



## Assessment Information

[CoreTrustSeal Requirements 2017–2019](#)

Repository:

Geoscientific Data & Discovery Publishing Center

Website:

<http://geodb.cgs.gov.cn/en>

Certification Date:

20 August 2020

This repository is owned by:

**Development and Research Center of China Geological Survey;  
National Geological Archives of China**

**CoreTrustSeal Board**

W [www.coretrustseal.org](http://www.coretrustseal.org)

E [info@coretrustseal.org](mailto:info@coretrustseal.org)



# Geoscientific Data & Discovery Publishing Center

## Notes Before Completing the Application

*We have read and understood the notes concerning our application submission.*

True

*Reviewer Entry*

**Reviewer 1**

Comments:

**Reviewer 2**

Comments:

## CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

### Background & General Guidance

### Glossary of Terms

## BACKGROUND INFORMATION

### Context

*R0. Please provide context for your repository.*

*Repository Type. Select all relevant types from:*

Domain or subject-based repository, Institutional repository, Publication repository, Other (Please describe below)

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## ***Brief Description of Repository***

Geoscientific Data & Discovery Publishing Center (GDD) is a data publishing repository for geological scientific data mainly from solid and surface Earth Sciences, which involves a wide range of disciplines, including geology, geophysics, geochemistry, hydrology, engineering, environment, remote sensing and technology, being often multi-disciplinary. It is affiliated to the National Geological Archives of China (NGAC) [1], which is an affiliated institution of the Ministry of Natural Resources [2] and China Geological Survey [3] and holds the largest collections of geological data in China and has a total of about 170,000 files of raw, result data formed by geological survey in China for more than a century. The data published and preserved in GDD are from not only geoscientific data in China but also those in other countries worldwide. A methodology of peer-review of datasets for data papers published with Digital Object Identifiers (DOIs) to link data repositories and journals is practised at GDD. The data publishing workflows are charted in the document named Data publishing process listed in the GDD website [4]. All data papers and datasets are fully and openly accessible through the GDD website [5] [6]. The data papers in GDD are also periodically published in Geology in China [7] (Supplementary Vol.), which is a famous Chinese journal in geological field and is sponsored by the China Geological Survey [3]. Bi-monthly issues of the data journal have and will be regularly published in Geology in China since 2017.

[1] <http://www.ngac.org.cn/>

[2] <http://www.mnr.gov.cn/>

[3] <http://en.cgs.gov.cn/>

[4] [http://geodb.ngac.org.cn/en/page/technical\\_documents/documents](http://geodb.ngac.org.cn/en/page/technical_documents/documents)

[5] <http://geodb.ngac.org.cn/en/>

[6] <http://dcc.ngac.org.cn/en/>

[7] <http://geochina.cgs.gov.cn/geochinaen/ch/index.aspx>

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## ***Brief Description of the Repository's Designated Community.***

The data published in GDD include geological map, hydrogeological map, geochemistry database, mineral deposit information database, geological disaster map, etc., which could be used as the basis for fundamental and applied scientific researchers and education. Therefore, the Designated Community includes (1) scientists engaged in basic research and applied scientific research in the field of geology, geophysics, geochemistry, engineering, hydrology, geological disaster and related fields of science, (2) engineers engaged in construction of railway, tunnel and bridge, (3) undergraduate, graduate students and teachers of higher educational schools and institutions. GDD mainly targeted on the scientific organizations, separate researchers, universities, mining company, policy makers and students in the different fields of sciences across the nationwide and around the world.

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## ***Level of Curation Performed. Select all relevant types from:***

D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## ***Comments***

GDD is assisting the author in preparing the data for long-term storage, free distribution and dissemination. All the original datasets submitted by the authors are stored unchanged; only copies of these datasets are edited and then published. All changes made to the original datasets submitted by data authors are recorded in a special log.

In GDD, there are two files required to be submitted, including database (dataset) and corresponding paper used to describe the metadata information of database (dataset) and the contents on data acquisition, processing, quality control and application values. The data publishing process of GDD mainly includes six basis links, i.e. submission of entity data and paper, editor's initial review and peer reviews, entity data and paper publication, permanent preservation, data citation and influence evaluation. In the specific process, the editor conducts an initial review on the paper and database (dataset) submitted by the author, including a review of confidentiality, publicity and error range. Then, at least two peer reviews are

carried out by experts in related fields to review and evaluate the completeness, scientific rigor and application value of the data. Afterwards, the peer review comments are submitted to the responsible editor-in-chief, who will decide whether or not to publish the dataset and paper. All these progresses are done in agreement and consultation with the producer (author) of the data.

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

***Outsource Partners. If applicable, please list them.***

Core parts of the GDD lifecycle are under the control of the applicant. However, GDD has several collaborators to help improve the data publishing system as follows:

Turui (Beijing) Information Technology Co., LTD is our outsourcing partner maintained by contract. According to needs, the partner is responsible for GDD website system design and deployment, as well as metadata database and DOI registration system maintenance.

Renhe Huizhi (Beijing) Information Technology Co., LTD is our another outsourcing partner maintained by contract. The partner utilizes its mature technology on layout file, HTML and XML files design, production and deployment of papers to help improve the expression of papers in GDD.

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

***Other Relevant Information.***

None

*Reviewer Entry*

**Reviewer 1**

Comments:

**Reviewer 2**

Comments:

# ORGANIZATIONAL INFRASTRUCTURE

## I. Mission/Scope

*R1. The repository has an explicit mission to provide access to and preserve data in its domain.*

### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### *Response:*

GDD is based on a creative open data service and positioned on an establishing international data sharing service platform, which will open the data licence and provide findable, accessible, interoperable and reusable (FAIR) data for users to readily download through public channels free of charge [1]. GDD strives to become a long-term archiving, data publishing and open sharing center in the geoscientific field of China. It shall enable the scientists, engineers, policy-makers and general public access to quality-assured scientific data, data service, data technology, as well as data products and information. As a data publishing product, it ensures long-term data stewardship [2] and all data follow agreed-upon national and international data standards and conventions. In order to facilitate access and tracing by users and establish a data center that can operate stably for the long term, a permanent access address is provided to the published dataset.

