Ranganathans Canons applied to ontology engineering: a sample application scenario in biomedical ontologies

**Linair Maria Campos[1], Maria Luiza de Almeida Campos[2],
Maria Luiza Machado Campos[3]**

[1] Universidade Federal do Rio de Janeiro – CISI/COPPE
Av. Horacio Macedo S/N - Ilha do Fundão – Bloco H-201- Rio de Janeiro - RJ

[2] Universidade Federal Fluminense - PPGCI/UFF
Rua Tiradentes 148 - Ingá - Niterói - Rio de Janeiro - RJ

[3] Universidade Federal do Rio de Janeiro - PPGI- IM/NCE
Av. Athos da Silveira Ramos, 274 - Ilha do Fundão - Rio de Janeiro - RJ

linair@cisi.coppe.ufrj.br, maria.almeida@pq.cnpq.br, mluiza@ufrj.br

Abstract. *Ranganathan, an Indian mathematician and librarian, has proposed a set of comprehensible canons to provide guidance to the process of building concept hierarchies. It is our proposition that Ranganathans canons can contribute to fulfill the gap between the high-level domain conceptualization guided by top level ontologies and the classification of such concepts within facets, needed when building ontologies taxonomical structures. In order to show the utility of Ranganathans canons applied to ontology structuring, we have analyzed the structure of a biomedical ontology: Gene Ontology (GO). As result, we have found that many of the existing inconsistencies on GO hierarchies could be avoided if Ranganathans canons were adopted.*

1. Introduction

Ontologies have been increasingly used since the early 90s, especially in complex domains such as Biomedicine, where the multitude of concepts and the need to deal computationally with resources described by them, has urged the adoption of standard vocabularies. However, the fast growing nature of the body of knowledge being described and the necessity of fast solution to the issue of concepts standardization has given rise to vocabularies such as the Gene Ontology [Gene Ontology Consortium, 2001], which has been created without a sound methodology and has been largely adopted as a *de facto* standard, despite its many structural problems, as mentioned in literature [Smith e Kumar, 2004][Smith, Williams, e Schulze-Kremer, 2003], which affects its efficient utilization.

Notions underlying concepts nature, materialized in classes of top level ontologies, have given significant contribution to ontology structuring, as it allows domain concepts to be identified and grouped together in basic categories according to pre-defined basic features. These top level classes are usually chosen according to principles discussed in areas such as Philosophy, Cognitive Sciences and Psychology, providing a sound basis for identifying the nature of concepts in a less ambiguous way. However, once groups of concepts with the same nature are identified, there is still the need to organize them in arrays (horizontal series of sibling concepts) and chains (vertical series of concepts, organized hierarchically): this is one of the challenges of ontology structuring.

Underpinned by more than fifty years of hands on experience in information classification and structuring of big vocabularies, Shialy Rammarita Ranganathan, an Indian mathematician and librarian, has proposed a set of principles, or *canons*, [Ranganathan,1951;1963;1967a, 1967b], tailored to provide guidance to the process of working on the *level of ideas*. This level is the space where the concepts of a given domain are organized, building a system of concepts [Campos e Gomes, 2008]. These canons were meant to be used for bibliographic classification, in the context of the development of documentary languages with taxonomic structures such as thesauri and controlled vocabularies.

It is our proposition that Ranganathans' canons can contribute to fulfill the gap between the high-level domain conceptualization guided by top level ontologies and the systematic classification of such concepts within facets, needed when building ontologies taxonomical structures. In order to show the utility of these canons applied to ontology engineering, we have applied a set of those canons to the widely adopted, *de facto* standard, Gene Ontology (GO) [Gene Ontology Consortium, 2001].

As a result, we have observed the timeliness and relevance of his work, as a series of existing inconsistencies on GO hierarchies, could be avoided if Ranganathans canons were adopted.

The remainder of this article is structured as follows: in section 2 we present related work. In section 3 we discuss Ranganathans' canons. In section 4 we analyze the Molecular Function branch of GO in accordance with Ranaganathans canons. Finally, in section 5 we present our conclusions.

2. Related work

Ontology structuring has impact on knowledge reasoning, especially when based on subsumption relations. Reasoners expect ontologies to comply with certain rules of classification, and ill-formed hierarchies can lead to false results. On the other hand, implicit structuring strategies used to form subhierachies can hinder human comprehension of ontology classification rationale leading to ambiguity when using and extending ontologies.

