

Key Technology Study of Agriculture Information Cloud-Services

Yunpeng Cui, Shihong Liu

Key Laboratory of Digital Agricultural Early-warning Technology, Ministry of Agriculture, Beijing, The
People's Republic of China 100081
cyunpeng@163.com, lius@mail.caas.net.cn

Abstract: The rural information service in the rural area in China developed rapidly these years. But the agriculture information resources need to be further development and utilization, the key of agriculture information service at the present stage is to summarize the real demand of the rural users, choose suitable technology, find the right solution, solve the information sharing and fusion problems fundamentally. AISC(Agriculture Information Cloud-Services) is a cloud service platform, which tried to import different datasets in different data sources that constructed by different organizations, and combine all these datasets into one big dataset logically. Based on the “big dataset”, many applications such as semantic information retrieve, intelligent information push, self-organized knowledge base construction etc. were developed, so the service efficiency is improved.

Key words: Agriculture Information, Cloud-Services

1. Introduction

Along with the information technology revolution influence continues to expand, the informationization course in China progress steadily, Meanwhile informationization in China started to gradually penetrate from city to rural area. As the rural information infrastructure project (“CunCunTong Project”) completed successfully, the rural area oriented information service in China become possible.

However, at present, in China, Rural information service development is not balanced, especially in the west regions of China, farmers' information literacy and information consciousness still need to be improved; The agriculture information resources need further development and utilization, the way that rural users obtaining and receiving information should be enriched too; The standard system of agriculture information should be established to regulate and promote the sharing, exchange and integration of agricultural information resources, and ensure the quality of information resources; the personalization of agriculture service also need to strengthen so that service can meet the requirements of different rural users.

So, the key of agriculture information service at the present stage is not database and system construction, but is to summarize the real demand of the rural users, choose suitable technology, find the right solution, solve the information sharing and fusion problems fundamentally, explore the effective mechanism and model of information service, develop available information products, provide quality and efficient information services, help them solve the problems in their production and living.

The goal of the study is to establish a comprehensive agriculture information service cloud through global catalogue exchange technology, universal data element and data item presentation

and mapping technology, implement the fusion of distributed agriculture data sources, develop a personalized, easy-to-use comprehensive agriculture information service cloud(AISC) and its portal, provide precise information service to rural users.

2. The technical architecture of agriculture information service cloud(AISC)

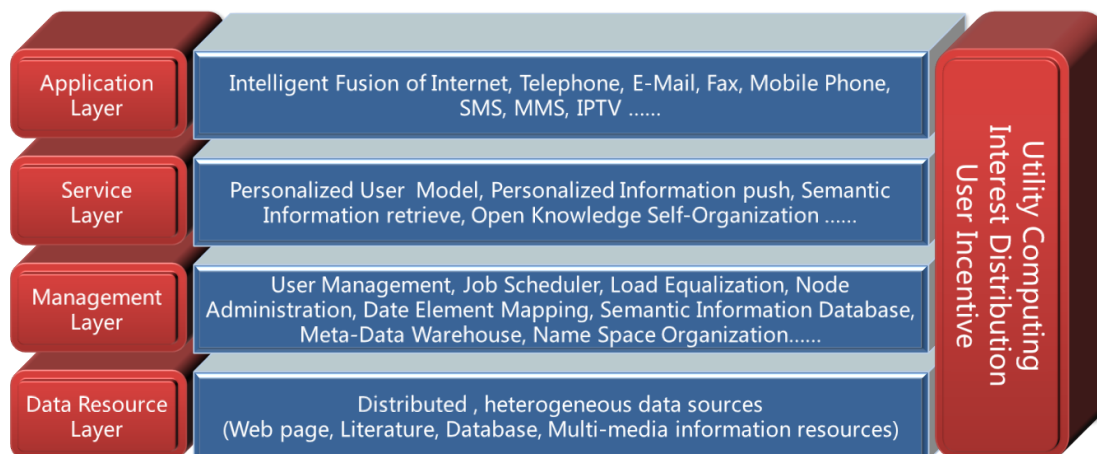


Figure 1 Technical Architecture of AISC

As shown in figure 1, AISC has four layers, from bottom to top are data resource layer, management layer, service layer, application layer and respectively.