GDD is affiliated to the NGAC, which is responsible for emphasizing and supporting the sharing of scientific data according to Guiding Opinions on Further Strengthening the Socialization of Geological Data [3] issued by the Ministry of Natural Resources, People's Republic of China (PRC), and has made nearly all catalogue and abstracts of the collected data public [4], launched the National Geological Data Innovation Application Competition [5], and produced a lot of special geological data products [6].

[1] [http://geodb.ngac.org.cn/en/page/sharing\\_policy/aims\\_scope](http://geodb.ngac.org.cn/en/page/sharing_policy/aims_scope)

[2] [http://geodb.ngac.org.cn/en/page/sharing\\_policy/data\\_preserve](http://geodb.ngac.org.cn/en/page/sharing_policy/data_preserve)

[3] [http://www.mnr.gov.cn/dt/ywbb/201901/t20190122\\_2389232.html](http://www.mnr.gov.cn/dt/ywbb/201901/t20190122_2389232.html) (Attachment No.1 is the English version of the beginning and summaries [See ANNEX A])

[4] <http://www.ngac.org.cn/Data/List>

[5] <http://www.ngac.org.cn/competition/>

[6] <http://www.ngac.org.cn/Special/List>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **II. Licenses**

***R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.***

### ***Compliance Level:***

3 – The repository is in the implementation phase

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

##### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

### ***Response:***

In GDD, all data holdings are completely open and are available for scientific and educational use without restriction. It is for this reason that GDD has used the Geoscientific Data Sharing licenses [1]. Referring to the entity databases or datasets, it mainly includes the following four points:

(1) The data submitted by the authors can be freely accessed to the whole society through the Internet system and downloaded by users free of charge, all users can download the data after register on website by email [2].

(2) Users need to label data sources in reference documents or appropriate locations according to the specified citation format when using “data”.

(3) Users of value-added services or those who distribute and disseminate “data” in any form (including through computer

servers) need to sign an agreement with the editorial office of GDD to obtain permission.

(4) Authors who extract some records from “data” to create new data need to follow the 10% quotation principle, that is, the data records extracted from this dataset are less than 10% of the total records of the new dataset, and they need to annotate the data sources for the extracted data records.

[1] [http://geodb.ngac.org.cn/en/page/sharing\\_policy/sharing\\_policy](http://geodb.ngac.org.cn/en/page/sharing_policy/sharing_policy)

[2] <http://www.ngac.org.cn/sso-ui/src/eng/register.html>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

### **III. Continuity of access**

***R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.***

#### ***Compliance Level:***

3 – The repository is in the implementation phase

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

##### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

#### ***Response:***

GDD is responsible for the guaranteed long-term storage of geoscientific data and ensuring access to data holdings.

1. Policy guarantees for data management and sharing in national level. The Chinese government has formulated relevant laws and regulations to promote data sharing, such as the Law of the People’s Republic of China on the Progress of Science and Technology [1] and State Administrative Measures for Scientific Data Management [2]. This stable policy environment is a guarantee for the public repositories sustainable development in China, including GDD.

2. Institute guarantees for the sustainable development of GDD



GDD is one of the core data services provided by NGAC. As a public service organization in China, NGAC is responsible for the sustainable development of databases according to the law Regulation on the Administration of Geological Data, Decree No. 349 of the State Council of PRC [3].

### 3. Funding guarantees for long-term operation of GDD

As a core service provided by the NGAC, GDD is funded by the China Geological Survey (CGS) in the long term. Currently, the collaboration limit for relevant funding support is until 2021. These funding supports will be continually renewed every 3 years according to Chinese funding period policy.

NGAC will be responsible for the sustainable development of the repository and will guarantee the accessibility and long-term availability of the data. As a public service organization in China, NGAC is unlikely to cease operation or substantially change its scope or mission.

### 4. Technology guarantees for the long-term data management and services in GDD

GDD ensures the long-term geoscientific data and data services for the international science community and other stakeholders, and provides management regulations and protection measures for the continuous access to data:

(1) Operational management of GDD. To ensure the continuous operation of the GDD platform, besides the stable cyberinfrastructure in NGAC, the data center has established a corresponding organizational structure and consulting agencies, such as its expert committee [4] and user committee.

(2) Data management of GDD. Centralized management is carried out for all metadata in this platform. Data management must be carried out according to relevant specifications. Systemic updates and expansion of the data are carried out in the data center every year.

(3) Service management of GDD. The data center website provides friendly data co-sharing capabilities and ensures that the Internet system will run continuously 24hours×7days. The services provided by the data center to users can be classified into data enquiry, browsing, downloads, and data information and knowledge dissemination services.

[1] [http://www.gov.cn/flfg/2007-12/29/content\\_847331.htm](http://www.gov.cn/flfg/2007-12/29/content_847331.htm) (Attachment No.2 is the English version of the beginning and summaries [See ANNEX A])

[2] [http://www.gov.cn/zhengce/content/2018-04/02/content\\_5279272.htm](http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm) (Attachment No.3 is the English version of the beginning and summaries [See ANNEX A])

[3] [http://f.mnr.gov.cn/201706/t20170605\\_1509431.html](http://f.mnr.gov.cn/201706/t20170605_1509431.html)

[4] <http://geodb.ngac.org.cn/en/page/expert/neighborhood>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## **IV. Confidentiality/Ethics**

***R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

The data repository of GDD contains only open data that can be freely distributed and used for research and educational purposes. Data with disclosure risks will not be published in the repository. GDD provides open access to all data, and ensures that data can't be used maliciously.

All submitted data are reviewed by the repository's management staff, who will communicate directly with the depositor to ensure that the copyrights, personal privacy, and legal rights of data are protected.

GDD respects the privacy of users and submitters and works to protect all personally identifiable information that we collect.

