

# II Congresso Brasileiro em Representação e Organização do Conhecimento: Knowledge Organization in Rio 2013—An Editorial

Richard P. Smiraglia

Smiraglia, Richard P. **II Congresso Brasileiro em Representação e Organização do Conhecimento: Knowledge Organization in Rio 2013—An Editorial.** *Knowledge Organization*. 42(2), 105-112. 2 references.

**Abstract:** The Second Brazilian Conference on Knowledge Organization and Representation took place in Rio de Janeiro, in May 2013. It was a lively gathering with international attendance through three days of presentations by keynote speakers as well as a large number of contributed research papers. The proceedings contain forty-three papers or extended abstracts from the speakers. Citation analysis, co-word analysis, and author co-citation analysis were employed to analyze the relationship between the domain of this conference and that of ISKO at large. The highly productive new ISKO chapter in Brazil is shown to be contributing to the further evolution of the core of knowledge organization as a science by both embracing and extending the application of core concept theory in the extension of the domain. Like their counterparts in other regions, they also are helping to refresh the intension of the domain with a large portion of local applied research, especially in bibliometrics analyses of scientific productivity, in the extension of knowledge representation and organization to archival practice, and in the integration of concepts of ontology along major epistemological trajectories.



**Keywords:** knowledge organization, domain analysis, ISKO Brazil, epistemology

## 1.0 The Second Brazilian Conference on Knowledge Organization and Representation

The Second Brazilian Conference on Knowledge Organization and Representation took place in Rio de Janeiro, at the Getulio Vargas Foundation—FGV, on May 27, 28 and 29, 2013. It was a lively gathering with international attendance through three days of presentations by keynote speakers as well as a large number of contributed research papers. The proceedings have been published as Dodebei and Guimarães (2013), containing forty-three papers or extended abstracts from the speakers. As one of the keynote speakers presenting on the epistemology of knowledge organization, I was able to demonstrate geo-political differences in the expression of knowledge organization as a science through meta-analytical domain analysis (see Smiraglia 2014 for a more recent iteration of that work). Brazilian authors, interestingly, had produced about one-third of the papers at the most recent International ISKO Conference held in Mysore, India in 2012. An interesting research question then, is, how does Brazilian knowledge organization align with the various parameters of the international knowledge organization

domain? In this editorial I present some informetric indicators based on the papers from the Rio conference (and a few photos to remind readers that KO is global).

## 2.0 Simple bibliometrics as indicators of domain epistemology

The forty-three papers contained 476 references, for a mean of 11 references per paper; the range was zero to 48. These are simple epistemological indicators—empirical science would tend to rely on a small range of recent citations, but more humanistic approaches would rely on extensive literature reviews. As is typical of the knowledge organization domain in general, we see both here. Twenty-four papers contain 1-11 references, and fifteen papers contain 12-25 references; only one paper contains 48 references. Geographical orientation shows that papers from Brazilian authors span the entire range, and the few papers from the United States, Uruguay, and Spain fall in the mid-range between 9 and 12 references. The contributions to the conference were overwhelmingly Brazilian, with 80% of the papers. Most of the papers from Europe or North America represented contri-



Three views of Rio de Janeiro during the 2<sup>nd</sup> ISKO-Brazil Conference

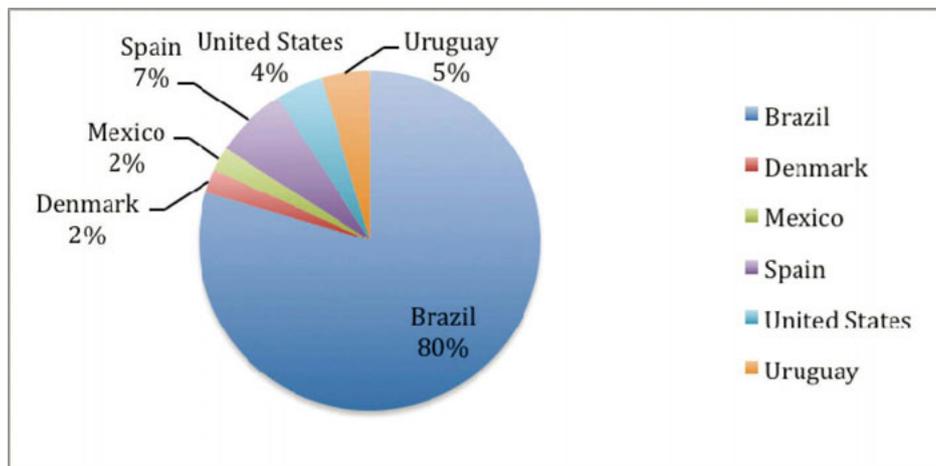


Figure 1. Geographic orientation of the papers.

butions by invited speakers. The geographical orientation is visualized in Figure 1.

