

Visualization analysis of the evolution of indexing and the subject migration——based on the view of internationalization

Zhao Rongying¹ and Wu Shengnan² and Ma Lina³

¹zhaory999@yahoo.com.cn ²454802740@qq.com ³malina5ok@163.com

Center for the Studies of Information Resources of Wuhan University , Research Center for China Science Evaluation, Wuhan 430072(China)

Introduction

Indexing is a serial of records organized according to certain system, or a kind of retrieval tool guiding documentation issues and unit of knowledge [1]. Indexing first appeared in the west ,which was used in European religious works. In the eighteenth century, the theme indexing emerged in Europe. And at the end of nineteenth century, content analysis indexing was widely used in the world.

The appearance of indexing in China is late. Generally speaking, the name indexing , which was written by Fu shan, is regarded as the symbol of the Chinese indexing. In the 1920's, as the introduce of western indexing theory and compile technical, the compilation and study of Chinese indexing was booming. «Indexing and Indexing Law» published in 1930 as well as «The Theory of Indexing» published in 1932, marked that Chinese modern indexing theory and technology has developed rapidly. In the 1950's, computer technology was used to compile indexing. Since then, the indexing made by machine was turning up, which makes profound revolution in the aspect of indexing compiled theory, technology, and the carrier form.

Data Sources and Research Methods

This paper respectively analyzed the indexing research at home and abroad. The international indexing study is based on the Web of knowledge published by Institute for Scientific Information. The data is from 2000 to 2010, with the search condition

that Theme =“indexing”. Every datum record includes authors, title, abstract, descriptors and identifiers. After some necessary processing and excluding, we got a total of 5527data. The time of collecting data is November 12, 2010.

In this study, according to visual co-citation analysis and clustering analysis by using of Citespace II. The author reveals the evolution and development, the representative figures and representative publications, the core journals and related subjects, as well as the research hotspots about indexing.

The Analysis of the Indexing Research At Home and Abroad

Visual Analysis of Representative figure and their Works in the Indexing Field

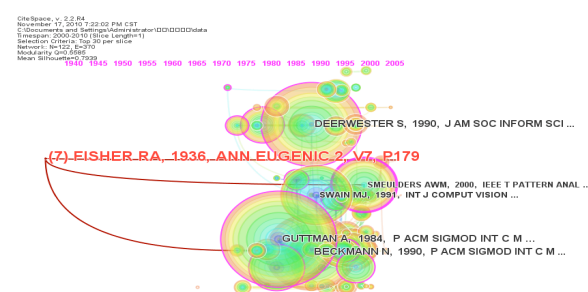


Figure 1 . The time series diagram of cited Knowledge map of the international indexing research

There are 122 nodes and 1135 connections in figure 1. Each node represents a article. The size of the node reflects the number of literature cited. The greater node is ,the more the time to be cited. The connections between nodes stand for quoting between literatures. The thicker the line is, the more

the number to be cited between them is. In figure 1, we can find the top 5 literatures of high cited frequency.

The Category Distribution Analysis of the International Indexing Research

In figure 2, we can find that Computer Science makes the most research on the indexing, and Engineering as well as Information Science & Library Science also does more study.

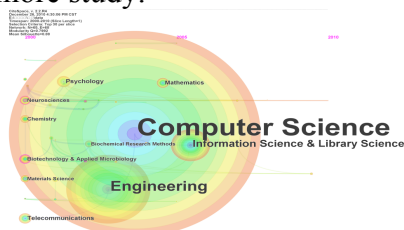


Figure 2. The category distribution knowledge map of the international indexing research

Through the further analysis of Computer Science and Information Science and Library Science, we would find that the literature distribution of indexing research exist the obvious subject migration phenomena, which is shown in the following figures.



Figure 3 the published articles of indexing in Information Science and Library Science



Figure 4 the published articles of indexing in Computer Science

The phase analysis of the evolution of indexing and the subject migration

From the analysis of the highly cited literature and the category of the indexing research at home and abroad, we can find that the study on indexing has not only confined in one aspect or a certain subject. Information Intelligence and Library Science, as the traditional subject on indexing research, published much less

articles than Computer Science, which means the main subject on indexing research has transferred from Information Intelligence and Library Science to Computer Science. Nowadays, the study of indexing is increasing. It is not only the demand of social development, but also the needs for the development of indexing. Reviewing of the development history of the indexing research, we can see that indexing research roughly experienced three phases: the period of Bibliography Indexing, the period of database indexing and indexing under the network environment.

The Knowledge Map of Indexing Research Hotspots

Key words are the core and essence of the article, although their frequency is not much in an article, they are highly condensed summary of the article topic, so the high frequency words are often used to regarded as the hot spots of indexing research^[2].

Hotspots Analysis of International Indexing Research

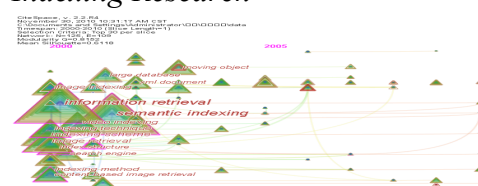


Figure 6. The knowledge map about main fields and hotspots of international indexing research

In figure 6, we can see that the hotspots of international indexing research are numerous and concentrated. And we can also conclude that the research hotspots of indexing are mainly distributed in semantic indexing, information retrieval and image retrieval, database and search engine, etc.

Conclusion

(1)Based on the analysis of the disciplines distribution and the published articles of indexing, we can see that the study of indexing appears the phenomena of subject migration. The indexing research can be

traced back to 1936. Computer Science is the most cardinal subject on international indexing research, the number of published articles is 5253, and the peak is in 2006; whereas the number of published articles in Information Science and Library Science field is only 847, and the peak in this field is in 2001. From above, we can see that the main subject on indexing research has transferred.

(2) This paper believes that the development of indexing research can be classified into three stages through the analysis of disciplines distribution and highly cited literatures at home and abroad: the period of Bibliography Indexing, the period of database indexing and indexing under the network environment. In the evolutionary process of indexing research, we found that there are two turnings in this process: ① along with the expansion of the computer application, the indexing has transferred from manual type to database indexing; ② in the network environment, the rapid growth of information makes database indexing not to satisfy people's needs, so a new indexing- indexing under the network environment. In this process, the indexing research realized the migration from Information Science and

Library Science field to Computer Science and formed the represented research team including Guttman and Deerwester. As the key researcher of indexing field, they immeasurably contributes to the development of indexing study.

(3) Through the visualization analysis of research hotspots of indexing, we can see that the international indexing research is focus on semantic indexing, indexing technology, indexing method, information retrieval, based on image content of research, an XML document retrieval, databases, search engines and automatic indexing and so on.

Acknowledgments

The authors would like to acknowledge the support of National social Science Foundation of China(09BTQ022) and enlightening comments from reviewers.

References

- Peng FZ. Introduction to Bibliography[M]. Beijing: *People's Medical publishing house*, 1982:35.
- Hou Hai-yan, Liu Ze-yuan, Chen Yue etc. Mapping of science studies : *the trend of research fronts [J]*. *Science*