



The Predicaments and Countermeasures of the Chinese Commercial Bibliographic Databases*

Wang Fengnian

Tsinghua University, Beijing, China

In China, the vast majority of the bibliographic databases is commercial, such as China National Knowledge Infrastructure (CNKI), Wanfang Database, Longyuan Journal Net, CQVIP Company, however, there are also non-profit open access (OA) databases, such as journal database jointly established by Chinese Academy of Social Sciences (CASS) and National Social Science Fund. The commercial bibliographic databases have to face many difficulties: intellectual property disputes, the benefit distribution between the hardcopy periodical and the commercial bibliographic database, the lack of quality assessment about the commercial bibliographic databases, the need of improving digital technology as well as the lack of a unified database regulation, which restricts the development of commercial bibliographic databases. This paper puts forward the countermeasures from the perspective of how to enhance the governmental management; how to protect the intellectual property right; how to improve the technical standard of the commercial bibliographic databases; how to build interest distribution between the hardcopy periodical and the commercial bibliographic database; how to improve the quality of commercial bibliographic databases; and how to improve the industrial chain of the commercial bibliographic databases.

Keywords: the commercial bibliographic databases, the predicaments of the commercial bibliographic databases, the strategies of the commercial bibliographic databases

Introduction

Currently, the vast majority of bibliographic databases in China is commercial, such as China National Knowledge Infrastructure (CNKI), Wanfang Database, Longyuan Journal Net, CQVIP Company; however, some of the bibliographic databases are open, such as the National Philosophy and Social Science Journal Database (supported by National Social Science Fund), a specially commissioned project supported by National Social Science Fund and implemented by the Survey and Data Center of CASS.

Big commercial bibliographic databases, such as CNKI, Wanfang Database, Longyuan Journal Net, CQVIP Company, have been developing rapidly, as well as the flourishing of other small databases; however, from the academic research and the user's application, there is a lot of serious criticism. Despite the problems existing in commercial bibliographic databases, the important role they play in research and teaching is obvious, therefore, taking CNKI, Wanfang Database, Longyuan Journal Net, CQVIP Company as examples, this article proposes some measurements based on the analysis of their problems.

* This work was supported by Asia Research Center in Tsinghua University.

Wang Fengnian, Doctor, Associate Professor, the Editorial Department of Tsinghua University Journal, Tsinghua University.

The Evolution of the Bibliographic Databases in China

In China, bibliographic databases began in the early 1980s and became commercialized and industrialized after the 1990s.

In 1992, Chinese Scientific and Technical Journal CD-ROM Library developed full digital magazines.

In July 1995, Daheng Computer Disc Magazine launched the bibliographic database.

In 1986, Chongqing VIP Database, China's first database of the electronic journals, was created by Chongqing VIP information Co., Ltd, who later in 1989 launched CBISTIC/CEPC Periodical Base in China, which was not only the country's first Chinese bibliographic database, but also was the largest self-built Chinese bibliographic database entitled CQVIP Company today.

On December 11, 1997, China Academic Electronic Magazine (CAJ-CD) was incorporated. It was the predecessor of CNKI that started online services in June 1999. In 1999, Tsinghua University and Tsinghua Tongfang Co., Ltd. released China National Knowledge Infrastructure (CNKI), which was not only one of the world's largest continuously updated Chinese periodical full-text databases, but also was the first comprehensive bibliographic database in China. In 2003, CAJ-CD officially changed its name to National Knowledge Infrastructure, CNKI. Until 2013, CNKI contained the full texts of 5,300 academic journals published since 1994.

In December 1998, Longyuan Journal Net, one of the largest Chinese journal nets in the world, was in trial operation. It was officially opened in June 1999. In 2010, it launched the mobile version and the handheld reader version, a comprehensive iPad, iPhone and other reading services of handheld terminals. Through computers, mobile phones or handheld readers, personal or institutional users can directly gain access to Longyuan Journal Net.