The geographic distribution was sufficiently homogenous that it was not useful to further analyze the numbers of citations. The next step was to look at the dates of works cited. This was difficult, as is usual in ISKO proceedings, because no particular style was required of the authors or imposed by the editors, so of the 476 refer-

ences only about 200 were manipulable in a spreadsheet; the rest had to be manually typed and delimited. Editors of ISKO proceedings should take note; you are contributing to the poor indexing of our domain when you do not require authors to adhere to standard citation practices (APA is not acceptable for indexing; one of the Chicago style sheets is preferable).

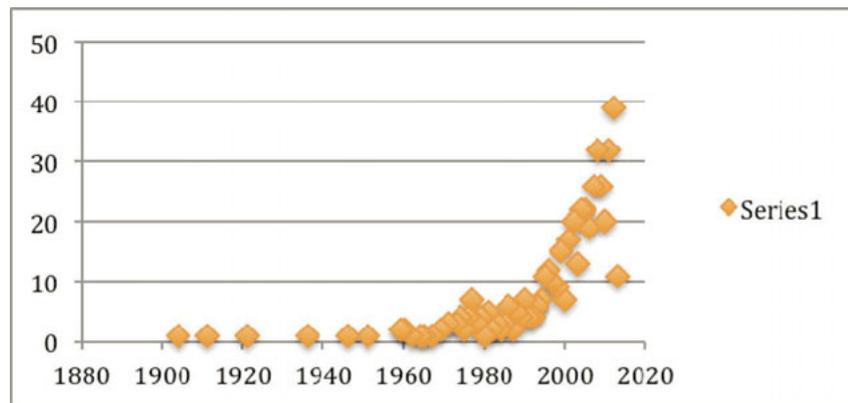


Figure 2. Dates of works cited.

The dates of works cited are shown in Figure 2. The range was from 1904 to 2013, and the clear majority were to works cited after 2000, with only sparse citations to works prior to 1860. This tells us that the works cited conform to works in information in general (a domain dating formally from about 1961), and that the majority of works cited are recent, which is an indicator of empirical epistemologies among contributors.

Again, because of the homogeneity of the contributions, no further analyses of dates of works cited were undertaken. The papers were sorted by primary medium; the distribution is in Table 1.

monograph	210
journal article	155
conference proceedings	31
other	24
dissertation(PhD)	18
website	13
thesis	8
report	3
database	2

Table 1. Distribution of media among cited works.

The humanistic bent of the participants is clear in this distribution with the majority of citations to monographic works rather than to scientific papers. We might consider adding journal articles, conference papers, dissertations and theses to gain a rough estimate of what might have been empirical research, in which case we have 212 for a nearly even split, which is consistent with other studies of knowledge organization.

### 3.0 Determining the intension and extension using co-word analysis

The thematic orientation of the domain can emerge in two explicit ways using co-word analysis; through analysis of keywords in the titles of contributed papers (sometimes abstracts are used as well), and through parallel analysis of keywords in the title of papers cited by the contributing authors. By comparing these two analyses we can see the extension of the domain as represented by the explicit intentions of the authors, as well as the intension of the domain as filtered through the granular prior research on which their work relies. Both lists of titles were entered into WordStat™ software. Term frequencies are calculated first and these are shown in Table 2.

We see two simple lists side-by-side. Interestingly, we see elements of the phrase “knowledge organization” in Portuguese, Spanish and English. The core terms here are the core terms in the knowledge organization domain. On the right, we see the thematic correspondence of the literature on which the authors relied. It is the same, of course, as the core thematic content of their papers, but more linguistically diverse. This table shows us the extension of ISKO Brazil’s domain as of the time of this conference. There are some pretty basic concepts, but they are used in three languages. This tells us that the basic extension of the domain is fixed, and it also shows us how there is dissemination across linguistic boundaries.

Co-word analytical tools in Provalis Suite™ include clustering algorithms and multi-dimensional scaling. Both lists were entered into WordStat™ from which three-dimensional visualizations were produced. The visualization of the list of title keywords is first in Figure 3.

