Theme:

Title:

Establishing the basis for systems interoperability: the terminology challenge

Author(s): AMORIM, Sérgio R. Leusin de, D.Sc., Peixoto, Lucia; MADEIRA, Luiz C.M.Sc.NUNES, Roberta M. Sc

Institution(s): Universidade Federal Fluminense, Niterói, RJ, Brasil.

E-mail(s): leusin@civil.uff.br ;lucia@civil.uff.br; madeira@civil.uff.br

Abstract: *System interoperability proposals have been, until now, largely based on text contents. IFC, STEP, XML are languages or standards that need a consensual signification behind their terms. Each term can have many different meanings, depending on the context in which it is inserted and to include a full description of each one in every document is too complicated and irksome. It would be necessary, somewhat, to be able to select the meaning respecting the context, and this would probably consume too many resources.*

Any text-based system must run over a terminology which is well accepted and perfectly understood by its users. This is one of the main obstacles encountered in reaching full interoperability. Not only because we seldom have, in any language, a technical dictionary for the construction field, which could work as an official reference, but also due to the difficulties of translating between languages. Distinct building cultures and organisations make it difficult to obtain a proper translation.

One way to overpass this hindrance is to develop a reference construction terminology, including concepts description and terms interrelationship. The concepts must be associated with a conceptual map, which reflects terms associations and pertinence. This map can be linked to the OCCS Overall Construction Classification System, in order to make the correlation easier.

The relationships established through this conceptual map can be easily translated to any system, especially to XML schemes and data descriptions, making it possible to fulfil a real interoperability.

Building this terminology is the aim of a Brazilian government- sponsored project, so called CDCON, now in course under the co-ordination of a group of universities and participation of building contractors, suppliers associations and systems developers.

Keywords: *Interoperability, terminology, design process, design detailing*

Introduction

Although in the last few years significant efforts [1] have been made in order to implement an effective inter-operability of construction-oriented systems, practical results have been disappointing to a certain degree, as already indicated by NEDERVEEN, WOESTENENK and also TOLMAN, among others.

Little progress has been made towards the effective implementation of systems that would adopt the parameters required to allow full system integration, with a clear flow of information. In practical terms, multiple translations are still needed in each system swapping, resulting in losses and errors.

The actions developed until now have been ruled by high level integration, where a good part of the operations take place at the moment when the user accesses, as in the systematic case proposed by TOLMAN. It is also highlighted that "*integrated processes and teams have been indicated as one of the five key drivers of change for building industry* (EGANS, 1998).

As in any other text-based computer language, these proposals, supported by a XML language and its derivations, need to be translated into current or "normal" language at the user end. For that purpose, they

[1] See http://iaiweb.lbl.gov/ , www.econstruct.org and also http://www.icis.org/

depend fundamentally on two supporting points: the existence of standard terminologies for construction in each idiom and a typology of construction objects or construction taxonomy.

Specification, negotiation and vocabularies

WOESTENENK proposes a common vocabulary for construction, based on the "Built object" concept, an idea which has been developed into a central dictionary as suggested on the website www.econstruct.org ... But, there is a problem for its usage since in the course of a process development its components go through a progressive definition in the management of project and construction processes, the cases being rare in which a perfect specification occurs from the conception starting point – inception. On the contrary, this definition is usually managed by various intervening factors during the course of the productive process. It means that specification is not a finished object which is incorporated to the project at a given moment; it is the result of lengthy communication and negotiation processes between different agents.

The definition of the objects that compound the work deepens during the project development process, in a progressive and incremental manner. In this way, the architect first defines the existence of the wall then outlines some of its dimensions and basic components, until he reaches a list of effective components. Furthermore, it will not be always necessary to reach the final level of definition; one may want a kind of definition that is able to leave greater margins of choice to those responsible for the execution of the project. For example, it is possible, in some projects, to establish that one component must be installed without specifying colour or model, or to define boundaries for some performance characteristics, leaving the final choice of a solution to the builder. Although this situation might seem odd to an European, the flexibility in the execution is common and even necessary in less developed countries, as well as indispensable in performance-based projects.

