

Domain Analysis Redux: An Introduction

Richard P. Smiraglia* and María J. López-Huertas**

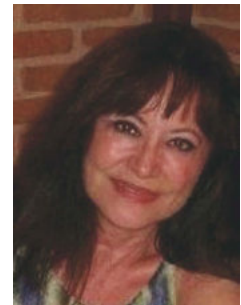
* Knowledge Organization Research Group, School of Information Studies,
University of Wisconsin, Milwaukee, NWQB2569, 2025 E. Newport St.,
Milwaukee WI 53211 <smiragli@uwm.edu>

**Facultad de Comunicación y Documentación, Colegio Máximo de Cartuja,
Universidad de Granada, 18071 Granada, Spain <mjlopez@ugr.es>

Richard P. Smiraglia is a professor, Knowledge Organization Research Group in the iSchool at the University of Wisconsin-Milwaukee. He has explored domain analysis for evolution of knowledge organization, epistemological analysis of the role of authorship in bibliographic tradition, the evolution of knowledge and its representation in knowledge organization systems, and the phenomenon of instantiation among information objects. He is an associate researcher of the e-Humanities Group, of the Royal Netherlands Academy of Arts and Sciences (KNAW), and Editor-in-Chief of this journal.



María is a professor at the Faculty of Communication and Documentation of the University of Granada, Spain. She teaches knowledge organization theory and methods and knowledge organization systems design, construction and implementation. Her research is focused on said topics, with special emphasis on interdisciplinarity and transdisciplinarity. She was formerly president of ISKO.



Smiraglia, Richard P. and López-Huertas, María J. **Domain Analysis Redux: An Introduction.** *Knowledge Organization*. 42(8), 553-556. 14 references.

Abstract: The knowledge organization (KO) community is about a decade from the catalytical work that set out a research agenda for information science, and specifically for knowledge organization, to embrace domain analytical methods for ontology extraction. A specific research agenda provided a framework of methodological approaches, based on epistemic stances in the knowledge organization community. This special issue of *Knowledge Organization* contains several papers invited from current scholars in KO who have been among the most productive scientists to embrace domain analytical methods.

Received: 8 November 2015; Accepted 15 November 2015

Keywords: knowledge organization, information science, domain analysis, domains, research

1.0 Domain Analysis Redux

We are about a decade out from the catalytical work of Hjørland and Albrechtsen (1995; 1998), which set out a research agenda for information science and specifically for those of us in knowledge organization (KO), to embrace domain analytical methods for ontology extraction. From 2002 onward Hjørland's specific research agenda gave us a framework of methodological approaches, based on his understanding at the time of epistemic stances in the knowledge organization community. This special issue of

Knowledge Organization contains several papers invited from current scholars in KO who have been among the most productive scientists to embrace domain analytical methods. This special issue has been guest edited by Dr. Maria-Jose Lopez-Huertas, whose work has been critical in several domain analytical areas ranging from Spanish KO research to feminist ontologies.

As any reader can see from the preceding paragraph it is impossible to write in the domain of domain analysis for KO without repeatedly citing the catalytical papers by Hjørland and Albrechtsen. But much has actually been ac-

completed in the intervening decade and a bit. In my recent primer on domain analysis (Smiraglia 2015) I summarized the domain analytical work within KO. About 100 papers have been published that are identifiably domain analysis for KO. They mostly use informetric methodologies, which we should note here are empirical. But there turned out to be few “subject gateways,” which might have been expected of practitioners. It seems this practice of the late 20th century has given way to digital libraries. And, surprisingly, discourse analysis and critical theory have become important contributors. A new list of eleven approaches, derived from analysis of existing research, appears in my book (97):

- Subject pathfinders
- Special classifications and thesauri
- Empirical user studies
- Informetric studies
- Historical studies
- Document and genre studies
- Epistemological and critical studies
- Terminological studies
- Database semantics
- Discourse analyses
- Cognition, expert knowledge and AI

And yet with only 100 studies over a decade, little theoretical knowledge has been generated. In fact, only a small group of domains has been studied more than once (98):

- Two studies: astronomy, cooking, Chinese information science, digital libraries, the Dublin Core Metadata Initiative, the *Encyclopedia of Milwaukee*, gender studies, nursing, race, and tripanomatides;
- Three studies: archives, image searching, LGBT communities, physics, and social media.
- Four studies: music
- Twenty-two studies: knowledge organization

It is good that we have almost two dozen domain analytical studies of our own domain. Navel-gazing is useful for a science as useful as ours. But none of the other domains listed above have yet generated theoretical understanding from domain analytical work. It is important that scholars in KO replicate and work toward hypotheses for theory-building in many domains.

In the case of our own domain, there are now more than twenty studies (and with this special issue even more). This now means meta-analysis would be an appropriate new methodology to join the list of “approaches to domain analysis.”

