

# ISKO 11's Diverse Bookshelf: An Editorial\*

Richard P. Smiraglia, Editor-in-Chief

\* The author would like to acknowledge the assistance of Elizabeth Milonas with acquiring the text of the 1998 classification described in this paper.

## 1.0 When in Rome

As we all know, Knowledge Organization (KO) is a pretty broad domain. Although the concept-theoretic approach to classification is at the core along with several other important pieces of what we call classification theory, both the intension and the extension of the domain are represented by broad trajectories. Arguably, the biennial conferences represent way stations within the matrix of the domain—points in time when we pause to take stock of our current research. Also, because each conference is hosted and planned by a regional chapter, each then reflects peculiar parameters of the intersections of intensional and extensional trajectories. Perhaps because the domain of knowledge itself is so immense, so also is our corporate attempt to grapple with the theoretical and applied aspects of its organization. Furthermore, because of the breadth of our domain, many possibilities exist for its representation, depending on the constitution of the research front (or fronts) at any moment in time. That is, research in the domain stretches in all directions from its solid theoretical core down many much more granular roadways. Thus by analyzing the activity and contents of these metaphorical way stations—that is, by bringing the tools of domain analysis to bear on our own biennial conferences—we are able to visualize the moment in time represented by the accumulated scholarship generated by each conference. 2010's 11<sup>th</sup> International ISKO Conference in Rome offered the latest opportunity for analysis on a broad scale.

To take advantage of the wonderful Italian weather, ISKO's 2010 conference was moved from the usual August to February; the venue was the Sapienza University (officially Sapienza - Università di Roma) and the conference took place 23-26 February 2010. The

conference was organized and hosted by ISKO Italy and the Faculty of Philosophy of Sapienza University. Each morning as attendees arrived, we were treated to the garden pictured in Figure 1, and especially interesting was the fountain and the statue of St. Francis. Of course, the mystery was the turtle at St. Francis' foot, which looks quite like part of the statue but turned out to be real. The peaceful gardens were just a hallmark of the contemplative nature of the conference.

Officially the 11<sup>th</sup> International ISKO Conference, the theme was "Paradigms and Conceptual Systems in Knowledge Organization." The proceedings and the conference program together listed 65 presentations, of which 64 were actually presented and 61 had papers included in the proceedings (or, 4 papers were presented but not included in the proceedings, and 1 paper included in the proceedings was not presented). Although space is insufficient for a full analysis, following from my editorial following ISKO 10 (Smiraglia 2008), I will use this space to paint a brief bibliometric portrait of the domain at the core of this conference. Data for this analysis come from the PDF of the proceedings; all citations for all papers were pasted in an Excel spreadsheet, where the citations were variously delimited for the following analyses. The original file is available on my blog: <http://lazykoblog.wordpress.com/>.

## 2.0 International Presence and Thematic Foci

The conference was truly international, of course, with authors affiliated with institutions in 19 countries (the top tiers of the distribution appear in table 1). The largest number came from the United States





Figure 1. St. Francis and the turtle

(29.2%), with large contributions from Italy, Germany, Spain, and the United Kingdom, and respectable showings by authors from Canada, Denmark, France and Hungary as is usual, but the second largest contribution came from Brazil (13.8%), whose chapter was just formed in 2009. All papers were presented in English and appear in English in the proceedings.

Country of Origin	Frequency	Percent
United States	19	29.2
Brazil	9	13.8
Italy	8	12.3
Germany	5	7.7
Spain	4	6.2
United Kingdom	3	4.6
Canada	2	3.1
Denmark	2	3.1
France	2	3.1
Hungary	2	3.1

Table 1. Distribution of countries of affiliation.

The proceedings were organized using a systematic table of contents, which employed three to five classes from an abridged scheme for KO literature (KO literature 1998). The lead term was used to designate a main theme for each paper and these were analyzed (the top

tiers appear in table 2). The largest thematic category was online technologies, with other large clusters occurring for KOS, language, and structure. Somewhat surprisingly, neither conceptology nor cultural warrant fell into this top tier; applications filled out the tail.

Theme	Frequency	Percent
online technologies	7	10.8
KOS	6	9.2
language	5	7.7
structure	5	7.7
conceptology	4	6.2
nonbook	4	6.2
science	4	6.2
special KOS	4	6.2
construction	3	4.6
cultural warrant	3	4.6
problems	3	4.6

Table 2. Distribution of papers by main theme.

