

A PHILOSOPHY RESEARCH DATABASE TO SHARE DATA

RESOURCES

Jili Cheng

Institute of Philosophy, Chinese Academy of Social Sciences, CASS, Beijing, 100732

Email: chengjl-zxs@cass.org.cn

ABSTRACT

Philosophy research used to rely mainly on the traditional published journals and newspapers for collecting or communicating data. However, because of financial limits or lack of capability to collect data, required published materials and even restricted materials and developing information from research projects often could not be obtained. The rise of digital techniques and Internet opportunities has allowed data resource sharing of philosophy research. However, although there are several ICPs with large-scale comprehensive commercial databases in the field in China, no real non-profit professional database for philosophy researchers exists. Therefore, in 2002, the Philosophy Institute of the Chinese Academy of Social Sciences began a project to build "The Database of Philosophy Research." Until Mar. 2006 the number of subsets had reached 30, with more than 30,000 records, retrieval services reached 6,000, and article-reading reached 30,000. Because of the concept of intellectual property, the service of the database is currently limited to the information held in CASS. Nevertheless, this is the first academic database for philosophy research, so its orientation is towards resource-sharing, leading users to data, and serving large number of demands from other provinces and departments.

Key words: Philosophy, Professional research, Database, Resource sharing

1 INTRODUCTION

Scientific data is the most practical scientific resource of the information era. It is also the foundation needed for the development of the national innovation system, and data sharing is the best method to maximize the efficiency of the resources. However, resource data sharing has been troubled by many conflicts, such as the contradiction between freedom of information and confidentiality of information, the conflict between circulation of information and intellectual property, the contradiction between discrete data and polymerization, the conflict between local and national interests, the contradictions between commercial and nonprofit, and the contradiction between specification and comprehension.

Currently, advances in information technology and promotion of the country's scientific data sharing projects provide a good basis and reference to the solution of these conflicts. Therefore, the mandate of building a philosophy and social science data sharing system in China should be placed on the country's agenda. In addition to system platform/standards/sharing strategies, the primary need for a data resource sharing mechanism is the construction of data resources. Practitioners in the field philosophy and social sciences should resolve conflicts between specification and comprehension and commercial and nonprofit research-oriented tasks in order to build a database for philosophy and social sciences.

Philosophy is a subject focused on abstract thinking, so its academic achievements are almost always reflected in books and articles. Therefore, academic research in this regard depends largely on access to the literature. At present, Chinese researchers who engage in the study of philosophy are primarily in several major sectors: central and local research institutes, colleges, government schools for vocational education, vocational schools, and local propaganda departments. Because of the diversity of their organizations and geographical distribution, data acquisition and exchange rely mainly on the traditional newspapers and journals. Therefore, many smaller departments are so limited in financial resources and the lack of ability to obtain information that they cannot even obtain required public information and have no hope of obtaining restricted materials and information on research projects that even larger research institutions find difficult to collect.

Digital technology and the emergence of the Internet provide new opportunities for data resource sharing in the study of philosophy. The philosophy information published in many newspapers and the literature has already been entered into databases. At present, internal databases provided by Internet services are mainly integrated large commercial databases such as: the full text database of the Chinese Journal Network, the full text database of reprinted newspapers and journals, the Wanfang database, and the Longyuan data. However, there is still a big difference between these databases and professional databases with real research-oriented services for philosophy study. Listed below are major problems with commercial databases.

1. The content of the databases is comprehensive in subject matter and not specifically geared to philosophy; that is the coverage of specific references is not good enough and the accuracy rate is low;
2. The information included has been published in newspapers and periodicals. Few books, book reviews, monographs, research reports, multimedia, internal reference data, and other types of documents are included;
3. Few forms of academic activities, such as profiles of subjects, progress of disciplines and projects, forums and conferences, vocational education, social surveys, statistics, and personal data of academic researchers are included; and
4. Finally, these commercial databases are not free or have compensatory or low-reimbursement for services. Thus, even fewer possibilities exist in the professional area to conduct coordinated and managed activities of resources sharing.

On other hand, the existing traditional and digital libraries that contain comprehensive books, book reviews, and monographs, etc. are only a subset of the professional literature such as found in the databases mentioned above, and they need still more additions from other types of databases. In view of this, the best choice at present is to construct a research and sharing-oriented professional philosophy database. Such a database should contain resources not in present databases, such as:

1. contents and indexing for databases on theses and bibliographies;
2. data sets, such as the project's progress;
3. information supporting the possibility of developing professional activities for resource sharing and coordination that: are be non-commercial and nonprofit, have strong professional research, possess the capacity to gather information, have industrial appeal, covers the scope of internal system trials, and so on.

2 BUILDING THE PHILOSOPHY DATABASE

As an internal center of philosophy research, the Philosophy Institute, CASS began a philosophy research database in 2002 after the above goals had been settled. The implementation strategy for the project is:

According to research needs, set the types of database, data structure, scope of collection, and processing procedures and then establish the core database efficiently and rapidly. According to the institute's 50 years of research experience, more than 400 core newspapers were selected from thousands of possible as professional newspapers. The bibliographic database of the Institute and the Court has been regarded as the core bibliographic database. Fifty core sites were selected from 1,000 philosophy websites according to study standards, professional academic requirements, subject characteristics, and first produced/reproduced capabilities for information, while relying on internal reference materials and relevant files collected over 50 years.

