The Cross-agency Sharing Scheme and Data Quality Evaluation —— A Case study of Geo-spatial big Data

Yi Bai1,2*, Song Li1, Wei Zhang1

1School of Geography and Resources, Guizhou Education University, Guiyang 550018, China
2Guizhou Provincial Key Laboratory of Geographic State Monitoring of Watershed, Guiyang 550018, China;
*Corresponding author’s e-mail: baiyigznc@foxmail.com

Abstract. In order to improve the sharing degree of big data and solve the pain points that are unwilling to share and can not be shared, cross-agency analysis and research should be carried out. Aiming at the construction characteristics of Geo-spatial big data, the main problems and solutions in Geo-spatial data sharing were analyzed. Starting from the management system, technical standards and physical architecture, a big data cross-agency sharing scheme was proposed. Then the quality of 3D building model data were evaluated. The results showed that constructing systematic cross-agency sharing solution will be helpful to promote the rational, standardized and efficient application of Geo-spatial big data.

1. Introduction
The basic connotation and infrastructure of all kinds of big data can not be separated from Geo-spatial information[1]. The existing body of research on open sharing of big data suggests that it mainly focus on individual aspects such as cooperation mechanism, data protection, platform construction, evaluation method[2-4].

With the advent of the big data era, especially the large number of government departments involved in the management and use of Geo-spatial big data, researchers have shown an increased interest in defining the responsibility and ensuring the efficiency of data sharing. There is an urgent need to address the safety problems caused by open sharing of Geo-spatial big data. There are two primary aims of this study: (1) To investigate Geo-spatial big data cross-agency sharing scheme. (2) To ascertain the data quality of open shared. Both qualitative and quantitative methods were used in this investigation.

2. Research methodology
2.1. Cross-agency sharing
"Overall government" (Bardach.E., 2011) is a new concept of contemporary government management, which has become a hot practice and academic research field in the government reform of developed countries[5]. Based on the summary of many attempts of our government to promote inter-departmental administrative coordination, domestic scholars put forward that cross-agency data coordination is a new direction[6][7]. To solve the pain points that data are unwilling to share and can
not be shared, it is not only need start with single aspect of cooperation means, technical operation or effect evaluation, but also need to carry out comprehensive cross-agency analysis and research.

2.2. Multi-level Fuzzy Comprehensive Evaluation

Multi-level fuzzy comprehensive evaluation method is a scientific method for quantitative description of fuzzy things. It can fully take into account all kinds of available information and quantify and distinguish its fuzziness[8]. It is suitable for evaluating the data quality in the open sharing of Geo-spatial big data.

The variable $U$ represent $m$ factors for evaluation, it is expressed as follows:

$$U = (U_1, U_2, ..., U_m)$$  \hspace{1cm} (1)

The variable $V$ represent collection of $n$ comments, it is expressed as follows:

$$V = (V_1, V_2, ..., V_n)$$  \hspace{1cm} (2)

In GIS quality evaluation, $V = \{\text{excellent, good, medium, qualified and unqualified}\}$. If use $r_{mn}$ express the membership degree from factor $m$ to comment $n$, the fuzzy relation between the factors and comments domain can be expressed as:

$$R = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1m} \\
    r_{21} & r_{22} & \cdots & r_{2m} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{mn} & r_{m2} & \cdots & r_{mm}
\end{bmatrix}$$  \hspace{1cm} (3)

If $U$ is composed of $k$ layers ($k \geq 2$) and the first layer (the highest level) has $m$ factors, that is $U = (U_1^{(1)}, U_2^{(1)}, \cdots, U_m^{(1)})$, then the data model of Multi-level fuzzy comprehensive evaluation can be expressed as:

$$B = A \circ \begin{bmatrix}
    A_1 \circ R_{11} \\
    \vdots \\
    A_k \circ R_{1k} \\
    \vdots \\
    A_m \circ R_{m1} \\
    \vdots \\
    A_m \circ R_{mk}
\end{bmatrix}$$  \hspace{1cm} (4)

3. Geo-spatial big data sharing Scheme based on cross-agency sharing

3.1. Main issues

The Geo-spatial data, especially the basic surveying and mapping data, are collected and managed by many departments, so there are some problems such as repeated and outdated construction, and the multi-source heterogeneous data conflict with each other. In the platform construction, the problem is lack of an authoritative sharing platform. In the data sharing system, the shared data is mostly difficult to use or outdated data.

