



Assessment Information

[CoreTrustSeal Requirements 2020–2023](#)

Repository: Geo Big Data Open Platform
Website: <https://data.kigam.re.kr/?lang=en>
Requirements version: CoreTrustSeal Requirements 2023-2025

This repository is owned by: **Korea Institute of Geoscience and Mineral Resources**
Created at: 04 December 2023, 04:50

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background Information

Re3data Identifier

Please fill you Re3data identifier from the website: <https://www.re3data.org/>

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

<https://www.re3data.org/repository/r3d100014025>

Links:

Repository type

Please select your repository type.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

- Specialist repository

Links:

Overview

Provide a short overview of key characteristics of the repository, reflecting the repository type selected. This should include information about the scope and size of data collections, data types and formats. Further contextual information may also be added.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

The Geo Big Data Open Platform of the Korea Institute of Geoscience & Mineral Resources is a data-based repository that allows anyone to easily access the latest geoscience information scattered in Korea. It was established for the purpose of quickly organizing and providing domestic and foreign geoscience research information pouring out of a super-gap society to utilize the solution of national social problems and create an open science research ecosystem in the geoscience field.

Links:

Designated Community

A clear definition of the Designated Community demonstrates that the applicant understands the scope, knowledge base, and methodologies—including preferred software/formats—of the group(s) of users at whom the curation and preservation measures are primarily targeted. The definition should be specific so that reviewers can assess whether that community is being served in the responses to other requirements.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Geo Big Data Open Platform

Response:

- KIGAM, the operator of the Geo Big Data Open Platform, is dedicated to the assetization and big dataization of research data by establishing a comprehensive national system for managing and preserving geoscience research data. It aims to foster a research ecosystem that emphasizes data utilization and promotes data-driven R&D through the implementation of a sharing and utilization system. By leveraging research data, it addresses national social challenges related to land balance development, national safety, and livelihood issues, while also contributing to the emergence of new industries and the advancement of technologies through the integration of ICT technologies such as data-based AI and IoT. Moreover, the platform offers customized geoscience information tailored to the needs of users.
- Designated communities in the Geo Big Data Open Platform continuously share data and information.
- KIGAM researchers and users who have obtained permission to register datasets can register datasets.
- Users consist of government (affiliated) organizations, local governments, public institutions, research institutes, schools, industrial companies, and general organizations, and are used for research or thesis purposes, and for business purposes of local governments, public institutions, and industrial companies.

Links:

Levels of Curation

Please fill you level(s) of curation.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

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

Links:

Levels of Curation - explanation

Please add the description for your Level(s) of Curation.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

- As of June 12, 2022, the Geo Big Data Open Platform provides 132,125 items including Investigation■Exploration data, Sample■Analysis data, Geoscience Thematic Map, and Literature & Data.
- The Geo Big Data Open Platform manages research data in stages according to the data life cycle, ensuring data integrity and reliability through data change version control.
- The Geo Big Data Open Platform has backup and recovery procedures in place to prevent and protect research data from loss and damage. It ensures data integrity and accessibility during the backup process.
- The data in the Geo Big Data Open Platform is evaluated for preservation level based on the digital asset preservation framework, which includes four levels: Level 1 (data protection), Level 2 (data recognition), Level 3 (data monitoring), and Level 4 (data recovery).

Links:

Cooperation and outsourcing to third parties, partners and host organisations

Please describe any cooperation and outsourcing to third parties, partners and host organisations.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

Geo Big Data Open Platform

"KIGAM is the parent organization of the Geo Big Data Open Platform, a national research data platform specializing in geoscience. As an affiliate of the National Science Foundation (NST), which operates under the Ministry of Science and ICT, KIGAM plays a vital role in managing and preserving geoscience research data at a national level. The Geo Big Data Open Platform is funded through the government budget under the Science and Technology Research Organization Act." (See section R5)