In order to tackle those issues, researchers such as Guarino and Welty [2004] have proposed the use of philosophical notions, such as rigidity, identity, and unity, materialized on top ontologies, to guide the identification of concepts nature and, thus, to provide foundational principles to evaluate the conceptual correctness of specialization relationships [Guizzardi, 2005]. In this sense, Guarino and Welty have used those notions to underpin the *OntoClean methodology* [Guarino & Welty, 2004, 2002a, 2002b] aiming at building "clean" taxonomical structures on ontologies. Inspired by the work of Guarino and Welty, Guizzardi also has proposed a theory presenting a set of postulates aimed to aid the construction of well grounded conceptual models [Guizzardi, 2005; Guizzardi, Wagener and Sinderen, 2004]. The idea behind these approaches is to axiomatize a set of rules that can be applied systematically to taxonomies and on doing so, prevent structural errors. For example, based on the axiom that rigid concepts cannot be subsumed by anti-rigid concepts, the concept `human` (rigid) cannot be subsumed by the concept `student` (anti-rigid).

Smith (2005) observes, complementary, the role of definitions to identify attributes "in a consistent manner, thus assuring their transitive inheritance through a type hierarchy", and points that the definition of a concept within an ontology should encompass the definition of all its parents. Besides, all intermediate classes in the hierarchy where the concept is situated should also be defined, in order to ensure transitive inheritance of essential characteristics. With the intention of establishing guidance for building well formed hierarchies, built according to a sound classification systematic, Smith proposes a set of axioms for a "Theory of Biological Classification". According to Smith (2005), those axioms were motivated by the theory of classes found in Aristotle's writings. As an example, we can mention the axiom that addresses the issue of polihierarchy and states that a species should never have two parents:

$$\text{lowestspecies}(A) \wedge \text{lowestspecies}(B) \wedge A \neq B \rightarrow \neg\exists x(\boldsymbol{inst}(x, A) \wedge \boldsymbol{inst}(x, B))$$

In Information Science, Dahlberg (1978a, 1978b) also stresses the importance of definitions as they make explicit the contents of concepts and provide the elements that forge the relationships between them. Dahlberg, through her Concepts Theory, proposes also that definitions reveal a set of common characteristics which are useful to build any system of classification or thesaurus [Dahlberg, 1983]. Dahlberg, however, focuses on proposing principles to organize concepts in broad categories, and, although highlighting with examples the importance of definitions when structuring hierarchies, she does not provide detailed guidance on how to organize them systematically in subclasses.

To exemplify the problems caused by the lack of a systematic approach on structuring an ontology taxonomy, Smith points out several issues in the Gene Ontology (GO) hierarchies. Although his axioms can help to identify solutions to those problems, members of the GO community were not very receptive to the proposal, perhaps due to the complexity that comes with it: "(…) When challenged with such problems, the members of the GO and associated communities standardly insist that their concerns are those of practicing biologists, and that they are thus not concerned with the sorts of scrupulousness that are important in logic" [Smith, 2005].

Guarino and Weltys' (2004) as well as Smiths' (2005) proposals have the focus on identifying concepts nature and on providing rules and axioms to help the identification and grouping of concepts with same nature in a consistent way. The inspiration behind the idea of identifying concepts nature, which has its roots in Philosophy, has been used since the 60s in the context of library classification by Ranganathan, who also adopted fundamental categories to help to identify and group vocabulary concepts according to their high level nature. In this context, Ranganathans' categories provide a more intuitive, transparent, although less formal (and consequently more ambiguous) way to approach categorization.

It is worth noting that if, for one hand, the use of formal axioms can improve ontology structuring, when applied by ontologists with some expertise in logics and with some background in Philosophy, on the other hand it can represent a challenge for domain experts to deal with the inherent complexity of such philosophical notions and the formalisms used to express them [Yu, 2006].