The full classification scheme was acquired and each paper was assigned a broad “theme class” based on the name of the class to which each paper had been assigned. This yielded a set of ten descriptors. The frequency distribution of “theme classes” is shown in Table 3.

Theme class	Frequency	Percent
classification systems and thesauri	17	26.2
theoretical foundations	16	24.6
knowledge representing by language and terminology	15	23.1
applied classing and indexing	6	9.2
on special subjects cs&t	4	6.2
knowledge organization environment	3	4.6
classing and indexing	2	3.1
form divisions	1	1.5
on special objects cs (taxonomies)	1	1.5

Table 3. Theme Classes

Nine of the ten classes were occupied; the class that was not used was "4. On Universal Classification Systems and Thesauri." The majority of the papers fell in three categories, classification systems and thesauri, theoretical foundations, and knowledge representing. Cross-tabulation by country was problematic because so many cells had low numbers, but what was discernible were certain trends: papers from Brazil and Canada, for instance, were in applied classing and theoretical foundations; papers from Germany and Italy were in classification systems and knowledge representing, papers from Spain were in theoretical foundations, and papers from the US were in all classes. We can observe, without statistical significance, that most papers are in classification systems, knowledge representing, and theoretical foundations; and most of those papers come from the US, Brazil, Italy, and Germany. This is illustrated in Figure 2.

No country had only papers in non-populous categories. Applied papers came only from Brazil and

the US. Classing and indexing, and KO environments, and "on special subjects (taxonomy)" had few papers. Therefore, it seems there is not likely any geographical difference, but rather that this distribution simply represents the interests of the participants in this particular conference.

### 3.0 Citations define the domain

There were 967 citations in 65 papers. The number of citations per paper ranged from 3 to 45, with an overall mean of 14.88. The median was 7.5; the mode was 4.6—most papers had 5 citations, in other words, although there was a wide range. The mean per country was analyzed.

Country	Mean
South Africa	45
Sweden	30
Bulgaria	20
Brazil	17.44
Hungary	17
Netherlands	17
United Kingdom	16.67
Iran	16
Italy	15.75
United States	14.53
Spain	14.25
Canada	14
France	12.5
Denmark	11
Finland	11
Germany	8
Poland	8
India	5
Slovenia	0

Table 4. Mean citations per country

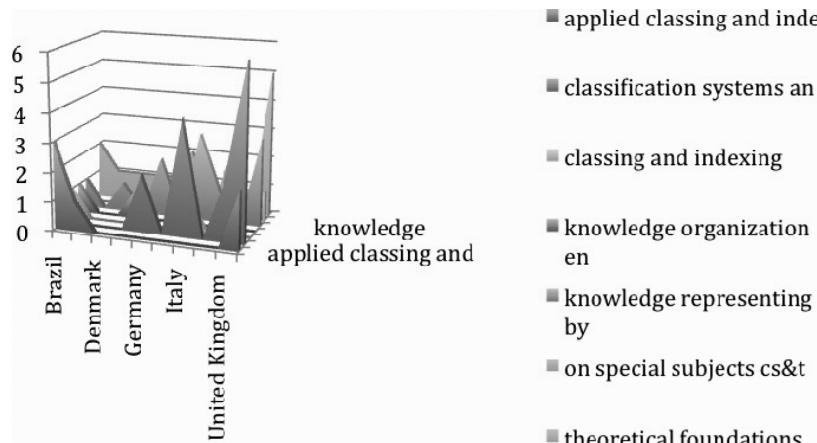


Figure 2. Plot of country of origin by thematic class

Observed differences are statistically significant, but as before, there are too many countries for the test to be meaningful. A glance at the means column shows that most country means are very near the overall mean of 14.88. The mean number of citations per theme was also analyzed (Table 5).

Theme	Mean
cultural warrant	26.33
science	24
KOS	21.83
order	18
language	16.4
terminology	15.67
special KOS	15.25
taxonomies	15
problems	13.33
structure	12.83
online technologies	12.57
nonbook	12.4
conceptology	11
construction	11
applied classing	8.5
indexing	8.5
psychology	8.5
compatibility	7

Table 5. Mean citations per theme

The means ranged from 7 for compatibility to 26.33 for cultural warrant. These differences are not statistically significant. However, we can see that there is a tendency in some thematic clusters toward humanistic citation practice and a more scientific pattern in others. This should be mirrored in the analysis of citation age. Table 6 arrays citation age per country and per theme side by side. Bear in mind as you look at this table that there is no connection between the countries on the left and the themes on the right.

