Idea Collider: From a Theory of Knowledge Organization to a Theory of Knowledge Interaction

by Richard P. Smiraglia and Charles van den Heuvel

Knowledge Organization Innovation: Design and Frameworks

EDITOR'S SUMMARY

A new and novel addition to core concepts of information science and technology is the Idea Collider, adapting the conceptual basis of the particle accelerator. The Idea Collider is presented as a theory of information and an information retrieval tool that deconstructs text to the underlying elements – the concepts, ideas, knowledge entities, taxons and knowledge bases – and then reconstitutes groupings of data. The authors champion the idea of a "multiverse of knowledge," in contrast to the "universe of knowledge" that is an important metaphor in traditional classification theory. A historical review of outstanding theoretical landmarks in classification theory explores the universe of knowledge metaphor, faceted classification theory and the universe of concepts, moving from a holistic view of existing knowledge to an elemental deconstruction that can account for all knowledge, past and future. The Idea Collider is proposed as a theoretical approach to identifying the essential parts of knowledge, dissociating those elements from culture- and time-bound dimensions and making them available for a bottom-up reassembly process.

KEYWORDS

classification theory

analytic models

knowledge

concepts

information retrieval models

We believe there is sufficient empirical knowledge, most of it from within the discipline of information science, to populate an elementary theory of knowledge. We see this theoretical base as an alternative to the long-sought universal classification. Rather than a universal *order*, we suggest that there is a fundamental approach to the interfaces and interactions of knowledge. We bring together in one theoretical model evidence from social classification and bibliographic classification, from ontology and taxonomy, from semiotics and from evidence about relationships among information objects, primarily the relationship known as instantiation.

This view raises interesting questions about the relationship between syntax and semantics in various knowledge organization systems (KOS). In a recent paper [1] we introduced the thought experiment of an Idea Collider, analogous to the Large Hadron Collider at the European Organization for Nuclear Research (CERN). We see a parallel between breaking physical elements into physical particles and breaking ideas into knowledge particles in order to arrive at a more complete understanding of concepts. We also think the relationship between syntax and semantics is critical, that neither is sufficient unto itself to support the reassembly of particles exploded from ideas broken apart in a collider. We are seeking through our theoretical model a more pragmatic approach to the reconstruction of knowledge along structural (syntactical) lines, rather than the traditional semantic-only approach.

Richard P. Smiraglia is professor, Information Organization Research Group, School of Information Studies, University of Wisconsin-Milwaukee. He brings empirical epistemology to the problems of the order and structure of recorded knowledge, and he is an expert in the phenomenon of instantiation among information objects. He can be reached by email at smiragli<at>uwm.edu.

Charles van den Heuvel is head of research on the history of science and theme leader, "Circulation of Knowledge," at the Huygens Institute in The Hague (Netherlands). He is an historian of art and science, including information science and an expert in the life and works of Paul Otlet. He can be reached by email at charles.van.den.heuvel<at>huygensinstituut.knaw.nl.

In contrast to historical attempts to formulate a classification theory, which we will discuss below, our theoretical model is not based in the notion of one, single universe of knowledge, but rather on one of multiple constellations or even multiple universes. These multiple constellations or universes comprise various levels of entities, all of which represent knowledge in an elemental state. Some knowledge entities reside in documents as material objects; others reside in intellectual works. In either event, knowledge entities are made up of ideas which are combinations of concepts. All knowledge entities function as signs, and we have gathered together the empirical evidence of this semiotic nature of knowledge. Groups of entities comprise taxons, which in turn generate canonical knowledge bases. Everything in the universe of knowledge is subject to perception. Our model explains and operationalizes these interactions.

In this article we have chosen a rather pragmatic approach to exploring these ideas, that of analyzing existing theories that could be used or adapted to operationalize such a model of knowledge interaction. This approach is comparable to the task that the Classification Research Group formulated for itself as the result of the Scientific Information Conference in London of 1948 to explore the development of classification theory by setting up experiments with information retrieval in various classification systems in which the universe of knowledge metaphor played a central role.