One can see immediately that the map is sparsely populated, which is a reflection of the meta-level terms that were used by the authors of these papers. There are three clusters, representing: knowledge organization and representation (the word “artigos” meaning “articles” occurs in this cluster indicating the bibliometric nature of

Conference paper keyword	Frequency	Cited paper keyword	Frequency
CONHECIMENTO	21.50%	KNOWLEDGE	4.60%
ORGANIZAÇÃO	16.50%	INFORMATION	4.30%
REPRESENTAÇÃO	13.90%	INDEXING	3.60%
ANÁLISE	6.30%	INFORMAÇÃO	3.60%
CIENTÍFICA	6.30%	ORGANIZATION	3.60%
PRODUÇÃO	5.10%	SUBJECT	2.90%
INFORMAÇÃO	5.10%	ANALYSIS	2.60%
PESQUISA	5.10%	ANÁLISE	2.10%
ARTIGOS	3.80%	THEORY	2.10%
INDEXAÇÃO	3.80%	CIÊNCIA	1.90%
REFLEXÕES	3.80%	CONOCIMIENTO	1.90%
SOCIAL	3.80%	BRASIL	1.70%
		INDEXAÇÃO	1.70%
		SCIENCE	1.70%
		CIENTÍFICO	1.60%
		CLASSIFICATION	1.60%
		INFORMACIÓN	1.60%
		REPRESENTAÇÃO	1.60%
		RETRIEVAL	1.60%
		CONHECIMENTO	1.40%
		DIGITAL	1.40%
		TEÓRICO	1.40%

Table 2. Keyword frequencies, paper titles and cited reference titles.

much of the research at the conference), reflections on research and social implications of information comprise the central cluster, and indexing and analysis of scientific productivity comprise the third and leading cluster at the left.

Figure 4 shows the MDS visualization of keywords from the cited titles. There is a much broader mix, and we can see that there are English and Spanish clusters amid the Portuguese. This is a fascinating result, because it shows that there is a genuinely global and multilingual community at work here. Still, the content of the clusters hovers at the same meta-level as in Figure 3; information documentation and knowledge organization at the research front, with reference to classical classification indexing and information retrieval from the English-language sources, and theoretical science information, indexing, documentation and linguistics from the Spanish-language sources. All of it is anchored at the center by the Brazilian meta-level theoretical work in knowledge organization and representation.

#### 4.0 Triangulating with author co-citation analysis

Basic bibliometric measures told us something about the epistemological stance of the contributors to this conference, and co-word analysis showed us a meta-level orientation on key concepts of the domain. There has been little granularity visible in this analysis of this conference; rather, the core seems quite solid and recursive. Author co-citation analysis is often used as a domain analytical tool to help us visualize how a domain is perceived both from outside and from within. The perceived similarities of writings cited together can demonstrate both a domain's intension and extension, essentially offering a methodological triangulation of the co-word analysis. In this case we will take a snapshot of each.

We begin by determining the most-cited authors in the domain, from the cited references in the conference papers. Of the 476 cited references, sixty were to works by authors cited more than once. A frequency distribution of authors cited more than once was constructed. Authors cited four times or more appear in Table 3. It is an interesting list, mixing as it does authors from the domain