However, although the identification of concepts nature has a major role in structuring well formed and consistent hierarchies, there is more to it than that. There is still the need of detailed classification principles to help organize concepts in subclasses, and, besides, if possible, that those principles could be more easily assimilated by the community responsible for creating and maintaining the ontologies. In this sense, to the best of our knowledge, few proposals provide a systematic set of principles to address the problem. Even so, some are directed specifically to a given subject, such as *folk biological classification of organisms* [Berlin, Breedlove, Raven, 1973] or *construction works* [ISO DIS 12006-2, 1999], while others [Ekholm, 2002] are focused on identifying objects properties, which even though helps on identifying facets of interest, does not present a solution to the issue of organizing them in a more thorough way.

Bodenreider and others (2004) point that there are some principles of good classification that (biomedical) ontologies are expected to be compliant and that, as they believe, "rest on a wide consensus among those working on biomedical terminologies". Such principles can be summarized as: (i) each hierarchy must have a single root; (ii) children should have exactly one parent; (iii) non-leaf classes must have at least two children; (iv) each class must differ from another class in its definition. In particular, each child must differ from its parent and siblings must differ from one another. The authors, however, do not present evidences on how long their proposal has been adopted and how exactly the consensus was reached. Besides, proposing that a hierarchy must have a single root seems to limit the possibility to express different aspects of a domain, which, in a different perspective, could be easily presented as facets. Also, their proposal does not provide guidance to other important aspects of classification, such as the need to define homogeneous hierarchies, as pointed out by Smith (2005).

As observed, although there has been some concern with the adoption of systematic classification practices, preventing *ad hoc* built taxonomies, many of those practices seem to be recent and still need to mature. Also, some of them lack a more intuitive and transparent explanation, in order to allow a better understanding by end users and so, to avoid their rejection.

It is our proposal that Ranganathans' canons of classification, in use for more than fifty years, provide a methodological path that can join the convenience of a comprehensive explanation and a more complete and mature set of guidelines, which can be easily adopted to help building more consistent classificatory structures. The usefulness of these canons can be seen in a sample scenario for analyzing Gene Ontology main classificatory structure, as presented bellow.

3. Ranganathans' canons

In the present paper, we highlight two sets of Ranganathans' canons, which provide guidelines to the organization of classes of concepts: canons for the creation of arrays and canons for the creation of chains. Chains are vertical series of concepts, which can be organized hierarchically according to generic-specific relations, or according to part-of relations. Arrays are horizontal series of concepts, organized as siblings in relation to a parent concept.

In the specific case of the organization of ontologies taxonomical structures, we have selected a subset of Ranganathans' canons, of particular relevance to our purposes, and which we shortly present in the following sections, based on Campos e Gomes (2008) and also Gomes Motta e Campos (2006).

Some of the canons aim at organizing arrays, as for instance, the canons of Differentiation, Concomitance and Exclusivity, while others aim at organizing chains, as, for instance, the canons of Modulation and Subordinate Classes (Ranganathan, 1967a). The canons provide principles that facilitate the creation of classes in a more consistent way, and, according to Ranganathan (1967a), their violation may result in ill-formed classificatory structures. The selected canons are explained next.

3.1. Canons for organizing arrays

Ranganathans' **Differentiation canon** states that a principle of division used as a classificatory basis should originate at least two classes. For example, let us consider the array used to classify catalectic activities of enzymes. That array can have a principle of division according to the kind of enzymes (hydrolise, isomerase, among others) and another principle of division according to the kind of reaction catalyzed by the enzyme (free radical formation, first spliceosomal transesterification, among others).

If the principles of division used to organize the arrays are explicit, it makes the classification of new concepts easier, as the comprehension of the rationale used to form the hierarchy helps to figure it out where is the right place for the concept within the ontology structure:

> (...) in a classificatory scheme, concepts that are subordinated to a more general concept can be grouped more accurately according to the principle of division that guided this grouping. Principles of division bring transparency to the vocabulary and so improve searches, locating and relating the concept according to its inner characteristics. [Novellino, 1996, p.1].

Ranganathans' **Concomitance canon** states that two different principles of division should not result in the same array. For example, if we adopt the criteria of year of birth and age to classify a set of individuals, we will have as a result arrays constituted by the same elements.

