

**Jihee Beak** — University of Wisconsin-Milwaukee

**Jeanette Glover** — University of Wisconsin-Milwaukee

**Daniel Martínez-Ávila** — Universidad Carlos III de Madrid, Spain

**Suellen Oliveira Milani** — Sao Paulo State University, Brazil

## **International Comparative Domain Analysis in Knowledge Organization Research Topics in Four Countries - Brazil, South Korea, Spain and the United States**

### **Abstract**

This study aims to identify and compare the domains of knowledge organization from four countries: Brazil, South Korea, Spain and the United States. Four researchers from diverse backgrounds study investigate knowledge organization (KO) on an international scale using domain analysis of keywords from journal articles. Each country selected two journals in LIS and evaluated each article to find those related to KO. The findings show there are some similarity in an international level and difference in a national level of knowledge organization domain. 21 overlapped topics across four countries have been identified. In addition, the findings show some examples of unique research topics of KO domain from each country. This international comparative domain analysis study can contribute to promote academic communication amongst KO researchers and bring more international collaborative research opportunities.

### **Introduction**

International scholarly communication involves many aspects related to science and production of literature from different historical and theoretical backgrounds. This also includes different methodological approaches, points of views in regards to language barriers, different concepts for the same term or different terms for the same concept and so on. Hjørland (2002, 446) postulates that “(l)anguage and terminology are very important objects for IS [Information Science] because they affect our thinking and thus the questions we put to databases as well as the texts we search.” Based on this, this study finds it necessary to examine Knowledge Organization (KO) domains that have been represented and researched in different countries and written about in various languages.

Domain analysis is a sociological-epistemological standpoint that was formulated at the beginning of the 1990s as an alternative to the dominant cognitive view (Hjørland 2008). Domain analysis is now one of the main research approaches in LIS, as observed by López-Huertas in 2008. Domain analysis has been discussed primarily by Dr. Birger Hjørland and Dr. Joseph Tennis, but other authors are also interested in this approach, such as: Smiraglia (2011; 2013), etc.

Through international comparative domain analysis, this study would offer elements to support the increase of the academic communication amongst KO researchers and bring more international collaboration research opportunities. In turn, it

is thought that the international scenario of KO research and international collaborations will benefit from the findings of this study.

### **Literature Review**

The domain analytical approach was popularized in LIS by Hjørland and Albrechtsen in 1995. However, as Smiraglia noticed,

(w)hereas their approach to domain-specificity has been largely embraced in the knowledge organization community, their call for domain-analytic research has been less apparently successful. Limited empirical research of a domain-analytic nature has emerged in KO as a domain, although bibliometric and informetric analyses continue to play a prominent role in information science at large. (Smiraglia 2012, 115).

In a thorough review of the literature, Smiraglia also stated: "In addition to traditional bibliometric techniques, co-word or term analysis can provide triangulating evidence about the emergence of trends in scholarly domains" (Smiraglia 2012, 118).

Domain analysis provides a comprehensive understanding of domains, while also helping scholarly communication by suggesting possible research collaborations. The domain would then not just act as an offering of tools for mapping a scientific field, its' disciplines and sub-disciplines, but would also reveal the characteristics of a discursive community. Hjørland and Albrechtsen (1995) introduce their paper, which is a milestone concerning domain analysis, explaining this domain's feature:

The domain-analytic paradigm in information science (IS) states that the best way to understand information in IS is to study the knowledge-domains as thought or discourse communities, which are parts of society's division of labor. Knowledge organization, structure, cooperation patterns, language and communication forms, information systems, and relevance criteria are reflections of the objects of the work of these communities and of their role in society. (Hjørland and Albrechtsen 1995, 400)

When it is discussed about scholarly communication amongst international researchers, it seems to be limited to those that reside in the same region. For example, ISKO could be the most influential and international conference for researchers in KO. Smiraglia (2011; 2013) has analyzed countries of affiliation of the first author of each paper from last two ISKO conferences: 11th conference in Rome in 2010 and 12th conference in India in 2012. Although 12th ISKO showed more diverse authors' country affiliations from such as India, Taiwan, Algeria, Iran, and Singapore, it might be due to the location of the conference. In addition, given that there are ISKO chapters from Brazil, Canada and United States, China, etc., scholarly communications in KO seem to be slanted by researchers from North and South America or Europe than from Asia, Middle East, or North Africa.