The scientific data stored by GDD include the maps on natural geological, environmental and mineral geological field survey, which do not involve sensitive geographical coordinates, elevation point and other sensitive spatial data. GDD has an experienced team for processing to examine and response to disclosure risks for all data.

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **V. Organizational infrastructure**

***R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

GDD is maintained by data managing and development department of NGAC, which is commissioned by the Ministry of Natural Resources of the PRC and accepts geological data and provides services to users according to the law Regulation on the Administration of Geological Data, Decree No. 349 of the State Council of PRC [1]. NGAC has 39 full-time employees to complete these tasks, including 11 full-time employees who are working on GDD with other 13 part-time employees. In order to improve the work ability and professional development of employees, several training courses and expert lectures are arranged every year, covering geological data confidential, journal editing and publishing, data management and database construction. In addition, all the employees participate in relevant academic conference and forum for academic exchanges, such as the WDS Asia–Oceania Conference 2019 held on 7–8 May 2019 in Beijing, China, and CODATA 2019, etc.

There is annual funding from the Ministry of Finance of the PRC for the maintenance of fundamental instruments. China Geological Survey (CGS) also provides adequate IT resources and funds for database construction and GDD maintenance projects to NGAC. A new project named Development and Utilization of Library Geological Data is also in place for 2019–2021 (including direct funding of 25 million Chinese Yuan per year), which mainly aims to improve and perfect the geological data management system, develop long-term storage capacity of the data and establish a credible data system to ensure the effectiveness and safety of the database.

[1] [http://www.gov.cn/gongbao/content/2017/content\\_5219175.htm](http://www.gov.cn/gongbao/content/2017/content_5219175.htm) (Attachment No.4 is the English version of the beginning and summaries [See ANNEX A])

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

Accept

## **Reviewer 2**

Comments:  
Accept

## **VI. Expert guidance**

*R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).*

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:  
4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:  
4 – The guideline has been fully implemented in the repository

### ***Response:***

GDD is hosted by the NGAC that has a good in-house experts for advice. The NGAC staff have experienced abilities in data management, data storage, metadata development, data release procedures, document management, data formats, etc. and they are leaders in the field of geoscientific data in China. As a new model of data sharing of the NGAC, GDD is provided by in-house staff engaged in information technology, scientific research and data science fields.

In addition, GDD has an active external science domain advisory committee, called the Data Journal Scientific Committee, which comprises more than 25 experts from related fields of earth science, data science, information technology and geospatial data. Its main responsibilities include executing top-level planning for the GDD's overall development; providing suggestions and consultancy for the data repository's long-term stewardship, development and new technologies as needed; inspecting and assessing the data quality in GDD. GDD has an annual face-to-face meeting and regularly communicates with the editorials by telephone or email.

The Designated Community includes scientists, engineers, graduate students and teachers of higher educational institutions, who are engaged in basic research and applied scientific research in the field of geology, geophysics, hydrology, geochemistry, geological disaster and related fields of science. GDD regularly communicates with the Designated Community through messages or personal meeting (for example, during conference), by the need to resolve the arising questions on data sharing and utilization.

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## DIGITAL OBJECT MANAGEMENT

### VII. Data integrity and authenticity

*R7. The repository guarantees the integrity and authenticity of the data.*

***Compliance Level:***

3 – The repository is in the implementation phase

*Reviewer Entry*

**Reviewer 1**

Comments:

3 – The repository is in the implementation phase

**Reviewer 2**

Comments:

3 – The repository is in the implementation phase

***Response:***

In accordance with the requirements of the Regulation on the Administration of Geological Data, Decree No. 349 of the State Council of PRC [1], domestic geological work must ensure the authenticity and integrity of data, otherwise the work unit will be subject to administrative penalties such as revocation of work permits.

As a data publishing repository, all the data depositors or producers are checked on the basis of the submitted information by depositors. When the data source is ensured, then the entity data, metadata and paper are required to be submitted to GDD as a file folder. Then all the files in the file holder will be processed as a whole by editor and peer experts, and in the processing, any changes to data and metadata are documented and could be traced to process data versions. The original submitted data is documented as v0, and any change occurs, the data will be documented v1, v2, v3, etc.. When the paper, data and metadata are approved by authors and editors, the papers could be published in the journal of Geology in China and the data could be openly shared and permanently preserved in GDD. The stored files are regularly checked by the administrators to ascertain that the files remain unchanged. All the workflow steps assisted to guarantee the integrity and authenticity of the data in the repository.

[1] [http://www.gov.cn/gongbao/content/2017/content\\_5219175.htm](http://www.gov.cn/gongbao/content/2017/content_5219175.htm) (Attachment No.4 is the English version of the beginning and summaries [See ANNEX A])

*Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## VIII. Appraisal

*R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.*

***Compliance Level:***

4 – The guideline has been fully implemented in the repository

*Reviewer Entry*

**Reviewer 1**

Comments:  
4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:  
4 – The guideline has been fully implemented in the repository

***Response:***

According to the Regulations on the Management of Geological Data of Chinese Order No.349 of the State Council [1], NGAC is responsible for the collection, management and service of geological data. In order to better manage and make full use of these data, in recent years, NGAC has created GDD that provides users with important datasets and metadata. To ensure the data compliance, the submission guidelines [2] and the submission policy [3] in GDD indicate the scope and selection criteria of submitted data. For example, each author shall confirm that he/she owns the intellectual property rights to the database (dataset).

GDD attempts to implement ISO 19115-1-2014 [4] metadata standards as much as possible. In combination with the characteristics of the production, storage and service of the geoscientific data, the GDD metadata standard for geoscientific data [5] was formulated. GDD has strict control over data quality and ensures that all data are available and

valuable in the current situation. We check if files related to geological catalogue are all archived in place. Data quality is not only provided by data owners, but also reviewed by experts. The quality control of data receiving and publishing includes self-evaluation of authors in papers, peer review and editorial approval. The GDD web pages [6] provide data depositors with the information and forms necessary to lodge their digital data with the data center. In order to ensure the correctness of the data, users can provide us with comments or suggestions by phone or email for data conditions.

