Data Sharing Practice in Agricultural Science Domain

Sijing Li¹, Guomin Zhou²

¹ Agricultural Science and Technology Press, China, lsj@mail.caas.net.cn
² Agricultural Information Institute of the Chinese Academy of Agricultural Sciences, China, zhougm@mail.caas.net.cn

Abstract

Faced with the rapid development of agricultural science and technology, it has new requirements of the basic condition and environment construction to support agricultural science and technology development. In order to speed up the development of agricultural science and technology, it is essential to build a fine essential work environment, such as scientific data sharing. Therefore, the agricultural scientific data sharing work has very strong age background and enormous actual demanded.

In 2003, the Project of Agricultural Scientific Data Sharing Center (Pilot project) in China had initiated and started the sharing practice of scientific data in agricultural science domain. The Center (Pilot project) has built the agricultural scientific data center including 1 main center, 7 sub-centers and 28 data nodes, which is built on the data sharing platform based on Internet. It has integrated nearly 60 major agricultural databases, realized Internet data sharing. In addition, it has built a stable agricultural scientific data resource integration, management and service team, realized the sharing and service of agricultural resources of the whole society and built an agricultural scientific data center with a coverage of the whole nation and linked with other parts of the world, so as to provide data support for agricultural science and technology innovation and agricultural science and technology development.

On the basis of the formulation of standards, the project group has pooled all agricultural specialized major database groups and database sets in secondary indexing, grading and classification, metadata registration and filing, remodeling and recombination of non-structural data, organization and release of networking organizations, data backup and other work. Agricultural scientific data sharing platform network has provided good online data service capacity. The network can provide data navigation, search, download and other functions, and partial databases can also provide data visualization functions.

Keywords: Agricultural science data, Data sharing, Sharing service

Introduction

Agricultural science data are the essential data produced in the engagement of agricultural science and technology activities, and systematically processed and sorted data products and relevant information based on different demands. Agricultural science data are important essential resources of agricultural science and technology innovation. The construction of agricultural scientific data center can provide support and guarantee of agricultural science data information resources for agricultural science and technology innovation and agricultural science and technology management decision-making.
In 2003, the Project of Agricultural Scientific Data Sharing Center (Pilot project) had initiated and started the sharing practice of scientific data in agricultural science domain. The project has the AII-CAAS (Agricultural Information Institute, Chinese Academy of Agricultural Sciences) as its main body with the participation by partial specialized institutes of CAAS, the Chinese Academy of Fishery, Chinese Academy of Tropical Agricultural Sciences, etc.

Social Demand

The development of agricultural science and technology requires the support of agricultural scientific data sharing environment. Meanwhile, it also releases enormous demand of agricultural scientific data. Since the Chinese economic development entered a new phase, the development of agricultural science and technology has faced with new challenges. Although agricultural science and technology has basically met the demand of agriculture and rural economic development, yet there have been great gaps as compared with agricultural promotion, transformation and span development in the new age. Objectively, these challenges require a great development in agricultural science and technology. In fact, China has gradually strengthened agricultural science and technology input in recent years and promoted agricultural science and technology innovation capacity. It can be said that China has faced with a new spring of agricultural science and technology development. Faced with the rapid development of agricultural science and technology, it has new requirements of the basic condition and environment construction to support agricultural science and technology development. In order to speed up the development of agricultural science and technology, it is essential to build a fine essential work environment, such as scientific data sharing. Therefore, the agricultural scientific data sharing work has very strong age background and enormous actual demanded.

Overall Design

The Agricultural Scientific Data Sharing Center (Pilot project) is a set of comprehensive standard system of agricultural scientific data sharing under the overall framework of the “National Scientific Sharing Data” construction and based on the “Overall Program of the National Agricultural Scientific Data Sharing Standard System”. The Agricultural Scientific Data Sharing Center (Pilot project) has built the agricultural scientific data center including 1 main center, 7 sub-centers and 28 data nodes, which is built on the data sharing platform based on Internet. It has integrated nearly 60 major agricultural databases, realized Internet data sharing. In addition, it has built a stable agricultural scientific data resource integration, management and service team, realized the sharing and service of agricultural resources of the whole society and built an agricultural scientific data center with a coverage of the whole nation and linked with other parts of the world, so as to provide data support for agricultural science and technology innovation and agricultural science and technology development.

It is essential of integrate existing agricultural scientific data based on the principle of “highly logic concentrated and appropriate physical distribution”. The integrated agricultural scientific data will provide sharing service in the form of networking. The construction of the National Agricultural Scientific Data Sharing Center (Pilot project) includes 6 aspects of work. The first is the formulation and revision of agricultural scientific data sharing standards and regulations. The second is the digitalized processing and standardized arrangement of major agricultural databases. The third is networking organization, storage and sharing of agricultural
scientific data resources. The fourth is the development of key technology research and software systems of agricultural scientific data sharing platform. The fifth is the operation management of agricultural scientific data sharing center and the maintenance of portal website. The sixth is development of agricultural scientific data sharing service.

**Data Resource Distribution and Integration Means**

In the past 20 years, the agricultural domain in China has constructed several hundreds of agricultural scientific databases of different scale, different quality and different application degrees. At present, in the domains of crop science, animal production, veterinary medicine, agricultural resources, pratacultural science and agricultural ecological environment, biological information and biotechnology, farm produce processing, food quality monitoring and standard, agricultural economics and policy, tropical crops, fishery science, etc., it has created and accumulated large quantities of primary data in scientific experiment, observation, exploration, test, examination, investigation (survey), as well as series of data and comprehensive data information after processing. According to incomplete statistics, the agricultural scientific data accumulatively stored by the three academies directly affiliated to the Ministry of Agriculture (the Chinese Academy of Agricultural Sciences, the Chinese Academy of Fishery Sciences and the Chinese Academy of Tropical Agriculture Sciences) accounted for over 60% of the national total, and has preliminarily formed the national agricultural scientific data center population and sharing service network.

Agricultural scientific data need be integrated through a set of effective standards and regulations. At present, China has already studied and formulated the framework of agricultural scientific data sharing standard system, completed a series of standards of data manufacture and data management, such as metadata standards of agricultural scientific data, data element standards of agricultural scientific data, subject category classification code standards of agricultural scientific data, database format standard, database use instruction standard, user authentication regulation and other standards and regulations in agricultural scientific data sharing and service, which have laid a solid foundation for an overall building of agricultural scientific data sharing platform and made the integration and sharing of agricultural scientific data under the instruction of certain standards.

On the basis of the formulation of standards, the project group has pooled all agricultural specialized major database groups and database sets in secondary indexing, grading and classification, metadata registration and filing, remold and recombination of non-structural data, organization and release of networking organizations, data backup and other work.

**Sharing Service Status**

Agricultural scientific data sharing platform network has provided good online data service capacity. The network can provide data navigation, search, download and other functions, and partial databases can also provide data visualization functions. Since its operation, the access amount of the main website and sub-websites has exceeded one million times. The ranking of the Alex website in the world has increased to the 190 000th, the total sharable data volume integrated and remolded has reached 32 GB, the registered users have reached 4000 and it has already provided an accumulative 2.47GB of online downloaded data. The users are mainly distributed in the research institutions, educational organizations and enterprises and institutions. Data provided off-line have reached 5.9 GB.