Similar studies have been done by McIlwaine & Williamson (1999) who analyzed trends in subject analysis research for the years 1988-1998 based on an analysis of 575 publications. In a follow-up study, McIlwaine (2003) again surveyed trends in KO in the years 1998-2003. The data used was drawn from journals and conference proceedings but most of the analysis relied on the author's knowledge of the field. López-Huertas (2008) provided a detailed and insightful review of what she perceived as being the current research trends in KO over "the last ten years" based on data collection from the Web of Science database (WoS). Saumure & Shiri (2008) conducted a trend survey of KO research in the pre- and post- web eras, from 1966-

2006. The authors observed that KO research has remained focused throughout the period covered on mainstream topics like cataloging and classification, which is similar to the conclusion by López-Huertas (2008). However, they characterized the pre-web era more by indexing and cataloging issues. A shift in the focus in the post-web era was noticeable with topics like metadata generation and harvesting by computers and interoperability issues. López-Huertas (2008) also thought that once traditional issues are recast in the framework of the web era, especially in the era of the semantic web, we can give new life to the traditional research issues.

Most recently Ibekwe-SanJuan and SanJuan (2010) applied an automatic topic mapping system to knowledge organization publications records published between 1988-2008. The authors collected the data from journals publishing KO articles from the WoS. The authors generalized their results by showing that topics in the first decade (1988-1997) were more traditional whereas topics in the second decade (1998-2008) was marked more by a technological orientation and the appearance of more specialized topics driven by the Web environment. These results were consistent with the previous studies by López-Huertas (2008) and Saumure & Shiri (2008).

Academic journals are another venue for scholarly communications amongst researchers. Researchers in countries with non-English primary languages tend to publish in those languages and in their own national journals. Although many international journals require English abstracts or keywords, that information might not be sufficient and accessibility issues might remain. These issues could be a result of various interpretations of a domain, which could be a result of gaps in language and/or translations. This can hinder international scholars from active communication and potential research collaboration.

To make clear the domain in which will be analyzed in this paper, this study takes into account the methodological paper from Tennis (2003), who presents two analytical devices, built on Hjørland's work, to support domain analysis. Firstly, Tennis (2003) recalls the eleven approaches proposed by Hjørland (2002). Then, Tennis presents the two axes to shape Hjørland's approaches which may support the choices made by the domain analyst: "Areas of Modulation, which sets parameters on the names and extension of the domain, and the second axis is Degrees of Specialization, which qualifies and sets the intension of the domain" (Tennis 2003, 192). This is a descriptive study and, more specifically, a terminological study as presented by Hjørland (2002) in his ninth approach.

## **Method**

For this study's purpose, four countries - Brazil, South Korea, Spain and the United States have been selected based on the authors' language capacity. For the purpose of our study, each country selected two LIS journals from that particular country. The selection of journals related to KO was a problem as far as limiting the selected journals to KO journals or related to KO. Some countries like South Korea do not have specialized journals for KO, whereas there are several distinct KO-specialized journals written in English such as Knowledge Organization, Cataloging and Classification Quarterly, etc. Therefore, this study relies on the four authors' expertise to select two journals from each country. In addition, the four authors asked KO scholars in each country to recommend two journals that have a good representation of

KO research. Through the purposive sampling, this study analyzed the following journals (See Table 1). The scope of data is limited to five years of each journal, 2007 to 2011.

