

Visualizing Library and Information Science from the practitioner's perspective

Yusef Hassan-Montero* and Víctor Herrero-Solana*

*yusefhassan@gmail.com, victorhs@ugr.es

Scimago Research Group, University of Granada, Faculty of Library and Information Science,
Colegio Máximo de Cartuja s.n., 18071 Granada (Spain)

Introduction

The knowledge domain visualization studies are mostly based on scientific publishing data, such as citation and co-citation data. These studies provide a valuable view of the knowledge domains' intellectual structure and evolution. However, scientific data do not enable one to visualize the other side of the coin: how a knowledge domain is viewed by its practitioners.

In this work we present an overview of the Library and Information Science profession. For this purpose, professional subjects and their relationships are visually represented as a network.

Materials and Methods

The data sample was downloaded from 'Exit' web site (<http://www.directorioexit.info/>) by means of a crawler, on November 24th, 2006. 'Exit' is a free online directory of 'experts in information handling', where the admission of each new member is evaluated by a committee. The data sample contains 646 member records; 72.5% of the members are from Spain.

Each record is self-filled by the experts. The record contains no more than ten fields (name, institution, country, email, web page, subject, etc.). The expert can select at most five subjects from a predefined list with 48 different subjects. Figure 1 shows the distribution of experts for each subject, from four (*Statistical Information*) to 204 (*Knowledge Management*).

Co-occurrence analysis enables us to calculate relationships between subjects. That is, the relation between two subjects would be determined by the number of members that have selected both subjects in their professional profiles. Co-occurrence relationships among subjects were represented as a NxN matrix, where N is the number of different subjects (i.e. N=48).

The matrix was visually represented as a network display, where each subject is shown as a node, and their co-occurrence relationships as links between nodes. In order to show a clearer and more comprehensible graph, the less significant links were pruned. The pruning method used was Pathfinder algorithm (Schvaneveldt; 1990), with the parameters $q=N-1$ and $r=\infty$. Also, Kamada & Kawai's algorithm

(1989) was used to situate the nodes in space. In the final network display (Figure 2), link thickness indicates the relationship weight (co-occurrence frequency), and the node's label size reflects the number of members.

Discussion

The network display (Figure 2) shows a prominent root node with the most popular subject (*Knowledge Management*). From this node emerge five long branches: an Information System Design branch (*Information Architecture, Usability, Web Design, etc.*); a Digital Libraries branch (*Open Access, Electronic Journals, Academic Library, etc.*); an Information Retrieval and Bibliometrics branch (*Bio-medical Information, Patents, Databases, etc.*); an Archives branch (*Historical Archives, Company Archives and Records Management*); and finally, a multifaceted branch that enclose *Planning and Management, Marketing and Library*.

Also, there are a few nodes without branches, connected directly with the root node (e.g. *Law Information, Journalistic Information, Document Management Software, etc.*). The nodes in the same branch have strong relations among them, but in many cases these relationships are remarkable. For example, while *Academic Library* is in the Digital Libraries branch, other kinds of libraries (*Public, National and School Libraries*) are in the Marketing-Management branch.

Another noteworthy node is *Ontologies and Thesauri*, because it is not connected with *Classification and Indexing*. Similar is the case of *Competitive Intelligence*, because it is not in the Marketing-Management branch. Another interesting case is *Search Engine Optimization*, which is strongly connected to the Usability branch instead of to Information Retrieval.

Conclusions

The LIS domain shows at least five major groups of subjects: 1) Information Retrieval & Bibliometrics, 2) Digital Library, 3) Information Architecture & Usability, 4) Marketing & Management, and 5) Archives & Record Management. *Knowledge Management* is the most popular subject and the root of the network.

The presented visualization provides a means of knowledge acquisition and understanding of the LIS domain from a professional perspective. Also, this visualization could be used as a graphic interface for the 'Exit' directory.

Schvaneveldt R. (Ed.) (1990). Pathfinder associative networks: studies in knowledge organization. Norwood, NJ : Ablex; 1990.

References

Kamada T, Kawai S. (1989). An algorithm for drawing general undirected graphs. Information processing letters, 1989, 31(1), 7-15.