The mean citation age varies from 6.2 in India to 14 in the United States; most countries hover around the mean citation age of 11.79 years, and ANOVA shows these differences by country are not statistically significant. Thematically, the mean age of citation ranges from 4.7 for psychology to 35 for order, and ANOVA shows that these differences are statistically significant. Together with the data in Table 5 we can make an assertion that there are different epistemological approaches living together in KO, ranging from humanistic to scientific, and these are reflected in both the mean number of citations used and in the age of works cited.

### 3.1 Most Cited Authors

A list of names of the authors most cited by conference contributors was compiled to generate a visualiza-

Country	Mean	Theme	Mean
Netherlands	15.52	order	35.1
United States	14.0495	taxonomies	17
Iran	13.81	nonbook	16.75
Brazil	13.6656	KOS	14.38
United Kingdom	13.5667	science	13.99
Bulgaria	12.8	structure	13.0717
Germany	11.866	terminology	13.005
South Africa	11.48	conceptology	12.008
Italy	10.9925	special KOS	11.345
Denmark	9.98	language	9.944
Hungary	9.5325	cultural warrant	8.9967
Canada	9.375	problems	8.9833
Finland	9.18	construction	8.8067
Spain	8.115	applied classing	8.525
Sweden	7.93	online technologies	7.1086
Poland	7.62	compatibility	6.54
France	6.85	indexing	5.44
India	6.2	psychology	4.735

Table 6. Country and Theme by Citation age

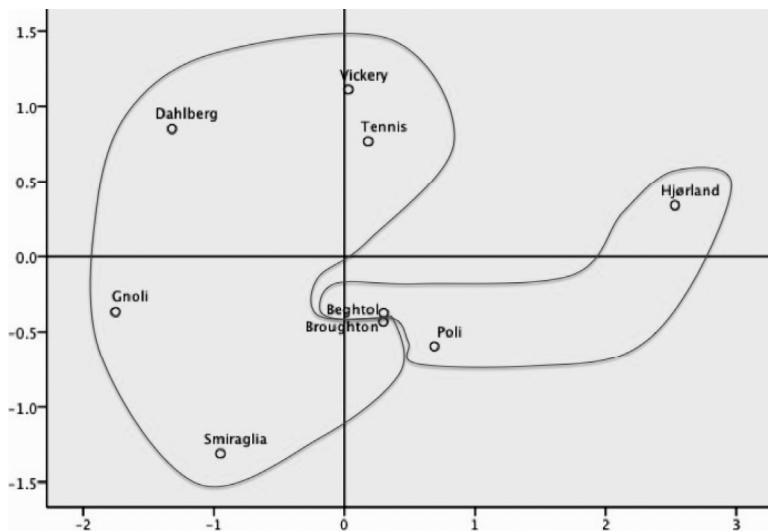


Figure 3. Author Co-Citation within the Rome Conference (stress = .24 R<sup>2</sup> = .68)

tion of their perceived research front. There were 972 citations in 65 papers; three presentations had no citations. The citations were arrayed by author (or first author) and duplicates removed, which yielded 891 citations, which is an indication of the remarkable breadth of literature cited by these authors. Single-occurrence authors were next removed from the list, leaving 146 authors. These 146 authors had been cited 535 times; 356 unique citations had been removed, representing again a remarkable breadth of scholarship. The remaining authors were arrayed in a frequency distribution by the number of multiple citations and the upper tier of this distribution appears in Table 7.

Author	Frequency
Hjørland	32
Gnoli	22
Poli	15
Smiraglia	14
Beghtol	13
Tennis	11
Dahlberg	10
Vickery	10
Broughton	8
La Barre	8
López-Huertas	8
Mai	7
Priss	7
Ranganathan	7

Table 7. Most cited authors

These authors' names were used to generate two co-citation analyses. First, the proceedings were analyzed to find co-citation of these authors among conference contributors. This matrix was plotted using SPSS and appears in Figure 3.

