

Organizing Information

Ruminations and Conversation

Data vs. Information vs. Knowledge

- Data - individual facts, statistics or items of information [Random House Dictionary]
- Information - something received or obtained through informing [Webster's]
 - Hence information is the content of a message
- Knowledge - 'a matter of degree' [Russell]
 - The highest degree is found in facts of perception, and in the cogency of very simple arguments
 - The next highest degree is in vivid memories
 - When a number of beliefs are each severally in some degree credible, they become more so if they are found to cohere as a logical whole. [Human Knowledge: Its Scope and Limits]

An Information Theoretic Spin

- The amount of information in a message is associated with the probability of its occurrence within the ensemble of all messages of the same length derivable from a given set of symbols [Svenonius paraphrasing Shannon]
- This basically equates novelty with information

A Linguistics Spin

- hypernyms/hyponyms - x is a kind of y
- meronyms - x is a part of y
- synonym - x is y

What's in a Language

- Symbols - atomic units of presentation/representation
- Vocabulary - compositions of symbols into representations of concepts
- Semantics - the meaning of a construct
 - Relational semantics - between terms
 - Referential semantics - techniques used to limit the meanings (referents) of terms
 - Category semantics - the facets or grammatical categories partitioning the terms
- Syntax - the rules defining the construction of expressions from expressions
- Pragmatics - use and application

An Intellectual Foundation for Organizing Information [Svenonius]

- An ideology, formulated in terms of purposes (the objectives to be achieved by a system for organizing information) and principles (the directives that guide their design);
- Formalizations of processes involved in the organization of information, such as those provided by linguistic conceptualizations and entity-attribute-relationship models;
- The knowledge gained through research, particularly that expressed in the form of high-level generalizations about the design and use of organizing systems; and
- The key problems to be solved if information is to be organized intelligently.

Ontology

- Philosophically, the science or study of being.
- A theory regarding the entities, especially abstract entities, to be admitted into a language of description [Webster's]

Ontology, con't.

- A language for knowledge representation [Cyc Project]
 - `#$Skin` - A (piece of) skin serves as outer protective and tactile sensory covering for (part of) an animal's body. This is the collection of all pieces of skin. Some examples include `#$TheGoldenFleece` (representing an entire skin of an animal) and `(#$BodyPartFn #YulBryner #Scalp)` (representing a small portion of his skin).
 - `isa: #AnimalBodyPartType`
 - `gens: #BiologicalLivingObject #AnimalBodyPart #SheetOfSomeStuff #VibrationThroughAMediumSensor #TactileSensor`

Taxonomy

- A (usually hierarchical) organization of categories or kinds
- Vehicle
 - Truck
 - Pickup
 - Panel
 - Semi
 - Passenger
 - Coupe
 - Sedan

Thesaurus

- A formal organization of vocabulary
- An example from the NASA Thesaurus:
 - abilities
 - UF proficiency
 - skills
 - GS abilities
 - . aptitude
 - RT effort
 - human performance
 - intelligence