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

Linair Maria Campos  
(CISI/COPPE/UFRJ)
Maria Luiza de Almeida Campos  
(UFF – Ciência da Informação)
Maria Luiza Machado Campos  
(UFRJ – Ciência da Computação)

Setembro /2011
Summary

- Introduction
  - Complex domains x standards vocabularies
  - Problems to organize concepts hierarchies
- Goal
- Related work
- Canons for organizing arrays
- Canons for organizing chains
- Conclusion
Ontologies have been used in complex domains like Biomedicine.

Many ontologies...
(anatomy, phenotype, biochemical...)

They provide a standard vocabulary, that machines can deal with.

However...
Genome annotation

- **Where?** Nucleotide-level annotation
- **What?** Protein-level annotation
- **How?** Process-level annotation

**GOAT demo**

- **Name**: Saccharomyces cerevisiae U1 spliceosomal RNA
- **Gene-product type**: small nuclear RNA (snRNA)
- **Molecular function**
  - mRNA splicing
  - 3'-splice site cleavage, exon ligation
  - MAT a1 (A1) pre-mRNA splicing
  - lariat formation, 5'-splice site cleavage
  - mRNA cis splicing
  - mRNA splice site selection
  - mRNA trans splicing
  - regulation of mRNA splicing
  - negative regulation of mRNA splicing
  - positive regulation of mRNA splicing
  - snRNP recycling
  - spliceosome assembly
  - splicing AT-AC intron
  - splicing GT-AG intron

- **Biological process**
  - Add term
  - Add term

- **Cellular component**
  - Add term

- **Endoribonuclease activity**

- **Additional information**
  - Add term(s)
### Problem

Biomedical ontologies have a huge number of terms. It is hard to find the most suitable term among similar and to understand classification rationale. It is difficult to articulate related ontologies. Ontologies usually do not provide information about their ontological commitment; poor documentation; difficulty in organizing ontology taxonomical structure.
Goal

- Present some of Ranganathans canons, as support to organizing ontology concepts in arrays and chains.

Canons are principles, guidelines to think about the classificatory structure of a given domain as a metamodel of those classificatory principles.
Who is Ranganathan?

- Indian mathematician and librarian
- Work published since 1920s
  - Theory of Faceted Classification

Classes

Facets

Properties

Categories (PMEST)

Characteristics of division
More recently...

- The use of philosophical notions, such as

![Diagram]

Nicola Guarino and Chris Welty [2004]; Giancarlo Guizzardi [2005]…

- The role of definitions

Barry Smith [2005]

- The role of definitions to identify attributes “in a consistent manner, thus assuring their transitive inheritance through a type hierarchy”

Dahlberg [1983]

- Concepts theory: “definitions reveal a set of common characteristics which are useful to build any system of classification or thesaurus”
Related work - considerations

- A systematic approach is useful to help organize concepts, once their nature has been identified.

- If such approach is easily understandable by end users, the better.

- We propose that Ranganathan's canons can be applied to help organize ontologies taxonomical structure in a comprehensible way.
Ranganathans canons

- Canons as principles to organize concepts in arrays and chains

- Canons of:
  - Differentiation
  - Concomitance
  - Exclusivity
  - Modulation
  - Subordinate classes
Canons for organizing arrays

- Arrays are horizontal series of concepts, organized as siblings in relation to a parent concept.
Differentiation canon

- A principle of division (classes) used as a classificatory basis should originate at least two classes.
- For example, catalytic activities of enzymes:
  - According to type of enzyme (hydrolase, isomerase, etc.)
  - According to type of reaction catalyzed (free radical formation, first spliceosomal transesterification, etc.)
Concomitance canon

- Two different principles of division should not result in the same array.

John: age 51
Mary: age 23

John: born 1960
Mary: born 1988
Exclusivity canon

- Elements belonging to an array should be mutually exclusive, i.e., disjoint in relation to elements belonging to another array.
Canons for organizing chains

- Chains are vertical series of concepts, organized hierarchically.
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”.

According to the Canon of Modulation, it should exist a term like ATP-dependent helicase.
In a hierarchy of classes, the classes nature should be the same, according to the principle of division that guides the organization of the array.
Canon for Subordinate Classes

- **Binding**: The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule.
In a succession of classes each one must follow immediately the other without being interrupted by the inclusion of one or more classes of different nature.

The definition of *binding* indicates that the principle of division of this chain has to do with interaction between *molecules*, which put term like "*bacterial binding*" and "*extracellular matrix binding*" in disagreement with the nature expected for terms in the "binding" chain.

Relevance and timeliness of Ranganathans work!

Foundations of a metamodel for concepts organization.
### Analysis of GO molecular function

<table>
<thead>
<tr>
<th>Canon</th>
<th>Total violations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiation</td>
<td>5</td>
</tr>
<tr>
<td>Concomitance</td>
<td>0</td>
</tr>
<tr>
<td>Exclusivity</td>
<td>7</td>
</tr>
<tr>
<td>Modulation</td>
<td>7</td>
</tr>
<tr>
<td>Subordinate Classes</td>
<td>7</td>
</tr>
</tbody>
</table>
Conclusion

- challenges of ontology construction
  - the creation of the backbone taxonomy

- it is important to provide the grounds for an effective dialogue between people with different backgrounds
  - principles that can rapidly and easily be assimilated and adopted on ontology structuring

- Ranganathans’ canons are a comprehensive set of principles, that can be easily assimilated
  - provide effective guidance
  - Helps to avoid many of the problems found in ontologies structures

- Ranaganathans’ canons are still current and relevant nowadays
There are more Canons, indeed...
Want to know more?

www.ontotaxo.uff.br
Ranganathans Canons applied to ontology engineering:  
a sample application scenario in biomedical ontologies

Linair Maria Campos – linair@csi.coppe.ufrj.br
Maria Luiza de Almeida Campos – maria.almeida@pq.cnpq.br
Maria Luiza Machado Campos – mluiza@ufrj.br

Setembro /2011