In addition, there is the traditional problem expressed by the Italian saying, *Traduttore, traditore* (translator, traitor), since cultural differences lead to different interpretations of the same word, and so it will be necessary to contextualise in order to be understood. Even if we consider that "*technical terminology has the higher probability of giving unequivocal translation equivalence", this "correspondence", is by no means complete, but once the technological equivalents established, they can offer a relatively lower level of difficulty"* (MXITRO, 1997).

The latent question is how to establish this technological equivalence. Although it is of a much simpler nature than a literary translation, there are a great number of situations in which it is necessary to contextualise a term in order to get the proper translation.

For example, "tijolo" is the literal Portuguese translation, in Brazil, for the English word "brick", but this term refers to really many different products. Besides, it is possible that complementary definitions will be needed in one culture which has product regulations, while other situations require to stipulate whether the brick has the component "a" or "b" and even some of its dimensions. Furthermore, developing construction taxonomy means to classify objects according to certain logic. This involves two aspects: the adoption of a logical reference and the possibility of this logical reference to adapt itself to a specific region or country, which means to be valid in face of different requirements.

Logical references mirror determined ways of organizing a process, in the case the construction process. But this organization goes through many variations from country to country, sometimes inside a single country. This is a point rarely approached by the works about inter-operability, so far. As far as construction is concerned, there are important organizational and cultural differences in central countries, particularly European, and in developing countries. Even inside relatively homogeneous blocks, these differences may alter word meaning and relevance, either in nomenclature or in product characteristics. This has consequences for the meaning of the employed terms and thus for the clarity and adequacy of descriptors and specification terms.

It is possible, in theory, to adopt an "optimal logic", or organizational reference model, but, besides the difficulties of establishing a standard, this model would limit system acceptance and as a result the proposed systems would only be universally effective once the organizational model had been granted full market leadership. On the other hand, the incompatibility between the proposed systems and the current models is one of the reasons for its little usage, in a kind of vicious circle.

Another aspect to be considered, as regards the proposal's adequacy for international markets, is the diversity of normative tools available and their acceptance by each country's market. Although a great effort has been made towards its harmonisation, there is still a long way to go for full equivalence and even longer for their acceptance and effective usage at markets outside central countries. At present and in the near future, a situation prevails where the deficiency or even the lack of a normative framework leads to different specification requirements. These might be fewer in the presence of a solid normative system, but could increase in its absence.

Proposing a structured terminology

This assembly of considerations has led to CDCON Project development proposal, in Brazil. It came as a response to a FINEP Edict, at the HABITARE Programme, presented by construction oriented research teams at three federal universities, UFF, UFSC, UFRGS, gathered in ANTAC – Associação Nacional de Tecnologia do Ambiente Construído – National Built Environment Technology Association-(for more information on this association, access www.antac.org.br).

Its objective is to consolidate a terminology with associations and logical relationships between terms, defined by the approach to construction processes, to guarantee a system inter-operability basis. At first limited to the building sub-sector, this project includes in its guidelines both attention to the peculiarities of the Brazilian situation and respect for compatibilities between inter-operability proposals developed in the international arena.

Differing from a dictionary, where one finds only a list of words and its concepts, this project, besides consolidating a national terminology, now scattered and sometimes conflicting, intends to conceptualise each term and establish their associated relationships, synonymies and partitive relations, respecting organizational logic and the practices present in Brazilian construction.

Although compatibility with international systems is one of our objectives, in its still early days to know whether and how it will be achieved.

Perhaps it is possible to establish this compatibility through simple tabulation or perhaps more sophisticated components will be needed.

The term associations and other interdependency relations, such as synonymies and partition, allow the contextualisation of terms, by inserting them into conceptual maps[2]. In a first moment, it will available in graphical interface, and, afterwards, direct access to servers for corporative users will be implemented. The graphical interface is important for the presentation of the interrelationship of the concepts, since it facilitates a quick understanding by the user. It may be focused on a term and its directly associated words or it may present various association levels, taking the shape of a conceptual map.