3.2. Solutions

In view of the above problems, the sharing of Geo-spatial big data can be realized by using cross-agency sharing system, by data integration and data interface. Geo-spatial big data has spatial attributes, it must be acquired, developed, counted and analyzed on the basis of unified standards. The perspective of "human-logic-physics" can not only start from the specific connotation and spatial attributes of big data, but also take into account the comprehensive attributes of its open sharing.
involving management, technology and data. Aiming at the specific standard object of geographic information, from the point of view of "human-logic-physics ", it corresponds to three dimensions: data, technology and management. The three dimensions are closely related to each other to form a soft environment that supports the open sharing of Geo-spatial big data. The research perspectives and solutions are illustrated as Figure 1.

Figure 1.Geo-spatial big data cross-agency sharing scheme.

It includes the following aspects:

1) **Top-level design for cross-agency collaboration and sharing.** "Top-level design", using system theory to plan a task as a whole, it is the overall consideration of all levels and elements of the project. It can improve the information technology, ensure information sharing.

2) **Building a common platform for the use of geographic information.** The use of a unified platform to manage data resources in multiple applications conforms to the SOA design idea and can maximize the simplification and interaction of the application system. Such as Web Service JMS. According to the type of data and the purpose of use, users can use the data free of charge or paid. Non-confidential data from each department should be remitted to one department for management.

3) **Establishing metadata systems and data standards.** Metadata is the data that describes objects such as information resources or data, it is an important basis for data sharing. It provides specifications, general description methods and retrieval tools for various forms of digital information units and resource sets.

4) **Open data service interfaces.** There are great differences and also intersections in business among various departments. Such as real estate informatization involving land management, urban planning and other business, involving development enterprises, urban real population, economic data and so on. Establishing a unified and secure data interface is an effective way to share data in various departments.

4. Application of Geo-spatial big data cross-agency sharing Scheme

In order to further explain the practicability and effectiveness of the shared system, taking the 3D building model data as an example, the construction department is the responsible department for the sharing. The 3d building model data standards and sharing behavior standards are constructed. Then describe the standards in details.

4.1. Technical standards

Describe the content, quality and other information. Specifically include:

**Geometric model standard:** geometric model is used to express the spatial position and spatial relationship of the model. Format standards are *.dx, *.dwg, *.3ds, *.obj, *.wrl, *.Open Flight. Its features include geometric shape, color, material, illumination, camera, animation, real-time fast visual rendering.
Texture data standard: the visual effect of the building is represented by the texture given to the model surface. Including static texture image, dynamic texture image. Generally use image format, and the format standard are * .jpg, * .bmp, * .png, * .mpeg, or * .gif.

Attribute data standard: building property information related to specific purposes, format standards include video, audio, text and other multimedia information.

Data switching standard: VRML, is a network 3D data representation standard; X3D, is an extensible 3D description language; GML, is a spatial data expression and switching standard for storing and publishing geographic information of various features; KML, uses XML syntax and format to describe geographic information 3D data language to support network dynamic transmission.

4.2. Shared standards of conduct
The responsibilities of the relevant institutions for data sharing are defined, and the rules and procedures for operation and maintenance, release and use, and quality are established.

The duty of the data sharing department is: the construction department, as the sharing party of the big data of 3D building, should build and share the data when receiving the shared instruction.

Operation and maintenance mechanism: data sharing department is responsible for storing of information, including data version maintenance, storage location maintenance, database system backup, audio and video data backup, backup data validity check, and data recovery.

Publishing and usage standards: including data directory services, basic information query, authentication authorization services, remote callable service interfaces, information publishing system management and maintenance.