Table 1. Selected LIS journals

| Country       | Journal's title   |
|---------------|---|
| Brazil        | <ul style="list-style-type: none"> <li>• <i>Perspectivas em Ciência da Informação</i></li> <li>• <i>Transinformação</i></li> </ul>  |
| South Korea   | <ul style="list-style-type: none"> <li>• <i>Journal of Korean Library and Information Science Society</i></li> <li>• <i>Journal of Korean Bibliography Society for Library and Information Science</i></li> </ul> |
| Spain         | <ul style="list-style-type: none"> <li>• <i>Scire</i></li> <li>• <i>Profesional de la Información</i></li> </ul>  |
| United States | <ul style="list-style-type: none"> <li>• <i>Journal of the American Society for Information Science and Technology</i></li> <li>• <i>Journal of Documentation</i></li> </ul>                                      |

A total number of articles analyzed is 2488. The study needed to sort out KO-related articles from each journal because of the general scope of the selected journals. With 94.08 % of intercoder reliability, 468 articles were identified as KO-related research (See table 2).

Table 2. # of articles

|  | Brazil | South Korea | Spain  | USA    | Total  |
|--|--------|-------------|--------|--------|--------|
| Total number of articles                     | 309    | 652         | 273    | 1254   | 2488   |
| # of KO-related articles                     | 69     | 116         | 77     | 206    | 468    |
| Percentage of KO in journals of each country | 22.33% | 17.79%      | 28.20% | 16.42% | 18.81% |

## Results

The study analyzed keywords from the 468 KO-related articles. Table 3 shows the number of unique keywords from each country.

Table 3. # of unique keywords

|                                  | Brazil | South Korea | Spain | USA |
|----------------------------------|--------|-------------|-------|-----|
| # of independent keywords        | 210    | 420         | 329   | 413 |
| # of keywords more than one time | 11     | 71          | 34    | 103 |

For the comparison of keywords among four countries, the study examined keywords that occurred more than one time. About the top 10 keywords from each country shows to some extent the similarity and differences of keywords among the four countries (See table 4). Given two axes of domains for domain analysis suggested by Tennis (2003), some keywords such as classification, cataloging, knowledge/information organization tend to represent extension of KO domains of four countries, while others identify intensions of KO domains. Especially the KO domain represented by American journals tends to be broader than the KO domains from the rest of three countries. For example, keywords from American journals include

Classification, Knowledge Organization, and Categorization, whereas keywords from Korean or Spanish journals include more specific keywords such as Korean/Dewey/Nippon Decimal Classification, RDA, FRBR, Legal information system, Web 2.0, etc.

Table 4. Top frequent keywords from each country

| Brazil                           | South Korea                   | Spain                          | USA                   |
|----------------------------------|-------------------------------|--------------------------------|-----------------------|
| Knowledge Representation         | Korean Decimal Classification | Semantic Web                   | Classification        |
| Knowledge Management             | Dewey Decimal Classification  | Ontologies                     | Information Retrieval |
| Ontologies                       | RDA                           | Thesauri                       | Information Science   |
| Automatic Indexing               | Nippon Decimal Classification | Knowledge Organization         | Systems               |
| Classification Systems           | Korean Cataloging Rule        | Information Architecture       | Retrieval Science     |
| Information Organization         | FRBR                          | Knowledge Management           | Web Model             |
| Information Science              | OPAC                          | Information Retrieval          | Information Knowledge |
| Knowledge Management Instruments | Library Catalog Metadata      | Information Systems            | Organization Search   |
| Online Catalog                   | Subject Headings              | Internet                       | Categorization        |
| Ontology                         |                               | Knowledge Organization Systems | Knowledge             |
| Semiotics                        |                               | Legal Information Systems      |                       |
| Terminology                      |                               | Metadata                       |                       |
|                                  |                               | SKOS                           |                       |
|                                  |                               | Spain                          |                       |
|                                  |                               | University Libraries           |                       |
|                                  |                               | Web 2.0                        |                       |

Figure 1 and Table 5 show the overlapped keywords among the four countries. There is only one keyword, *Online catalog*, appearing in all countries. 21 overlapped keywords show common research topics in KO across the countries. Comparing to the top frequent keywords from each countries in Table 4, there are some differences from overlapped keywords. For example, although *online catalog* is the only topic that belongs to the KO domain of all four countries, it is not a topic that is frequently studied by all countries. This illustrates that the domains of each country focuses on different intensions in KO.