[1] [http://www.gov.cn/gongbao/content/2017/content\\_5219175.htm](http://www.gov.cn/gongbao/content/2017/content_5219175.htm) (Attachment No.4 is the English version of the beginning and summaries [See ANNEX A])

[2] <http://geodb.ngac.org.cn/en/page/submission/instructions>

[3] [http://geodb.ngac.org.cn/en/page/sharing\\_policy/author\\_submission\\_policy](http://geodb.ngac.org.cn/en/page/sharing_policy/author_submission_policy)

[4] <http://www.dcc.ac.uk/resources/metadata-standards/iso-19115>

[5] [http://geodb.ngac.org.cn/en/page/technical\\_documents/documents](http://geodb.ngac.org.cn/en/page/technical_documents/documents)

[6] <http://geodb.ngac.org.cn/en>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **IX. Documented storage procedures**

***R9. The repository applies documented processes and procedures in managing archival storage of the data.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

Our data storage policy has been documented in an internal file, which is used as our guide for storing data.

#### 1. Data storage process

The data submitted by data producer (author) include entity data and corresponding paper, according to the requirements of data publishing for temporary storage [1]. GDD carries out technical review and peer review on the data to ensure compliance and quality, referring to the relevant submission guidelines [2] and document named Data publishing process [3]. 1) Database (dataset) that has passed the review will be considered as the initial data. When the initial database (dataset) is stored in GDD, the data administrator needs to login into the GDD system and input the relevant metadata information through the interface. The metadata information is also stored in the metadata database. The initial database (dataset) and its copies are placed in three different kinds of media (Blu-ray disc, tape and hard disk) for the long time preservation without any changes, while another copy is published to the shared file server using for external viewing and downloading. A unique DOI is registered for each database (dataset). 2) The paper for describing the database (dataset) that has not been reviewed could be published in preprint system if the author agrees, referring to Data publishing process [3]. The data administrator also needs to login into the GDD system and input the metadata information through the interface. Once the paper is reviewed and accepted by GDD, it will be made into the XML and PDF files and released by the administrator for viewing and downloading. The electronic files are also stored in three different kinds of media for the long time preservation. A jump link access between the paper and entity data webpage is provided when browsing.

#### 2. Data storage security regulation

When submitting the data, the data producer (author) is always reminded that the data will be published and shared. If the author disagrees, the data uploading and publishing cannot be performed to ensure the safety of data reference. For users, GDD provides the unified user authentication management system and audits the user identity who downloads the dataset to ensure the safety of data use. Only the administrator could perform data editing when logging into the intranet to prevent the data from being tampered with.

The data storage and backup systems are deployed in the core computer room of the China geological survey (CGS), and 24-hour security monitoring and inspection guarantee is also carried out. Data security is reviewed and maintained by specific personnel to ensure the security of data storage. When a new database (dataset) is being released in GDD, it is stored in both the data storage system and the backup system to ensure the response ability of the system and data to disasters. Initial data and all copies are regularly checked to verify that they have not been altered or corrupted. For short-term disasters, the restore and backup is used. Data restoring includes two type of data processing: database and file. The former adopts regular backup files and can be quickly restored using database management tools, while the latter can be directly copied from the backup media for restoration.

[1] [http://geodb.ngac.org.cn/en/page/sharing\\_policy/data\\_preserve](http://geodb.ngac.org.cn/en/page/sharing_policy/data_preserve)

[2] <http://geodb.ngac.org.cn/en/page/submission/instructions>

[3] [http://geodb.ngac.org.cn/en/page/technical\\_documents/documents](http://geodb.ngac.org.cn/en/page/technical_documents/documents)

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

It is encouraged to provide a public version of the internal documentation by the time of next renewal.

##### **Reviewer 2**

Comments:

Accept



## X. Preservation plan

***R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.***

### ***Compliance Level:***

3 – The repository is in the implementation phase

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

##### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

### ***Response:***

GDD is a geological scientific data repository sponsored by NGAC under administration of the Ministry of Natural Resources, P.R.C, which is responsible for the collection, custody, long-term preservation and service of national geological data, according to Implementation Measures of Regulation on the Management of Geological Data [1]. The preservation plan of GDD relies heavily on the policies of NGAC [1]. Therefore, GDD formulated a Data Submitting Processing Specification (see attachment No.6 [ANNEX A]) detailing how to ensure long-term data preservation. It requires that the submitted data should be in common format such as doc, docx, pdf, xml, html, jpg, png, mapgis, arcgis, geomap, zip, url, xls, etc. The physical data together with database (dataset) abstract and metadata information are published at the Geoscientific Data & Discovery Publishing System [2] for open access, linked with the corresponding paper. The Geoscientific Data & Discovery Publishing System [2] uses modern technologies for data providing long-term preservation and permanent free services.

Data publishers (depositors) should submit entity data, metadata and paper together for storage, ensure that datasets have clear ownership of intellectual property rights, and can fully handle the open sharing of datasets and other issues. The dataset shall be formed under the conditions of strict data production and processing methods, and the dataset publishers shall adopt effective quality control measures to ensure the authenticity of the dataset. All received datasets are reviewed by experts, registered in a specialized database, and placed in the appropriate database.

GDD has the rights of reproduction, distribution, network dissemination, translation, compilation and other rights of manuscripts and datasets, and signs relevant agreements with authors. Each dataset is assigned a unique DOI, and the identifier will track citations to scientific data. DOI will persist for the lifetime of the original deposit as part of an archive. Data publishing supported by DOI technology can effectively protect the intellectual property rights of data producers and evaluate the effectiveness of data citation.

GDD is a data repository that can operate stably for a long time, even if the physical address of physical data changes. It is also possible to keep the logical address of the data unchanged by means of re-analysis, so that users can access and trace the data for a long time. With the continuous updating of GIS software platform, the dataset needs to be converted to another format or upgraded version, which will be more convenient to use, and then the dataset is modified, and its copies are made for backup storage. Users are given information about the conversion of dataset.