In 2000, Wanfang Data was co-established by Wanfang Data Company, Institute of Scientific and Technical Information of China (ISTIC), Beijing Gold Technology Investment Limited Company, Chinese Cultural Industry Investment Fund, Chinese Science Publishing & Media Limited Company (CSPM), Institute of Sichuan Science and Technology Information, and Sichuan Scientific and Technical Literature Publishing House. Wanfang Data covers large network databases of meeting minutes, papers, academic achievements, and conference articles. Currently Wanfang Data has established a service center with 24 databases holding independent copyright.

The Predicaments of the Commercial Bibliographic Databases

Technical Confusion

Because there is not any unified technical standard to the commercial bibliographic databases in China, to gain access to different commercial bibliographic databases, users have to use different readers to browse and different software to convert format.

CNKI, Wanfang Database, Longyuan Journal Net, and CQVIP Company use different reading formats to protect their own interests, especially to keep the amount of the users, which has limited the sharing of academic resources. For example, CNKI adopts CAJ and PDF formats, so users have to download CAJ and PDF viewers to read. CQVIP Company adopts two formats to get the full text available for users: VIP format and PDF format, the latter is an internationally commonly used format. To view VIP format, users have to install the company's special "VIP browser"; to view PDF format, users have to install Adobe Reader. "Digital

Periodicals Subsystem” of Wanfang Database adopts HTML and PDF file formats. In China, in addition to HTML format, there are other formats, such as PDF, CAJ, DCC, CEB, TEB, etc. To read different format files, users have to download and install different reading softwares for different formats.

Apart from technical barriers, the technology used by the commercial bibliographic databases needs improving. For example, although getting confirmed of “subscribe success” from the commercial bibliographic databases, users cannot gain access to any articles. The commercial bibliographic databases made the hardcopy periodicals digitized; however, if users want to download and copy any articles in commercial bibliographic databases, they have to install text-recognition software to convert the graphic text into text characters. The conversion is far from precision but with a lot of errors.

Interest Conflicts Among Commercial Bibliographic Database

In the early years, it was costly to build commercial bibliographic databases and there was a remote chance of making profits. As the main resource of commercial bibliographic databases, hardcopy periodicals did not require paid view, and even did not care the payments from commercial bibliographic databases, because it was an effective way to raise awareness if the hardcopy periodical contents appeared on Internet. With the interest increase in commercial bibliographic databases, hardcopy periodicals request reasonable economic returns. In China, the commercial bibliographic database industry makes profits by summarizing and digitizing various hardcopy periodicals. Commercial bibliographic databases pay certain shares to hardcopy periodicals. Because it is hard to get the accurate data for how many times of reviewing or downloading, hardcopy periodicals won't be able to provide the requirements of a reasonable share, which causes the unfair collaboration between the hardcopy periodical and the commercial bibliographic database. In fact, hardcopy periodicals have been in a passive situation.

When readers download articles from different commercial bibliographic databases, such as CNKI Wanfang Database, Longyuan Journal Net, CQVIP Company, even for the same article with the same bytes, different commercial bibliographic databases might be charged with different rates.

The Lack of Quality Control

The commercial bibliographic databases have been collecting a large number of hardcopy periodicals, which, without any doubt, contain many high quality academic articles; however, besides the forefront of academic papers, some of them are about simple knowledge or popular science articles. The commercial bibliographic databases should apply their view system the same as hardcopy periodicals, but who is going to review? How to financially support the review? If there is not any strict review system, how to guarantee the quality of the commercial bibliographic databases? And how to ensure high quality articles to be retrieved in the commercial bibliographic databases?

The Lack of Scale

Although the bibliographic databases have been developing by leaps and bounds, there are many drawbacks in the commercial bibliographic databases: not well-integrated, not well-organized. There are not any government agencies, academic institutions to regulate the bibliographic databases. In fact, there are not effective statistical methods. Although the circulation of the hardcopy periodicals gradually decreases, although users can get access to the commercial bibliographic databases at any time and for a long term, some commercial bibliographic databases are viewed less than hardcopy periodicals.

