Comparative Classification

Joseph T. Tennis

University of Washington

Information School

Comparative Classification

First, I want to acknowledge that I am speaking to you from lands of the Coast Salish people.

Second, thanks to Joseph Busch for putting this panel together.

Third, I want to acknowledge the depth of expertise in this room, and I hope my offering can be of use.

Comparative Classification

To know one language is not know none.

In the spirit of Max Müller (1823-1900)

Comparative Classification

We compare to understand the internal workings and contextual factors that make a classification *what it is and how it works*.

This is essential for us to understand whether we are building and deploying *good* or *useful* classification schemes.

The internal workings of a scheme are its semantics and structure.

The contextual factors include technological context, time, culture, purpose, and use (among other things).

Comparative Classification

I see three major *basic kinds* of comparison available to us:

1. Comparison of a scheme over time
2. Comparison of two or more schemes at the same point in time
3. Comparison of a scheme *repurposed*

NB: These can be combined

Comparative Classification

I have said this same thing elsewhere as part my conception of *second-order classification theory,* viz.

"1) how schemes change over time and how we update them, 2) how installed schemes interoperate, and 3) how systems change when they change context (reapplied or reengineered)," (Tennis, 2015 p. 246).

Comparative Classification

Scheme Change Over Time

Comparative Classification

Scheme Change Over Time

EUGENICS used to be a biological science in the DDC (Tennis, 2012)

GYPSIES, NOMADS, and OUTCAST RACES were an *other* class in DDC because before the editors separated geopolitical divisions and language and race and ethnicity there was no place for them. (Tennis, 2016a)

Schemes can perpetuate bias and systemic discrimination as well (Higgins, 2016)

Comparative Classification

Scheme Change Over Time

SUBJECT ONTOGENY

Subject ontogeny is the life of the subject in a scheme, like the DDC. Examining how a subject is treated over time tells us about the anatomy of a scheme. For example, GYPSIES as a subject has been handled differently in different editions of the DDC. (Tennis, 2002).

COLLOCATIVE INTEGRITY

If an indexing language changes over time, how does that affect the power of the scheme to collocate? Is there a threshold below which a scheme becomes useless?

SEMANTIC GRAVITY

Linked to collocative integrity, semantic gravity is the weight of the outdated class number in cataloguing practice. Often libraries will keep an old number because they think it helps users.

EPISEMANTICS

Episemantics considers semantic effects outside of the indexing language. The idea of episemantics is to account for meaning as it changes over time outside of the scheme, and relate that to the scheme. (Tennis, 2016b)

Comparative Classification

Comparing two or more schemes
   Ideal type comparison
   Comparison in order to interoperate (conversion)

Comparative Classification

More than one scheme laid side by side against an ideal type (Ranganathan, 1967) e.g.,
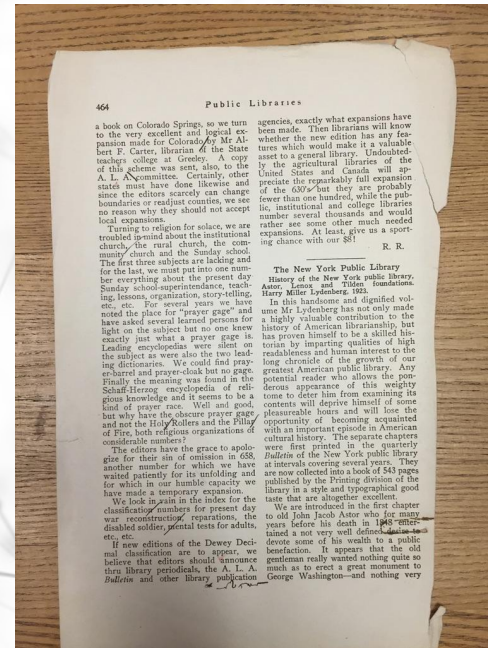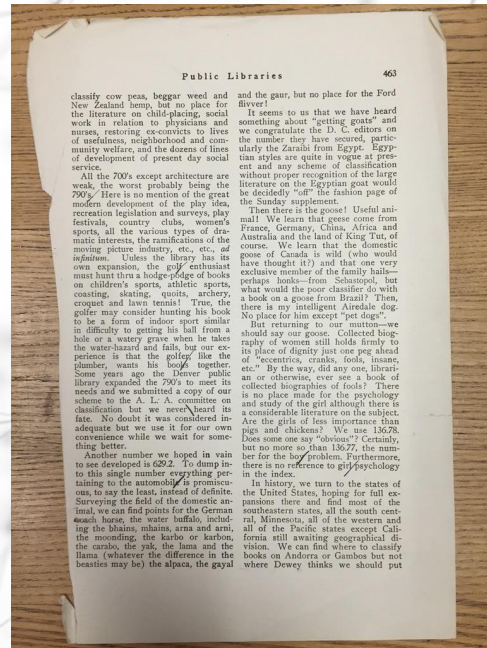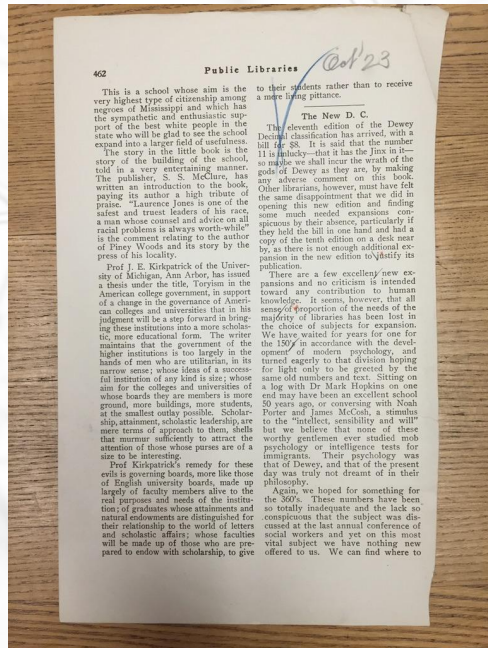
Rigidity (the degree to which we are free to express subjects)