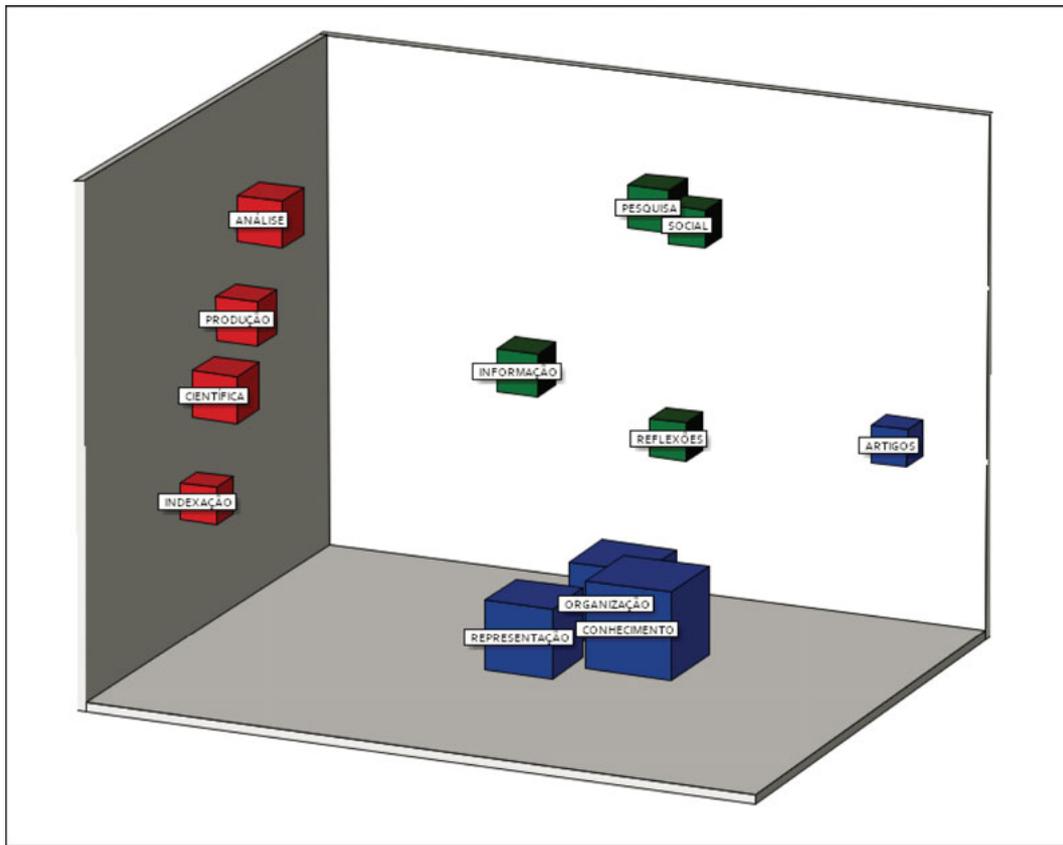


Figure 3. MDS visualization of title keywords (Stress = 0.18243 R<sup>2</sup>=0.9302).