Figure 1. Overlapped keywords among the countries

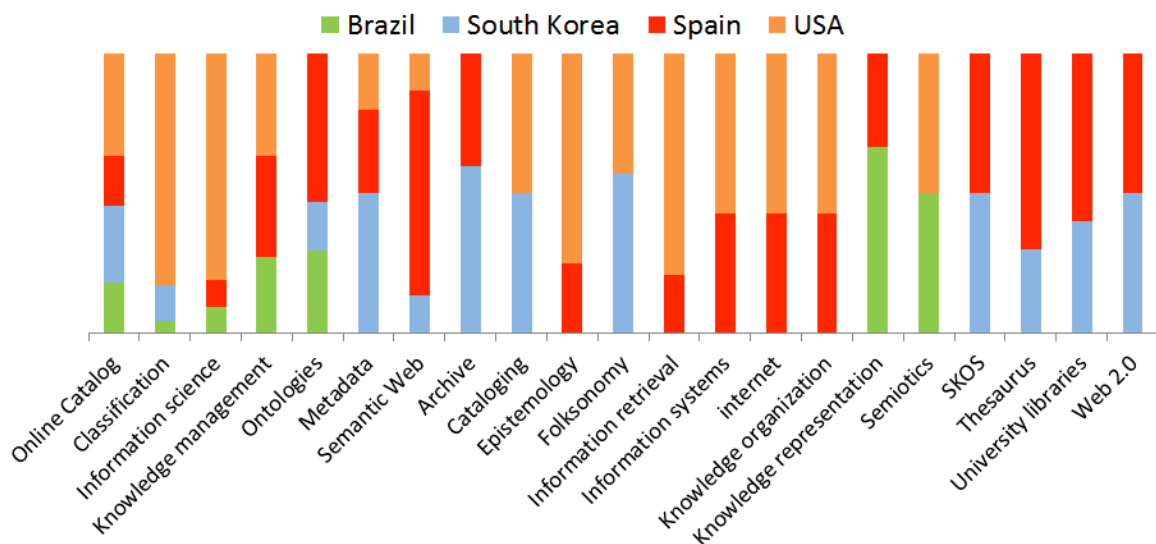


Table 5. Overlapped keywords with frequency

|                      | Brazil | South Korea | Spain | USA |
|----------------------|--------|-------------|-------|-----|
| Online Catalog       | 2      | 3           | 2     | 4   |
| Classification       | 2      | 6           |       | 39  |
| Information science  | 2      |             | 2     | 17  |
| Knowledge management | 3      |             | 4     | 4   |
| Ontologies           | 5      | 3           | 9     |     |
| Metadata             |        | 5           | 3     | 2   |
| Semantic Web         |        | 2           | 11    | 2   |
| Archive              |        | 3           | 2     |     |
| Cataloging           |        | 3           |       | 3   |
| Epistemology         |        |             | 2     | 6   |

|                          |   |   |   |    |
|--------------------------|---|---|---|----|
| Folksonomy               |   | 4 |   | 3  |
| Information retrieval    |   |   | 5 | 19 |
| Information systems      |   |   | 3 | 4  |
| internet                 |   |   | 3 | 4  |
| Knowledge organization   |   |   | 6 | 8  |
| Knowledge representation | 4 |   | 2 |    |
| Semiotics                | 2 |   |   | 2  |
| SKOS                     |   | 3 | 3 |    |
| Thesaurus                |   | 3 | 7 |    |
| University libraries     |   | 2 | 3 |    |
| Web 2.0                  |   | 3 | 3 |    |