Ranganathans' **Exclusivity canon** states that elements belonging to an array should be mutually exclusive, i.e., disjoint in relation to elements belonging to another array. For example, the term `multidrug transporter activity` should not be subordinate to both arrays `transmembrane transporter activity` and `drug transporter activity`. Even if those arrays are organized according to different principles of division (in the above example, according respectively to the principle of the local – `transmembrane` – where the transport occur and according to the kind of element – `drug` – which is transported).

3.2. Canons for organizing chains

As classificatory principles for chains, we highlight the following canons of Ranganathan: Modulation and Subordinate Classes [Ranganathan, 1967a] as explained next.

The **Canon of Modulation** states that within a hierarchical classificatory structure of concepts there should be a gradual specificity when organizing concepts in chains, allowing thus a "conceptual consistence between the classes of concepts" [Gomes, Motta e Campos, 2006]. For example, let us consider the terms `helicase`, `ATP-dependent RNA helicase` and `ATP-dependent DNA helicase`. According to the Canon of Modulation, the last two terms should not be directly subordinated to the term `helicase`. It should exist, between the first and the last two terms, a term like `ATP-dependent helicase`.

The **Canon for Subordinate Classes** states that in a hierarchy of classes, the classes nature should be the same, i.e., they should conform to the perspective adopted as principle of division that guides the organization of the array. For example, in GO, the definition of the term `binding` indicates that the principle of division of its array has to do with interaction between molecules. This makes us believe that terms like `bacterial binding` (Interacting selectively and non-covalently with any part of a bacterial cell) and "`extracellular matrix binding`" (Interacting selectively and non-covalently with a component of the extracellular matrix) fall in conflict with such principle, for a *bacterial*, which is an organism, and an *extracellular matrix*, which is a cellular component, both have different natures from "interaction between molecules", which is a process. According with Gomes, Motta e Campos (2006), this canon complements the Canon of Modulation, and, if it is not violated, the affiliation sequence of the chain to the array is correctly assured.

4. Analyzing GO (Molecular Function) and observing Ranganathans' canons

Our considerations about the utility of Ranganathans' canons are made in the context of the analysis of the *ontological commitment* of the Molecular Function branch of GO.

Ontological commitment can be defined briefly as an agreement shared by a community about the consensual meaning intended for the ontology, not only considering its comprehension by humans, but also considering its computational processing by software agents. We assume that this commitment is not always precisely explicit, however it can be identified, although partially, by means of the existing ontology documentation, the analysis of concepts definition and the metadata associated to ontologies terms. On retrieving the ontological commitment, we expect to have as results the criteria observed as classificatory principle for the organization of first level hierarchies of GO's Molecular Function branch, together with the problems observed as consequence of the adopted classificatory approach.

The choice of GO´s Molecular Function branch was due to the fact that this branch has intermediary complexity, if compared to GO´s Cellular Component (less complex) and Biological Process (more complex), as we could notice when analyzing the terms definition and also the composition of the first level hierarchies (considering the hierarchies' depth, number of terms and existing relationships). In this sense, this choice is adequate to our purposes, for if, on one hand, it provides richness of issues to explore, on the other hand it minimizes the complexity, already big, of the analysis of the ontological commitment of GO.

The analysis of the ontological commitment is made upon the analysis of the ontology's hierarchical structure, the terms nomenclature, and especially of the terms definition, which is of main relevance to the formation and comprehension of domains classificatory structures, once they provide rich semantics about intended meaning of concepts.

For the sake of space, however, we are not going to reproduce the definition of subordinate terms, but only data about the first level term being analyzed, and the final result obtained, i.e., the criteria observed as classificatory principle, along with the summary of problems found, considering the array being analyzed. In order to help understanding the analysis of the `binding` array, we present on Figure 1 its immediate subordinate classes. The `binding` array contains more than 1000 subordinate classes.