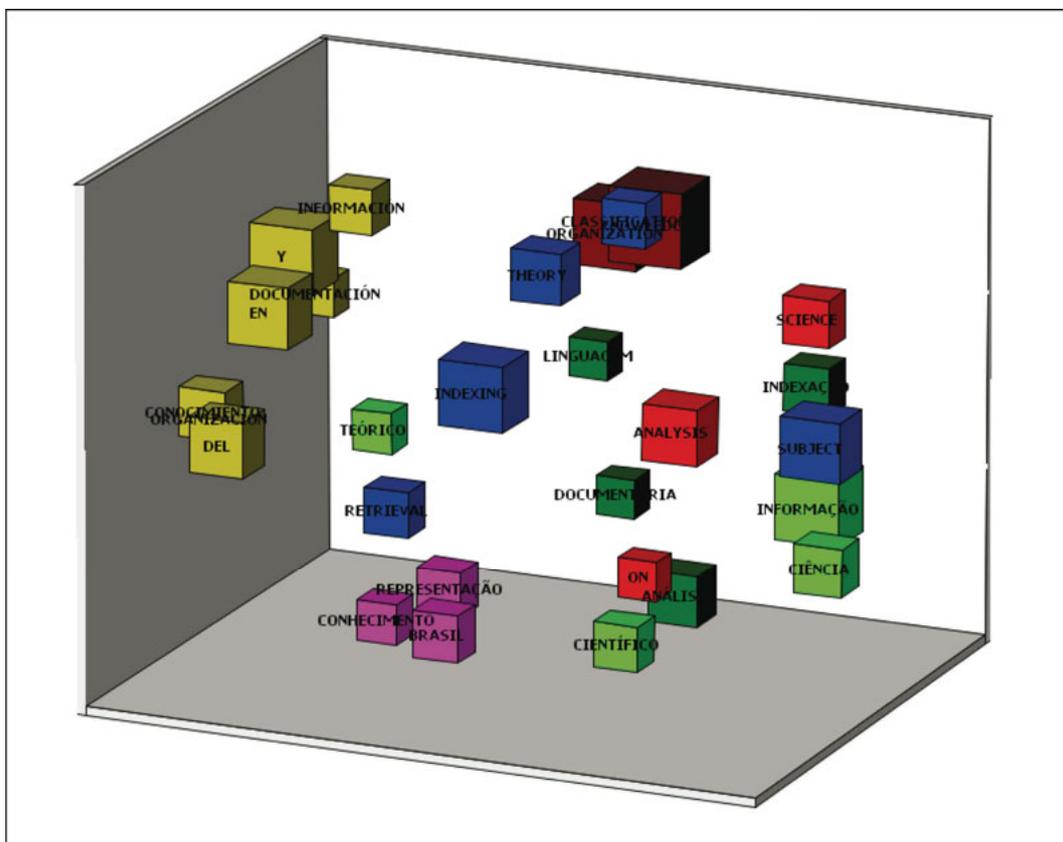


Figure 4. MDS visualization of cited title keywords (cited title keywords Stress = 0.26859 R<sup>2</sup>=0.7401).

of knowledge organization globally. Two corporate “authors” that were much cited by conference papers were the Book Industry Study Group and the International Standards Organization. Those were removed from the list prior to constituting author co-citation matrices.

Author	Frequency
Hjørland	14
López-Huertas	10
Dahlberg	7
Fujita	7
Beghtol	6
Smiraglia	6
Smit	6
Araújo Júnior	5
García Gutiérrez	5
Gil Leiva	4
La Barre	4
Mai	4
Olson	4

Table 3. Authors most cited in the conference papers.

Author co-citation analysis is a means of visualizing the ways in which citing authors perceive a domain. So in this case, if we consider this list of authors as somehow representative of, or core in, the domain of the Second ISKO-

Brazil conference, we can learn from a visualization how all authors in the larger knowledge organization domain perceive the interrelationships among them, particularly with regard to the thematic content of their published research. A co-citation matrix was compiled for these thirteen authors from the Web of Science™ and IBM-SPSS™ was used to construct an MDS visualization of these authors in intellectual space. This visualization is shown in Figure 5.

We see two clusters at something of a distance from each other. The cluster on the right clearly represents traditional knowledge organization, and the cluster on the left mixes influences from Spanish, Brazilian and American authors. Bearing in mind that these data came from the Web of Science™, we can interpret this map to represent a perception among global authors in knowledge organization. Figure 6 contains a visualization of actual author co-citation in the papers of the Second ISKO-Brazil conference. It is rather a different map.

Here we see more or less the same groupings but differently positioned, and with an epistemological axis implicitly ranging from the empirical on the left to the rational and historicist on the right. This map represents the manner in which the authors at the conference perceive their own work alongside or within the intension of the domain of knowledge organization. There is a clear integration of traditional knowledge organization concepts. But there also is original work from the Americas,

