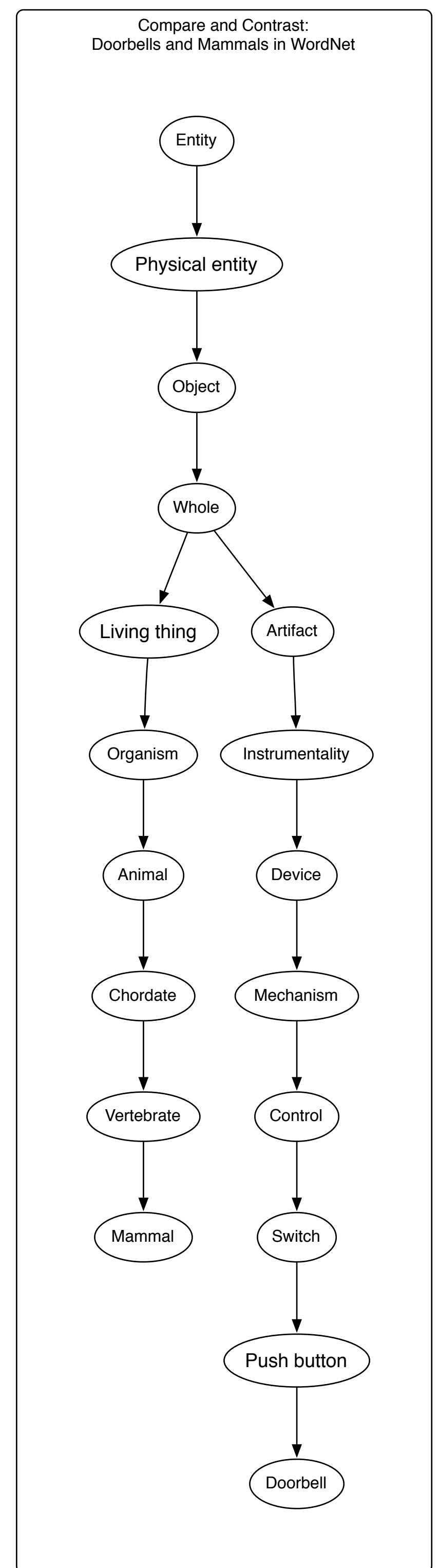
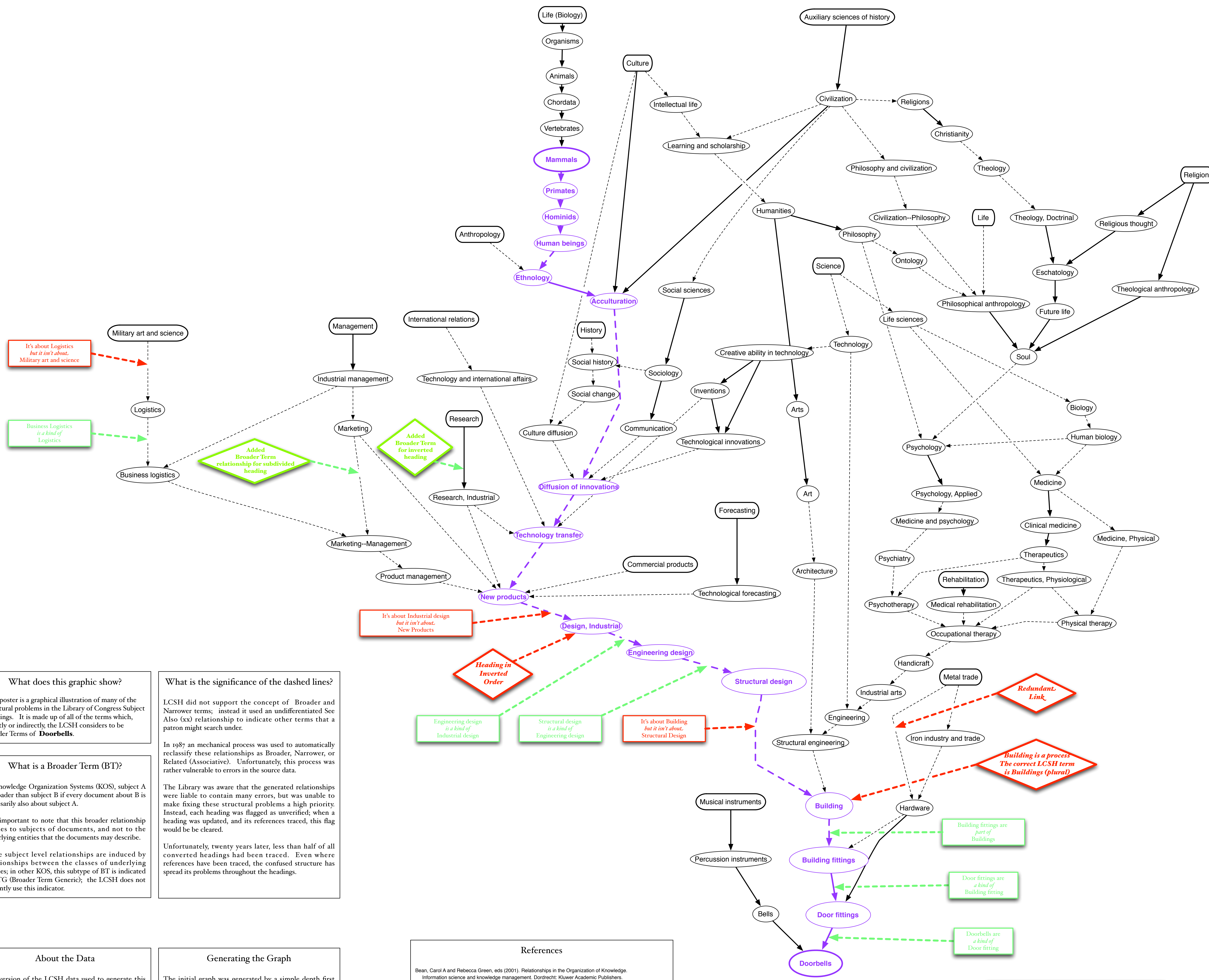