A final challenge to our domain is to embrace interdisciplinarity. Scholars in many other domains are using our

methodologies to analyze and extract essential ontologies, to trace semantic shift, and to create user interfaces to systems that are based on domain analytical research. But can we accept their papers for our journal when their authors are unaware of the work in our own domain? Dare we contribute papers to the journals in their domains? How can we build useful bridges across these sciences that all are studying domain analysis for KO? These are challenges for our domain for the future. But these also are the topics of the papers contributed to this special issue.

2.0 Moving the domain forward

In the issue’s opening article, Hanne Albrechtsen goes back in time from her thesis on domain analysis for classification of software in 1992 until the time she coauthored the seminal article on domain analysis (Hjørland and Albrechtsen, 1995). She tells us first-hand of her research experiences that then led to the theory of the analysis of domains. She claims that (560): “Domain analysis was scaled up from the KOS journey on software reuse to a comprehensive methodological framework in information science.” She talks about the very first years in which she was involved in research on software, the multidisciplinary way of working in the research teams involved, the search for a classification to organize this domain and the important role of Prieto Díaz (1991) who came up with a faceted classification for it and who first used the term domain analysis. For her domain analysis is a method. She reflects on the concept of domains, arriving at the idea that they are “fields of work” that have to be constructed. As she puts it, “domains are not terrains out there, waiting to be described and analyzed by the initiated few. Fundamentally, we may all create them.”

In the second paper, Regina Marteleto and Lidiane dos Santos Carvalho explore the theoretical and methodological constructs developed by Birger Hjørland, (2002, 2004), Hjørland and Albrechtsen (1995) and Pierre Bourdieu (1972, 1975). The main idea is to bring together and to compare the contributions of said authors in order to “investigate structures of production, organizations and communication of knowledge from a critical point of view” (561) from which to arrive at domain analysis. They chose “health,” with special stress on Brazil, as a testing ground for the study. The research is conducted by using the following categories of analysis: historical and institutional, relational, and of production, organization and dissemination of knowledge. They put forward questions related to the interdisciplinary nature of the topic at hand and they wonder how to understand the concept of extent of knowledge claimed by Joe Ten-

nis (2012) which sets it as a social and scientific field. The authors after establishing a dialogue between knowledge organization, information studies and theories and methods of social sciences, came up with the possibility of indicating paths of the health field as a knowledge domain and a scientific field. The health domain shows gaps related to the social conditions of knowledge production and it is claimed that the theories and methods of knowledge organization and the sociology of knowledge could favor the study of complex knowledge domains as health.

The following article by María López-Huertas calls attention to interdisciplinary knowledge and how this knowledge fits into domain analysis theory and methods. The departure point is the idea, poured in the article by Hjørland and Albrechtsen (1995), that the theory of domain analysis is mainly oriented towards disciplinary spaces and because of that, it raises doubts about the ability of this method to represent inter and transdisciplinary fields. To find out possible deviations, it compares the essential characteristics of interdisciplinary knowledge based on the contributions by Klein (2010, Repko (2008), Nowotny (2001) and Gibbons (1994), among others, to those of disciplinary knowledge taken from Sugimoto and Weingart (2015). After a reflection on the nature and meaning of said characteristics, it was quite evident that there is no correspondence between both kinds of knowledge to a large extent. So it was clear that domain analysis would need a reformulation in order to extend the parameters of the theory. The second part of the paper is devoted to a reflection on the methods for domain analysis (Hjørland, 2002), considering that these were designed keeping in mind the disciplinary context, as it was the seminal theory. Only those considered the most appropriate methods for interdisciplinary contexts were studied: indexing and retrieving specialties, terminological studies, constructing special classifications and thesauri, bibliometrical studies, empirical user studies, document and genre studies and epistemological and critical studies. This last reflection suggests that the methods of domain analysis should be extended and or reinterpreted in order to incorporate the peculiarities of the ID and to incorporate additional methods if needed.

Richard Smiraglia authors the fourth paper. He looks into the production of domain analytical knowledge for knowledge organization along two decades (2004-2014). He produces a domain analysis of domain analysis with specific reference to knowledge organization. In his words (602), “this study reports an analysis of the effort by scholars to respond to the call for the use of domain analysis as a methodological paradigm in KO.” The paper’s objective is to contribute to theory-building through domain analysis

in knowledge organization. It approaches the topic by studying several angles of it that range from the methodological approaches, finding that most contributions are infometric or terminological and that the discourse analysis is growing over time, to the identification of the research front, core authors and data about most productive authors, countries, etc. It also gives results obtained from the analysis of co-citation, inter-citation and cited references that allow Smiraglia to arrive at the most cited authors, to claim that there is some evidence of discourse among the core authors and to state that there is a strong influence of the foundational contributions of domain analysis for knowledge organization. On the other hand, the core community’s discourse on domain analysis is oriented towards ontological discovery for knowledge organization, epistemologically towards bibliometrics in information science at the time that influences from other fields are also recognized. As a result of this domain analysis of domain analysis for knowledge organization, it can be said that there are enough empirical studies along the past two decades to begin to make theoretical statements, that there is a discourse in the group of the studied scholars, a response of the Hjørland’s call for domain analysis and recognition of Dahlberg’s contributions. In Smiraglia words (610), “domain analysis for KO is a very vibrant field of research and development not only for KO as a science but for humanity at large.”