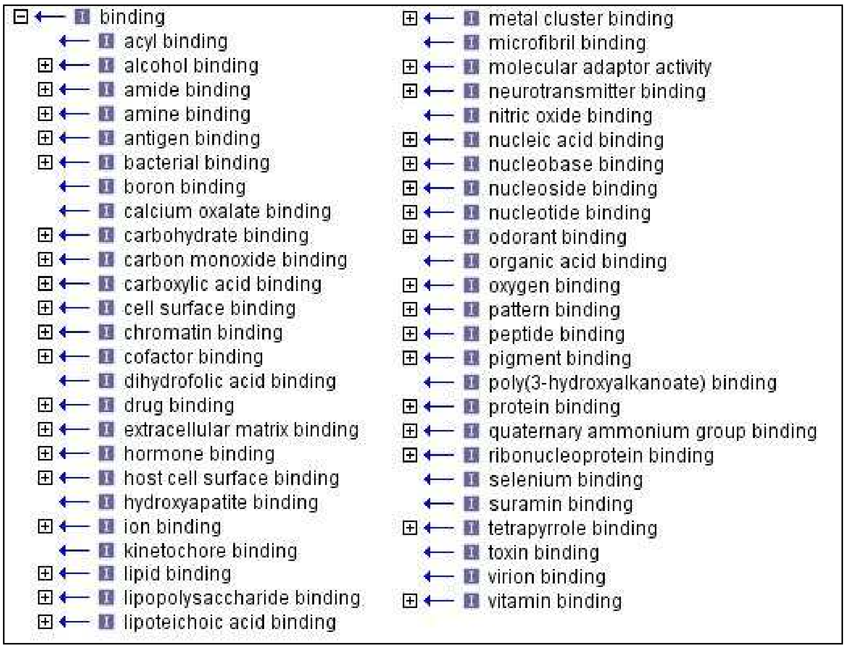

Figure 1. Subordinate classes of GO´s Molecular Function `binding` array

Results of the Analysis of the `Binding` array

Identification: GO:0005488 - Binding

Definition: *The selective, non-covalent, often stoichiometric* [1], *interaction of a molecule with one or more specific sites on another molecule*).

[1] The quantitative relation of the products and reactants of a chemical reaction in the proportion they appear in the chemical equation which describes the reaction [Smith et al., 2000].

Narrow Synonym: ligand

Classification[2] criteria observed:

- Chemical elements (ex: `boron binding`), organical compounds (ex: `lipid binding`), non organical compounds (ex: `nitric oxide binding`), kind of ion (ex: `ion binding`), organical radicals (ex: `acyl binding`), clusters of atoms (ex: `metal cluster binding`), role of molecules (ex: `antigen binding`), cellular locations (ex: `cell surface binding`).

In order to obtain the classificatory principles, besides analyzing the ontology hierarchy and terms definition, it was necessary to refer to specialized literature, in order to understand the nature of certain terms. For example, in the case of the term `Acyl binding` (Interacting selectively and non-covalently with an acyl group, any group formally derived by removal of the hydroxyl group from the acid function of a carboxylic acid), we came to the conclusion that "acyl" refers to a group of atoms (or radical), due to the fact that GOs[3] term definition refers to the term `acyl group`, which can be understood as a group of atoms or radicals (see definition of `group` and `acyl` in Oxford Dictionary of Biochemistry and Molecular Biology) [Smith et al., 2000] [4].

It is worth remembering that the analysis of GOs hierarchical structure, if carried out by a domain expert, or if applied more thoroughly and deeply in the ontologies hierarchies, could bring richer results, possibly with a wider range of observed problems.

Problems observed:

- **Violation of the Canon of exclusivity**

  Example:

  - `**norepinephrine binding**: Interacting selectively and non-covalently with norepinephrine, (3,4-dihydroxyphenyl-2-aminoethanol), a hormone secreted by the adrenal medulla and a neurotransmitter in the sympathetic peripheral nervous system and in some tracts of the CNS.`
  - This class is subordinate both to `alcohol binding` (binding/alcohol binding/norepinephrine binding) and to `amine binding` (`binding/amine binding/` norepinephrine binding).

- **Violation of the Canon of modulation**

  Example:

  - `**nitric oxide binding**: Interacting selectively and non-covalently with nitric oxide (NO).`

---

[2] For each organization principle observed, we have put in parenthesis, and italics, a term exemplifying those principles.

[3] Interacting selectively and non-covalently with an acyl group, any group formally derived by removal of the hydroxyl group from the acid function of a carboxylic acid.

[4] This dictionary appears as bibliographic reference within comments in GOs terms.

- Nitric oxide is a drug, according to the definition of the term `drug binding` (Interacting selectively and non-covalently with a drug, any naturally occurring or synthetic substance, other than a nutrient, that, when administered or applied to an organism, affects the structure or functioning of the organism; in particular, any such substance used in the diagnosis, prevention, or treatment of disease) and literature [Gerlach and Falke, 1995]. Therefore, according to the canon of modulation, the term `nitric oxide binding` should be subordinated to drug `binding`.

