Modeling Ephraim Chambers' Knowledge Structure from a Naïve Standpoint

Scott McClellan, MRC/CCI, Drexel University, sm4522@drexel.edu

Mat Kelly, MRC/CCI, Drexel University, mrk335@drexel.edu

Jane Greenberg, MRC/CCI, Drexel University jg3243@drexel.edu

DREXEL UNIVERSITY
Metadata Research Center
College of Computing & Informatics

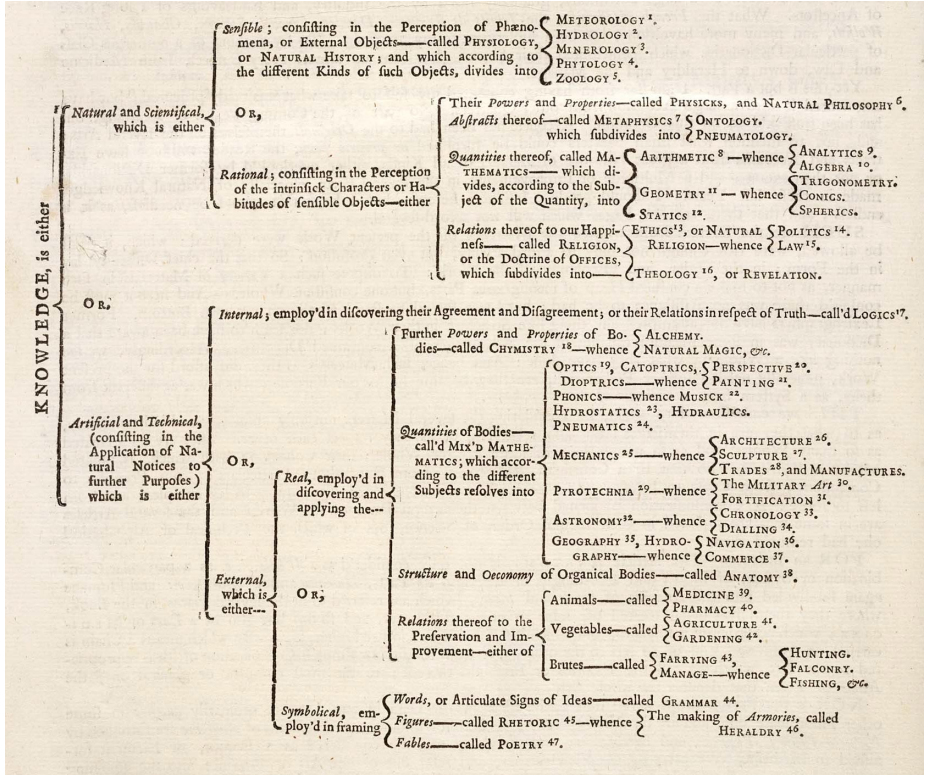
Overview

- What is Chambers' *Cyclopaedia?* And why is it important?
- Naïve vs Informed Modeling
- Modeling
  - Thesaurus
  - Ontology
- Implicature
- Conclusions

Chambers' *Cyclopaedia*

- Published 1728
- "Preface" lays out knowledge structure
  - Taxonomic tree
  - Domain vocabulary
- Taxonomic graphically represents abstract knowledge
  - Lowest nodes of the tree are (mostly) the domain vocabulary headwords
- Domain vocabulary
  - Structured sets of related terms

Taxonomic Tree



Domain Vocabulary (Example)

3 MINEROLOGY, or the History of EARTH; 1°, Its Parts, as *Mountain, Mine, Moſs, Bog, Grotto*; and their Phænomena, as *Earthquake, Volcano, Conflagration*, &c. Its Strata, as *Clay, Bole, Sand-* &c. 2°, Foſſils or Minerals, as *Metals, Gold, Silver, Mercury*, &c. with Operations relating to 'em, as *Fuſion, Refining, Purifying, Parting, Eſſaying*, &c. *Litharge, Lavatory, Pinea*, &c. Salts, as *Nitre, Natron, Gemma, Allum, Armoniac, Borax*, &c. Sulphurs, as *Arſenic, Amber, Ambergreaſe, Coal, Bitumen, Naphtha, Petrol*, &c. Semimetals, as *Antimony, Cinnabar, Marcaſite, Magnet, Biſmuth, Calamine, Cobalt*, &c. Stones, as *Marble, Porphyry, Slate, Asbeſtos*, &c. Gems, as *Diamond, Ruby, Emerald, Opal, Turcoiſe*, &c. *Emery, Lapis*, &c. whence *Ultramarine, Azure*, &c. Petrifactions, as *Cryſtal, Spar, Stalactites, Trochites, Cornu Ammonis*, and the like.

Naïve vs Informed Modeling

- What is meant by naïve:
  - Less knowledge about underlying subject
  - Less access to a subject matter expert
  - Less familiarity with system of expression
- Spectrum
  - Modelers vary in degrees of expertise and naivete
- Crossover skills
  - Language
  - Adjacent Studies

Thesaurus/Ontology

	Thesaurus	Ontology
Pro	• Expresses basic hierarchy well • Easier to reconcile logical inconsistencies, e.g., duplicate terms • Describes domain vocabulary well	• More robust class and sub-class descriptions • Expresses complex connections between and across classes • Incorporates taxonomic tree structure
Con	• Facets sometimes difficult to describe • Relationships tend to be less expressive	• Model relies on greater understanding of logic • Tend to be more interpretive than descriptive

Implicature

"Our talk exchanges do not normally consist of a succession of disconnected remarks and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction" –Paul Grice, *Studies in the Ways of Words*, 26

- Attempts to understand how participants in a conversation derive meaning from each others' utterances based upon situation and environment
- Modified Occam's Razor: Try not to allow meaning to proliferate
- Lack of a physical second actor complicates the theory

Implicature Continued

- Applying the theory to Chambers' vocabulary
  - Descriptive connectors
  - Typographical features
  - Shared language (for English speakers)
  - Well adapted for more descriptive knowledge organizations (e.g., thesaurus)
- Problematic Points
  - Subtle shifts in language usage across time
  - Lack of deictic markers in places
  - Less useful in low-context situations (e.g., taxonomic tree)

Conclusions

- Information needs of the end user define the best model
- Encoding in Simple Knowledge Organizing System (SKOS)
- Integration into the Metadata Research Center's Helping Interdisciplinary Vocabulary Engineering (HIVE) application
- Continued research with the 19th Century Knowledge Project and persistent identifiers for computational vocabulary work

Acknowledgements

- Joseph Tennis & Christopher Hodges, top-level SKOS encoding
- Sam Grabus & Jane Lippman, extended SKOS encoding
- Peter Logan, PI 19th Century Knowledge Project
- Grice, P. (1989) *Studies in the ways of words*. Harvard: Cambridge.