The goodness of fit statistics probably reflect the small numbers of co-citation occurrences. In fact, several of the authors had no co-citation. Author co-citation analysis is a means of visualizing how the citing authors perceive similarities among the co-cited authors. The clusters indicate perceived arenas of common interest among the co-cited authors. This plot has two large clusters, one containing Poli, Beghtol, and Hjørland, and the other covering more ground, obviously. One interpretation of this plot is that the cluster on the right represents theoretical points of view, and the cluster on the left points toward applications, representing a sort of intellectual tension between KO as a domain and KO systems. This interpretation takes into account the aforementioned dichotomous nature of the domain in which both humanistic and scientific characteristics are operating simultaneously.

A second author co-citation analysis was performed using the same set of most-cited authors from the conference, but this time by using co-citation data from Web of Science. In this case we are producing a visualization of how the domain at the core of the conference is viewed by KO scholars overall—in other words, this is a broader visualization of how the domain sees itself. This plot appears in Figure 4.

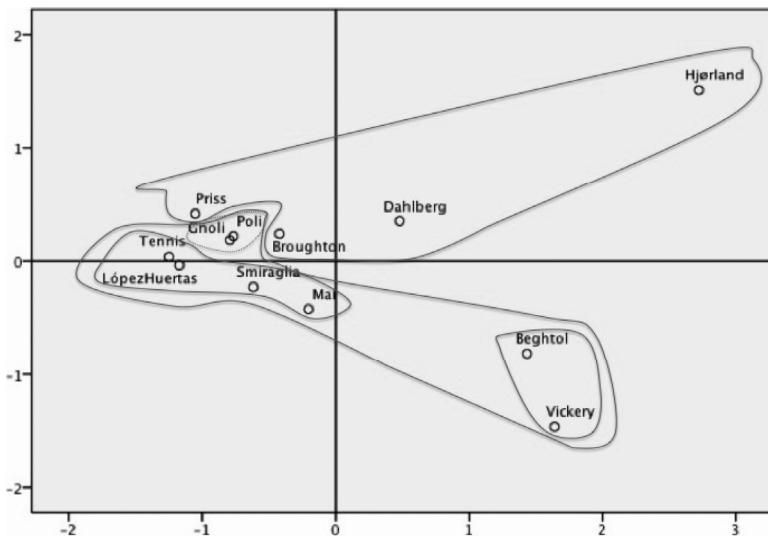


Figure 4. Author Co-Citation of Core Conference authors from WoS  
(stress = .02 R<sup>2</sup> = .99) 68)

This time the goodness of fit indicators suggest the data fit the model very well. Once again there are two clusters. A smaller cluster contains Hjørland, Priss, Broughton and Dahlberg. The other cluster is anchored by Vickery and Beghtol, with the remaining authors in the research front. This time Smiraglia and Tennis are linked to Mai and these three are linked to Lopez-Huertas, Gnoli and Poli are linked, as are Beghtol and Vickery. One is tempted to make the same interpretation of this plot as before. That is, we see facets and concepts in the upper cluster, and everything else in the lower. The association of classic theoretical authors on the right could be contrasted in each cluster with the presence of an active research front on the left. In that research front we see, as before, the dynamic tension between KO as a domain and KO systems, between theoretical approaches and applied, and between humanistic approaches and scientific.

#### 4.0 Co-Word Analysis

Titles of all 65 papers were entered into WordStat. A frequency distribution of title keywords was run; this appears in table 8.

TERM	FREQUENCY	% SHOWN
KNOWLEDGE	48	10.90%
ORGANIZATION	6	1.40%
INFORMATION	5	1.10%
APPROACH	5	1.10%
CLASSIFICATION	5	1.10%
DOMAIN	5	1.10%
INDEXING	5	1.10%

TERM	FREQUENCY	% SHOWN
SEMANTIC	5	1.10%
SCIENCE	4	0.90%
ARCHIVAL	3	0.70%
CONCEPTS	3	0.70%
PISTEMOLOGICAL	3	0.70%
FACETS	3	0.70%
FICTION	3	0.70%
PARADIGMS	3	0.70%
PERSPECTIVE	3	0.70%
PRAGMATISM	3	0.70%
RETRIEVAL	3	0.70%
SEARCH	3	0.70%
SOCIAL	3	0.70%
SPECIFIC	3	0.70%
SUBJECT	3	0.70%
THEORETICAL	3	0.70%
THEORY	3	0.70%
THESAURUS	3	0.70%
WEB	3	0.70%
WORKSHOP	3	0.70%