Resilience (ability to add topics and hence numbers)

Parsimony (how short are the class numbers)

Comparative Classification

A single scheme against an ideal type (Reece, 1923)

Here is a review of DDC by a librarian

Comparative Classification

Comparison for Interoperation (cf., Zeng, 2016)
   Both semantics and structure are examined
   Various techniques to *interoperate*

Comparative Classification

Comparison for Interoperation
    Direct Scheme to Scheme e.g.,
    Mapping (conversion, crosswalk) (ALA, 2022)
    Derivation (Zeng, 2016)*

    Intermediary Tools e.g.,
    Switching (Lancaster, 1986)
    Supra-Schemes (Neville, 1970)
    Universal Sources Schemes (Soergel, 1974)
    Core Ontologies (Doerr et al., 2001)
    More at Zeng (2016)

Comparative Classification

Repurposed and Reengineered e.g.,

Techno-Structure Reengineering: wrapping pre-semantic web schemes in RDF/XML (Soergel et al., 2004)

Collection assessment repurposing: quantitative measures of collections using scheme numbers

In both cases we compare by reimagining the purpose of the scheme, and from there interrogate the original intent beside the reimagined one.

Comparative Classification

Comparing classification opens our eyes to our assumptions about the mechanics and motivations of our knowledge organization work.

I look forward to further conversation on the topic.

Comparative Classification

Thank you

Joseph T. Tennis

jtennis@uw.edu

Google scholar: "jt tennis"

https://ischool.uw.edu/people/faculty/profile/jtennis

References

Higgins, M. (2016). Totally Invisible: Asian American Representation in the Dewey Decimal Classification, 1876-1996. Knowledge Organization 43(8): 609-621.

Soergel D, Lauser B, Liang A, Fisseha F, Keizer J, Katz S. Reengineering thesauri for new applications: the AGROVOC example. Journal of digital information. 2004;4:1-23.

Tennis, J. T. (2012). The Strange Case of Eugenics: A Subject's Ontogeny in a Long-Lived Classification Scheme and the Question of Collocative Integrity. Journal of the American Society for Information Science and Technology 63(7):1350–1359, 2012.

Tennis, J. T. (2015). Foundational, First-Order, and Second-Order Classification Theory. Knowledge Organization. 42(4), 244-249.

Tennis, J. T. (2016a). Conceptual Provenance in Indexing Languages. In Vicki Lemieux (ed.). Building Trust in Information: Perspectives on the Frontiers of Provenance. Springer Proceedings in Business and Economics (IEEE). 93-99.

Tennis, J. T. (2016b). "Methodological Challenges in Scheme Versioning and Subject Ontogeny Research." In

Knowledge Organization 43(8): 573-580.