The Historical Quest for Classification Theory: Universe-Multiverse

"Ideas," Ernest Cushing Richardson states [2, p. 13], "are real things calling for a place in the order of things." To describe this order of things Richardson used the universe of knowledge metaphor. A close reading of Richardson's discussion of this metaphor reveals that his "order of things" is dealing with existing knowledge. The places of things or ideas are dependent at every instant on the position of other things or ideas. Ranganathan also used the "universe of knowledge metaphor" to describe the working of his Colon Classification, but in contrast to Richardson's view, Ranganathan's universe not only accommodates past and present, but also unknown future knowledge [3, p. 100]. The flexibility of the faceted structure and Ranganathan's concept of hospitality in chain (that allowed for the addition of classes in the future) made it a popular system, especially within the British Classification Research Group, which stressed the need for a faceted structure in classification to facilitate information retrieval [4].

Ernest Cushing Richardson has been considered the first real pioneer of classification theory, while Ranganathan was put forward by the CRG and several other authors (including himself) as the first theorist of faceted classification theory. Since Richardson, authors such as Bliss, Ranganathan, Miksa and Beghtol have discussed the universe of knowledge metaphor, which has become an important theme in the historiography of classification theory [1].

Thomas Daniel Wilson introduced, next to the top-down "universe of knowledge systems," the bottom-up "universe of concepts (or concept identification) systems." [5, p. 63] The first type is based on the view of a finite universe of knowledge that can be subdivided until the individual concept is reached (a view similar to Richardson's). The second assumes an infinite universe that gradually is built up by grouping elemental concepts.

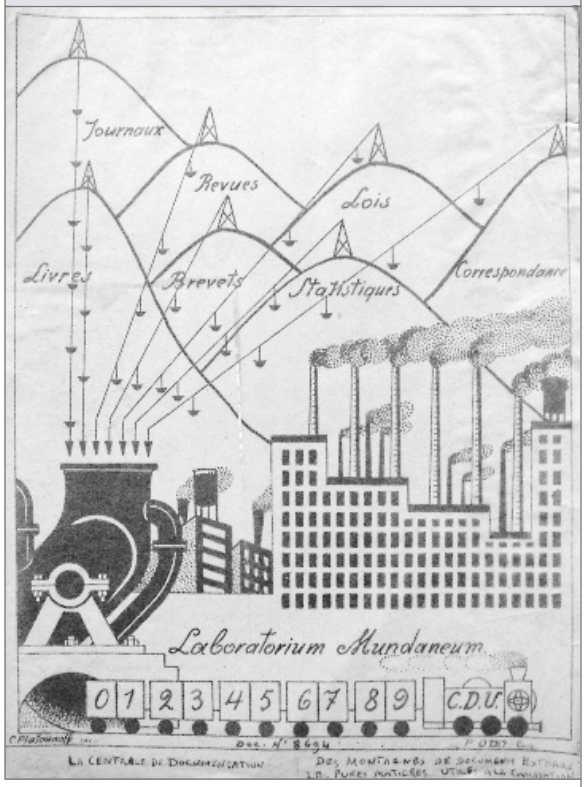
Beghtol [6, pp. 132-33] discussed Ranganathan's Colon Classification as the first bottom-up faceted system. However, it can be argued that the Universal Decimal Classification (UDC), heavily ignored in the comparative analyses of classification systems by Miksa, Beghtol and the British Classifications Research Group (which were limited to the Dewey Decimal Classification (DC) and the CC), can be read in a similar way.

Donker Duyvis made clear that Ranganathan's CC was indebted to this multi-dimensional classification and that the genesis of Otlet's UDC in the first place had to be situated in the analysis of concepts [7, p. 99]. The aim of this remark is not a rehabilitation of the UDC in classification theory – Donker Duyvis stressed the point that the UDC is not a theory – but a better understanding of the mechanical device that Otlet had in mind for breaking down and re-assembling documents, which we consider a first historical prototype of our Idea Collider.

Before we discuss this device in more detail, we need to point out that the framework for which it was conceived had the same limitations as all classification systems discussed so far. The distinction between the universe

of knowledge and the universe of concepts is a distinction between more or less defined knowledge entities, not between universes. All theoretical endeavors with classification have focused on the order of and on information retrieval within one single knowledge universe. We start from the notion that space is populated with constellations of a very different nature and behavior and from the possibility put forward by several astrophysicists that there are parallel or multiple universes (a so-called multiverse). This notion implies a reconsideration of the universe of knowledge metaphor. CERN's Hadron-Collider hurls the tiniest elements against each other until they fall apart, not only with the aim of understanding their composition, but also of answering questions about the origin, nature and behavior of the universe, or perhaps the multiverse. The Collider inspired us to set up the thought experiment of an Idea Collider to enable the re-examination of the universe of knowledge metaphor within the framework of multiple constellations or the multiverse.