Table 8. Frequency Distribution of Title Keywords

What is most interesting in this distribution is the granularity indicated; only 8 terms are used more than three times, only 26 terms more than twice. In fact, not shown here are the 233 remaining terms in the long tail, including 216 that were unique. This suggests an amazing breadth for the research front of KO as it is represented by the papers in the Rome biennial conference. Of course, KO is a domain that crosses all disciplines and domains with its concern

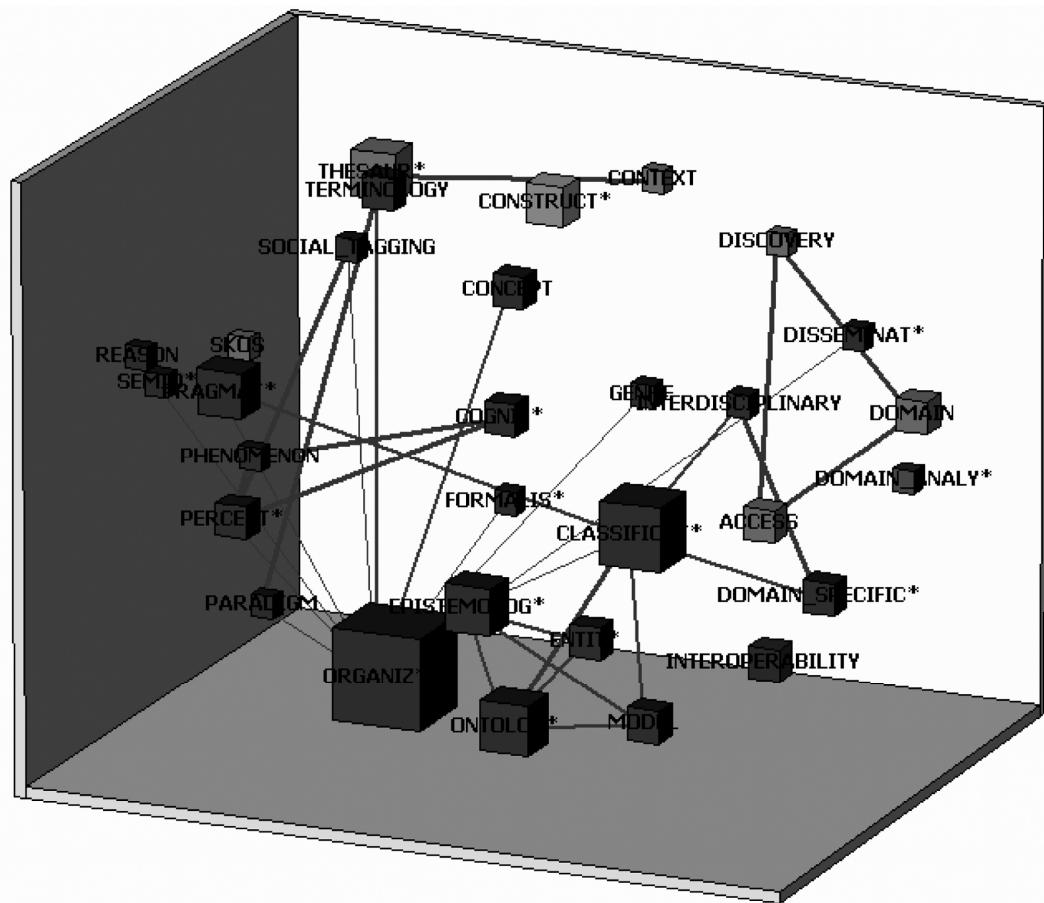


Figure 5. Co-Word Analysis (stress = .25 R<sup>2</sup> = .64)

for conceptual order of knowledge. Keywords from table 8 were arrayed taxonomically using the theme classes from table 3 and the ensuing “dictionary” was used to conduct co-word analysis of the 65 titles. WordStat then generated an MDS plot of the terms in the taxonomy. This plot appears in Figure 5.

Goodness of fit is about at the same level as figure 3 above and again is probably a function of the granularity of terminology. We see knowledge organization anchoring the domain with close links to ontology, epistemology and classification. We then see three clusters of activity: 1) information retrieval in the cluster to the upper right; 2) terminology, social tagging and thesaurus construction in the cluster at the upper left; and, 3) unsettled theoretical issues in the cluster at the left. This visualization provides a useful triangulation of the intellectual parameters and tensions in KO that we've already observed in the citation and author co-citation analyses. That is, humanistic approaches reside alongside scientific, and theoretical issues occupy the research front alongside applications issues. The farther we get from the core the more clearly we see applied issues of interest to re-

searchers, but the core itself receives constant attention, which in turn grounds the domain intellectually.