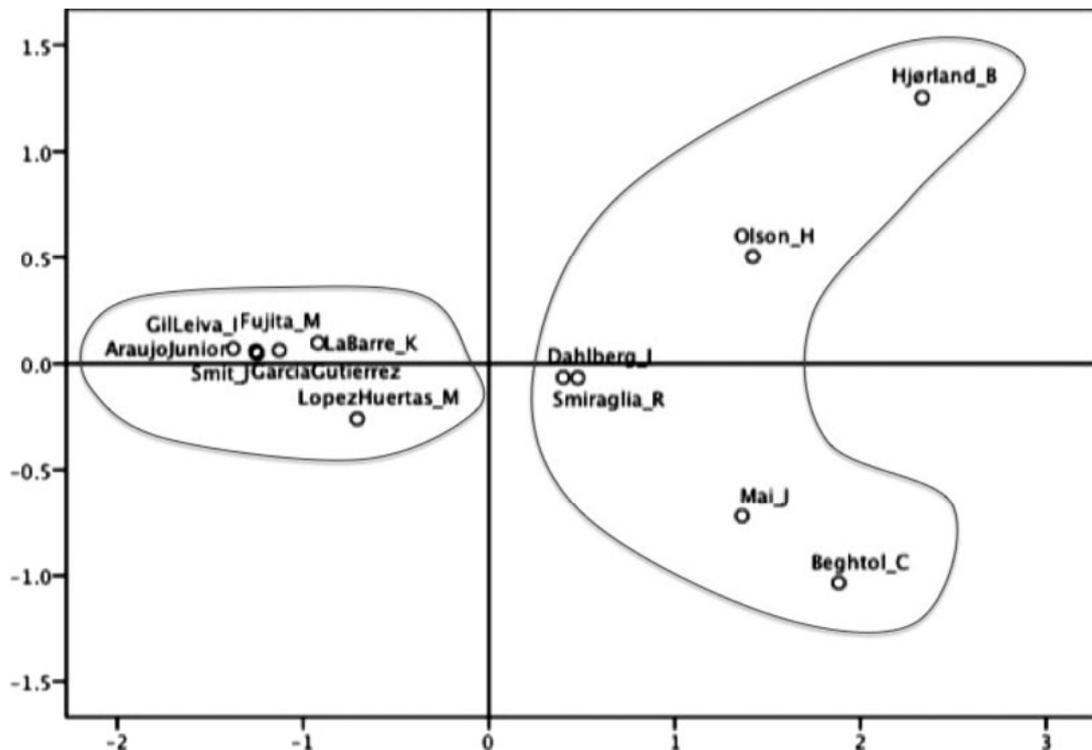


Figure 5. MDS visualization of author co-citation from Web of Science™ (Stress = .04148 R² = .99453).

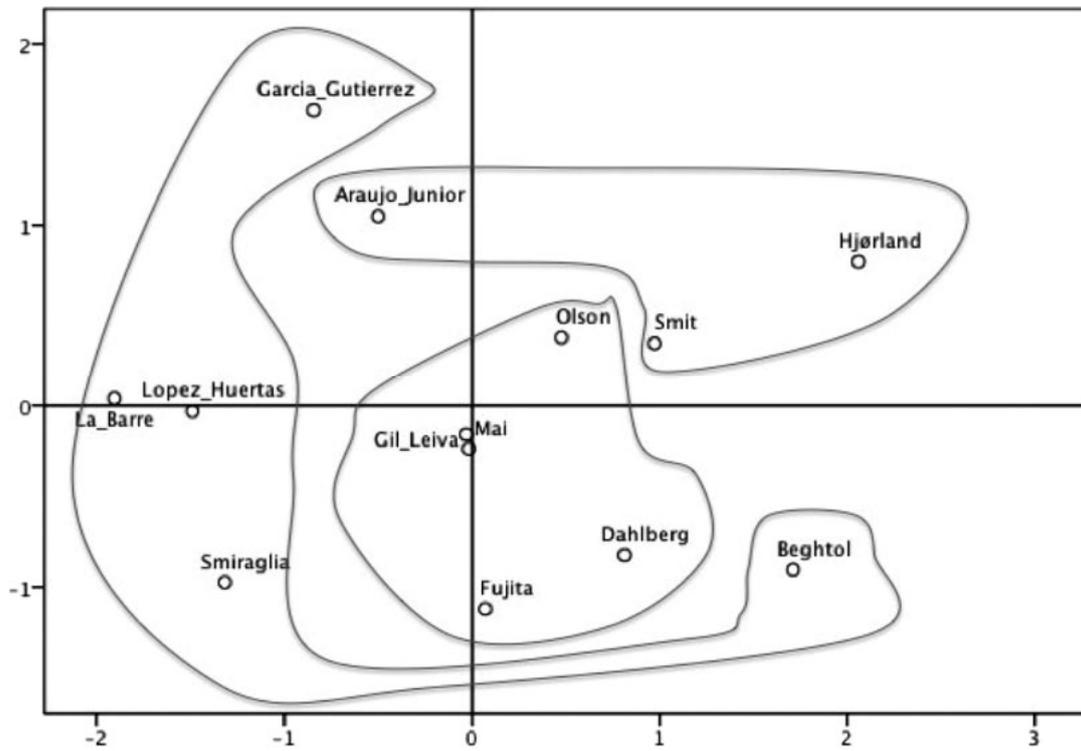


Figure 6. MDS Visualization of author co-citation within the conference (Stress = .12311 R<sup>2</sup> = .90868).