- **Violation of the Canon of subordinate classes**

  Example:

  - A `trisaccharide` (Interacting selectively and non-covalently with any trisaccharide. Trisaccharides are sugars composed of three monosaccharide units) is an `oligosaccharide` (Interacting selectively and non-covalently with any oligosaccharide, a molecule with between two and (about) 20 monosaccharide residues connected by glycosidic linkages), therefore its subordination to the class `sugar binding` (Interacting selectively and non-covalently with any mono-, di- or trisaccharide carbohydrate) violates the canon of subordinate classes, i.e., `trisaccharide binding` should be subordinated to the class `oligosaccharide binding`.

When tabulating the problems found, we highlight the importance of the canons of exclusivity, subordinate classes and modulation, as having the greater number of violation occurrences (7), followed by the canon of differentiation (5). There were not found evidences of violation of the canon of concomitance (although it is worth remembering that the analysis conducted did not cover thoroughly the complete deepness of GO´s hierarchies).

Table 1. Total occurrences found on analyzing GO first level classes

Canon	Total violations
Canon of Differentiation	5
Canon of Concomitance	0
Canon of Exclusivity	7
Canon of Modulation	7
Canon of Subordinate Classes	7

The analysis of first level hierarchies of GO´s Molecular Function branch shows a diversity of problems, which are materialized in a variety of non uniform classificatory principles observed, which seems to indicate a lack of adoption of well defined classificatory principles, gap which could be fulfilled by the adoption of Ranganathans' canons. In particular, we have observed the violation of Ranganathans' Exclusivity Canon, which points to the relevance of understanding existing perspectives to think about the nature of the domains concepts. In contrast, we have not observed[5] the violation of the Canon of Concomitance. This finding could be evidence that some classificatory principles are more intuitively assimilated than others, but could be as well due to the characteristics of the domain, or yet, due to the deepness of the analysis conducted.

5. Conclusion

Ontology construction, although a maturing research field, still faces many challenges, especially in domains with a rich variety of complex concepts, and whose knowledge advances dynamically. In Biomedicine, for example, the need to organize concepts in a systematic way has to cope with the pragmatic nature of its community, with no deep knowledge of Ontology related disciplines such as Logics and Philosophy, but with urge to improve their ontologies.

One of the challenges of ontology construction is the creation of classificatory structures, or the backbone taxonomy, with its subclasses organized in a systematic way. The challenge presents itself not only due to the complex nature of the domains, but also due to the interdisciplinary nature of ontology building, which demands knowledge of experts in knowledge organization, such as Computer Scientists and Information Scientists, but especially, end users who detain the knowledge of the domain and the intended meaning of concepts contained in their ontologies. Considering that it is important to provide the grounds for an effective dialogue between people with different backgrounds and, at the same time, to provide principles that can rapidly and easily be assimilated and adopted on ontology structuring, it is important to overview existing and previously successfully adopted initiatives. In this sense, Information Science can provide relevant contribution, as it has mature hands on experience of information organization for more than fifty years, classifying subjects from many different areas.

Ranganathans' canons, some of those (but not all of them) were presented in this paper, can bring an important contribution to ontology structuring as it provides a comprehensive set of principles, that can be easily assimilated, and that provide effective guidance to avoid many of the problems found in ontologies structures, as we could observe by analyzing Gene Ontology. Although Ranaganathans' canons have been originally proposed long ago, they are a mature set of guidance that have been used successfully by more than fifty years, and are still current and relevant nowadays, as we have presented on our application scenario.

References

Berlin, B., Breedlove, D. E., & Raven, P. H. (1973) "General principles of classification and nomenclature in folk biology", American Anthropologist, n.75, p.214-242.

Bodenreider O, Smith B, Kumar A, Burgun A. (2004) "Investigating subsumption in DL-based terminologies: a case study in SNOMED CT", In: Hahn U, editor. KR-MED 2004; 2004. Whistler, Canada: AMIA; p. 12–20.

[5] It is worth noting that with the help of a biologist, with deep knowledge of the subject, other violations could have been found.

