acceptable in the course of time, but beyond family communities, such a quantity of names is clearly unsuitable for the identification of every child. From this point of view, the formation of compound names where the full name of the person consists of the concatenation of a family name and a first name can be a bit of help, but we know that this solution cannot ensure that everyone has a unique name. With the help of composite person

names, we can only identify people clearly if we determine the domain of people, in other words, we create a naming context with a precise boundary. What does this mean? In narrow contexts (family, small community, clan, village) there could be hope that all people can be identified with the help of family names and first names, but in a wider context this hope is eliminated because homonymy will become more and more common.[1]

The name *John Smith* is no longer able to identify a person unambiguously if there are several families in the village called *Smith* that have given the first name *John* to one of their sons. This example highlights the importance of the naming context. Within the family as a naming context, the good functioning of the namespace can be guaranteed; everyone can have a unique name. If this context is expanded, the homonymy-free state can still be maintained for a while, but expansion can reach a limit where it cannot be assured that one particular name denotes only one person. At this point, the namespace becomes useless or, more precisely, partially useless, as it may have some segments that can still function well, but the functionality of the namespace can only be restored if the scope of the naming context is narrowed to the required size.

The example of the person namespace indicates that names in themselves are not able to identify the entities denoted, only with the use of the context and the names together can we hope for the successful operation of namespaces.

4. The components of the namespace

The next important question is what the components of namespaces are. It is obvious that whatever one means by namespaces, names are always among their elements. On its narrowest possible interpretation, a namespace could be defined as a simple collection or a list of names, where the type of objects they denote is also specified. According to this approach, a namespace is a unique list of names that can be used for some purpose. This means that in a geographical namespace the name *Old Hill* could only appear once, and the only thing we could say about it is that it possibly points to one or more geographical locations. Similarly, a family namespace interpreted in the narrowest sense could only be expected to include all possible family names in a manner that every name only occurs once. Such

[1] Synonymy does not cause any problems here, either. One person may have several names; if they are unique, each of them may be suitable for identifying the given person.

a namespace must be disjoint and exhausted simultaneously (Bittner et al. 2004). Disjointness ensures the uniqueness of the names, and exhaustedness guarantees that all possible usable names are included. A namespace defined so narrowly could be used to control the process of selecting names in a data system by requiring that at certain input points only elements of a given namespace (family namespace) can be chosen. Such a namespace, however, would not have too many practical benefits, it would not help to reach the goal of uniquely identifying the set of things to be described with namespace elements. For such a purpose, a namespace like the one above would only be suitable if the number of names were larger than that of the things to be denoted, or if the names could be reused with the help of a hierarchical name composing technique. For people, organizations, geographical locations and many other things that could not be guaranteed, so in such contexts no homonymy-free state can be achieved.

Returning to the question of what the additional elements of the namespace can be, the answer must take into account the purpose for which the namespaces are used. If we want to use namespace elements to identify things of a certain type, then we have to include in the namespace the entities that the names refer to. These can be called *name holders*. In afile system files are the name holders that can be identified by file names, and in a domain name system IP addresses (and content packages packed there) are the name holders that can be identified by domain names. A name holder may be a person who can be referred to with one or more names but name holders may be organizations or geographical locations, too.

If namespaces contain not only names but also name holders, they must include something else, too: *the relations between names and name holders*. In fact, managing these relations is the real sense of the use of namespaces.

5. The definition of the namespace

Namespaces can be defined on the basis of the quality of the system of relations between names and name holders (dynamically changing in space and time): a *namespace is a function that returns one and only one name holder for each possible naming*. This definition seems simple, but the question is how a namespace defined in the manner above can be used in practice.

In the case of person namespaces it has already turned out that the names themselves are not capable of unambiguously identifying people (name holders): several people may be named *John Smith*. The question

is how the namespace can operate as a function under these conditions. In order to be able to proceed, the concepts we have used so far need to be be clarified further. The first clarification can already be seen in the definition of the namespace above, which does not contain the term *possible name* but *possible naming*. What is the difference between these two terms and why is it necessary to introduce the new category (naming)?

It is true for all of the namespaces referred to so far (even if in some cases it may be surprising) that the names that appear as elements of namespaces can be interpreted as *proper names*. For person names, names of geographical locations, organisations this is self-explanatory, it is easy to see and accept it for domain names as well, but in the case of file systems this qualification may be surprising for the first time. But it is true. File names play the same role for files as person names do for people or names of geographical locations do for geographical locations. They refer to a specific existing entity (file, person, geographical location) in order to distinguish this particular thing from the other similar, specific things (the given file from the other files, the given person from the other people, etc.). When using names for this purpose, we use special names: proper names.