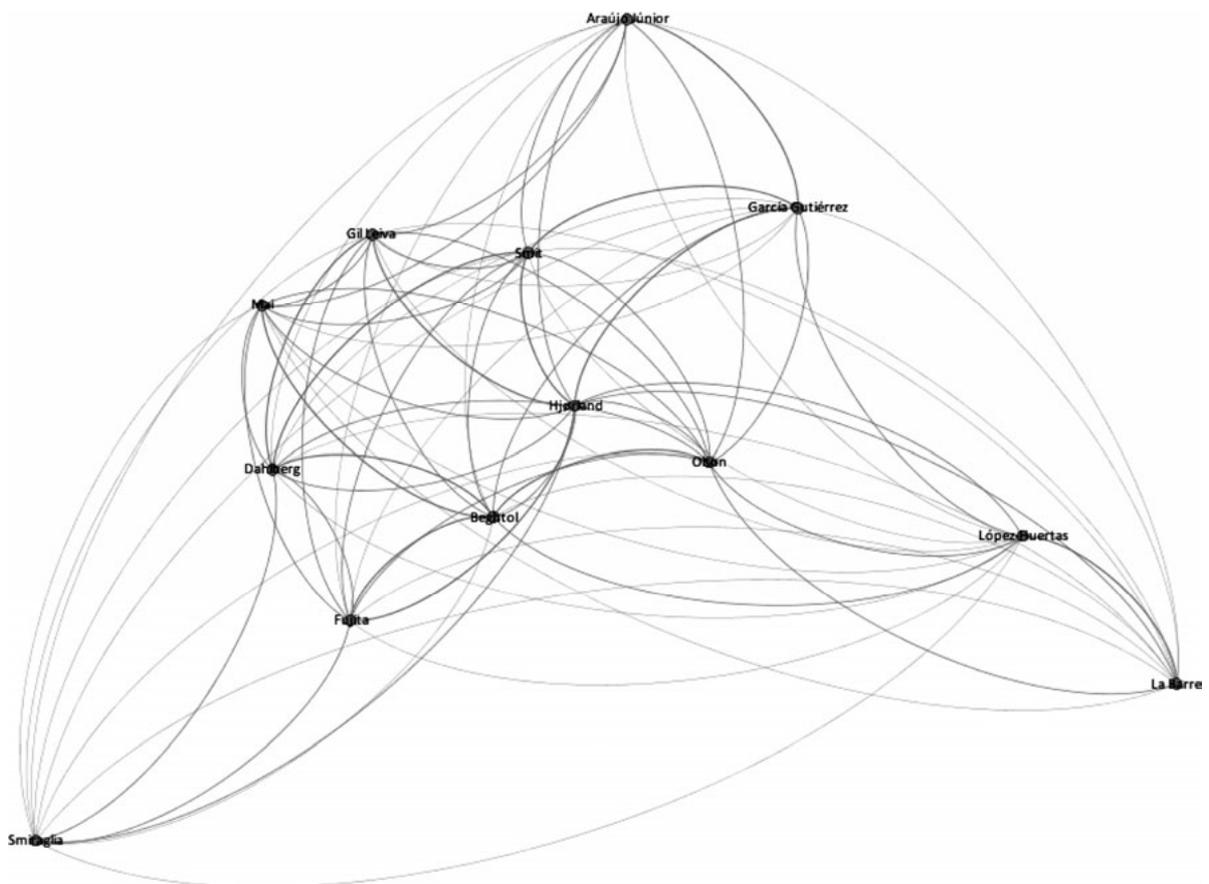


Figure 7. Gephi network diagram of interconference author cocitation.

particularly with regard to indexing, documentation, and informetrics. Another visualization is possible, a network diagram showing the direct co-citation connections among these authors. This is produced using Gephi™ and appears in Figure 7.

Here we see more clearly the perceived connections between the thematic nodes, with Smiraglia (epistemology) and La Barre (facet theory) at extreme points but linked extensively to the core. The density of the core represents stable concept theory from traditional knowledge organization, but it is inextricably linked to the local emphases on knowledge representation and indexing.

Taken altogether we can see that the highly productive new ISKO chapter in Brazil is contributing to the further evolution of the core of knowledge organization as a science by both embracing and extending the application of core concept theory in the extension of the domain. But like their counterparts in other regions, they also are helping to refresh the intension of the domain with a large

portion of local applied research, especially in bibliometrics analyses of scientific productivity, in the extension of knowledge representation and organization to archival practice, and in the integration of concepts of ontology along major epistemological trajectories.

## References

- Dodebeí, Vera and José Augusto Chaves Guimarães, eds. 2013. *Complexidade e organização do conhecimento: desafios de nosso século*. Estudos Avançados em Organização do Conhecimento v.2. Rio de Janeiro, Brasil: Sociedade Brasileira de Organização do Conhecimento (SKO-Brasil); Marília: Fundação para o Desenvolvimento do Ensino, Pesquisa e Extensão (FUNDEPE).
- Smiraglia, Richard P. 2014. The epistemological dimension of knowledge organization. *IRIS - Revista de informação, memória e tecnologia* forthcoming.