FIGURE 1. Paul Otlet – Laboratorium Mundaneum – Dissection of Documents and Re-order in UDC ©Mundaneum Mons

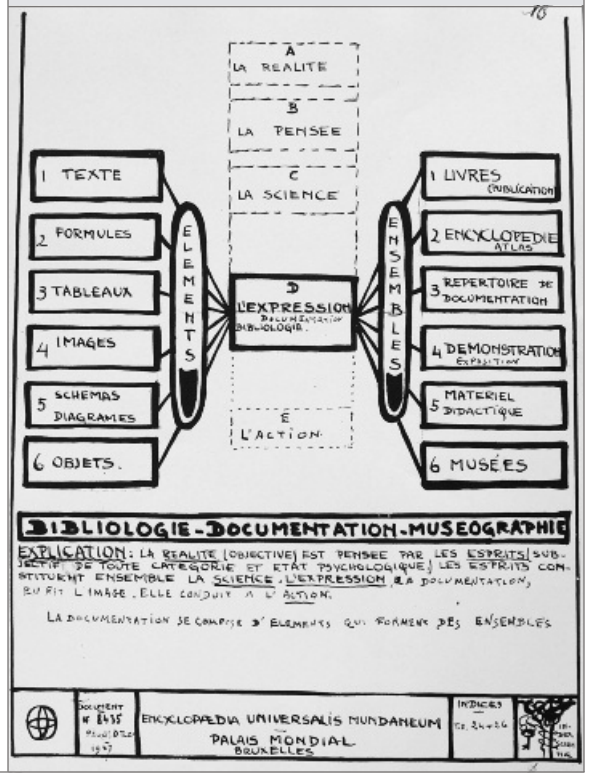


The Laboratorium Mundaneum and the Idea Collider

The idea of breaking down knowledge is not new. Paul Otlet envisioned and visualized the breaking down of all sorts of documents as a huge factory, named Laboratorium Mundaneum, in which all sorts of documents were crushed and sifted mechanically in a huge grinder into their most essential parts. These parts then ended up in an ordered way in 10 wagons representing the knowledge classes pulled by the Universal Decimal Classification locomotive that distributed "these classed facts" to the world [8, p. 133]. Figure 1 (the Laboratorium Mundaneum) only shows a part of the process that Otlet had in mind. Another visualization (Figure 2) not only shows the breaking down of knowledge into particles ("elements"), but also their re-combination into new knowledge entities and formats, called "ensembles." Apart from this bottom-up process from elements to ensembles, this image also shows Otlet's visualization of the transformation of perceptions of reality – from thoughts via expression to action. Both processes cross each other in expressions.

These images are certainly interesting historical antecedents of what Shera in his work on classification as the basis of bibliographic organization described as the disassociation of the idea content from its physical embodiment [9, p. 81]. Moreover, the idea that a knowledge element could get a new function in another ensemble tallies with the pragmatic approach of Shera to classification, which held that any single unit may be meaningful in any number of different relationships. "Thus it is the external relations, the environment,

FIGURE 2. Paul Otlet – Elements and Ensembles – EUM 8435
©Mundaneum Mons



of the concepts that are all important to the act of classifying" [9, pp. 83-84]. However, as we mentioned above Otlet's model is limited in that it focused on only one universal order of knowledge.

For the Idea Collider, on the other hand, we are interested in the nature and behavior of knowledge entities in various constellations or universes to formulate an alternative to a universal classificatory order, in order to create (temporary) interfaces that allow for interactions of knowledge between various universes. Ranganathan explained that thought is multi-dimensional, but that we still prefer all things to be arranged in one dimension. Therefore he sees classification as a transformation of a many-dimensional universe into a "uni-dimensional, uni-directional one." [10, p. 96] The mathematician and philosopher Alfred North Whitehead introduced the concept of "non-metrical projective geometry" that he defined as an example of "the science of cross-classification." [11, p. 137-38] Ranganathan's idea of transformation of multi-dimensions into one (applied within one classification) could be applied to create interfaces between classifications as envisioned by Whitehead. Since these constellations would have different orders and behave in different dynamics, such interfaces would be temporary events of knowledge interactions. These events can be contextualized in the framework of the theory of instantiation networks [12].

