José Augusto Chaves Guimarães - UNESP, Brazil
Joseph T. Tennis - University of Washington, USA

Constant Pioneers:
The Citation Frontiers of Indexing Theory in the ISKO International Proceedings

Abstract
Presents a citation analysis of indexing research in the ISKO Proceedings. Understanding that there are different traditions of research into indexing, we look for evidence of this in the citing and cited authors. Three areas of cited and citing authors surface, after applying Price’s elitism analysis, each roughly corresponding to geographic distributions.

Introduction and Background
Reflecting on past research in indexing is beneficial for both the theory and the practice of indexing for a number of reasons. First, we want a greater understanding of the contours of indexing theory. Second, we want to know what has been said so we can think about it and then act with knowledge. Third, we want to be able to evaluate indexing theory, and finally we want to set the stage to develop theory further. The process of examining indexing theory this way is called metatheory (Ritzer, 1991).

In order to reflect on past research on indexing, we gathered authors researching indexing and publishing in the ISKO context from 1990-Present, and extracted citations in these papers. We wanted to see what the landscape of indexing theory looks like over time through the literature cited.

It is important that we better understand the literature of indexing because we can see that there are at least three traditions: (1) subject cataloguing, (2) indexing, (3) and analyse documentaire (Guimarães, 2008). These three traditions draw on separate literatures as well as national traditions. So it is important for indexing theory in the international ISKO context to study the landscape of the literature.

Subject cataloguing had its origins in North America, specially from Cutter’s principles of alphabetical cataloguing, the tradition of LC subject headings and the thoughts of the School of Chicago, whose emphasis is mostly focused on the catalogue - a product of information treatment in libraries. Such conception is defined by Fiuza (1985, p. 257) by “... representing, in library catalogues, the subjects contained in the collection”. We can count Cutter, Kaiser, and Coates as core thinkers in this tradition. Currently, authors such as Hope Olson and Sanford Berman, among others, have dedicated to the aspects of subject cataloguing. This approach has found in periodicals as Cataloguing and classification quarterly, a special vehicle, considering it being mainly related to questions of organization of bibliographical registers and the bibliographical control in general.

The second approach - Indexing – comes from a British tradition where indexes, while products of the information subject treatment, are pinned down from the usage of indexing languages – specially thesauri - in specialized information centers or libraries, with a deep concern on the theoretical nature about the construction of such languages, is much influenced by the work of Classification Research Group. The contributions of Foskett, Austin, Farradane, Metcalfe, Aitchison, Gilchrist and Lancaster stand out. The journal The Indexer can be considered a special vehicle for spreading this conception, considering that it predominantly devoted to issues concerning history, organizations, systems, standards, methods, practices and technology indexing.
Another dimension, called analyse documentaire, comes from the French tradition, whose focus is centered on the subject treatment process itself, by means of the explanation of the procedures involved in identifying and selecting concepts. Those studies had a strong emphasis from the late 60's, with deep interface with Logics, Linguistics and Terminology, from the pioneering work of Gardin (1966a, 1966b, 1967, 1970, 1973, 1974, 1981) and of Coyaud (1966). Journals like Documentaliste and Journal of Documentation, can be considered as representative of such theoretical approach, especially by considering their tradition more focused on the documentation matters as well as on the methodological questions involving content analysis. In the analyse documentaire approach it is possible to observe how important the theoretical questions surrounding the document are, by revealing the influence of both the Diplomatics discussions (Tessier, Bautier, etc.) from the École de Chartes from the second half of the 19th Century until the beginning of the 20th Century and the Documentation concepts (Otlet, Briet) in the first half to the 20th Century.

Now we can ask an empirical question: do we see these three dimensions of indexing in the ISKO proceedings?

Methodology

Drawing on all of the proceedings of ISKO International Conferences, we searched for papers that had the word indexing in the title or abstract. We then recorded the author’s name and extracted the citations for each of these papers (excluding self-citations and standards), and from this data we created citation networks for each ISKO proceedings volume and also cumulatively for all the proceedings. We found a total of 784 cited authors.

For the creation of the citation networks we used the software PAJEK which allows for the construction of graphic network displays for the citations. From this data, we can describe the landscape and contours of indexing theory against the backdrop of these national and intellectual traditions of subject cataloguing, indexing, and analyse documentaire, and specifically talk about the ISKO environment.

