

Current Status and Development of Chemical Database System

Yunshen Zhang

Department of Petrochemical Engineering, Puyang Vocational and Technical College,
puyang457000, China

Keywords: Chemical engineering, information technology, database, present situation, development.

Abstract. The research and innovation of science and technology need scientific and technological for support. Research in the chemical and chemical fields these years requires a more specialized database system to match and support. The purpose of this paper is to study the current status of chemical and chemical databases from abroad and domestic, point out the current chemical database problems based on the actual situation of scientific research in China, and give the next step in the development of specific measures and objectives of the improvement.

1. Introduction

Chemical industry is in charge of the key lifeline of China's whole national economy which can not be ignored for science and technology's development. Therefore, we must pay attention to giving chemical research a more solid foundation. Fortunately, far-sighted scientists and leaders in the field of science put computer technology chemical research and development, pre-design and implementation of control and other links in 1978, so in these 30 years, computer information technology in the field of chemical applications has not been overlooked. In the process of chemical research, it often need to take a large number of chemical and chemical properties of objects, related structures, maps, data to support research, so we need to have a strong database extraction system. Content of database system is very extensive, in the broadest sense, it includes "bibliographic databases" and "factual databases". The so-called "literature database" is the integration of journals, conference reports, patented technology materials and so on, including bibliography, documents, etc.; The so-called "fact database" can save more data and information, this poses a higher demand for the use of artificial intelligence in chemical and chemical databases. Over the years, artificial intelligence technology took technical information and operational details of the actual problems together, formed a database form which has a strong link with the market. There are additionally such as "chemical fact database", "numerical database" and so on.

2. The Development of Chemistry Engineering Database System in the World

The United States attached great importance to the construction of the database research from the sixties of last century; developed countries in Europe and Asia, such as Japan as early as the mid-seventies, follow the footsteps of the United States, began to build a strong scientific research database system. The rise of the CD-ROM database in the late 1980s was the beginning of a database on-line retrieval system. The popularity of computer information technology and the rapid development of Internet started from the nineties of last century, and the "network database" system through the Internet technology to access began to develop, this is a concrete manifestation of the generalization of scientific information.

Nowadays, the Internet chemical database appeared along with the Internet process. American DIALOG system and ORBIT system, ESA system in Europe, and JOIS system in Japan have achieved the sharing of scientific and technological information all over the world. People have begun to focus on chemical and chemical database which is important and based on the Internet, they are more convenient than the traditional database retrieval methods, and also more international.

NCI Database Browser Data Retrieval System ([http: 129 .43 .27 .140 ncidb2](http://129.43.27.140/ncidb2)). This is the database system is mainly search for important compounds, as well as the structural properties of drugs and other information, of which there are a total of 251 compound structures to meet the needs of user search, users needing information can easily get it through the Internet, the entire search service and data are available free of charge.

The database search path includes: molecular formula, relative molecular mass, CAS coding, NSC coding, name, compound structure, complexity, the scope of the number of atoms and so on. The users can achieve multiple search targets according to their needs, so that they can get comprehensive information of specific compounds. The comprehensive information of these compounds includes but not limited to: CAS, relative molecular mass, complexity, molecular formula, anti-HIV activity, hydrogen bond acceptor, number of donors, number of donors, and predictive value of relevant biological activity. Users can also directly search for the substructure of the desired compound, dimensional efficacy, and similarity etc. on NCI, you can download the data information after getting it.

ZINC Database System. This database system is a three-dimensional database system provided by a professor named Shoichet at the University of California at San Francisco in the laboratory, its features is able to simulate screening directly, the system includes 46000000 compounds. Its search approach is the compound SMIEL format, and can specify the nature of the compound range, it is also possible to perform functional group searches directly in this database.

Zinc database system has a great contribution to drug virtual screening and design, the reason is that the search results are basically the same as the original input structure, the important properties of the compounds are shown together in the results at the same time, for example, "solubility", "number of hydrogen bond acceptors", "charge" and so on, and simultaneous display dimensional structure and three-dimensional structure of the compound can get all the compounds of structural molecules, search information can be converted into four formats of the files, and it can also be converted into table form.

Chemfinder Database System. Cambridge Soft Inc. in the UK created the Chemfinder database system, which is partially charged. Users can search for "basic properties and structure of compounds" free of charge. Users need to register when they use of this chemical database system the first time, the most common information about 75,000 compounds is then available. Convenience and speed are features of the Chemfinder database system and are a significant advantage over other systems. The path that the user searches on this database system can be by entering the chemical name, CAS number, molecular formula and so on.

SRC Database System with Physical and Chemical Properties. Syracuse provides this database free of charge, which contains 25,000 compounds of the basic physical and chemical properties. The weakness of this database system in use is that the search path is only achievable by CAS coding of the compound. Of course, it has a link service that looks for CAS numbers online. Through this database system, users can find the name, molecular formula, molecular weight, melting point, PKA, henry constant and a series of information. This database system is simple and convenient, and gives estimates that lack of experimental values for users' reference.