GDD has set up a data preserving and processing team responsible for data compilation, management, processing, storage, etc. The documentation specifying requirements relevant to preservation as well as measures to ensure them must be clearly formulated before published on GDD website, in order to make the data depositors and designate community better understand the responsibility.

[1] [http://f.mnr.gov.cn/201702/t20170206\\_1436345.html](http://f.mnr.gov.cn/201702/t20170206_1436345.html) (Attachment No.5 is the English version of the beginning and summaries [See ANNEX A])

[2] <http://dcc.ngac.org.cn/en>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

By the time of next renewal, the measures described in the final paragraph of the response should be in place.

##### **Reviewer 2**

Comments:

Accept

## **XI. Data quality**

*R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.*

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

## ***Response:***

GDD has appropriate expertise to address technical data and metadata quality, and ensure that sufficient information is available for end-users. GDD requires that all datasets have comprehensive discovery metadata, according to document named DOI registration metadata standard for geoscientific data [1]. The datasets are formed and stored in the center with available formats of accdb, xlsx, jpg, pdf, doc and wp, wl, wt files of MapGIS and ArcGIS. We will organize experts to conduct data and metadata review, and then record the metadata in regular format, and check if the metadata conforms to the dataset. Then, the data will be organized and well classified so that the end-users can easily use. The data query system binds metadata to data for publication. Users can search the title or related keywords of data in the portal. The website will provide metadata information and linked data descriptors of relevant content. During data sharing and application, experts and users will evaluate the data, which helps to identify errors in the dataset and improve data quality. The staff all have professional skills in programming and data management. All staff members have an education background of geoscience, computer science and technology or data science, and most has more than 10 years working experience in data management. Moreover, there is also a team of experts [2] to guide our work, ensure the correctness of the direction of work, and ensure reliable data quality.

[1] [http://geodb.ngac.org.cn/en/page/technical\\_documents/documents](http://geodb.ngac.org.cn/en/page/technical_documents/documents)

[2] <http://geodb.ngac.org.cn/en/page/expert/neighborhood>

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## **XII. Workflows**

***R12. Archiving takes place according to defined workflows from ingest to dissemination.***

### ***Compliance Level:***

3 – The repository is in the implementation phase

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:  
3 – The repository is in the implementation phase

#### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

## ***Response:***

The main work flow starts from the data owner to the repository, and the query system, referring to document named Data publishing process [1].

(1) Data submission. When an owner believes that their data has reached production level, they ask GDD to publish it.

The data owner generates the data and we will inform them about the data publishing format, metadata content information and other considerations required by the data center. Data requirements include:

- a. Independent intellectual property rights. Data publishing applicants ensure that the datasets have clear ownership of intellectual property rights and can handle the open sharing of datasets with full authority.
- b. High quality data. Datasets described in papers should be formed under strict data production and processing conditions. Data publishing applicants should take effective quality control measures to ensure the authenticity of datasets.

(2) Data review. GDD will check whether the metadata and paper provided corresponds to the dataset correctly, and then organize experts to review the entity data and paper. If qualified, the data information will be converted to the format of the DOI system identification; if not, the reason will be explained to the applicant and returned for modification.

(3) Data storage. After the above work is completed, the repository allocates storage space, pushes the directory to the database, and integrates the archived data into the query system. Metadata and related paper are also archived with entity data.

(4) Data service and sharing. GDD requires users to register and login through the user management system [2] for data downloading, and receiving data usage instructions and citation formats for data reuse. At the same time, data editors will inform data applicants of the publication results, and users can contact GDD through the contact method in the website [3].

GDD has set up future process improvement plans, such as planning more peer experts to participate in the review process, hiring responsible editors, and strengthening the use of technical means to participate in data audit. The documentation specifying all archival workflows must also be clearly formulated before published in the GDD website.

[1] [http://geodb.ngac.org.cn/en/page/technical\\_documents/documents](http://geodb.ngac.org.cn/en/page/technical_documents/documents)

[2] <http://www.ngac.org.cn/sso-ui/src/eng/register.html?systemKey=2c918084711cd94e017153f0389003c2>

[3] <http://geodb.ngac.org.cn/en/page/contactus>

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **XIII. Data discovery and identification**

***R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### ***Response:***

To facilitate users' discovery of data, GDD provides search, abstract, keywords, download and access capabilities for all the data holdings [1]. The specific data discovery mechanisms supported by GDD are created as follows:

1. GDD provides one-stop search function on the data portal homepage to search for data [2], and users can input a keyword to find the targeted dataset.
2. GDD provides advanced search by titles, keywords, authors, abstracts and DOIs.
3. Users can also browse data list through data catalogue and Downloads in the navigation bar on the homepage of the GDD website [2].

GDD publishes data in two forms simultaneously, i.e. online dataset [2] and paper [3], and applies for DOI registration for entity data and papers. Published entity data are registered with a DOI in the Tsinghua University Library (10.23650) [4], which is a registered member of DataCite and an authorized DOI Registration Agency (RA). The papers are registered with China DOI (10.12029), which is a DOI RA officially authorized. All data in the GDD have undergone a rigorous process to ensure their validity and integrity before being assigned a DOI. And DOIs are assigned to every dataset released in GDD. The data citation and DOI are presented to users in the dataset landing pages within the "citation format" [1].