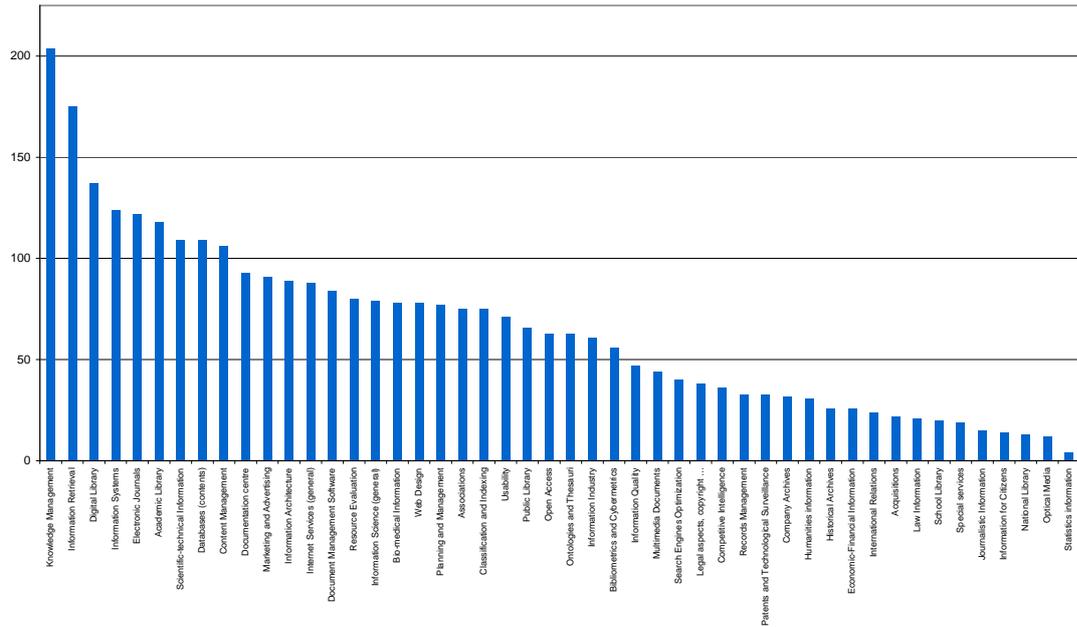


Figure 1: Number of members for each subject.

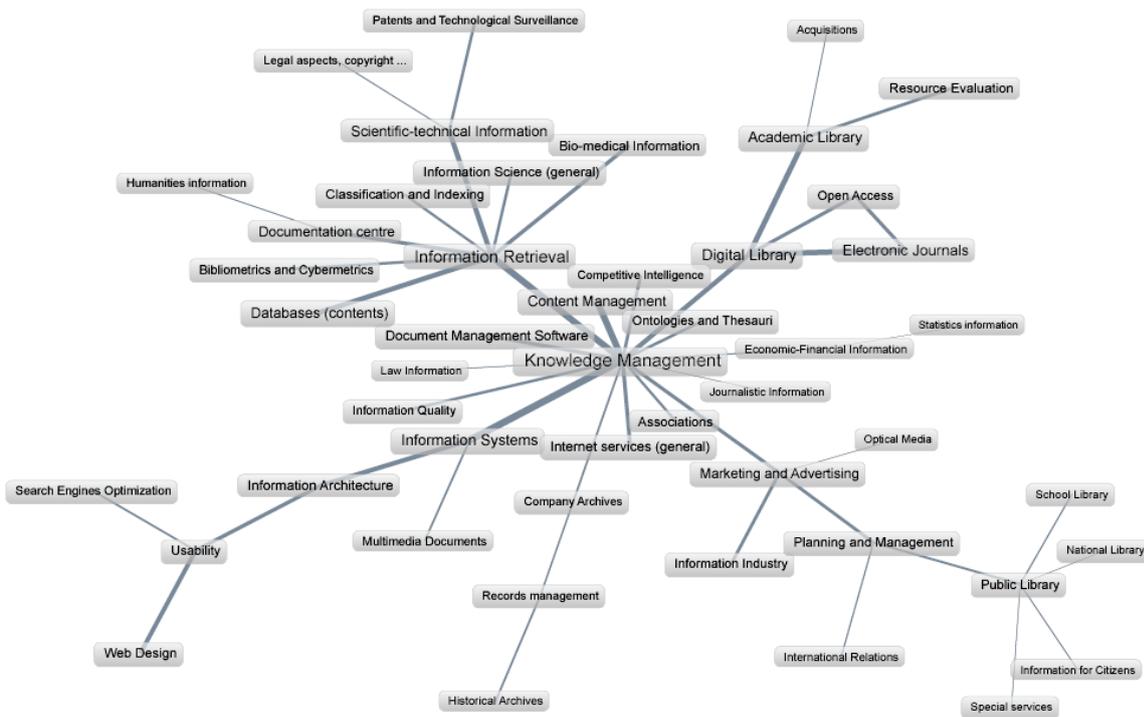


Figure 2: Network display.