CAIS Paper: The Concept of Concepts: A Case Study from *American Documentation*

Richard P. Smiraglia (University of Wisconsin, Milwaukee)

Abstract: A core entity of information science is the "concept." Agreement on the basic definition as a mental construct representing a concrete instance, conceals divergence in understanding of the nuances. A case study of the domain's nascent era represented by *American Documentation* reveals some of the contours of the terms evolution.
Résumé:

1.0 The concept of the concept

The concept is the central atomist element of knowledge organization systems, and according to Dahlberg (2006, 12) it is the central entity in the science of knowledge organization itself. Concepts, "predicating the properties of or making statements about referents" are the central elements of knowledge, which can be represented, transferred, shared, etc. An appeal to WordNet™ (http://wordnetweb.princeton.edu) reveals the concept as a noun representing "an abstract or general idea inferred or derived from specific instances." A direct hypernym is "idea, or thought" which is "the content of cognition."

Smiraglia and van den Heuvel (2013) explored the diverse consideration of the search for elementary particles of knowledge across the history and evolution of knowledge organization as a component of information science, arriving at a taxonomic structure in which "works are made up of ideas, ideas are made up of concepts, [and] concepts, the atomic elements ... are represented by signs" (p. 374). Hjørland (2009) says the concept is a socially negotiated construct, dependent on perception, not representing universals, but rather functioning like nodes in linked networks among various discourses. Ranganathan (67) used the term "isolate" to differentiate the meaning of a concept from its representation in a document as a subject. Similarly, Bliss defined a concept as a subjective social construct associated with the form of likeness that constitutes a class (29, 1-22). Although this is of necessity a brief overview, it synthesizes the conclusions of a more thorough review by Smiraglia, van den Heuvel and Dousa (2011) that concepts are the ideational metaphors represented by signs in knowledge organization systems. Relationships among the signs, much like the linked networks among discourses in Hjørland's conception, or attributes of meaning arising from Bliss' and Ranganathan's, were suggested by Meincke and Atherton (76, 23) as the essential elements of knowledge organization systems using: "the assignment of representational vectors as in a multidimensional knowledge space with components on basic concept vectors."

Because of the central role of meaning in all iterations of information as either phenomenon or domain, the concept of concepts is critical. Several summaries have appeared, including two by Hjørland in his epistemological lifeboats (2008 and Hjørland and Nicolaisen 2010), concluding essentially that the concept is a mental construct used to classify, and furthermore, that the concept of the concept is difficult and is differently understood. The purpose of this paper is to reveal the historical contours of the use of the

concept of concept in information science by delving into the evolution of the use of the term. A first step is presented here as a case study of *American Documentation (AD)*, the first formal journal of information science in North America. For this study, all issues of *AD* in digital form were searched for the presence of the stem "concept." All occurrences of the term were recorded. Non-significant occurrences, such as the use of the term in passing (e.g., "the concept of information science") were recorded as well. Significant occurrences were tabulated together with the phrases in which they appeared.

2.0 The Contents of *American Documentation*

American Documentation (*AD*) appeared quarterly from 1950 through 1969 after which it was replaced by the *Journal of the American Society for Information Science*. Eighty issues appeared over the two decades of *AD*, containing 1104 separately indexible articles, reviews, reports etc. (advertisements were not analyzed for this study). *AD* spans a fascinating period in twentieth-century history. When it began publication George VI was King of Canada and Harry Truman was President of the United States; at the other end of the era Richard Nixon was President of the United States, Elizabeth II was Queen and the moon landing had just taken place. In between lay the Cold War, the era of 1950s prosperity and 1960s adventure, the Viet Nam experience and the age of Aquarius, to mention just a few of the more memorable points. The titles from the 1104 indexible entries were analyzed for term frequency; the most frequently occurring title words appear in Table 1.

DOCUMENTATION	61
RETRIEVAL	60
SYSTEM	55
INDEXING	54
TECHNICAL	44
SCIENTIFIC	42
RESEARCH	38
SCIENCE	38
LIBRARY	33
MACHINE	32
COMPUTER	30
INDEX	30
SYSTEMS	28
ORGANIZATION	27
SEARCHING	27
INDEXES	24
DATA	22
ANALYSIS	21
STORAGE	21
CLASSIFICATION	20
BIBLIOGRAPHIC	19
NEW	19
SUBJECT	19
CARD	18
ABSTRACT	17
DOCUMENT	17

Table 1. Most frequent title keywords

Documentation, retrieval system, classification and many aspects of automated indexing dominate the spectrum. Indexing and classification, and their common central component the "concept," were key.