Intellectual Property Right Disputes

Many bibliographic databases, including the commercial bibliographic databases, are not the original publishers and authors. There is no doubt that all of them should obey the intellectual property right, but some of the commercial bibliographic databases sell the papers coming from the hardcopy periodicals without getting permission either from the hardcopy periodicals or from the authors, which makes the copyright issues more complex. CNKI, Wanfang Database, Longyuan Journal Net, CQVIP Company have collected a huge quantity of papers. It is impossible for them to get permission directly from every author. Therefore, there have been disputes among the commercial bibliographic databases, hardcopy periodicals and the authors, for example, CNKI and Wanfang Database were sued for intellectual property right. In 2008, the Haidian Court concluded the case of 284 plaintiffs with Master and Doctoral degrees prosecuted Wanfang Database, and 234 of them got the court in favor, the winning rate was 80% (ZHANG, 2008). The copyright ownership of the dissertation is unclear, which is one of the reasons why Wanfang Database was sued. If the intellectual property right is not effectively solved, it will become a serious obstacle to the commercial bibliographic databases to develop further.

The Lack of a Unified Regulation

The commercial bibliographic databases, such as CNKI, Wanfang Database, Longyuan Journal Net and CQVIP Company, do not have unified regulations, which certainly prevents from spreading their influence and internationalization. For example, the layout of articles in the commercial bibliographic databases still uses a dual-column design, which is inconvenient for readers to return to the top of the page to read the right column after reading the left column.

The Reasons for the Predicaments of Commercial Bibliographic Databases

Management System

In China, hardcopy periodicals can only be published after a CN number is applied for and approved of. In a like way, commercial bibliographic databases can only be created after digital publishing permission is approved of. There are only 50 authorized organizations for digital publishing by the General Administration of Press and Publication (the predecessor of the State Administration of Press, Publication, Radio, Film, and Television of the People's Republic of China), which causes the small scale of the commercial bibliographic databases.

Evaluation System

To evaluate hardcopy periodicals, there have been different evaluation systems, such as “core journals”, “top journals”, and “authoritative journals”... These evaluation systems have their shortcomings with shocking results to journal evaluation (ZHU, 2015). To establish an objective and impartial evaluation system, it is more difficult and complex for commercial bibliographic databases than for hardcopy periodicals because different reviewers and readers have different reading habits and background knowledge, causing the evaluation results with some bias (WANG, 2015). So far there is not any evaluation system for the commercial bibliographic databases in China.

Reading Habits

Different age groups have different media expectations and demands (WANG, 2014), which has triggered readers to have different reading habits. Some readers get used to the banner reading, while others have the dual-column reading habits. Most hardcopy periodicals take the dual-column layout in order to reduce the

fatigue for readers; while some commercial bibliographic databases take the banner layout for the reasons of technical symbols and image processing convenience. Some readers are accustomed to the dual-column layout, which reduces the reception of the commercial bibliographic databases.

Measurements to Promote the Commercial Bibliographic Databases

To Implement Digital Publishing Strategies for the Government

To support digital publishing, including commercial bibliographic databases, preferential policies and financial aids have been issued by the Publicity Department of the Central Committee of the CPC, the Ministry of Finance, the General Administration of Press and Publication, the Ministry of Science and Technology, and the Ministry of Industry and Information.

On August 16, 2010, the General Administration of Press and Publication (the predecessor of the State Administration of Press, Publication, Radio, Film, and Television of the People's Republic of China) issued the "Ideas on Accelerating the Development of China's Digital Publishing Industry", which clearly illustrated the objectives, tasks, and measurements of digital publishing industry.

On July 15, 2014, at the Annual Digital Conference, Chinese Academy of Press and Publication issued a report titled, *Industry Annual Report of China's Digital Publishing 2013-2014*, which made clear that "the early digital publishing model organizations had formed a relatively clear digital publishing development path in terms of the management systems, production forms, business models, and strategic layouts that had been put into practice, which played a leading role as models, and achieved effective results"¹. In August 2014, the Fourth Conference of the Central Leading Group comprehensively initiated and approved the "Guidelines on Promoting Convergence of Traditional Media and New Media Development", which put forward "to deepen the integration of contents, channels, platforms, management, administration, to strengthen the content development, to innovate editorial processes, and to optimize the information services in order to make the content advantage win the development advantages". From 2013 to 2014, the amount of financial support from the central government on the digital transformation projects was over 1 billion Yuan.