The data resource layer include all the distributed, heterogeneous data sources which join the cloud, include web pages, literatures, databases and multi-media information resources etc., the data sources will be integrated into a transparent logic “big single data source”, which can provide data to upper layers/

The management layer is consist of different functional components such as global user management, job scheduler, load equalization, node administration etc., these components can manage the cloud effectively,

3. Core Technology of agriculture information service cloud(AISC)

To construct the AISC, the following problems must be solved:

(1) The mapping technology between universal agriculture information data elements and data items in databases

Because the AISC is constructed based on the distributed information datasets, so it’s very important to make the datasets connect logically and interoperable. The solution is to construct a warehouse that contains all the mapping relationship between the universal standard agriculture information data elements and the data items of all the distributed database.

So we must first construct the universal standard agriculture information data elements, we can construct them by brain storm, and also we can develop a tool to extract related concepts from articles. As the basis of the future intelligent retrieve, we establish the semantic relationship between data items, and save the relationships in the data warehouse.

(2) Study of agriculture information dataset metadata and service metadata standards

To implement the universal information fusion, we must have a set of standard metadata to describe the datasets consistently, so a core metadata standard is required. We compiled agriculture information datasets core metadata standard^[1], which include 75 metadata elements, for every element there are 9 attributes(see Table 1) to describe and restrict it.

On the other hand, the agriculture information core metadata can only implement the

universal datasets description, if we want to import all the datasets into AISC and make them to be accessed as a logical single data source, we must have service metadata support. The service metadata provide the universal, standard description for different services, such as HTTP Service, FTP Service, Data Connection Service, Data Access middleware etc., with these service metadata, computer can read the values of service metadata, locate and access the dataset, so retrieve data from different data sources. figure 2 shows an example of service metadata, which is dataset connection service metadata, through it computer can get the information like IP address of database host, access port number, database name etc., and then computer can locate the host and connect to the database.

Table 1 the attributes of the elements of agriculture information resources dataset core metadata

Name of attribute	Description
Chinese name	Chinese name of the element
English name	English name of the element
Identification	The unique identification of the element, string.
Definition	The specifications description of the meaning of the element.
Type	The type of the element, the available types include: composite(the element contains sub elements),integer, float, text, date, time, datetime etc.
Range	The allowed range of the value of the element
Optional	The element is required or optional
Maximun appearance	The maximum appearance of the element, such as 1 (only once) 、 N (unlimited times) etc.
Note	Supplementary specifications of the element