LCSH is to Thesaurus as Doorbell is to Mammal

Contact Info

Simon Spero (ses@unc.edu)
School of Information and Library Science
University of North Carolina at Chapel Hill



What does this graphic show?
This poster is a graphical illustration of many of the structural problems in the Library of Congress Subject Headings. It is made up of all of the terms which, directly or indirectly, the LCSH considers to be Broader Terms of **Doorbells**.

What is the significance of the dashed lines?
LCSH did not support the concept of Broader and Narrower terms; instead it used an undifferentiated See Also (xx) relationship to indicate other terms that a patron might search under.
In 1987 a mechanical process was used to automatically reclassify these relationships as Broader, Narrower, or Related (Associative). Unfortunately, this process was rather vulnerable to errors in the source data.

What is a Broader Term (BT)?
In Knowledge Organization Systems (KOS), subject A is broader than subject B if every document about B is necessarily also about subject A.
It is important to note that this broader relationship applies to subjects of documents, and not to the underlying entities that the documents may describe.
Some subject level relationships are induced by relationships between the classes of underlying entities; in other KOS, this subtype of BT is indicated as BTG (Broader Term Generic); the LCSH does not currently use this indicator.

The Library was aware that the generated relationships were liable to contain many errors, but was unable to make fixing these structural problems a high priority. Instead, each heading was flagged as unverified; when a heading was updated, and its references traced, this flag would be cleared.
Unfortunately, twenty years later, less than half of all unconverted headings had been traced. Even where references have been traced, the confused structure has spread its problems throughout the headings.

About the Data
The version of the LCSH data used to generate this graph came from the Fred 2.0 LCSH data set, which was harvested from authorities.loc.gov in mid December 2006.
This data has been enhanced by attaching inverted headings to their focus term if and only if the focus term is already established, and by connecting established subdivided headings to the headings they immediately subdivide.

Generating the Graph
The initial graph was generated by a simple depth first search, following only broader term relationships.
The nodes and links extracted were used to generate dot format input for the AT&T graphviz layout engine.
This poster was generated by importing the generated file for the LCSH term Doorbells into OmnGraffle 5.1, which also uses the Graphviz layout engine, then manually adjusting the layout to add annotations and text.

References

Bean, Carol A and Rebecca Green, eds (2001). Relationships in the Organization of Knowledge. Information science and knowledge management. Dordrecht: Kluwer Academic Publishers.

Cruse, D. A. (1986). Lexical semantics. Cambridge textbooks in linguistics. Cambridge University Press.

Dykstra, Mary (1988). "LC Subject Headings Disguised as a Thesaurus". Library Journal 113.4. p42.

Gansner, Emden R. and Stephen C. North (1999). "An Open Graph Visualization System and Its Applications". In: Software - Practice and Experience 30. Pp. 1203-1233.

Spergel, Dagobert (1974). Indexing languages and thesauri: construction and maintenance. Information sciences series. Melville Pub. Co.

Svenonius, Elaine (2000). The Intellectual Foundation of Information Organization. Cambridge, Mass.: MIT Press.

Inferences about Doorbells

Everything about Doorbells is also about Buildings, because a Doorbell is a kind of Door fitting, a Door fitting is a kind of Building fitting, and Building fittings are parts of Buildings.

Doorbells are a kind of Building fitting.

Door fittings are a kind of Building fitting.

Building fittings are part of Buildings.

Doorbells, aren't a type of Buildings.

It's about Doorbells, but it isn't about Structural Design.

Legend

Terms

- Top Level Term (circle)
- Term (circle)
- Term on path from Doorbells to Mammals (circle)

Sentence Judgements

- Sentence judgement: Not a valid Broader Term relationship (red diamond)
- Sentence judgement: Valid Broader Term relationship (green diamond)
- Other Error (red diamond)
- Other Correction (green diamond)

Relationships

- Broader Term - Traced (solid arrow)
- Broader Term - Untraced (dashed arrow)
- Narrower Term (circle)