

# Development and Current Situation of Agricultural Scientific Data Sharing in China

Hua Zhao and Jian Wang

Agricultural Information Institute of CAAS, Beijing 100081

**Abstract.** Since agricultural scientific data are valuable resources in the field of agricultural science and technology, it is important to share scientific data in agricultural science. The paper analyses current situation and existing problems on data sharing in our country, based on expounding the connotation of agricultural scientific data. Then some countermeasures and suggestions are put forward to promote the level of the agricultural scientific data sharing.

**Keywords:** Scientific Data, Agriculture, Data Sharing.

## Introduction

In recent years, with the rapid development of science and technology, the amount of scientific data accumulated by human society is more than the sum of the past five thousand years, the number of scientific data collection, storage, processing and dissemination has got a sharp increase. The scientific community has unanimously recognized that scientific data have a close relationship with knowledge innovation, which are the important resources for the sustainable development of modern science. Scientific data management and sharing have acquired unprecedented attention in various research fields. Many International organizations, such as the World Data System, Committee on Data for Science and Technology, Research Data Alliance, have been established to promote scientific data management and sharing on a global scale. Similarly, China has always paid great importance to agricultural science research. As important resources in agricultural science and technology activities, scientific data need to be shared in the fields of science, technology and education, such as government departments, researchers of agricultural scientific research institutes, persons of management-decisions and other research groups to promote agricultural science and technology innovation. In this paper, we analyze the current situation and the existing problems of agricultural scientific data sharing in China, and put forward improvement measures and suggestions for promoting the continuous development of and maximize value of agriculture scientific data.

## 1 The Connotation of Agricultural Scientific Data

According to national scientific data sharing project, scientific data refer to the primitive, basic data generated by the human from activities of science and technology

which they are engaged in during the period of perceiving and changing the world. They also contain processing data products and relevant information according to the different requirements of system [1]. Scientific data include both massive data from large-scale observation, exploration, investigation, experiment and comprehensive analysis of the long-term accumulation and integration carried out by the social public welfare sector, and a large amount of data generated by scientists through their practice of long-term taking part in the national science and technology planning projects.

Agricultural scientific data refers to basic data resulting from agricultural science and technology activities, as well as systematically processing data products and related information according to the different requirements, which include agricultural scientific experiment data, essential data and achievements transformation data of agricultural science and technology [2]. Agricultural scientific data is closely connected with agricultural scientific research, for the reason of agriculture science which belongs to a comprehensive discipline, involves many different fields such as biology, technology, environment and economy. Agricultural scientific data has three obvious characteristics [3]: (1) Complexity, dispersion and intersectionality. Scientific data on agricultural biodiversity, agro-ecological environment, soil and fertilizer research field have a clear overlap with research data from other disciplines; (2) large amounts and diverse types. Agricultural scientific research data contain both large-scale observation, detection and investigation data, and experimental data, which have extensive sources, many types, complex structure, wide spatial and temporal distribution; (3) diverse data service levels. Agricultural research data service objects include policy makers, scientific researchers, technology extension workers, agricultural business owners, farmers and so on.

As special agricultural information resource, agricultural scientific data have three aspects of important significance to be shared[4]: (1) For researchers, agricultural scientific data is the basis of their engaging in agricultural scientific research; (2) For the governmental functional departments, agricultural scientific data provides reference to the decision-making for agricultural science and technology management; (3) For the whole society, agricultural scientific data support the realization of agriculture informatization, as precious resources of agriculture and rural economic development. Realization of agricultural scientific data sharing is not only the foundation but also driving force to develop agricultural science and technology.