Analysis and Findings

The most highly cited authors, those in the 98th percentile are (in rank order): Hjørland, B., Lancaster, F.W., Vickery, B. C., Aitchison, J., Beghtol, C., Larson, R. R., Ranganathan, S. R., Soergel, D., Reisthuis, G. J. A., Cochrane, P. A., Croft, W. B., Saracevic, T., Gilchrist, A., Svenonius, E., Chan, L. M., Hearst, M. We see a large representation of the British and North American thinkers, and a relative lack of the French tradition. In fact, only five citations total for Gardin, the only representative of the French school present in these ISKO proceedings.

We then applied Price’s Theory of Elitism (Price, 1963) to the 784 cited authors. This theory supposes that the elite of a certain citing environment is composed by the square root of the number of cited authors: 

Elite X = √X (Braga, 1974).

In this sense, and considering the original group of cited authors, it was possible to arrive at a group composed of 28 cited authors: Albrechtsen, Atchison, Bates, Beghtol, Broughton, Chan, Coates, Cochrane, Croft, Gilchrist, Hearst, Hjørland, Lancaster, Larson, Markey, McIlwaine, Mitchell, Olson, Politt, Ranganathan, Riesthuis, Salton, Saracevic, Soergel, Svenonius, Tobias, Vickery, Vizin-Goeetz. The original group of 127 citing authors was transformed in a group of 87 citing authors who have cited the elite authors.

Using this group of elite cited authors and the corresponding citing authors, we built a citation network using PAJEK. (See Figure 1).
In the mentioned group it is possible to observe that 3 authors - Albrechtsen, Cochrane and Politt belong at the same time to the categories of most productive (citing authors) and most cited authors. The remaining 25 most cited elite authors surround them, while the citing authors (N=84) surround them.

It is also possible to observe a high level of relationship among the cited authors, revealing a high density level in the network. It allows us to conclude that subject indexing presents a strong cohesion in terms of theoretical referents. That is, very few cited authors stand alone.

Only one of the 28 most cited authors – Larson - does not establish relationships with the others, and so constitutes a separate network with his citing authors (Bartolo, Kin, Buckland, Norgard, Chen and Trimble).

The universe of most cited authors (Price’s elite) reveals different theoretical approaches. In this sense, authors like Aitchison, Broughton, Coates, Gilchrist, Lancaster, McIllwaine, and Vickery represent a traditional “English approach on indexing” (Guimaraes, 2009), with a strong relationship and influence of the Classification Research Group – C.R.G. On the other side, authors like Hjorland, Soergel, and Riesthuis, are more connected to the theoretical and conceptual basis of knowledge organization - KO, representing the ethos and agenda of investigating the fundamental basis of ISKO creation. In this environment, Ranganathan acts as a connection between the mentioned approaches, since he has deeply influenced the English tradition of C.R.G. as well as Dahlberg’s and Soergel’s conceptions of KO.
The North-American tradition is also significant in the analyzed network. The authors Bates, Beghtol, Mitchell, Olson, Saracevic, and Svenonius establish a strong relationship between the theories of subject cataloguing and information retrieval and confirming the presence of North American authors in ISKO literature as observed by Smiraglia (2009).

The one outlier is Larson, connected with information retrieval, but not in dialogic relationship with any other literature in ISKO. And as noted above, the French tradition does not figure in the proceedings of ISKO.

Conclusion
This small contribution to the metatheory of indexing comports with similar studies of different indexing theories (Tennis, 2005); that is to say, there is a great diversity of thought indexing. We can add to what was described above, there are more than three traditions: a) subject cataloguing, based on the North American tradition of School of Chicago and mostly focused on building catalogues for public and school libraries; b) indexing, based on the British tradition of the Society of Indexers as well as of the Classification Research Group, mostly focused in building indexing languages for representing knowledge in special libraries and documentation centers, and c) analyse documentaire, based on the French tradition and mostly focused in explaining the logic and linguistic procedures involved when representing knowledge in special libraries and documentation centers (Guimarães, 2008). And two of these are present in ISKO. The other seems conspicuously absent. The next question then, is why, especially given the prominence of Otlet and Briet in our current conceptions of the field. We can add to this list the North Americans concerned with information retrieval (i.e., Larson), and those that combine subject cataloguing and information retrieval (i.e., Bates, Beghtol, Mitchell, Olson, Saracevic, and Svenonious).

What is clear from this part of the study is that indexing theorists are constant pioneers, constantly charting new territory by incorporating new literature into the field of indexing, and maneuvering in particular traditions. On the other hand, it is possible to observe that although those traditions are focused in different points, they establish dialogical relationships in the KO environment – maintaining their position in their citation field, but exploring new frontiers as the research advances into unknown territory.

References