The following article is written by K. Raghavan, K. Apoorva and Aarty Jivrajani. They analyze the domain of information retrieval in order to map the borders of the research literature on said domain over a period of 14 years. Definitions of domains and domain analysis are given in this research context, understanding the latter as (592) “the process of mapping the contours of a domain with a view to study its evolution and transformation over time,” and stressing its capacity to study its evolution over time, to know the trends of the research and to visualize the topics that make up the domain. To carry out the study, the authors use two data sets coming from the IEEE and EBSCO databases. One of the reasons of this choice is that the points of view of both are different; LISTA is focused on the LIS community and IEEE is more computer science oriented. To see the differences between the two is an objective of this study. The results show that a few areas are of common interest to the research communities represented in the studied databases. It is evident that information retrieval is moving towards new territories, according to the top twenty research themes identified in the study. The analysis of this group shows that information retrieval is changing its dimensions as a domain in both databases. The Web has been the main factor influencing this move.

The special issue ends with a paper by José Augusto Chaves Guimarães and Natália Bolfarini Tognoli. They explore the application of the domain analysis approach to the organization of archival knowledge based on its core processes. They take the principle of provenance as a domain analysis approach for archival science whose social assumptions are based on the context of records production that ties them to a specific context that gives them the necessary meaning for them to be organized. The authors claim that this (562) “characterizes a discourse community for which classification processes and description becomes effective.” It reinforces the idea that archival knowledge organization is gaining more and more ground in the archival community, although until the late twenty century the discipline did not recognize information as its object of study. Archival knowledge is referred to a set of documents produced by a person or an institution (the provenance). It is suggested that the production context is understood as domain for content extraction. In this sense, the concept of archival bond becomes an effective methodological evidence of provenance to be used in domain analysis for archival knowledge organization. In fact, provenance studies include a procedure that involves the study of the person or entity where the document originates and the study of the functions of them. Only after this process, it is possible to assign records groups to the document and to set its arrangement and classification. The authors claim that the provenance studies can be considered a specific domain analysis approach for archival knowledge organization, based on three axes: provenance, respect des fonds, and the merger of original order and “organicit e.” They also point out that the suggested method deeply differs from those included in Hjørland (2002).

References

- Bourdieu, Pierre. 1972. *Esquisse d'une th orie de la pratique: pr ced e de trois  tudes d'ethnologie kabyle*. Gen ve : Librairie Droz.
- Bourdieu, Pierre. 1975. “The Specificity of the Scientific Field and the Social Conditions of the Progress of Reason.” *Social Science Information* 14 : 19-47.
- Gibbons, M. et al. 1994. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary societies*. London, Sage.
- Hjørland, Birger. 2002. “Domain Analysis in Information Science: Eleven Approaches—Traditional as well as Innovative.” *Journal of Documentation* 58: 422-62.
- Hjørland, Birger. 2004. “Domain Analysis: A Socio-cognitive Orientation for Information Science Research.” *Bulletin of the American Society for Information Science and Technology* 30: 17-21.
- Hjørland, Birger and Hanne Albrechtsen. 1995. “Toward a New Horizon in Information Science: Domain Analysis.” *Journal of the American Society for Information Science* 46: 400-25.
- Hjørland, Birger and Hanne Albrechtsen. 1998. “An Analysis of Some Trends in Classification Research.” *Knowledge Organization* 26: 131-39.
- Klein, Julie T. 2010. “A Taxonomy of Interdisciplinarity.” In *The Oxford Handbook of Interdisciplinarity*, R Frodeman, J T Klein and C Mitcham, eds. Oxford, Oxford University Press, pp. 15-30.
- Nowotny et al. 2001. *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, Polity Press.
- Prieto-Diaz, Ruben. 1991. “Implementing Faceted Classification for Software Reuse.” *Communications of the ACM* 34 no. 5: 88-97.
- Repko, Allen F. 2008. *Interdisciplinary Research: Process and Theory*. University of Texas at Arlington.
- Smiraglia, Richard P. 2015. *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Chandos Information Professional Series. Oxford: Elsevier/Chandos, 2015.
- Sugimoto, Cassidy and Scott Weingart. 2015. “The Kaleidoscope of Disciplinarity.” *Journal of Documentation* 71: 775-94.
- Tennis, Joseph. T. 2003. “Two Axes of Domains for Domain Analysis.” *Knowledge Organization* 30: 191-95.