The term association network, derived from its associative, partitive and synonymous relationships, is a useful database for the knowledge systems management and search for products ("e-procurement") etc. It may be made available through direct access to a database or through its cession, with periodical updating, since the alterations are not supposed to be that frequent.

In this way, the project presents as results of both a reference system for its contents and the possibility of its usage as input to other systems, providing a unified reference base. Upon that, one may develop systems able to recognise products and services specifications and descriptors.

Another positive aspect is cost reduction for those systems located at the information consumer's end. Whether being an architect seeking a product specification or a supplier who wants to show his product catalogue, they can do it without setting up and updating the database, which would be shared by various organisations.

Terminology Structuration

The best way to structure terminology is to relate it to the predominant logic among the users of the language in question. In the case of construction, organised around spaces, processes and components, in that order, it is natural to privilege these aspects. These basic concepts, allied to others which may be of interest, will constitute the main categories or facets to organise terms. The using of categories in concept

[2]For samples and concepts of conceptual maps see CANAS

organisation, and thus, classification, is a resource to understand the nature of concepts and to create conceptual structures. Facets possess the property of allowing knowledge systematisation. These categories are easily visible starting from term characteristics, conceptually defined inside the approached area of knowledge. Category definition is nothing more than the using of faceted classification.

The Faceted Classification Theory was developed by Shiyali Ramamrita Ranganathan in the 30's (See CAMPOS, 2001). It is based on the need for the creation of classification schemes that may follow changes and knowledge evolution.

The facets scheme "breaks" the subjects into fundamental categories. This means that the nature of concepts is analysed and split into homogeneous groups, according to particular characteristics. These divisions compose the classification structure's spinal cord.

Terms are ordered in different classes and each class is a group of terms defined by classification relevant properties, the facets.

A facet is a comprehensive assembly of similar quality properties. The properties, which determine classes in a given area of knowledge, may be sorted out by a crescent order specification, from general to particular. Properties of superior rank are general and properties of lesser ranks are specific.

The classification aim is to distinguish the terms in a group, singularising them. In order to give a comprehensive character to the classification, each term in a group must be attributed to a class, and to give exactness to the classification, each term must belong to only one class. Without these criteria, there would be unclassified terms and multi-classified terms in the same rank. In both cases, classes would not be accurately defined (EKHOLM, 2000).

A faceted classification allows the free combination of a group of properties that is able to characterise a term, as well as to accept new terms for classification. The faceted classification differential factor is the quality capacity of creating new search strategies and their applications for intelligent systems.

The main difficulty is how to achieve compatibility of the objects concept, at that moment represented by terms, with the incremental progressive character of the construction development process.

This has been achieved, traditionally, through reference hierarchy, with the creation of classes and subclasses that reflect the depths of this project process. It implies, however, a rigid structure that has been pointed at by many authors as a restricting factor. Furthermore, that structure might be adequate to a particular environment and language but could become inadequate after translation.

A more dynamic approach, where the employed facets could be freely combined or privileged by the user, would allow the system to adjust itself immediately to the usage criteria. In that case, there would be not one but as many facets as defined. The visual inter-relationship is constructed from each facet's internal logic.