Quality control standards: including the mathematical basis, the geometric accuracy, and the topological relations of geometric data, the integrity, resolution ratio, and texture state of texture data. The relationship between attribute data and spatial data, attribute data, texture data should ensure integrity, consistency, easy to store.

In this case, when the shared instruction is received, the 3D building data sharing department provide data timely in compliance with exchange standards. If the data conforms to the technical standard and the shared behavior standard, the sharing is completed. Otherwise, rebuild the data.

4.3. Data quality evaluation
The data quality evaluation system of 3D building model is established as shown in Table 1:

Table 1. 3D building model data quality evaluation system.

| Evaluation objectives          | Factors                                | Principles                                                                 |
|-------------------------------|----------------------------------------|---------------------------------------------------------------------------|
| Data quality of 3D building models | Basic requirements                     | File name, data format, data organization                                  |
|                               | Geometric data accuracy                 | Mathematical basic precision, plane and synthesis precision, elevation precision, set precision, topological relation |
|                               | Texture Data accuracy                   | Texture data, resolution, texture status and building matching             |
|                               | Attribute data accuracy                 | Integrity of attribute fields, matching with actual attribute values, correctness of text and multimedia, integrity and consistency of attribute data |
|                               | Present situation                      | Factor acquisition time, factor update time                                |
|                               | Annex Quality                          | Correct and complete metadata files, correct and complete documentation     |

According to formula (3) and formula (4), the Delphi method is used to evaluate the data quality of open shared 3D building model from six aspects in Table 1.
The quality final evaluation set of 3D building model data is obtained by the first floor operation accordingly:

\[
B = \begin{bmatrix}
0.30 & 0.28 & 0.17 & 0.14 & 0.11
\end{bmatrix}
\]

It shows that the data of open shared 3D building model is of excellent quality and conforms to the goal of system construction.

5. Conclusion

This study analyzed the main problems and solutions in Geo-spatial big data sharing, and put forward a big data cross-agency sharing scheme. From the perspective of manager, user, data connotation, and sharing technology, the international representative management theory "Overall government" was applied. In order to establish a benign operation mechanism and give full play to the best benefits, open sharing suggestions were put forward. The practicability and effectiveness of the shared system were further explained by the case of 3D building model data.

The research showed that establishing a systematic and integrated open sharing mechanism, taking into account the production department, management department, users and technical standards, could promote the innovation of government governance, helped to promote the rational, standardized and efficient use of big data, and ensured the quality and efficiency of big data sharing.

Acknowledgments

This work was funded by the project of Guizhou Province Science and Technology Agency (Qiankehe Jichu [2018] 1120, Qiankehe Zhicheng [2018] 2776, Qiankehe Zhicheng [2019] 2834).

References

[1] Shan Z.G. (2015) Several interpretations of the platform for action on promoting the development of big data. Money China, 28: 30-35.
[2] Zhang X.J. (2016) A Study on the Policies and Regulations of China-US Government Data Openness and Personal Privacy Protection. Intelligence Theory and Practice, 39: 38-43.
[3] Ye R.G. (2017) Security assurance model and standardization framework of geographic information big data, Science Technology and Engineering, 17: 105-111.
[4] Geng Y.D. (2017) Promoting Transformation of Governance Model in China through Big Data. Social Sciences in Xinjiang, 3: 1-6.
[5] Eugene B. (2011) Getting Agencies to Work Together: The Practice and Theory of Managerial Craftsmanship. Peking University Press, Beijing.
[6] Xing H.L., Gao C.Y., Zhai L.L. (2020) Research on Dynamic Evolutionary Game Model of Data Resource Sharing among Big Data Alliance Members: Based on the Perspective of Sharing Positivity. Management Review, 32: 155-165.
[7] Gu M.C, Meng Q.G. (2017) On the Cross-agency Administrative Coordination from the Perspective of Data Integration. Journal of Northeastern University (Social Science), 2: 167-172.
[8] Yang A.M. (2008) Product classification model and its integration method. Science Press, Beijing.