A unique feature of proper names, which differentiates them from common names, is that proper names are rigid designators (Kripke 1980). Philosophers are in dispute about whether proper names have a meaning but there is a consensus among them concerning the fact that with a proper name we refer to a single individual, and this reference is rigid in the sense that under all possible conditions (in all possible worlds) the reference remains permanent between the proper name and the individual referred to. This is not the case with common names, where the scope and extension of names may change in different conditions (in different possible worlds). However, the question arises as to how names (especially proper names) can fulfil their function if they are very different from the denoted entities in an important philosophical quality. What are we talking about?

When talking about name holders and names that we use for identification, we always talk about specific individuals (people, group of people, institutions, geographical locations places, books, IP addresses, etc.) that can be placed in some kind of category. Their existence, concreteness and uniqueness cannot be disputed: they are all unique in the sense that they can be always localized in space and time.[2] However, when we refer to these

[2] A defining quality of individuals is that they are always connected to space and time. In the case of an individual we must always be able to determine (in principle) where he/she "is located" in space and time. A building, a mountain, a person, a book, a video cassette, a CD can always be clearly localized in spatial and temporal dimen-

individual name holders with the help of their names registered in a given namespace, we can face a serious problem because of the time-and-space-independent quality of the names. Names can exist in more than one place at the same time. As far as a person is concerned, we can tell a exactly where he/she is at a given time, but we cannot say that "his/her name is here now". This is exactly the source of the phenomenon of homonymy. When we want to identify a person with a name, we might face the same problem: the same name can be assigned to several people. The big question is then the following: how can namespaces be used for identification if proper names themselves are not suitable for that purpose?

6. Unambiguous names

We can ask the question a little bit differently: how can (proper) names be made unique in such a way that they become fit for identifying individual name holders? The solution might be that for the identification we use *namings* (naming events or naming states) and not proper names. This solution can be traced back to Kripke's interpretation of proper names (Kripke 1980). According to Kripke, a reference of a proper name can be determined not by any description associated with it but by the naming events associated with the name in a historical chain:

> "[...] when I pronounce the name 'Ernő Rubik', its reference is the person who is determined by the historical chain (or rather web) of the use of names associated with my utterance. At the beginning of the chain there is the introductory use of the name, which is followed by the forwarding (repeating) use of the name, by which the name ultimately came to me." (Zvolenszky 2015)

Giving and using a proper name means creating and maintaining a rigid designator in the practice of a naming community. The new elements here are the concepts of naming and of naming practice. By means of "pointing" and naming gestures in a community's practice, a name gets attached to

---

sions. Geographical entities are to be determined in the first place by specifying where their components are located in a geographical area at a given time. In principle, we could identify any person by specifying the segment of the space at a particular time where he/ she is, but this would not be economical in terms of its informational needs, since we should know about each and every one their temporal and spatial locations at all times. It is more practical to identify remarkable events from people's lives and to link them to space and time coordinates. Such a notable event is the birth and death of a person, which we can "catch" with the space and time coordinates of the two events.

a name holder and this maintains a permanent relation between the name and the entity referred to through a coherent network of naming events. It is not the (proper) name itself that identifies but the naming practice of the naming community, the naming chain, which in turn means a chain of unique events. This is important because individual name holders can be identified through individual naming events. This will eliminate the philosophical difference between names and name holders, since a naming as event and a name holder are both individual entities. Of course, the same proper name can be applied to different name holders through different naming events, simultaneously or at different times. This interpretation can thus handle the phenomenon of homonymy that we mentioned earlier. Naming is not merely a one-shot action but a continuous naming sequence (or just a naming state). This solution is able to ensure the identification of specific names because a naming (as an event or state) can already be tied to space and time coordinates. This spatial and temporal localization of the naming event can be interpreted as a way of creating the context for the given proper name that allows the latter to become appropriate for unique identification.