2.0 Usage of the term "Concept"

The term concept was discovered in 409 articles, although most consist of casual usages. and not used in 117. Thirty papers were considered to have text in which the idea of the concept was significantly developed. A robust discourse existed over the importance of the comprehension and use of the concept of "concept." These appear in Table 2 below.

1950	Wise, Carl S; Perry, James W	Multiple coding and the rapid selector: introduction
1953	Luhn, H P	A new method of recording and searching information
1953	Macha, H Rowland	A technique for evaluating information analysis methods and personnel
1954	Morris, Jack C	The duality concept in subject analysis
1955	Perry, James W; Allen, Kent; Berry, Madeline M	Machine literature searching X. Machine language; factors underlying its design and development
1957	Bar-Hillel, Yehoshua	A logician's reaction to recent theorizing on information search systems
1958	Stiles, H Edmund	Identification of the conditions for valid application of machines to bibliographic control
1958	Buchanan, Bruce	A system of social significance
1960	Minder, Thomas	The mathematical foundations of bibliographic organization: abstract
1960	Wagner, Frank S, Jr	A dictionary of documentation terms: Kommos
1961	Costello, J C Jr	Uniterm indexing principles, problems and solutions
1960	Buck, R Creighton	Studies in information storage and retrieval: on the use of Godel Indices in coding: 1. Introduction
1960	Costello, John C Jr	Some solutions to operational problems in concept coordination
1960	Holm, B E; Rasmussen, L E	Development of a technical thesaurus
1961	Johnson, H Thayne	A polydimensional scheme for information retrieval
1961	Hoffer, Joe R	Information retrieval in social welfare: experiences with an edge-notched information retrieval system
1961	Melton, John L	The semantic code today
1961	Maloney, Clifford J	Semantic information
1963	Hyslop, Marjorie R	A compatibility study of two information systems
1964	Hillman, Donald J	The notion of relevance (I)
1964	Jahoda, G	A technique for determining index requirements
1964	Ruhl, Mary Jane	Chemical documents and their titles: human concept indexing vs. KWIC-machine indexing
1965	Salton, Gerard	The evaluation of automatic retrieval procedures selected test results using the SMART system
1966	Fischer, Marguerite	The KWIC index concept: a retrospective view
1966	Tinker, John F	Imprecision in meaning measured by inconsistency of indexing
1966	Wahlin, Ejnar	Classification systems and their subjects: a general analysis of different kinds of classification systems characterized by different types of subject
1968	Graziano, Eugene E	On a theory of documentation
1968	Eller, James L; Panek, Robert L	Thesaurus development for a decentralized information network
1969	Lesk, M E	Word-word associations in document retrieval systems
1969	Borden, George A; Nelson, William F	Toward a viable classification scheme: some theoretical considerations

Table 2. Articles developing the concept of concept

A few terms occupy most of the ideational space in the titles of these articles: information organization, information retrieval, systems, technical methods, analysis, classification, and development. We can see that the discourse was continuous over the period as well as robust, and although many articles are specifically devoted to classification, the integration of classification with technological developments, most notably coding and machine literature searching, also helped promote the discourse.

3.0 Evolution of the meaning of a "Concept"

Space being at a premium in the present abstract, we can look at just a few notable examples from the preceding table. In the second issue of *AD* Wise and Perry's "Multiple coding and the Rapid Selector: Introduction" brought the concept forward in these ways (square brackets are used to indicate multiple occurrences), placing the notion squarely in its atomic role alongside entities and processes:

> entities, concepts and processes
> entities, concepts and operations ... [3]
> a vocabulary of 456,976 concepts
> the nature of the concepts coded
> direct a search to any one concept

In 1953 Macha's "A technique for evaluating information analysis methods and personnel" revealed these notions of indexible concepts and their properties:

> An indexible concept is any idea contained
> name given to indexable concepts
> indexable concepts existing in the literature
> indexable concepts both analysts have classed
> misclassed concept ... [4]
> indexable concept is an independent act
> indexable concept is tenable

From 1954 Morris' "The duality concept in subject analysis" surveyed the relationship between the duality of "direct and specific" usages in the catalogs of the day, a reflection of Ranganathan's distinction between the isolate and the subject and Bliss's between the likeness and its social construct:

> duality concept ... [8]
> Born's concept of "complementary"
> complex concepts of science

From 1955 a machine literature searching contribution from Perry, Allen and Berry— "Machine literature searching X. Machine language; factors underlying its design and development."

> of concepts and the invention of various forms
> documentation, concepts and relationships
> to conceptualize is essential

Here we see the advancing development of technology require greater precision in the definition of the concept and its role as symbol and as node in a network of relationships.