ChemBlink Database System. This database system provides two interfaces in English and Chinese, showing that International chemical and chemical database system has a large number of Chinese users, this also requires China to speed up the development and update in the chemical and chemical data library system. There are links that Chinese, English to switch to each other in the bottom of the system interface, there are more chooses on the use of the path, in addition, the system provides information on compound suppliers, so that users can buy them in time.

Based on the above-mentioned international chemical database system for the detailed study and the pros and cons analysis, we can learn from each other to help China's chemical and chemical database system to promote and develop.

3. The Development and Current Situation of Database System of Chemical Engineering in China

To the new century, after several technical upgrading, computer industry develops rapidly, many commercial computer software enter China in the world trade circulation environment, the use of computer technology penetrates into the chemical and chemical fields. Scientists realized that chemical database system replacement and the development of China's chemical industry will be closely related. Many chemical experts began to join in the chemical database system' research, development and application with their own professional combined with computer information technology to achieve "E-science" concept.

Referring to the practical application of the concept of "E-science", there must be online database system for users to use, now database a lot of chemical researchers and enterprises use include dozens of systems such as "Crystal structure database", "rare earth database", "a variety of chemical and chemical materials database". The amount can not say less, but there are still many problems to be solved. Our chemical and chemical database development can make a step closer only when these problems can be solved.

Problems of "Production Relations" and "Productivity". Scientific data should be detected by professionals themselves or get through journals and other academic channels, after the data collection, some professionals need to be fixed as a database professional. If the mobility of these professionals is too high, it will cause the database information flow. China has large deficiencies in this area at present, and also lack of obvious remedies behind. Solutions of the Western countries is "professional club" form, make "national", "industry", "members" three aspects support. While China is lack of experience, when there are tasks, we called experts to form a temporary group and dissolved it when finished, there is no long - term nature.

Problem of Innovation Degree. China's database builders are lack of the deep grasp on the law of scientific data, lack of awareness of the development of extrapolation model, they can not solve the innovation problem, just blindly collect data and work impetuously. In fact, the database requires a more in-depth knowledge of the quantitative structure of matter, need experts to set the heart to work in depth. If we ignore these, on the one hand it is difficult to ensure the quality of information in the database system, on the other hand, some column problems can not be structured and modeled, and it will reduce efficiency. The entire data and model derivation process, need deep professional foundation.

The Development of Internet Technology. China's Internet technology development and application can not be enough to solve these problems. In the international arena, especially in the developed countries, the high development of the Internet has promoted the solution of these problems, this also requires our country to further promote the development of the Internet network because it will be an important way to further enhance the development of chemical and chemical databases. Foreign "virtual laboratory" way is worthy of our reference, what the research leadership

needs to think about is how to promote the implementation of measures combining our realities, and can combine "academic democracy" and "mass line".

Employees to Control the Information. The majority of staff maintaining the normal operation of the database system should also be included in the necessary range. Researchers, production designers, and technology developers need to collect a lot of technical information in addition to scientific data. Such as equipment, materials, equipment, economic data and other information.

These problems are all need to be solved, in the real case, these descriptions are not enough to describe all the problems, but still sounded the alarm to us.

4. Measures and Suggestions on Developing Chemical Database System in China

To promote the development of chemical database system is an important way to enhance the soft power of scientific research in China. This paper proposes the following measures to promote the development of chemical database systems.

Firstly, combine the schools, enterprises, scientific research, production and other aspects to form a chain to help and complement each other.

In addition, looking for shortcomings when use it and integrate into our wisdom to make commercial software be further added and developed. We can choose some institutions and institutions having advantage in the information technology to establish "software development team" and create new areas, so that this commercial software can better fit the development and research of database system to form a better database system.

5. Summary

In short, the chemical database system is not only an auxiliary means of the Academy of Sciences, it already related to the entire national economic scientific development. It combined traditional and modern technology, made the network, artificial intelligence, integration of technology into integration so we need to face it. Problems need to be solved, it is more comprehensive scientific and technological than a "Chemical Database" system issue. Such a key problem is solved, technology is developed will be benefit for the whole field.

References

- [1] Zhang Yisen, *Advances in Computers and Applied Chemistry*, J. Computer and Applied Chemistry. (1954).
- [2] Zhou Jiaju, Xie Guirong, Yan Xinjian, *Handbook of chemical constituents of traditional Chinese medicine*, Chemical Publishing House, Beijing, 2004.
- [3] Xu Lu, Hu Changyu, *Application of Chemical Graph Theory*, Science Press, Beijing, 2000.
- [4] Hu Yan, Zhou Guosong, *Database Technology and Application*, Tsinghua University Press, Beijing, 2005.
- [5] Wang Zaiping, Wu Rongfeng, Ren Shiyong, *The Formation and Development of Database Industry*, J. Information Science. 13 (1992) 55-65.
- [6] Liu Zhenxi, Li Runsong, *An Introduction to Practical Information Retrieval Technology*, Tsinghua University Press, 2006.