The database is divided into six sub databases:

1. **The philosophy thesis database:** includes three subsets: academic papers, degree thesis database, and conference papers.
2. **The database of bibliography and book reviews:** refers to the philosophy bibliography subset plus a book review subset.
3. **The database for research progress:** divided into four subsets: disciplines progress, research projects, winning achievements, and academic activities.
4. **Graduate Education database.**
5. **Philosophical profile database:** scholars and institutions subsets.
6. **Other information:** important information falling outside the above scope.

Through links among these databases, full text retrieval across the databases can be realized, so that a complete database system has been built. This system also uses the information retrieval system that belongs to the 863 national plan, having a strong network-sharing function, rapid retrieval speed, good user interfaces, user classification with access control, etc. To avoid "information isolation" and to be readily accessible, the project also conducted a comprehensive selection for standardization of information, such as the selection of classification types, retrieval points, evaluation of different file formats, and so on.

After more than three years effort, the project has basically been completed, and the intended objectives and data size achieved. In March 2006, it contained 30 sub-databases, more than 30,000 data records, and had 6,000 retrievals. Meanwhile, when statistically comparing the data from our researchers with "The full text database of the Chinese Journal Network" and "The full text database of reprinted newspapers and journal," we discovered that this database is more suited to professional research. The main reasons for this are targeted services (specifically references), the collection of various carriers/record forms, flexibility of service, and so on.

"As a professional database system for the study of philosophy, it is a brand new form, which is different from the large-scale comprehensive databases and also from commercial databases. It focuses on professional, systematic, and comprehensive information. It is not the traditional 'paper database' but a 'database integration'

fully reflective of all aspects of the philosophical circles...”(Wang). It is the first internal Chinese philosophy database attempting to get a firm grip on the information included in the core subjects, frontiers of development, and cross and marginal disciplines and covering all second-class disciplines for data integration of different types, media, sources, and data formats, reflected in a timely manner

3 NEXT STEPS

Even more significance lies in the sharing of resources that is its basis of non-commercial, specific references with professional research abilities, information-gathering capabilities, and a range of internal system trials. In view of the above, the data resources sharing can be tested in a local range with a lack of sharing platform, standards/norm, and sharing strategy. The benefits of the internal system trial are not so sensitive to the contradictions involved in the categories mentioned. Also, the trial can be continued in a larger scope. Presently the database has provided trial services within more than 40 research institutes owned by CASS in the range of philosophy, social sciences, humanities, and area study. The trials will be undertaken in two fields:

- The database group of philosophy and social sciences;
- The professional field of philosophy research;

First, the database group of philosophy and social science is the basic resource for the development of philosophy and social sciences, as well as the basic condition for data sharing in the fields. On this basis, a sharing platform and user interface can be built, and more in-depth works of data mining and data warehousing can be carried out. However, the concept of “database group” is fairly broad. The requirements and standards are difficult to regulate for class-levels, chain-maps, and cast-links among individual database/subsets/related databases. Although the technology is relatively mature across heterogeneous databases, the specifications, functions, and scale of “database group” members should have some standards. As a research-oriented database, the “Philosophy database” has reached a certain scale in professional coverage, the amount of data included, degree of integration, and the relationships of class-levels, chain-maps, and cast-links among individual database/subsets/relative databases are quite clear. Therefore this database possesses not only independence of operation but also the ability to share data, which makes it an eligible member of the philosophy and social science databases group. Once the data-sharing platform for philosophy and social sciences has been established, the sharing trial for the “Philosophy database” can begin operation. As for the sharing strategies, this approach is also in line with “the pilot first” principle (<http://www.sciencedata.cn>) (The principle of construction for scientific data sharing projects).

Second, in the field of philosophy research, as mentioned earlier, research tasks and research staffs are distributed throughout several major departmental systems. They are the main users of the philosophy database services. This is a new challenge for data resource sharing in the scope of cross-systems, especially as conditions are not yet quite ready for the philosophy and social science database group, sharing platform, and sharing strategies trial. However, trade barriers will be greatly reduced by the Internet, so coordination of resources and mutual assistance in the field of professional activity is feasible and could take a variety of forms:

1. Cooperation could be started in data sharing, information security, and mutual division of work among units and databases similar in information resource strength and status.
2. Agreements can be reached concerning obligations for data collection and resources utilization among

units where large differences exist in resources and strengths.

3. To provide services to scattered individual needs, we can apply for registration in accordance with real names, determination of responsibilities, free use approval, the establishment of a personal credit files, etc.

This approach can also be used to accumulate sharing strategies for the future of the philosophy and social sciences data sharing platform. Early establishment of such professional research databases and the conducting of resource sharing trial activities make for a superior database.

4 CONCLUSION

From the practice, trials, and discussion of this project on the Philosophy database, it can be seen that although many factors affect the sharing of the data resources, it is essential to possess the databases and sharing strategies with sharing conditions. Starting from here, more explorations can be conducted and achievements can be reached.

5 REFERENCES

The Introduction of Scientific Data Sharing Engineering. Retrieved from the WWW, December 13, 2007:

<http://www.sciencedata.cn>.

Wang, T. J. The Project Report of the Database for Philosophy Research.

Xiao, X. M. (2005) To promote the construction of policy system for information resource sharing in China. *Information and Documentation Services (Bimonthly)*, No.6.