[1] <http://dcc.ngac.org.cn/en//geologicalData/details/doi/10.23650/data.D.2019.P24>

[2] <http://dcc.ngac.org.cn/en>

[3] <http://geodb.ngac.org.cn/en>

[4] <http://lib.tsinghua.edu.cn/dra/>

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## XIV. Data reuse

*R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.*

### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

**Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

### *Response:*

GDD aims to maintain the integrity, availability and reusability of its content in the long term. As a data publishing journal, GDD requires the data submitters to provide sufficient metadata and additional data description paper. The metadata standards are established by referencing the relevant metadata national standards and industry standards of China, intergrated with the characteristics of the production, storage and service of GDD. For each dataset, its full description is given in the corresponded paper. For example, for the catalogue of mineral deposit information, such attributions are indicated: mineral deposit's number, tectonic setting, litho-stratigraphic unit, intrusive rock, volcanicrock, metamorphic rock, geological structure feature, alteration of wall rock, dating methods, isotopic age, sampling location, sample No., analytical method of stable isotopes, and stable isotope geochemistry. And guidance for each data utility is also presented in every corresponded paper, which includes data source, data structure, processing steps, data description and data quality control, providing sufficient supporting information alongside the entity data, so that users of the repository can access these resources when re-using the data. By providing this additional supporting information in the related paper, GDD ensures that users can make an informed assessment of that whether individual dataset is fit for their specific purpose, and therefore they can make appropriate use of it.

#### *Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## TECHNOLOGY

### XV. Technical infrastructure

*R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.*

#### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

**Reviewer 1**

Comments:  
4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:  
4 – The guideline has been fully implemented in the repository

#### *Response:*

GDD is running based on NGAC's infrastructure, which has the amount of near-line storage of 900TB, 1.2PB backup capacity, more than 200 physical servers to support the needs of data management, processing and integration, as well as the need for Internet data services. One fourth of the infrastructure is reserved for GDD system.

With all the databases stored in designated professional organizations, GDD contains more than 100 sets of desktop client software licenses for data processing, and provides over 40 GIS server/portal licenses for Internet service clusters. All kinds of server clusters supporting Internet services are all highly available, including directory services, map services and document services.

All data servers are protected against intrusion by a reverse proxy in the DM Zone between two firewalls to ensure data security. All servers run under the operating supervision system. If there is a problem, the administrator can be alerted in

real time via SMS sent by the system.

*Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## XVI. Security

*R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.*

### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

*Reviewer Entry*

**Reviewer 1**

Comments:  
4 – The guideline has been fully implemented in the repository

**Reviewer 2**

Comments:  
4 – The guideline has been fully implemented in the repository

### *Response:*

GDD's IT infrastructure is deployed in the core computer room of the China geological survey, with corresponding power supply, network links, fire protection and 24-hour security monitoring and inspection guarantee. All the constructions including the corresponding computer room, information system and information security level protection management method, conform to the network security law of the People's Republic of China [1], and at the same time conform to the relevant provisions of the relevant industry norms and standards.

In the IT infrastructure security management work, there is a management mechanism for special personnel to manage, audit and maintain, to ensure the safety of the system in use, and at the same time, all the systems close the non-use services and ports, and regularly carry out system vulnerability scanning and repair. Firewall, user management system, tamper proof system and audit system are deployed in the network environment. The storage system and backup system are used to realize the full coverage backup of the system and data, and the off-line off-site backup is completed by using multiple media to ensure the multi-dimensional response ability of the system and data to disasters. In case, all data are



backed up at a different physical data center, with five sets of backups saved in three different kind of media: one set of Blu-ray disc, two sets of tapes and two sets of hard disks. In GDD, previous versions of datasets stay in database with date tags, and only that with the latest update is public to users.

[1] [http://www.cac.gov.cn/2016-11/07/c\\_1119867116\\_3.htm](http://www.cac.gov.cn/2016-11/07/c_1119867116_3.htm)

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## **APPLICANT FEEDBACK**

### **Comments/feedback**

*These requirements are not seen as final, and we value your input to improve the core certification procedure. To this end, please leave any comments you wish to make on both the quality of the Catalogue and its relevance to your organization, as well as any other related thoughts.*

#### ***Response:***

The catalogue of the core certification procedure is intergrated and comprehensive, but I still have a suggestion that the core certification procedure better uses different language versions, and consider the national conditions of different countries.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

##### **Reviewer 2**

Comments:

# ANNEX A

## ATTACHMENT

### No.1.

## Guiding Opinions on Further Strengthening the Socialization of Geological Data

Recently, the Ministry of Natural Resources issued the “Guiding Opinions on Further Strengthening the Socialization of Geological Data” (referred to as “Opinions”), and put forward clear requirements for the open sharing of geological data information, the construction of big data system on national geological data and the construction of geological data service capabilities.

Ministry of Natural Resources of the People’s Republic of China

22 January 2019

**Summaries:** The “Opinions” pointed out that it was necessary to further promote the open sharing of geological data information. The first opinion is to fully disclose the geological data catalogue, which requires that the abstract and catalog information of the newly transferred geological data should be compiled and published to the public within 90 days from the date of submission. The second is to increase the supply of public information resources and further improve the integrity system for geological data exchange. The third is to develop more digital products. The fourth is to improve and perfect the socialized service system. The “Opinions” also proposed to carry out the construction of a national big data system on geological data, and required that the construction of geological data service ability should be strengthened, including comprehensively strengthening the construction of geological collection facilities, continuously optimizing oil and gas industry’s archives for geological data disclosure, innovating and improving the service system and mechanism, and promoting the exchange and public utilization of geological data such as urban geology data, engineering geology data and scientific research data.

### No.2.

## Decree of the President of the People’s Republic of China

No. 82

# ANNEX A

“Law of the People's Republic of China on Progress of Science and Technology” has been passed by the 31st meeting by the 10th Standing Committee of the National People's Congress on 29 December 2017 after revision. The revised “Law of the People's Republic of China on Progress of Science and Technology” is now announced and will be effective from 1 July 2008 onwards.

President of the People's Republic of China Hu Jintao

29 December 2007

**Summaries:** Article 65 in the decree is the relevant content of policy guarantees for data management and sharing in national level.

**Article 65** The administrative department of science and technology of the State Council shall, together with the relevant competent departments of the State Council, establish information system of scientific and technological research bases, scientific instruments and equipment, scientific and technological literature, scientific and technological data, natural resources of science and technology, scientific and technological popularization resources, and other kinds of resources. The established system shall promptly announce the distribution and use of scientific and technological resources to the society.