For this purpose, each facet must be structured with its own hierarchy, which could be defined through concept aggregation levels. This approach is similar to that adopted both by OCCS tables and ISO DIS 12006-2

In this way, we have multiple hierarchies which may be freely associated, which allow us to have multiple entries, in associative structure. And so it is possible to get round the limitations of classification systems based on a sole hierarchy, which imply loss of contents given the difficulty, or perhaps impossibility, for a single logic to encompass all the kinds of objects relation. This idea corresponds to the general outlines for a "core terminology", described in Report ISO/TC59/SC13 (1998)

At the CDCON Project, from the premise of limiting the universe to the building context, the following categories have been adopted, representing the various playing facets:

Construction Processes (Process Facet): *Group of activities whose completion results in the Building product*

Construction Products (Component Facet): *Materials and products consumed at production level.*

Building Elements (Element Facet): *Products of a constructive process.*

Building Spaces (Space Facet): *Construction parts, delimited according to their spatial usage.*

__Building Typology__ (Using Facet): *Different uses for construction products.*

__Building Attributes__ (Attribute Facet): *Classification of the construction objects characteristics.*

Creation, availability and updating of the databases

After setting the guidelines for the terminology structure, the next question is how to proceed to create, make available and update databases, since we refer to quite a dynamic universe.

If we consider that terminology is a "nutrient" for other systems, we must establish means to the existence of tools for system retro-feeding during its usage to facilitate not only its updating, but also to allow roofing increase, through the procedure. In the CDCON case, a cooperative network of suppliers, builders, engineers/architects and government agencies have been established both for results availability and setting up database. Using a cooperative working tool through the public computer network makes it possible to expand participation beyond the boundaries of a single organisation and, at the same time, ensures immediate availability of results. The general scheme is presented in **Fejl! Ukendt argument for parameter.** . The availability of results "on line" allows them to be inserted dynamically into other systems, through server inter-action. We believe that the length of partnerships will propitiate a sufficient reaction to expand this basic reference using to the building sector's mainstream, in the future.

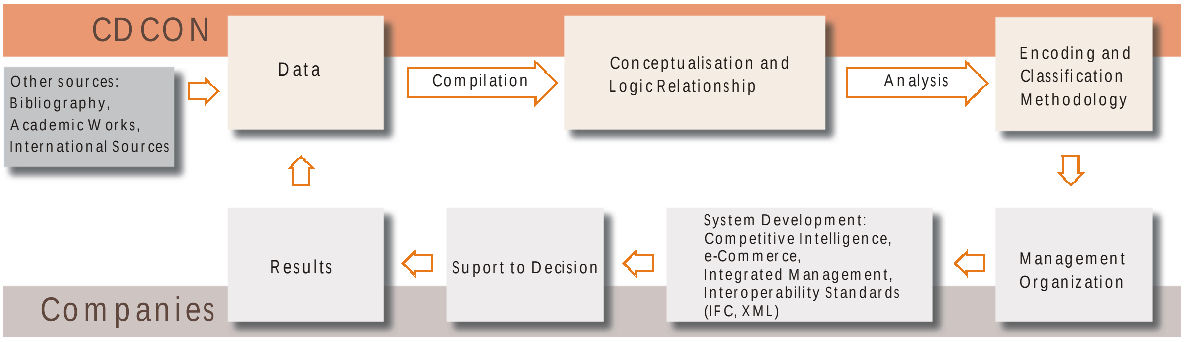

Figure Fejl! Ukendt argument for parameter. **Cooperative scheme for Database operation**

Proposal Outline:

One of the proposal's most interesting results is the definition of associated characteristics for each product type or subtype. By establishing a relation between products and services concepts and characteristics concepts, referred to dimensions, performance or even marketing, we create a standard map for their specification. That is, when we mention a construction element, the system returns the associated terms, including their attributes".[3].

Through the recurrence of associations and by the associated hierarchy to each facet, either for elements or construction components, we will have a parallel attribute hierarchy. For example, if we talk about "roofing", we will get some particular attributes, such as heat performance coefficient. A component associated to "roofing" will have the same need for describing this co-efficient and even if its own co-efficient is different from the whole, it will always be part of a composition. There will be, however, other characteristic attributes only for that component. This aspect makes possible to set up of standard descriptors for components, processes and further descriptive building items, indicating their availability to other systems, in a structured way, and, thus, making it much easier to assemble catalogue systems, with consistent and inter-changeable information, or CAD systems that encompass the necessary data to move and re-use their specifications for other systems that come in the wake of the construction production process.