Table 6 shows some examples of unique keywords from each country. For example, a domain of KO in Brazil shows more interests about automatic indexing and terminology. A domain of KO in South Korea includes children's library classification, kid's catalog, interoperability, next generation library catalog, and so on. Attention to children's libraries in KO is only represented by South Korea. A domain of KO in Spain also shows some unique topics such as information architecture, visualization, legal information system, RDF, and XML. These keywords suggest that a domain of KO in Spain is more interested in encoding schemes or visualization. Last, since keywords from American journals are general, it is hard to identify unique research topics. However, given some keywords such as systems, retrieval, or model, a domain of KO in the United States seems to be more closely associated with some aspects of information retrieval than KO domains from other countries.

Table 6. Unique keywords of each country

| Brazil                            | South Korea  | Spain  | USA   |
|-----------------------------------|--|--|---|
| Automatic indexing<br>Terminology | Children's Library Classification<br>FRBR<br>Interoperability<br>Kid's Catalog<br>Next generation library catalog<br>RDA<br>Subject Headings | Information Architecture<br>Information visualization<br>Legal Information Systems<br>RDF<br>XML | Systems<br>Retrieval<br>Science<br>Web<br>Model<br>Bibliographic Systems<br>Categorization<br>Knowledge<br>Latent Semantic Analysis<br>Topicality |

The number of overlapped keywords among the countries also implies that the KO domain of Spain tends to share more similar research topics with the United States and South Korea rather than Brazil. There is an explicit difference of domains between Brazil and Korea (See table 7).

Table 7. Overlapped keywords between two countries

| Countries      | # of overlapped keywords | Keywords   |
|----------------|--------------------------|--|
| Spain & USA    | 9                        | Information Science, Knowledge Management, Metadata, Semantic Web, Epistemology, Information Retrieval, Information System, Internet, Knowledge Organization |
| Korea & Spain  | 8                        | Metadata, Ontologies, Semantic Web, Archive, SKOS, Thesaurus, University Libraries, Web 2.0  |
| Korea & USA    | 4                        | Metadata, Semantic Web, Cataloging, Folksonomy   |
| Brazil & Spain | 3                        | Information Science, Knowledge Management, Knowledge Representation  |
| Brazil & USA   | 3                        | Information Science, Knowledge Management, Semiotics   |
| Brazil & Korea | 1                        | Classification   |

### Discussion

Divergences in the top frequent keywords for each country might be explained, in terms of domain analysis operationalization (see Tennis 2003), by the modulation of the KO domain defined by each journal and the degree of specialization of the KO domain construed by their authors with their descriptions. For instance, keywords from the American journals tend to be more general rather than providing KO specific terms, perhaps due to the more general scope of those journals and because broader KO keywords are not presumed in every article published on these journals. In this vein, the lesser intension of these keywords might be related to either a more general view of the KO domain, envisaged and shaped by the publications on these journals, or the more multidisciplinary scope of these journals.

In this vein, more general journals would be more likely to define a less intensive KO domain while not helping to clearly delimit its extension, being authors perhaps more likely to use general keywords in order to engage their KO research with other topics and audiences. In terms of analysis, these general keywords might not be the most helpful ones to compare to more specific keywords, such as those ones found on our selected Brazilian, South Korean and Spanish journals.