If digital publications were not certified by relative governmental departments, the digital publishing would not come true, therefore, digital publishing must be certified by the national authorities who issue formal qualifications to the commercial bibliographic databases that are equal to the hardcopy periodicals in the academic status. In calculating academic achievements, universities and academic institutions should take digital periodicals, journals and online databases, as the same with the hardcopy periodicals. In 2006, Shanghai Federation of Social Sciences listed internet article awards in assessing outstanding academic achievements.

To Improve the Digital Publishing Law

In 2007, as one of the four digital publishing projects by the General Administration of Press and Publication, the Institute of China's Publishing Research demonstrated the feasibility of digital copyright protection project. In 2009, the Ministry of Finance approved of this project, which focused on the key technology of digital copyright protection and explored the new digital publishing business models. In 2011, the General Administration of Press and Publication initiated a major science and technology project—digital copyright protection technology research and development project, which set up a series of digital copyright protection standards.

¹ China's Digital Publishing Industry Project of the Chinese Academy of Press and Publication: *Industry Annual Report of China's Digital Publishing 2013-2014*, July 15, 2014.

To Improve the Style of the Commercial Bibliographic Databases

Standard style specifications not only benefit document search, management and digital processing, but also promote the sharing of digital academic resources to fully achieve their academic values. To encode the text edition to electronic version, commercial bibliographic databases not only have to follow the same style specifications as hardcopy periodicals, but also to integrate the style specifications of arts and science. It is not necessary for the commercial bibliographic databases to follow the dual-column layout of the hardcopy periodicals, instead, they should adapt the layout design according to the electronic contents.

As for the formats, in May 2005, the International Organization for Standardization (ISO) approved of PDF format as the standard international format to read, download bibliographic databases. PDF not only supports XMI, but also stratifies the use of the contents, such as full copy, partial copy, read-only, print-only, not printing.

To Establish the Benefit Distribution Mode Between the Hardcopy Periodical and the Commercial Bibliographic Database

Industry Annual Report of China's Digital Publishing 2013-2014 showed that "digital publishing has been gradually breaking the previous embarrassment of the unprofitable situation. Some enterprises have gained returns in the digital publishing business".² CNKI, Wanfang Database, Longyuan Journal Net, CQVIP Company have been exploring the benefit distribution with the win-win model, for example, hardcopy periodicals allow CNKI, CQVIP Company, etc. to take articles into their databases, and a third-party will pay the digital copyright to the hardcopy periodicals. Hardcopy periodicals make profit by authorizing the third party to use the digital copyright; however, there have been the disputes in copyright, cost and benefit sharing. Commercial bibliographic databases and hardcopy periodicals should jointly establish a reasonable pricing system which could attract more users on the rational basis of the price, content and service. International bibliographic databases take different charging systems for organizations and individuals, which is lower for individuals than for organizations. Since the commercial bibliographic databases are not pure business, they should establish a non-profit pricing system, and even become public welfare in the future.

In the cooperation between the hardcopy periodical and the bibliographic database, in fact, the former, being the content provider to the bibliographic databases gets relatively lower profit. This requires commercial bibliographic databases to make it clear of when profit or deficit is made, and the amount of gain or loss. In this way, the mutual and profitable relation would be established between the hardcopy periodical and the commercial bibliographic database.

There are two popular types of systems currently practised in the world, Warehouse Management System (WMS) and National Licensing System (NLS). The latter may be more practical for China. NLS refers to the publisher who signs the national license with the Third Party which provides non-profit information service, so that the Third Party is allowed to use the resources from the hardcopy periodicals. It is free for all users to browse, download, search, print or photocopy from the Third Party. Because the Third Party, which is non-profit information service provider, undertakes the copyright liability, therefore, the legal disputing issues will be decreased for databases, hardcopy periodicals or authors. By shifting payment mechanism, the government will compensate economic losses to the bibliographic databases and the authors due to open access.