## 5.0 Conclusions: Granularity Describes ISKO 11's Bookshelf

With rare exceptions, granularity is the hallmark of our analysis of works cited by conference participants. The clear majority of resources cited were journal articles or book chapters (45%), which together with conference papers (15%) constitute nearly two-thirds of the citations. A large number of monographs are cited, but very few other resources (only 1%—“other” usually indicates an email, letter, or personal communication of some sort), and only 7% of the resources were Web resources. The domain is still clearly focused on standard published literature in peer-reviewed publications. The tension between humanistic and scientific approaches that we've observed several times is again apparent here in the large proportion (30%) of citations to monographs.

Thirty-eight articles were cited more than once in the conference proceedings. These ranged from arti-

cles by Dahlberg and Poli to Tennis and Buckland, and of course included several by Hjørland. Only two papers were cited four times—these were:

- Begtol Clare. 1986. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of documentation* 42: 84-113; and,
- Begtol Clare. 1986. Semantic validity: concepts of warrant in bibliographic classification systems. *Library resources and technical services* 30: 109-25.

There are 118 journals cited in the proceedings, of which 31 are cited more than once. The range of unique journals is fascinating, including journals such as the *Kathryn Sharp Review*, the *South African Journal for Higher Education*, and *Magyar Terminológia*. The journals cited more than once appear in Table 9. We are gratified to find *Knowledge Organization* at the top of the distribution, of course, and there are few other surprises in this list. It is noteworthy to see the inclusion of two Brazilian journals in the list, as well as two archival publications.

Journal title	# citations
<i>Knowledge organization</i>	50
<i>Journal of documentation</i>	29
<i>Journal of the American Society for Information Science and Technology</i>	20
<i>Axiomathes</i>	20
<i>Library trends</i>	9
<i>Information processing &amp; management</i>	8
<i>International classification</i>	7
<i>Archivaria</i>	6
<i>Cataloging &amp; classification quarterly</i>	6
<i>Journal of librarianship and information science</i>	5
<i>Information research</i>	5
<i>Journal of digital information</i>	4
<i>Journal of information science</i>	4
<i>Online information review</i>	4
<i>Scientometrics</i>	4
<i>Archival science</i>	4
<i>Ciência da informação</i>	4
<i>Library resources &amp; technical services</i>	3
<i>Perspectivas em ciência da informação</i>	3

Table 9. Journals cited more than once

Three hundred and twelve monographs were cited, but only 17 were cited more than once, and only 1 was cited more than 3 times. Vickery's *Faceted Classification: A Guide to the Construction and Use of Special Schemes* (London: ASLIB, 1960) was cited 6 times. There were 133 citations to conference proceedings, half of which cited sequential proceedings of conferences. These ranged from the European NKOS Workshops, to JCDL, to ISKO chapters, notably ISKO-España's Congreso and the North American Symposium on Knowledge Organization. ASLIB proceedings were cited 9 times; ISKO international conferences were cited 22 times. Of 74 web resources, only one was cited more than once: W3C. 2007. *SKOS. Use cases and requirements*, W3C working draft 16 May 2007, <[www.w3.org/TR/skos-ucr/](http://www.w3.org/TR/skos-ucr/)> was cited twice.

As noted at the outset, KO as a domain has great breadth both in its intension and its extension, which is a further reflection of its cross- or multi-disciplinary functional role. The map in Figure 5 is actually a pretty good representation of the whole domain at one point in time. Theoretical grounding is solidly at the core, although theoretical conversation constitutes a major component of the research front as well. Around the theoretical core, or perhaps it is better to say, anchored to the theoretical core are segments of the research front that reflect research on applications. Some of those intersect with other disciplines, some of them reflect the epistemological tension in the domain between humanistic and scientific methodologies. The breadth of the matrix represents the extension of the domain, and the depth its intension. Finally, at least at the point in time represented by this conference, there is immense granularity in the domain, which itself is a reflection of our cross-disciplinary role as scientists of the order of what is known.

## References

KO literature. 1998. *Knowledge organization* 25: 226  
Smiraglia, Richard P. 2008. ISKO 10's bookshelf: an editorial. *Knowledge organization* 35: 187-91.