## **2 Current Situation and Existing Problems of Agricultural Scientific Data Sharing**

### **2.1 Current Situation of Sharing**

Chinese government always attaches great importance to the development of agriculture, increases investment in agricultural science and technology year by year, and organizes to carry out a large number of agricultural data collections, initially shaping scientific data work pattern to department as the main body and the scientific research institutes and universities as complementation. China has rich agricultural scientific data resources. However, the current generation and distribution of scientific data are particular. Collection, management and maintenance of massive scientific data

are done by government investment. However, data resources are dispersed in different departments, institutions or individuals [5]. The national agricultural research system is composed of agricultural research institutions at all levels, the agricultural colleges and universities and agricultural enterprises, and accumulated data are also scattered in the three subsystems. Although agricultural scientific data resources are abundant, each agency goes its own way. Most of the existing data resources are only limited to construct the database in institution inside or among cooperative organizations to carry out data exchange and provide limited data service, which restricts the agricultural scientific data value and results in the waste of resources of science and technology, as well as financial and human resources.

In 2003, Ministry of Science and Technology started to launch national scientific data sharing project, which tried out developing scientific data sharing centers in the field of nine disciplines, including the agricultural scientific data sharing center. After 5 to 6 years of efforts, the agricultural data sharing center had been accomplished, which based on the agricultural sector and integrated many agricultural scientific data resources by using modern information technology. Meanwhile the service system had realized standardization management and high efficient utilization of agricultural scientific data resources, strongly supporting our country's agricultural science and technology innovation. At present, the agricultural data sharing center has got some achievements in many aspects, such as integrating scientific data resources, formulating sharing standards and scheme, developing sharing system, providing sharing service, constructing talent team and so on. The platform is playing an important role to promote scientific data management and sharing in our country. By the end of 2013, the platform had finished integrating 60 agricultural core databases from 12 categories of agricultural disciplines, which included a total of more than 700 data sets. The amount of data was up to 3217 GB, accounting for 80% of the total stock of agricultural scientific data resources in China. In addition to the national agricultural scientific data center, there are many other information sharing platforms, which are established by local agricultural scientific research units or other research institution with the overlapping of agriculture. All these platforms have provided good conditions for agriculture scientific data sharing in our country. Agricultural scientific data sharing of our country is steadily moving forward.

## **2.2 Existing Problems**

In the field of agriculture, scientific data sharing and management has been paid great attention by the government departments and scientific research institutions at different levels, and the various resources sharing platforms are being built and perfected constantly for promoting the level of data sharing and improving management efficiency. However, some problems still exist in the implementation of data sharing which are mainly reflected as follows.

### **2.2.1 Lack of Effective Sharing Mechanism**

Compared with the developed countries, the practice of the scientific data sharing in China is still in the exploratory stage, accompanied by lacking macro and micro level of data sharing policies and laws and regulations, and also fails to establish an effective mechanism of sharing data. It's mainly embodied that numerous scientific data are

produced in implementing agricultural scientific research projects invested by government, while the data become data producer own groups' private property, only used in a small scale, most data are difficult to be touched by other persons or groups outside producers. Proper value of data has not been fully exploited, which results in low using rate of agricultural scientific data. Against this kind of strange phenomenon, there is no appropriate restraint mechanism and guiding mechanism to deal with it.

### **2.2.2 Lack of Strong Data Sharing Consciousness and Effective Data Processing**

Scientific data in the field of agriculture has more types of users which are not limited to researchers than other areas and their levels of understanding data sharing are different. Agricultural researchers are not only the main users, but also data producers. However data sharing consciousness among agricultural researchers is not strong. It is prevalent that most researchers are willing to use data which are owned by someone else and unwilling to provide their own data to be shared by others. During their engaging in research activities, the majorities of researchers focus on the archiving of experimental data and observational data, and regard the scientific data only as a segment of scientific research, while they don't thoroughly take into account how to process the obtained data and increase the value of data.

### **2.2.3 Data Processing Technology Is Not Yet Mature**

Modern scientific research needs to extract information from data, and then extract knowledge from information, and achieve the goals of forming scientific conclusions as well. New techniques and tools, including data acquisition, data integration, data analysis, data visualization, data distribution, data interoperability and the technology of refining information and knowledge based on the data are required during current scientific research [6]. In the field of agricultural sciences in China, some technologies related to the agricultural data management, online application service system and the distributed database management are not mature yet, which need to be further strengthened.