The Idea Collider Research Team

We have formed a team devoted to the prototype of an information retrieval application based on our theoretical model. The Idea Collider will be an information retrieval tool that deconstructs texts in order to reconstitute grouped data. The collider would work by deconstructing texts; therefore, the theory describes in detail the ways in which texts are constructed and evolve culturally over time. However, we will also set up experiments with deconstructing non-text formats to establish meaningful patterns. For the Idea Collider:

- We are in need of understanding the nature and behavior of concepts and of the minimal syntactic and semantic component needed to reconstruct knowledge.
- We are in search of groupings that will be meaningful in relevant contexts or relationships (compare Shera [9, p. 85]).
- We are in search of multiple approaches to the relata rather than the provision of alternative locations for individual units (compare Shera [9, p. 88]).
- We need to capture processes (not objects) that can be represented as events in analogy with structures of thought (compare Shera [9, p. 80]) and temporary perceptions that allow for concurrence, comparison and association (compare Farradane [13, p. 81]).

Members of the Idea Collider research team are Richard Smiraglia, School of Information Studies, University of Wisconsin-Milwaukee (SOIS, UWM); Charles van den Heuvel, Huygens Institute and e-Humanities Research Group, Royal Netherlands Academy of Arts and Sciences, The Hague; Feifan Liu, Health Care Informatics, UWM; Margaret Kipp, SOIS, UWM; Victoria Rubin, Faculty of Information and Media Studies, University Western Ontario; Thomas Dousa, Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign; Ann Graf, SOIS, UWM; Edward Benoit, SOIS, UWM. ■

Resources on next page

Resources Mentioned in the Article

[1] Heuvel, C. van den, & Smiraglia, R.P. (2010). Concepts as particles: Metaphors for the Universe of Knowledge. In C. Gnoli and F. Mazzocchi (Eds.). *Paradigms and conceptual systems in knowledge organization: Proceedings of the Eleventh International ISKO Conference* (pp. 50-56). Wurzburg: Ergon-Verlag.

[2] Richardson, E. C. (1935). *Classification: Theoretical and practical* (3rd ed.). New York: H.W. Wilson. (First ed. published 1901)

[3] Ranganathan, S.R. (1957). *Prologomena to library classification* (2nd ed.). London: The Library Association.

[4] Classification Research Group (CRG). (1957). The need for a faceted classification as the basis of all methods of information retrieval. In J. H. Shera & M. E. Egan (Eds.). *Proceedings of the International Study Conference on Classification for Information Retrieval held at the Beatrice Webb House, Dorking, England 13th-17th May, 1957* (pp. 137-147). London: ASLIB; New York: Pergamon Press.

[5] Wilson, T.D. (1972). The work of the British Classification Group. In H. Wellisch and T.D. Wilson (Eds.) *Subject retrieval in the seventies: New directions* (pp. 71-72). Westport, CT: Greenwood.

[6] Beghtol, C. (2008). From the universe of knowledge to the universe of concepts: The structural revolution in classification for information retrieval. *Axiomathes, 18*(2), 131-44.

[7] Donker Duyvis, F. (1951). The UDC: What it is and what it is not. *Revue de la Documentation XVIII*(2), 99-105.

[8] Heuvel, C. van den. (2008). Building society, constructing knowledge, weaving the Web. Otlet's visualizations of a global information society and his concept of a universal civilization. In B. Rayward (Ed.). *European modernism and the information society* (pp. 127-153). London: Ashgate.

[9] Shera, J.H. (1951). Classification as the basis of bibliographic organization. In J. H. Shera & M. E. Egan (Eds.). *Bibliographic organization. Papers presented before the Fifteenth Annual Conference of the Graduate Library School July 24-29, 1950* (pp. 72-93). Chicago: The University of Chicago Press.

[10] Ranganthan, S.R. (1951), Colon Classification and its approach to documentation. In J. H. Shera & M. E. Egan (Eds.). *Bibliographic organization. Papers presented before the Fifteenth Annual Conference of the Graduate Library School July 24-29, 1950* (pp. 94-105). Chicago: The University of Chicago Press.

[11] Whitehead, A. N. (1967). *Adventures of ideas.* New York: Free Press. (Original work published in 1933)

[12] Smiraglia, R.P. (2008). A meta-analysis of instantiation as a phenomenon of information objects. *Culture del Testo e del Documento, 9*(25), 5-25.

[13] Farradane, J.E.L. (June 1952). A scientific theory of classification and indexing: Further considerations. *Journal of Documentation, 8*(2), 73-92.