² China's Digital Publishing Industry Project of the Chinese Academy of Press and Publication: *Industry Annual Report of China's Digital Publishing 2013-2014*, July 15, 2014.

NLS could effectively balance the interests among authors, readers and publishers. Some Nordic countries, such as Sweden, Iceland, Denmark, Norway, and Finland, have established and widely implemented NLS.

In March 2012, the National Philosophy and Social Science Planning Leading Group authorized the Survey and Data Center of CASS to undertake specially commissioned National Social Science Fund Project—National Journal Database of Philosophy and Social Sciences. This project collected more than 1,000 high level hardcopy periodicals, set up national, open-resource, public welfare digital database covering philosophy and social sciences, leading to the five-win-win situation among authors, readers, the hardcopy periodicals, agency communicators and management organizations. This project was a practical case of NLS in China.

To Improve the Quality of the Commercial Bibliographic Databases

Both “the three-review system” in China and the internationally accepted peer-review mechanism effectively control the quality of hardcopy periodicals, which sets examples for the bibliographic databases. Although there are many high academic papers in CNKI, Wanfang Database, Longyuan Journal Net, CQVIP Company, the quality of commercial bibliography databases remains uneven.

To Improve the Industrial Chain of the Commercial Bibliographic Databases

The government should improve the industrial chain among the bibliographic database, the hardcopy periodical, and the advertiser, just like *Industry Annual Report of China's Digital Publishing 2013-2014* declared, “industrial cooperation awareness of digital publishing has been intensified, and the coordination potency dimension of digital publishing has been enhanced. Cross-sectoral cooperation among traditional publishers, technology providers, internet service platforms has been deepened to take the complementary advantages. Industrial integration with further accelerating and powerful partnership with noticeable trend is the cross-sectoral driving force of digital product innovation”.³

To Establish the Evaluation System of the Commercial Bibliographic Databases

“Evaluation system could ensure evaluation fair and reasonable. In China, there are many gaps in the institutional evaluation. To develop and implement evaluation systems, many aspects are involved, and plans have to be made carefully” (YE, 2015). According to the principle of objectivity, fairness and authority, it is necessary for the commercial bibliographic databases to establish the evaluation system by making the quantitative evaluation criteria, organizing the qualitative evaluation by experts, and introducing advanced academic evaluation methods from abroad.

To Cooperate Among International Bibliographic Databases

By Internet, the domestic bibliographic databases should cooperate with international databases to jointly widen the content of the bibliographic databases. CNKI has cooperated with many professional Internets and electronic publishing organizations in more than 40 developed countries and regions whose users are distributed in more than 130 countries and regions. Among 1,200 institutional clients, there are international prestigious universities, government agencies and important companies, such as Harvard University and the White House. Jointing with almost one hundred domestic information institutes, Wanfang Database launched CEADB, which includes four bibliographic databases, “Chinese Enterprise, Company and Product Database”, “Chinese Research Institution Database”, “China Science and Technology Information Database”, “China's

³ China's Digital Publishing Industry Project of the Chinese Academy of Press and Publication: *Industry Annual Report of China's Digital Publishing 2013-2014*, July 15, 2014.

Higher Education Institution Database”. These four databases are updated once a month. The customers of CECDB include companies, institutions, business associations in China, university libraries from more than 50 countries and regions in North America, Western Europe, and Southeast Asia. American DIALOG, the world’s leading on-line system, providing online retrieval service to millions of worldwide users, takes CECDB as the first choice of China’s economic information databases.

References

- WANG, F. N. (2015). The evaluation mechanism of China’s digital academic journals. *Publishing Research*, 7, 62.
- WANG, P. (2014). The need analysis of the electronic reading users in the digital dissemination era. *Science-Technology & Publication*, 7, 82.
- YE, J. Y. (2015). The plurality and multitudinity of academic journal evaluation. *Journal of Tsinghua University*, 2, 186.
- ZHANG, L. (2008). 480 Doctors and Masters sued the Wanfang Database infringement. *Beijing Evening News*, September 22.
- ZHU, J. (2015). The relations between the research institution and the academic evaluation. *Journal of Tsinghua University*, 1, 7-8.