### **2.2.4 Deficiency of Data Resources Integration and Low Levels of Sharing Service**

Due to wide range of agricultural science research, the data resources may be scattered in various research fields, which expands the breadth and depth of resources integration. Standards of data classification, collection, organization and quality specifications in various areas are different, which increase the difficulty of data resources integration. Existing data sharing service ability has limitations and service pattern is relatively simple. There are a lot of works to do in order to increase the value of the scientific data and satisfy the innovation needs of agricultural science and technology [7].

### **2.2.5 Lack of Enough Talents in Data Sharing and Management**

Compared with other areas, it is relatively lack of compound talents who are proficient in expertise and information management in agricultural field. Most of the existing data sharing talents only have computer professional background, which are not familiar with areas related to agriculture. So there are many problems for data collecting, processing, preserving and sharing. It is difficult to adequately reflect the unique nature of agricultural scientific data and meet the needs of data sharing, the development of the talent team of scientific data sharing and management is hysteric.

### 3 Countermeasures and Suggestions

In view of the above analysis on existing problems about agricultural scientific data sharing, the author proposes the following suggestions for expecting to be helpful to the development of agricultural scientific data sharing and management in our country.

#### 3.1 Strengthening to Make All Aspects of Policies and Regulations and Establishing Open Sharing Mechanism of Agricultural Scientific Data

Large amounts of agricultural scientific data are generated through national investment, and have the strong public welfare, which belong to the state-owned assets. Much data are accumulated relying on the implementation of scientific research project and operation of agriculture-related departments. Therefore, it is necessary to overall plan and coordinate production, management and use of the scientific data between the various departments or institutions at national level, for making each research institution clearly recognize that management and sharing of scientific data are their obligation and responsibility. It is necessary to make multiple levels of relevant policies and regulations, including national level, the public sector level, as well as the specific research institutions so as to establish a good mechanism of scientific data sharing [8]. Many developed countries such as America and Britain, have ever established some policies and regulations to promote scientific data sharing. For example, the U.S. National Science Foundation (NSF) has required that the project applicant must simultaneously submit the "Data Management Plan" (DMP) at the time of submitting the project application since 2011. The U.S. National Institutes of Health (NIH) data sharing plan has required that since 2003, Project application for funding in more than 500 thousand dollars must be accompanied by explanations about data sharing or not, including sharing protocols, methods, and prospective institution [9]. Therefore, our management departments and institutions of science and technology related to agriculture can learn from foreign practices to make some policies and regulations suitable for the situation of our country for establishing and improving open sharing mechanism of agricultural scientific data.

#### 3.2 Changing Concepts and Raising Cognition

Researchers always focus on publishing research papers, they believe that scientific and technical literatures are closely correlated to their vital interests, while few of them care about processing and publishing of scientific data. The important thing is that publishing scientific data has nothing to do with the interests of the researchers, which leads to their lack of enthusiasm. In fact, scientific data is not only the basis of scientific research activities, but also the output of scientific research, as scientific literatures do. So by publishing scientific data, academic influence of researchers can also be improved. It is suggested that scientific data can be seen as an indicator, and can be absorbed into personnel performance appraisal and evaluation system of scientific researchers, with the purpose of encouraging agricultural scientific researchers to share data actively and enhancing their sharing awareness. It is essential to connect the scientific data with researcher' academic contribution and to motivate them to pay great importance for processing and publishing of the scientific data, ultimately to

change the old conception of advocating papers and ignoring data which has been existing in the mind of researchers[10].

### **3.3 Strengthening Studying on Data Management and Applied Technology**

Sharing of scientific data requires strong technical support, including database management and data mining technology, massive data storage and retrieval, data quality control technology. Processing, storage, management and sharing of agricultural scientific data also depend on the advanced technology. As a resource, agricultural scientific data have the great potential of deep processing and regeneration, it is a necessary way to study on data management and application technologies in order to process and regenerate agricultural scientific data. In the modern information technology, the key technologies to promote scientific data sharing include the interoperability of distributed database, supercomputer environment combined with a database, the combination of mathematical model and the database, standards and specifications of metadata and data documents, data mining , virtual reality and so on. In the field of agriculture, the application of these advanced technologies will be bound to play a crucial role in agricultural scientific data sharing, so study on these technologies must be strengthened.