The management unit of science and technology resources shall announce to the society the data sharing policies and use status of the managed science and technology resources, and arrange the use according to the policies; however, if laws and administrative regulations require confidentiality, follow their provisions.

The management unit of scientific and technological resources shall not infringe the intellectual property rights of users, and shall determine the charging standards in accordance with the relevant provisions of the state. Other rights and obligations between the management unit and the user are agreed by both parties.

## No.3.

### General Office of the State Council's Notification on the Issuance of Administrative Measures for Scientific Data Management

No. 17 [2018] of the General Office of the State Council

# ANNEX A

To people's governments of various provinces, autonomous regions, and direct-controlled municipalities, and various departments directly under the State Council,

The "Administrative Measures for Scientific Data Management" has been approved by the State Council and is hereby issued to your department. Please implement these measures carefully.

General Office of the State Council of the People's Republic of China

17 March 2018

**Summaries:** The General Office of the State Council issued the "Administrative Measures for Scientific Data Management" (hereinafter referred to as "Measures") to further strengthen and standardize the management of scientific data, ensure the safety of scientific data, improve the level of openness and sharing, and better serve the national scientific and technological innovation, economic and social development. The "Measures" clarify the general principles, main responsibilities, data collection and storage, sharing and utilization, confidentiality of China's scientific data management, focusing on specific management measures from five aspects. The first is to clarify the division of responsibilities of all parties, strengthen the responsibility of the main body of the legal entity, and clarify the responsibilities of the competent department, reflecting "who owns, who is responsible", "who is open, who benefits". Secondly, in accordance with the principle of "classified management to ensure safety and controllability", the competent departments and legal entities determine the confidentiality level and open conditions of scientific data according to law, and strengthen the supervision of scientific data sharing and utilization. The third is to strengthen the protection of intellectual property rights, regulate the behavior of users and producers of scientific data, and demonstrate respect for the intellectual property rights of scientific data. The fourth is to require scientific data generated by science and technology planning projects to be compulsorily submitted, and to carry out standardized management and long-term preservation through the scientific data center, and strengthen data accumulation and open sharing. Fifth, it is proposed that the legal entity should establish an incentive mechanism in terms of job setting, performance income, and job title evaluation to strengthen the scientific data management capability.

No.4.

# ANNEX A

## Regulation on the Administration of Geological Data (2017 revision)

No.349

Announcement of Decree No. 349 of the State Council of the People's Republic of China on March 19, 2002.

First revision according to the "Decision of the State Council on Amending Some Administrative Regulations" on February 6, 2016.

Second revision of the "Decision of the State Council on Amending and Abolishing Some Administrative Regulations" on March 1, 2017.

State Council of the People's Republic of China

1 March 2017

**Summaries:** The relevant contents in the "Regulation" giving evidence on to the application are provided in the following text.

**Article 1** The "Regulation" is formulated for the purposes of strengthening the administration of geological data, bringing into full play of it, and protecting the lawful rights and interests of submitters of geological data.

**Article 2** The "Regulation" applies to the submission, storage and use of geological data.

The term "geological data" referred to in the "Regulation" means original geological data and final reports in the form of text, diagrams, audio-visual products and electromagnetic medium and physical geological data such as drilling core, various specimens, polished thin sections and samples produced in the course of geological work.

**Article 3** The competent geology and mineral resources department of the State Council shall be responsible for the supervision and administration of the submission, storage and use of geological data throughout the country.

The competent provincial geology and mineral resources departments, autonomous regions and municipalities directly under the Central Government shall be responsible for the supervision and administration of the submission, storage and use of geological data within their respective administrative areas.

# ANNEX A

**Article 4** Geological archives of the competent geology and mineral resources department of the State Council and of the provincial competent geology and mineral resources departments, autonomous regions and municipalities directly under the Central Government (hereinafter referred to as geological archives), as well as geological data storage units entrusted by the competent geology and mineral resources department of the State Council (hereinafter referred to as geological data storage units), shall undertake the storage of and provision for use of geological data.

**Article 12** The geological data submitted shall be in conformity with the relevant provisions of the competent geology and mineral resources department of the State Council and the relevant technological standards set by provinces' departments.

No unit and individual may forge geological data or practise fraud when submitting geological data.

**Article 14** The geological archives and the geological data storage units shall establish systems for rearrangement and storage of geological data, be equipped with necessary storage, protective and security facilities and be staffed with professional technicians, so as to guarantee the integrity and security of geological data.

## No.5.

### Measures for Implementation of the Regulation on the Administration of Geological Data

Announcement No. 16 of the Ministry of Land and Resources of the People's Republic of China on January 3, 2003.

Amended according to the First Ministry Meeting "*the Ministry of Land and Resources' Decision on Amending and Abolishing Some Regulations*" of the Ministry of Land and Resources on January 5, 2016.

Ministry of Land and Resources of the People's Republic of China

5 January 2016

**Summaries:** The relevant contents in the "Implementation Measures of Regulation on the Management of Geological Data" giving evidence on to the application are provided in the following text.

# ANNEX A

**Article 1** These measures are formulated in accordance with the Regulations on Administration of Geological Data.

**Article 4** The National Geological Archives of China and the geological data storage units and geological archives of the competent land and resources administration departments of provinces, autonomous regions and municipalities (hereinafter referred to as geological archives) shall store and manage the submitted geological data and provide geological data service to the public in accordance with related laws and regulations, including performance of the following duties:

- (1) to receive and examine submitted geological data as entrusted by the competent land and resources administration departments;
- (2) to establish complete system for storage and use of archive geological data and to sort and store geological data in accordance with related regulations;
- (3) to establish and maintain geological data information service system;
- (4) to conduct comprehensive research on geological data and provide geological data service to the public in accordance with related laws and regulations;
- (5) Other duties specified by related laws and regulations.

**Article 12** The Ministry of Land and Resources entrusts the National Geological Archive of China or geological data storage units to undertake receiving, examination and acceptance of geological data.

The competent land and resources administration departments of provinces, autonomous regions or municipalities may entrust geological archives to undertake receiving, examination and acceptance of geological data.