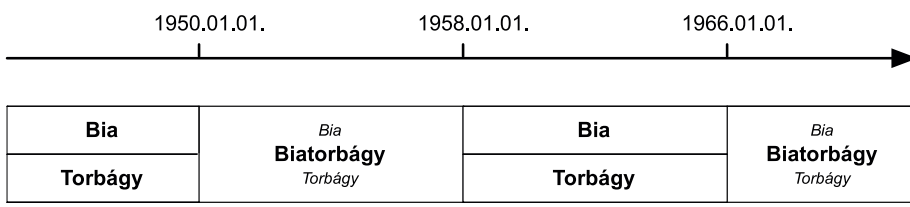
We have not yet mentioned the fact, but need to do it now, that names and naming events can belong to different types, so we can typify both names and naming events. As to person names, we can talk about nicknames or pseudonyms, official names, birth names, pen names, etc. These may all relate to different naming practices and communities, and of course it may be the case that several name using communities have the same name for someone. Viewing the naming practice as a context can also handle the phenomenon of two names being identical in two naming practices but still belonging to different name types within the two contexts.

After this long introduction, let us turn to the question of what follows from the theoretical considerations for the concrete practice of building namespaces. They have several important consequences.

Due to the fact that the namespace has two equally important components, the name and the name holder (the one that refers and the one referred to), both components must be handled in the namespace. In practice, this means that both name holders and names must be identified in some way, and (probably most importantly) there exists a third component, the relationship between them, and it should also be included in the namespace. The reason why this is important is because the phenomenon of synonymy is very common in the case of all types of proper names (person names, company names, geographical names, etc.), that is, a name holder

may easily have several proper names. This should be treated as follows: first, we identify the name holder and all the names associated with it in themselves, and, second, we record all the naming events as many times as the naming event (praxis) occurs in the case of the given name holder. Below we show why this is necessary.

It is enough to clearly identify the naming event as a context, there is no need for a deeper analysis. In a person name space, by connectingthe unique identifier of the name holder (*personid*) and the unique identifier of the associated name (*nameid*) in a manner that identifies the type of the connection as well we can capture the naming context itself. Let us look at the example of Biatorbágy (a Hungarian settlement). The independent villages Bia and Torbágy were united in 1950 and the name of the new settlement between 1950 and 1958 was Biatorbágy. In 1958 the two villages separated from each other again and officially got back the names Bia and Torbágy. Finally, they were reunited in 1966, and from then on the town is named Biatorbágy again. The two formerly autonomous settlements, Bia and Torbágy, were officially named by their original names twice, and the name of the united settlement was Biatorbágy in two periods. We thus have three names but six naming events (name usage practices). At the level of names, the names of Bia, Torbágy, and Biatorbágy have existed since their creation, they never ceased to exist, but at the level of naming events or naming practices they did not always exist. This example also shows that when we connect a name and a name holder in a naming event, it could happen that we build the same link between the same name and the same name holder in different times, but these connections can be instantiated and identified by different naming events.



Naturally, it is the names that are of primary importance for people and for them the multiple uses of the names (as it was shown in the case of Bia, Torbágy, Biatorbágy) do not need to be separated, but for a precise, machine based and scientific processing, these differences need to be handled.

In namespaces, we associate proper names with name holders through naming events. So far we have talked about namespaces as having only

proper names. After analyzing issues related to the management of proper names, we should also look at the issue of common names. There are language systems that contain common names and it can be useful to clarify their relation to namespaces.

We have already distinguished between the narrower and broader interpretations of namespaces. In the narrower sense, by the concept of namespace we mean a unique list of proper names, which can be used in the namespaces defined in the broader sense. We can also compile such collections from common names in order to characterize a certain domain of knowledge by giving the list of valid, useable common nouns (concepts) and the relations between these names (concepts). The latter is, however, a different kind of task and possibly yields different results. While proper names can only be organized in a semantically flat structure, common names can be formed into a semantically rich and complex structure. Such a structure of concepts can be used to represent our knowledge about the world.

7. Name structures – Knowledge Management Systems

A structured collection of names, i.e. a set of names and a set of relations defined on them is called a mathematical *structure*. In other knowledge domains, the concept of *Knowledge Organization System (KOS)* is used for the same phenomenon. We can distinguish a variety of knowledge organization systems, proceeding from simpler structures to more complex ones: the *term list*, the *classification system*, the *thesaurus* and the *ontology*. We can also differentiate among the latter according to the logic of their construction, and if we take into account the problem of control, we can distinguish between controlled and non-controlled KOS. Non-controlled KOS's are exemplified by the so-called *folksonomy*.[3]

Term lists simply list the names related in some sense, and apart from the lexicographic ordering, other relations cannot be defined on the elements of the set.[4] The elements of classification systems are linked by a subordination (containment) relation, resulting in a hierarchical structure of names. Since in many cases a single hierarchical relation is used instead of several, semantically different subordination relations when constructing classification systems, it is possible that a semantically inconsistent structure is created. The "weakness" of the classification systems is that only

[3] The term list, the classification, the thesaurus, and the ontology are controlled KOS.