### **3.4 Increasing Integration of Data Resource, Improving Service Capabilities of the Data Sharing Platforms**

Agricultural Scientific Data Sharing Center as a national platform has integrated some agricultural science data resources at present. Because of the continuity of agricultural research, the continuous generation of new data, meanwhile a large number of agricultural science data resources are scattered in different institutions or departments. These data resources must be integrated timely and efficiently. Therefore it is necessary to strengthen cooperation and communication with other platforms and institutions for implementing data exchange in order to realize adequate integration of the agricultural scientific data resources. In addition, to enhance data sharing service capabilities, the existing platforms will always focus on the needs of users, innovate service means, provide users with various forms, content-rich data-sharing services, including data storage and access to services, data citation services, and data analysis services in a deeper level. Under these service patterns, the platforms will advocate passive service combined with initiative services, paying more attention to increasing user's participation in particular.

### **3.5 Developing and Training Talent Team of Agricultural Scientific Data Sharing and Management**

Scientific data sharing is a complicated systematic project, which must be achieved through collaboration of multi-disciplinary talents. Sharing of scientific data needs a series of steps including data collection, processing, storage, preservation, sharing services and operating maintenance, in addition to the need for talents with expertise. Information management professionals are indispensable for scientific data sharing. Therefore, talent team of agricultural scientific data sharing and management must be

constructed in two aspects: On one hand, the information technology professionals and specialists in the field of agriculture and agricultural information management talents are introduced into the data sharing management team; On the other hand, it must speed up to train some compound talents who are proficient in both technologies and management.

## 4 Conclusions

Over the past decades, sharing and management of scientific data in the field of agriculture have been greatly improved and got some achievements. To some extent, scientific value, economic value and social value of agricultural scientific data have been realized. But there are still a lot of works to be done. The work of data sharing has a long way to go. There are many issues to be further considered and resolved, such as legal boundaries during effective management and sharing of scientific data, and how to use data more availablely. Therefore, management and sharing of agricultural scientific data will continue to be a major problem that faced in science communities, which will always run through the entire agricultural science and technology activities. It is becoming an important support to realize agricultural science and technology innovation.

## References

- [1] SDS/T 1003-2004. Scientific data sharing project technical standard [S]
- [2] Hu, H., Liu, S.: Discussion on the Content and Service of "National Agricultural Sci-Tech Data Sharing Platform". *Journal of Library and Information Sciences in Agriculture* 17(2), 214–217 (2005)
- [3] Zhang, L.: Study on the development of Agricultural Scientific Data Sharing in China. Doctoral Dissertation, pp. 19–22. Graduate school of Chinese academy of agricultural sciences, Beijing (2006)
- [4] Zhang, X., Li, S.: Study on Agricultural Scientific Data sharing service. *Chinese Science and Technology Forum* (5), 127–130 (2006)
- [5] Wang, Y., Hua, X., Wang, J.: Analysis of the Domestic and Foreign Scientific Data Management and Sharing. *Science & Technology Progress and Policy* 30(4), 126–129 (2013)
- [6] Guo, M.-H., Li, J.-C., Tian, J.-L.: The Management of Scientific Data Sharing in China. *Xi'an University of Architecture & Technology (Social Science Edition)* 28(4), 84–88 (2009)
- [7] Zhao, R.: Prospect and Practice on Agricultural Scientific Data Sharing Center. *Agriculture Network Information* (6), 4–12 (2009)
- [8] Song, L., Meng, X., Zhou, G.: Measures and Suggestions for Information Quality Management in Agricultural Scientific Data Sharing in China. *Journal of Agricultural Science and Technology* 11(6), 37–42 (2009)
- [9] Wang, Q.: Research on the opening Modes, Guarantee Mechanism and Optimization Strategies for Opening and Sharing Scientific Data. *Journal of The National Library of China* (91), 3–8 (2014)
- [10] Liu, C.: Study on Scientific data sharing mechanism in China. *Land and Resources Informatization* (1), 5–7 (2004)