**Article 17** Geological data archives and storage units shall have the following conditions:

- (1) hardware facilities conforming to class A archive standards;
- (2) staffed with required professional personnel;
- (3) with complete management system for receiving, sorting, storage, confidentiality and use of geological data;
- (4) capable to establish geological data information system and provide geological data network service to the public;
- (5) with funding guaranteed for management of geological data.

**Article 18** Geological data archives and storage units shall use modern information processing technology to improve the level of geological data processing and storage, establish geological data information

# ANNEX A

service network system, publish catalogue of geological data, conduct comprehensive research on geological data, provide basis for government decision making and provide public service.

## No.6.

### Data Submitting Processing Specification

#### 1. Mission and Vision

Geoscientific Data & Discovery Publishing Center (GDD) aims to obtain geoscientific data with high-quality assurance in order to achieve long-term data management and preservation.

On the basis of innovative open data services, GDD is positioned to establish a geoscientific data sharing platform for internationalized data services, which opens the data license and provides relatively integrated and authentic data for users to readily download through public channels free of charge.

#### 2. Data Sharing Policy

The data published by GDD (in Chinese and English) include metadata (in Chinese and English), entity data (in Chinese or English) and the papers published in “Geoscientific Data & Discovery Album” (in Chinese and English) of the journal *Geology in China*. The sharing policy is as follows.

(1) The geological data submitted by authors are open to the public for downloading by end users free of charge on the internet.

(2) When the data are cited, the end users shall indicate the sources of data in the reference or at other appropriate locations in the reference format.

(3) Users of the value-added service and those who distribute and communicate the data in any form (including through computer servers) shall sign an agreement in writing with the editorial office of GDD (in Chinese and English) to obtain a license.

(4) Authors who extract a part of the data to create new data shall follow the principle of 10% citation, i.e., the data record extracted from the dataset shall be less than 10% of the total records of the new dataset, and the source of extracted data records shall be indicated.

#### 3. Service for Authors

(1) Provide data intellectual property protection for authors (data producers); (2) provide more innovative



# ANNEX A

outputs to authors (data producers); (3) improve internationalized degree of data retrieval for authors (data producers); (4) ensuring the network's origin rights for authors' (data producers') data.

## **4. Data Retention Policy**

Geoscientific Data & Discovery Center (GDD) aims to maintain the integrity, accessibility and reusability of geoscientific data for a long time. (1) The author needs to provide sufficient metadata (data producers, departments, fund projects, etc.) and data background information so that users can understand and use the data; (2) the data submitted by the author needs to include the dataset and corresponding papers. The paper is a scientific and objective description of the dataset to ensure that the user uses the data to be well documented; (3) the data saving format supports doc, docx, pdf, xml, html, jpg, png, mapgis, arcgis, geomap, zip, url, xls, etc.; (4) GDD owns the ownership of data, is responsible for the collection, management, preservation, acquisition, dissemination of scientific data, and is responsible for protecting the intellectual property rights of data producers and protecting the signing rights of data producers, authors and the collaborators, etc.

## **5. Peer Review Policy**

In order to guarantee the quality of scientific data publishing and the reliability of data sharing and reuse, each scientific database (dataset) needs to go through peer review process.

(1) Select peer-review experts who have a higher academic level in the related field and are familiar with the content to be reviewed, and avoid peer-reviewed experts who have stakeholders with the database (dataset) authors.

(2) Three to five experts are generally selected to conduct independent reviews, and the opinions of peer experts are collectively submitted into the working group of the editorial department. The editorial working group reports the collated comments to the responsible editor (or editor). The responsible editor (or editor) finally decides whether to publish. The recommendations and results of peer reviews and the conclusions of whether the database (dataset) is published are conveyed to the submitting author.

(3) The required contents of peer review are shown in the following attached table.

(4) Peer expert review time should not exceed 20 days.

## **6. Submission Policy for Authors**

The submission requirements for authors of Geoscientific Data Publishing & Discovering Center (GDD) include authors' undertakings and submitting materials.

# ANNEX A

## *Authors' undertaking*

The authors who are interested in contributing data to GDD shall undertake the followings.

(1) Each author shall confirm that he/she owns the intellectual property rights to the database (dataset); all the authors of the contributions should be the creators of the contribution and are responsible for the correctness, authenticity and reliability of the submitted content. The authorship and order of authors will not be changed once the publishing agreement is signed. The authors agree to share with the publisher the following copyrights: editorial right, right of reproduction in different media, rights of data distribution, network communication, multilingual translation and print within registered data publication scopes, as well as the right to transfer the above said property rights. The right to publish the database (dataset) is exclusively owned by the data publisher.

(2) The author shall guarantee that the ownership of data is clear and the data owner entity is clear and free of disputes.

(3) Guarantee the authenticity and originality of the scientific data. The author shall guarantee that the scientific data is true and not false, the contribution is the creative work achievements of the author, and references have been listed for all citations of arguments, data or results of other authors.

(4) The contribution is not made to more than one publisher. One database (dataset), except for the updated version, can be published only once.

(5) Contribution has no confidentiality requirement on the publication of the contribution. In case that confidentiality requirement is involved, the contribution shall be reviewed and approved by the author affiliation for publication by affixing an official seal on it.

## *Submitting materials*

The author may choose to publish data paper(s) and entity data simultaneously or just publish the latter.

The author shall submit both entity data and metadata, or all of the entity data, metadata and data paper(s).

### (1) Metadata

Metadata is the key information for DOI registration of the geoscientific data publishing. The author shall fully complete the metadata table item by item.

### (2) Data paper

A data paper is an expository and argumentative article that describes the dataset (database) and argues

## ANNEX A

for the creativity and reliability of the dataset. Hence, a data paper has the features of both expository prose and a research paper.

### (3) Entity data

Entity data form the core content of data publishing. For GDD, data include scientific data in all earth science fields and show obvious diversity. Entity data are the essence of GDD and the basis for geological innovation.