[3] OCCS defines attribute as "any characteristic of a construction object. All of the entries in the other tables are objects, expressed as nouns (for things) or verbs (for activities). Attributes are adjectives and other descriptors and modifiers"

In addition, the relation structure of concepts allows the development of dictionaries that are not restricted to a univocal relation among terms, always passible of errors, giving a chance to multi-lingual systems to have a consistent approach, if not to all languages, at least to the ones that are supportive of similar proposals.

A question which remains open is the standard formatting for the distribution of the terms inter-relationships. One must consider the original format when adding information from various sources, such as product suppliers or representatives from different origins. There is no point in waiting for conversion into a common standard. The CDCON database tends to be considerably sized. Inter-operability is very well handled with the usage of XML standards. This technology, however, is still on the boil, it evolves in a remarkable way and thus it turns difficult its universalization and immediate acceptance, especially for very lengthy documents.

However, one must take into account, that implementation via Internet presupposes the inter-action between human beings and computers, which the pre-requisite is the existence of a common ontology (semantics), and also that there is not yet consensus for a common taxonomy (implementation). All efforts to go beyond information (extraction and evaluation) [ifcXML, 2001] are initiatives started by automatic translation. And, of course, translation demands full knowledge of the involved languages, meaning that as long as there is not a consensus about ontology, taxonomy and a "neutral" language, every effort towards availability will depend on great doses of replication.

We believe that the implementation of XML gateways, as intended by the European Industry for Building and Construction (TOLMAN et al, 2001), will be the most adequate way for Information Systems convergence. Another possible alternative is to break of the basis into sectors, in order to obtain portable documents in current technology, from now on.

Conclusions

Compiling a reference terminology is a basic step to enable systems to be developed by unified construction semantics, making full operability viable. National terminologies, if structured in a similar way, may make viable multi-lingual systems, wich would be more efficient and precise. The structured approach according to the usual construction logic allows terminology to reflect the practice of the incremental definitions, which characterise the conception and production processes in Construction, making it easier to develop systems that respect this fundamental characteristic and resulting in faster assimilation by users. However, it is not yet clear which technology would be the best suited for setting up terms' inter-relationships availability standards.

Bibliography:

CAMPOS, Maria Luiza de Almeida. (2001), *Linguagem documentária: teorias que fundamentam a sua elaboração*. EdUFF, Rio de Janeiro.

CAÑAS, Alberto. (1995), *Knowledge Construction and Sharing in Quorum.* Proceedings of the 1995 AI in Education Conference.

EGANS, Johns. (1998), *Report of the Construction Task Force to the Deputy Prime Minister*. http://www.dti.gov.uk/construction/rethink/report/index.htm.

EKHOLM, A.; FRIDQVIST, S. (1996), *A conceptual framework for classification of construction works.* Itcon, Sweden. http://www.itcon.org

ISO/DIS 12006-2. *Organization of information about construction works — Part 2: Framework for classification of information.* http://www.iso.ch

ISO/TC59/SC13 (1998), *Terminology for object modeling for construction information.* http://www.iso.ch

MIXTRO, Vera Lúcia, (1997). ***O Tradutor - Vítima ou Traidor?*** http://www.upis.br/revista_multipla/nr_002/vera.htm.

NEDERVEEN (2001) *Neutral object tree support for inter-discipline communication in large scale construction.* ITCON vol. 6 http://www.itcon.org

OVERALL CONSTRUCTION CLASSIFICATION SYSTEM – OCCS. (2001), *Preliminary Draft for Review and Comment*. http://www.occsnet.org/.

TOLMAN**,** Frits, (2001). *Econstruct: expectations, solutions and results***,** Itcon. http://www.itcon.org/2001/13/.

WOESTENENK, Kees, (1998). *A Common Construction Vocabulary*. STABU-The Dutch National Building Specification, Netherlands. http://www.natspec.info/Toolbox/AboutCIS/R&D/Comvocab.PDF

1.