Current key problems and Solutions for Geo-Scientific Data Sharing in China

CODATA,2006, Beijing

Youhua Ran, Xin Li, Jian Wang
ranyh@lzb.ac.cn
Lab of RS and Geo-Spatial of CAREERI, CAS
Environmental & Ecological Science Data Center for West China
http://www.westgis.ac.cn
http://westdc.westgis.ac.cn
2006.10

Contents

- Introduction
- Where the inspiration comes from?
- Key problems
- Practical solutions
- Discussion

Introduction

- Scientific data is the most active resource for scientific and technological innovation. More abundant, long series, high-quality data is pressing needed in the earth system science research.

- The scientific product, management, orderly sharing for mass geo-science data will greatly promote the continual development of Geo-science research and provide powerful support for science breakthrough.

- At the same time, scientific data are society shared knowledge resources and it invested by the country; it should be shared to serve the state's overall goals.

- As scientific data sharing is political and economic needs, it has obtained broad consensus by the whole society.

Introduction cont…

- In recent years, Chinese have done much work for geo-science data sharing. Many platforms for data sharing have opened and the corresponding services are working.
- Sharing platform: The China meteorological data service system, National MODIS data center Data-sharing network of Earth systematic Science, Resource and environment data center of the Chinese Academy of Sciences, Scientific database of the Chinese Academy of Sciences, Chinese Terrestrial Ecosystem Flux Research Network (ChinaFLUX) , National satellite weather center (NSMC) satellite climate share data network and the NSFC-WESTDC.
- The various data such as remote sensing images, meteorology, hydrology, ecology, soil, desertification, basic geographic mapping data and other basic data can be obtained through the online or off-line. The improved data environment also contributes to the development of Chinese geo-scientific research.

Where the inspiration comes from?

- The WESTDC visited the project of the NSFC ---“Environment and Ecological science Project” to provide data services and requirement investigation.
- We summarize these thoughts as following five aspects that are common in China:

Key problems

1. The data provided to share are very limited. Although many data can be downloaded on line or application off line, most data are limit to the satellite data and large-scale background data. It is incomplete categories and lack of systematic to be of little value for most research.
2. The data standards are inconsistent, metadata is missing and it is hard to evaluate data quality. Data provided by most platforms are not detailed documents. All are difficult for data user to evaluate the research results and for reference which affects the data authors' enthusiasm.
3. The consciousness of data sharing in science and technology field is poor, data applying procedure is tedious. (A example.). At the same time, most user may have experience cumbersome process such download and filling paper applications table, signature, stamped etc. There is a long time cycle even man-made barrier also restricts the sharing of scientific data to some extent.

Key problems cont…

4. It is understaffed that engage in the data sharing. Today scientific data sharing project are most carry out with researchers. But in academia of China, scientific data management, development and service was not considered as a science. We know many people do not want to do “non-science” in science environment.
5. The accessibility of many data sharing platforms is poorly. The building of data share platform is started by project way. After the project finished, the platform operating cost usually can not be guaranteed. Therefore this has a negative effect on the sustainable platform operation and on the confidence of users.

Solution ?

Potential solutions

Change the scheme of data integration to enrich database

Scientific data for its especial production mode and data output flexible and changeable. The interest group is complexity. So we must consider synthesized multi-interest and long-term planning to collect. Some suggestions as follows:

- To find data through literature.
- Cooperate with other data section and had best basic as a common data and technology standard .
- Active support data mining for historical data.

Potential solutions cont…

To perfect metadata and documentation

For geospatial data, such **metadata** information includes data abstract, format, projection, downloading etc. **Data documents** provided information that the data produce process include the data producing methods and ideas etc. This involvement invaluable information to supporting unanticipated users [Parsons et al.2005]. Moreover, detailed metadata information and data documentation are improving users access.

However, the work of metadata and data documentation are challenging. We suggest the personnel of documentation needed play by the corresponding field experts.

Potential solutions cont…

To emphasized data service and simplify application procedure

- Scientific research work has a traditional characteristic that is opening, for the people engaged in scientific data sharing, the need for a more active and open attitude.
- The major reason of inefficient, complex data application procedures of the science data services in China is lack of perfect function network platform used to provide data service and the poorly accessibility.
- Therefore, strengthens the geosciences data sharing talent team, mainly refers to the platform construction talents. Make full use of the advantage of the Internet to simplify data application procedures, and enhance the efficiency and quality of service.
- So the corresponding system should appear to standardize the data application procedure make the user more direct and convenient obtained data.

Potential solutions cont…

To enhance the sharing consciousness in science and technology field, which helps to the data sharing system forming

This requires data managers and data author should change their mind. Data sharing staffer should actively explore scientific data sharing.

The data sharing system should be formulated in the protection the rights of investor and data authors on the basis of state secrets, must balance the interests of all parties basis on the government-driven, When the time is ripe to be developed in the corresponding laws and regulations to standardize the action of scientific data sharing.

This is a long-term, indirect impact on data-sharing, but the work is very important.

Potential solutions cont…

To establish national data centers and realize sustainable data sharing

After experiencing decade's free development, The basic pattern of Chinese geosciences data sharing has formed.

Building of national data centres, the government should fixed long-term special funds and the corresponding laws to supporting it's operating. So it can leader the orderly data sharing environment forming.

In this foundation, we can reference the management scheme of scientific literature to scientific data publish and sharing.

Discussion

- Parsons had summarized well about establishing scientific data community.
- In the practical process, the reality and probability should be fully considered.
- The solutions proposed are interactional and can act on the data sharing commonly.
- However, we must understand that the contradiction between the supply and demand of scientific data will existent for ever. The development of geosciences data sharing still has many institutional obstacles.

The revolution has not succeeded, keep on working hard!

Thank you!