Campos, M. L. A. ; Gomes, H. E. (2008) "Taxonomia e Classificação: princípios de categorização". Datagramazero (Rio de Janeiro), v. 9, n. 1.

Dahlberg, I. (1978b) "A referent-oriented, analytical concept theory of Interconcept", International Classification, v. 5, n. 3, p. 122-151.

Dahlberg, I. (1983) "Conceptual compatibility of ordering systems", International Classification, v. 10, n. 2, p.5-8.

Dahlberg, I. (1978a) "Teoria do conceito", Ciência da Informação, v. 7, n. 2, p. 101-107.

Ekholm A. (2002) "Principles for classification of properties of construction objects", Distributing Knowledge in Building - CIB W78 Conference.

Gene Ontology Consortium. (2001) "Creating the gene ontology resource: design and implementation", Genome Research, v.11, n.8, p. 1425-1433.

Gerlach, H., Falke, K.J. (1995) "The therapeutic role of nitric oxide in adult respiratory distress syndrome", Current Anaesthesia & Critical Care, v. 6, n. 1, p. 10-16.

Gomes, H.E., Motta, D.F., Campos, M.L. (2006) "Revisitando Ranganathan : a classificação na rede", www.conexaorio.com/biti/revisitando.htm, Janeiro de 2011.

Guarino, N.; Welty, C., (2004) "An Overview of OntoClean", In: S. Staab, R. Studer (eds.), Handbook on Ontologies, Springer Verlag, p. 151-159.

Guarino, N.; Welty, C., (2002a) "Evaluating Ontological Decisions with OntoClean", Communications of the ACM, 45(2), p.61-65, 2002a.

Guarino, N.; Welty, C., (2002b) "Identity and Subsumption", In: R. Green, C. A. Bean, S. Hyon Myaeng (eds.), The Semantics of Relationships: An Interdisciplinary Perspective, Kluwer, p.111-126.

Guizzardi, G., (2005) "Ontological Foundations for Structural Conceptual Models", Doctorate Thesis in Computer Science, University of Twente, Enschede, Holand, ISBN 90-75176-81-3, 416 p.

Guizzardi, G.; Wagner, G.; Sinderen, M. (2004) "A Formal Theory of Conceptual Modeling Universals", In: Proceedings Of The Workshop On Philosophy And Informatics (WSPI), Cologne, Germany.

ISO DIS 12006-2 - International Standards Organization. (1999) – "Organization of information about construction works — Part 2: Framework for classification of information". Swiss.

Novellino, M.S.F. (1996) "Instrumentos e metodologia de representação da informação", Informação & Informação., Londrina, v.1, n.2, p.37-45.

Ranganathan, S. R. (1967a), Prolegomena to Library Classification, New York : Asia Publishing House.

Ranganathan, S. R. (1967b) "Hidden roots of classification", Information Storage and Retrieval, v. 3, n.4, p. 399-410.

Ranganathan, S.R. (1951), Philosophy of library classification. New Delhi: Ejnar Munksgaard.

Ranganathan, S.R. (1963), Colon Classification. Bombay: Asia Publishing House.

Smith, A.D.; Datta, S.P.; Smith, G.H., Campbell, P.N.; Bentley, R.; Mckenzie, H.A.; Bender, D.A.; Carozzi, A.J., Goodwin, T.W., Parish, J.H.; Stanford, S.C. (2000), Oxford Dictionary of Biochemistry and Molecular Biology, Oxford University Press Inc., New York.

Smith, B. (2005), “The Logic of Biological Classification and the Foundations of Biomedical Ontology”, In: Westerstahl, D. (ed.) Invited Papers from the 10th International Conference on Logic, Methodology and Philosophy of Science.

Smith, B.; Kumar, A. (2004) “On Controlled Vocabularies in Bioinformatics: A Case Study in the Gene Ontology”, BIOSILICO: Drug Discovery Today, v.2, p. 246–252.

Smith, B.; Williams, J.; Schulze-Kremer, S. (2003), “The ontology of gene ontology”, In: AMIA Annual Symposium Proceedings, p. 609-13.

Yu, A.C. (2006), “Methods in biomedical ontology”, Journal of Biomedical Informatics, n.39, p.252–66.