

Knowledge Classes or Canonical Representation

PhD Thesis Notes, Pedro Ferraz de Abreu, 1996

My approach to deal with the multiple-domain / multiple source problem in building knowledge bases is to establish a non-ambiguous, mutually exclusive classification of different types of knowledge, in other words, a *canonical representation*. The rationale is that by encapsulating each and all knowledge units in one of these categories, we create a virtual level of knowledge representation where the dominant traits are not domain-dependent, since we can define them at a syntactic level, instead of a semantic level.

This canonical representation was achieved by reviewing a large set of multi-domain vocabulary (more than one thousand items) and several field taxonomies (from different school curricula, job market demand and supply on domain qualifications, etc.). As a result, I identified the following categories: *Term*; *Concept*; *Definition*; *Model*; *Rule*; *Norm*; *Procedure*; *Methodology*; *Description*.

Term:

- Short word or sentence ;
- Represents an element of technical, scientific or cultural vocabulary; or a variable in an algebraic expression;
- May be defined in a simpler and less technical language (Glossary);
- Does not require extensive explanations or complex theoretical foundation;
- Definition may contain other terms only .

Concept:

- Word or sentence ;
- Represents an idea or abstraction (technical, scientific or cultural), or a knowledge domain (class, sub-class, domain);
- May be explained in lay language, eventually requiring more or less complex theoretical foundation;
- Explanation may contain terms or other concepts, of similar or lesser complexity.

Definition:

- One or more sentences;
- Represents the exact, non ambiguous explanation of a term or concept; or establishes an axiom, which should, in this case, be considered a term or concept;
- There may be more than one definition per concept, and they may or not contradict themselves (if they do, it implies the co-existence of several truth/belief systems);
- Explanation may contain other terms and concepts, other than the object being defined, of similar or lesser complexity.

Model:

- One or more algebraic expressions (set of variables linked by algebraic or logical operators);

- May establish an axiom (variables must also be considered terms).

Rule:

- Regular expression [IF precedent THEN consequent], in which precedent and consequent are a set of one or more conditions linked by the logical operator AND, where condition is a 3-tuple variable-operator algebraic-value;
- Represents a causal or dependency relationship between phenomena, identified through investigation and not arbitrarily set.

Norm:

- Regular expression [IF precedent THEN consequent], in which precedent and consequent are a set of one or more conditions linked by the logical operator AND, where condition is a 3-tuple variable-operator algebraic-value, and the consequent part may be a set of conditions or a set of procedures;
- Represents a causal relationship resultant from arbitrary determination.

Procedure:

- One or more phrases or images;
- Represents a sequence of one or more acts (operations, interventions) of one or more agents acting on one or more target-objects (people, things, entities, etc.);
- Is conditioned by rules or norms.

Methodology:

- Set of norms and procedures.

Description:

- One or more phrases, images or sounds;
- Factually represents things, people, entities, places, events, situations or states ;
- May contain models, terms, concepts and other descriptions.