The most rapid development took place in the 1960s with papers ranging from Minder's mathematical foundations to papers by Costello on uniterms and concept coordination. A particularly prescient paper by Melton is titled "The semantic code today."

> generic concepts of different classes
> the concept represented by the words
> relate concepts on various levels

Or Maloney's paper from the same year "Semantic information."

> of concepts of structural and mathematical
> each concept is an analytic code
> two concepts in the mind of the author
> compound concepts must be
> atomic concepts. This will
> primary concepts A
> ordered concepts

Finally, Graziano's 1968 article "On a theory of documentation"

> conceptual unity of the phenomena
> concepts at first are private;
> concepts may be shared between minds

and Borden and Nelson's 1969 article "Toward a viable classification scheme"

> events, concepts, referents, etc.
> a. conceptual structure for knowledge
> abstract concepts
> specific concepts
> to conceptualize the universe
> The concept of "superordinateness"
> concept, knowledge, tec.
> meaning, conceptual activity, rhetorical arguments, and philosophical inquiry
> of concepts from subordinate to superordinate

4.0 Toward Some Conclusions

Acknowledging the serious limitation that only one term has been used to date— "concept"—and not its direct hypernyms "idea" and "thought," we see a robust evolution of a core concept in documentation as the field marches toward its redefinition as information science. Discourse about the concept is continuous, robust, and evolutionary. It clearly becomes critical to the development of machine searching to have a concrete definition of a concept, however movable its parameters might be semantically. The term occurs in approximately one third of the published content of the journal, and in many cases is accompanied by rich analytical material. Not discussed here because of space

limitations are many applied instances, such as the use of the concept in the design of classifications for highly specific bodies of documentation.

This research has demonstrated the power of grasping the ideational discourse of our community across its history based on its own empirical evidence concerning even one of its central entities against the backdrop of the rich landscape of the field we now know as information science.

References

Bliss, H. 1929. *The organization of knowledge and the system of the sciences*. New York: Holt.

Borden, George A. and William F. Nelson. 1969. Toward a viable classification scheme: some theoretical considerations. *American documentation* 20: 298-302.

Costello, J. C,. Jr. 1961. Uniterm indexing principles: problems and solutions. *American documentation* 12: 20-27.

Dahlberg, Ingetraut. 2006. Knowledge Organization: A New Science? *Knowledge Organization* 33: 11-19.

Graziano, Eugene E. 1968. On a theory of documentation. *American documentation* 19: 85-90.

Hjørland, Birger, Jeppe Nicolaisen. 2010. Concept. In *The epistemological lifeboat.* http://www.iva.dk/jni/lifeboat/info.asp?subjectid=62

Hjørland, Birger. 2008. Concept in knowledge organization. In *The epistemological lifeboat.* http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/concept_in_kno...

Hjørland, Birger. 2009. Concept theory. *Journal of the American Society for Information Science and Technology* 60: 15-36.

Macha, H. Rowland. 1953. A technique for evaluating information analysis methods and personnel. *American documentation* 4: 35-50.

Maloney, Clifford J. 1961. Semantic information. *American documentation* 13: 276-88.

Meincke, Peter P.M. and Pauline Atherton. 1976. Knowledge space: a conceptual basis for the organization of knowledge. *Journal of the American Society for Information Science* 27: 18-24.

Minder, Thomas. 1960. The mathematical foundations of bibliographic organization: abstract. *American documentation* 11: 2-9.

Morris, Jack C. 1954. The duality concept in subject analysis. *American documentation* 5: 117-47.

Perry, James W., Kent Allen and Madeline M. Berry. 1955. Machine literature searching X. Machine language; factors underlying its design and development. *American documentation* 6: 242-55.

Ranganathan, S.R. 1967. *Prologomena to library classification*, 3d ed. Ranganathan Series in Library Science 20. Bombay: Asia Publishing House.

Smiraglia, Richard P. and Charles van den Heuvel. 2013. Classifications and concepts: towards an elementary theory of knowledge interaction. *Journal of documentation* 69: 360-83.

Smiraglia, Richard P., Charles van den Heuvel and Thomas Dousa. 2011. Interactions between elementary structures in universes of knowledge." In Slavic, Aïda and Civallero, Edgardo eds., *Classification & Ontology: Formal Approaches and Access to Knowledge: Proceedings of the International UDC Seminar -20 September 2011, The Hague, Netherlands*. Würzburg: Ergon Verlag, pp. 25-40.

Wise, Carl S. and Perry, James W 1950. Multiple coding and the rapid selector: introduction. *American documentation* 1: 776-84.