[4] Proper name spaces with a narrower meaning can also be qualified as term lists.

one subordination relation is used to express both generic (subclass/class), partitive (part/whole), and even instance-of relations. This vagueness is eliminated by the thesaurus, which defines multiple relations (generic and partitive subordinate-of, instance-of, synonym-of, etc.). This creates a more complex structure than the ones before. Ontologies move further towards deep structuring in the sense that they allow the introduction of any arbitrary, formally defined relation, which can create an even more complex structure than a thesaurus. The knowledge organization systems outlined here are to be considered controlled systems in the sense that both the elements of the systems and the relations between them can only be introduced into the system by authorised people. Folksonomies are different in this respect because they are built without any control. Folksonomies do not handle relations between elements, but any person who uses them for something may add elements. The applicability and desirability of folksonomies is greatly enhanced by the possibility that tags generated by the users can be connected to the traffic (usage) data resulting from user activities, which can be used as a quality assurrance filter.

Knowledge organization systems collect the terms needed to describe the world, and through the relations between the terms they facilitate access to the terms themselves. The different kinds of knowledge organization systems differ from each other in their relations only, and theoretically there is no ontological constraint on the usable terms. In principle, each system is capable of covering any knowledge domain (ontology) with its terms.

Knowledge organization systems are, in a sense, insensitive to whether they have to handle proper or common names; it often happens that a particular knowledge organization system contains both proper and common names.

The concept of namespace can be applied to knowledge organization systems, which appears understandable and manageable on the basis of the considerations above. It would be reasonable to distinguish the proper name spaces and common name spaces more from each other but it is more important to know about a given system what types of names it is built up from. If relevant, proper name spaces can be marked with an appropriate qualifier.

After this sketchy review of knowledge organization systems, we have to answer the question of what such systems can be used for. They are most often used to provide a suitable set of terminology for a knowledge domain – to support content description and to facilitate document retrieval.

8. Name structures – Document Descriptor Systems

Both namespaces and knowledge management systems can be interpreted as structures whose semantics depends partly on the relations applied in the system and partly on the ontological commitments made in the course of the compilation of the term set. These systems can be applied to all knowledge domains, there is no area in which they could not be used.

However, there are other knowledge management systems that aim to deal with knowledge from a given knowledge area only, based on a specific ontological commitment. To narrow our focus to the world of libraries, we can mention the Marc21, RDA or BIBFRAME systems. They also have semantics, they can also be considered complex structures, but their functionality is different from the KOS systems discussed earlier. They can use proper and common name spaces, or knowledge organization systems, but their main purpose is more specific: the description of a certain type of documents. They can achieve this goal by outlining a scheme representing the ontological commitment and the knowledge of the given profession, in which everything can be said about the type of document under consideration. This scheme contains all the entities that are needed for a professional description of the given knowledge domain, as well as the relations between the entities and the points where "external" namespaces and knowledge organization systems can be made use of. These document descriptor systems may be connected to namespaces but their purpose and content is different.

The real significance of namespaces is the unambiguous identification of names in the ontological segment that they want to describe. By completing this task, the inventory systems of libraries, museums, archives, etc., which associate the identifiers of namespaces with their own identifiers can be made interoperable because they are able to refer to certain entities of the world in the same way due to the namespaces they all use.

References

Bittner, T., M. Donnelly and B. Smith. 2004. Individuals, universals, collections: On the foundational relations of ontology. In A. C. Varzi and L. Vieu (eds.) Formal ontology in information systems. Proceedings of the Third International Conference (FOIS 2004). Amsterdam: IOS Press. 37–48.

Kripke, S. A. 1980. Naming and necessity. Cambridge, MA & Oxford: Blackwell.

Zvolenszky, Zs. 2015. Pofonegyszerű, kedves Kripkém? (Saul Kripke: *Reference and existence*. Oxford University Press, 2013). Műút 2015. (http://www.muut.hu/archivum/15361)