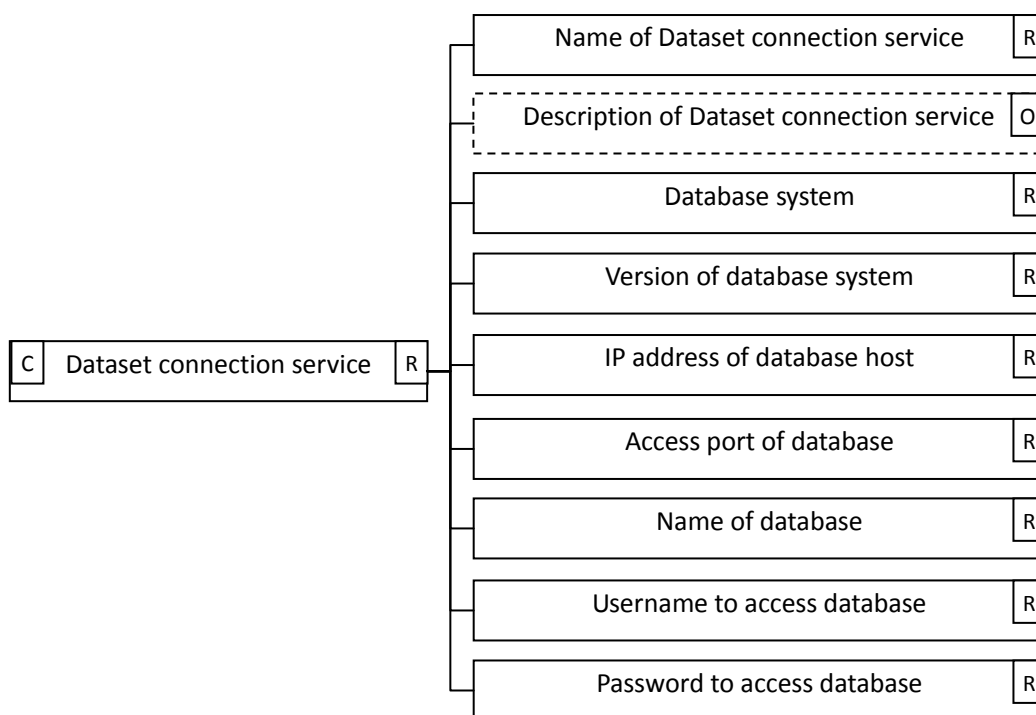


Figure 2 the example of dataset connection service metadata

So the agriculture information datasets core metadata can map different data items in different datasets and connect all the datasets into one single big dataset logically, through service metadata, computer find where the dataset is and access the dataset directly, so all the datasets are “imported ” into the cloud and can be managed in the cloud.

(3) Construction of ASIC platform and portal

We have logically bunched different information resources together in last step, so the problem now is how to use the resources, and how to let the information reach the user effectively.

So first of all, we should make agriculture information service individualized. We established user information requirement model to describe the users' information requirements dynamically through data mining from users' information behavior^[2], and use the model in the user information retrieve process^[3], and also push the subscribed information to users actively.

On the other hand, not all the farmers own a computer, or can connect to internet, so we must import different technology and create different models, so all the farmers can access the platform and service portal through the applicable way. We tried to import universal information technology into agriculture information service, so farmers can access the platform and get information with telephone, fax, email, SMS, IPTV and special information terminal equipment.

To improve information retrieve efficiency, we studied semantic information retrieve technology based on the semantic relationship establish in step (1), analyze the information resource in the cloud, aggregate the related information together automatically, and also use the technology in information retrieve and push process.

Though there is plenty of information in the cloud, it's not enough for farmers to solve all their problems, so we let them fill the knowledge base themselves. We studied the open knowledge self-organized technology, everyone can raise their question and everyone can answer any question raise by anybody, the people who raise the question can then evaluate the answer and score the answers, the best answer will be adopted by the platform as the final answer of the question and be saved in the knowledge base. When a user raise a question, he/she must choose tags to identify the question, the tags is in fact the keywords of the question which are provided by the platform, the tags in the platform are organized by a set of ontologies, so the questions and their answers is organized.

(4) Study of stimulate mechanism of information sharing

To stimulate the dataset providers contributing more resources to the clouds, there must be a set of mechanism to encourage them^[4]. The principle is "more contributions, more rights". Not all the information are free in the platform, the provider can also ask the users pay for special information, so attract the dataset providers and users participate in the service and communication.

4. Conclusion

The AISC is an attempt to promote agriculture information service. It tried to import different datasets in different data sources that constructed by different organizations, and combine all these datasets into one big dataset logically. Based on this "big dataset", we can develop many applications such as semantic information retrieve, intelligent information push, self-organized knowledge base construction etc., so improve the service efficiency.

ACKNOWLEDGEMENT

The research was supported by the special project from ministry of agriculture of the people's republic of China, named study of agriculture informatization standards system and special fund of basic commonweal research institute project of information institute of CAAS, and National 11th five-year technology based plan topic named study of Agricultural product quantity Safety Data

obtained standards (2009BADA9B02).

Reference

- [1] Cui Yunpeng, Liu Shihong. A metadata based agricultural universal scientific and technical information fusion and service framework. *Computer and Computing Technologies in Agriculture IV*. 2010 part1:56-61
- [2] Gao Xun. Research on web structure mining algorithm based On cloud computing. Master's degree thesis
- [3] HE Jian-jia, YE Chun-ming. Cloud computing-oriented data mining system architecture. *Application Research of Computers*. 2011, 28(4):1372-1374
- [4] Yin Xiaoming, Research on business model of cloud computing based on value net, Master's degree thesis