However, in spite of their different intensions, it was possible to identify a correspondence of topics represented by these different keywords across countries in our analysis. For instance, the most frequent keyword in the American journals, "Classification," is a broad concept of the two most frequent South Korean keywords "Korean Decimal Classification" and "Dewey Decimal Classification," and it is closely related to the Brazilian keyword "Classification Systems." Similarly, other American keywords such as "Web" and "Systems" are broader concepts of Spanish keywords such "Semantic Web" and "Web 2.0," and "SKOS" respectively. On the other hand, it is also worth noting that in the case of the Spanish journals, the greater generality of keywords did not seem to show a correlation with a lesser specialization of the journal. For instance, the very general keyword "Knowledge Organization," ranking fourth in



the Spanish top frequent keywords list after three other more specific keywords, appears only once in the most general Spanish journal of the two, *El Profesional de la Información*, while it is used six times in *Scire*, that is a KO specific journal, in spite of not being a distinctive keyword among all the articles published in this journal. In this vein, it seems to be suggested that the use of keywords is not being use to define the domain only within the journal corpus but also within the bibliographic databases in which journals are being indexed.

Concerning the unique keywords of each country, it seems that despite the different levels of specification detected in our analysis, journals of each country show different interests and construe a KO domain that, although having a common core around the classical Cataloging and Classification concepts, might show different extensions depending on each country. Brazilian keywords seem to suggest a greater interest in terminological and epistemological issues, i.e. they show a most epistemological conception of KO by these journals. South Korean keywords seem to be more related to specific systems, schemas and models, and therefore showing a more applied conception of KO. Spanish journals seem to show more interest in the Web and its applications and technologies. American journals seem to show a conception of the KO domain in which retrieval plays a very important part. All these aspects might also be considered a reflection of the way that the KO domain is being construed and defined by journals in each country, and the way these conceptions internationally communicated in databases to construct the global KO domain.

### **Conclusion**

This study tries to reveal the domain of KO represented in LIS journals from four countries. KO domains from four countries share common research topics of KO. It suggests that KO domain has developed important and fundamental research topics internationally. On the other hand, given there are many unique and nationalized research topics from each country, it means that KO domain is organically growing. However, there are many granular topics studied by each country. These topics have not been studied again. Therefore, it also implies that KO domain has to nurture and pay attention to potential research streams. This study also appears to be unique in comparison to the reviewed literature in that it placed more emphasis on which country is doing what in order to provide insight into research being done in the KO domain in these various countries.

### **References**

- Hjørland, B. 2002. Domain analysis in information science: Eleven approaches - traditional as well as innovative. *Journal of documentation* 58: 422-62.
- Hjørland, B. 2008. What is knowledge organization (KO)? *Knowledge organization* 35: 86-101.
- Hjørland, B., and Albrechtsen, H. 1995. Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science* 46: 400-25.
- Ibekwe-SanJuan F. and SanJuan E. Knowledge organization research in the last two decades: 1988-2008. *11th International Conference of the International Society for Knowledge Organization (ISKO), Paradigms and conceptual systems in KO*, Sapienza University of Rome, February. 23-26 2010, pp. 115-121.

- López-Huertas, M.J. 2008. Some current research questions in the field of knowledge organization. *Knowledge Organization* 35: 113-36.
- McIlwaine I.C. (2003). Trends in knowledge organization. *Knowledge Organization* 30(2): 75-86.
- McIlwaine I.C. and Williamson, N.J. (1999). International trends in subject analysis. *Knowledge Organization* 26(1): 23-29.
- Saumure K. and Shiri A. (2008) Knowledge organization trends in library and information studies: a preliminary comparison of the pre- and post-web eras. *Journal of information science* 34(5): 651-66.
- Smiraglia, R.P. 2011. ISKO 11's diverse bookshelf: An editorial. *Knowledge organization* 38: 179-86.
- Smiraglia, R.P. 2012. Epistemology of domain analysis. In *Cultural frames of knowledge*, ed. R. P. Smiraglia, and H. Lee. Würzburg, Germany: Ergon, pp. 111-24.
- Smiraglia, R.P. 2013. ISKO 12's bookshelf – evolving intension: An editorial. *Knowledge Organization* 40: 3-10.
- Tennis, J. 2003. Two axes of domains for domain analysis. *Knowledge Organization* 30: 191-5.