KIGAM has established MOUs with various organizations in the field of geoscience, both domestically and foreignly, fostering collaborations and knowledge exchange. Currently, KIGAM has signed MOUs with a total of 44 related organizations, including national research data platforms such as GEO DATA and DataON in Korea, as well as organizations from 24 countries spanning Europe, Asia, America, Oceania, and Africa. Notable international organizations such as UNESCO and CCOP are also among KIGAM's MOU partners. Furthermore, KIGAM actively participates in international organizations and programs such as ASEAN, CCOP, CTBTO, JCDP, JCL, IEA-GIA, IODP, and SCUFN-IOC, further enhancing its engagement in global knowledge exchange. (See section R6)

Links:

Applicants renewing their CoreTrustSeal certification: summary of significant changes since last application.

Please fill this field when you are renewing your CoreTrustSeal Certification.

This field can be marked with not applicable (N.A.) if you are acquiring a CoreTrustSeal certificate for the first time.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

N.A.

Links:

Organisational Infrastructure

R1 Mission & Scope (R01)

R01. The repository has an explicit mission to provide access to and preserve digital objects.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Korea Institute of Geoscience and Mineral Resources (KIGAM) has recently unified its geoscience information services, which were previously operated separately. In addition to providing research data generated during the research process, the institute has established the 'Geo Big Data Open Platform (<https://data.kigam.re.kr/?lang=en>)'. This platform facilitates the sharing of research data among researchers and enables collaborative research by allowing them to register their data on the platform.

The Geo Big Data Open Platform is an electronic system designed to handle tasks related to the opening and utilization of research data (<https://data.kigam.re.kr/manage/data/rule#item-1>). It serves as an open access library, providing accessible and reliable geoscience information for public use. Additionally, it functions as a global K-Geoscience data hub, facilitating easy access to geoscience data worldwide (<https://data.kigam.re.kr/about/info?lang=en>). The Geo Big Data Open Platform serves as a national data center in the field of geoscience, leading Open Science initiatives. Its primary role is to create an environment for research cooperation and collaboration (<https://data.kigam.re.kr/about/info?lang=en>).

The expected effects of operating the Geo Big Data Open Platform are as follows (<https://data.kigam.re.kr/about/info?lang=en>).

- Assetization of research data by managing geoscience research data at the national level and establishing a preservation system
- Creating a research environment for data utilization and promoting data-based R&D by establishing a sharing and utilization system
- Responding to national social issues such as balanced regional development, national security, and livelihood issues
- Creating new industries and developing new technologies through convergence with ICT technologies such as data-based AI and IoT
- Providing customized information by intellectualization of demand side geoscience information

Geo Big Data Open Platform is designed to collect and manage geoscience research data and provide customized research data to users. The scope of research data collection and management is as follows (<https://data.kigam.re.kr/manage/data/rule#item-5>).

- It collects and manages research data produced in the course of research, such as field survey and exploration data, geological samples and analytical data, and geological themes (<https://data.kigam.re.kr/about/info?lang=en>).
- It Improves data reliability through standardization and quality control of digital dataset (See section R10).

Geo Big Data Open Platform

- It provides the service in conjunction with the Data Management Plan (DMP).
- It establishes a data sharing cooperation system with domestic and overseas organizations (See section R6).
- It collects and manages geoscience datasets and allows users to easily access the resources they need for specific studies and projects. It helps streamline the research process and provide more accurate and reliable results. The datasets available include:
 - By topic: Geology, Mineral Resources, Petroleum & Marine, Geologic Environment
 - By type: Investigation■Exploration, Sample■Analysis, Geoscience Thematic Map
 - Datasets for Literature & Data

The Geo Big Data Open Platform offers a range of services, including (<https://data.kigam.re.kr/about/info?lang=en>):

- Free access to digital research data: We provide various types of geoscience dataset, produced during the research process, free of charge. Additionally, we offer geoscience research literature & data and a collection of domestic and foreign papers (approximately 130,000). Furthermore, Open API services are available for application system development.
- Geological Thematic Map integration service: Through this service, users can access integrated search and 2D/3D visualizing service for Geological Thematic Maps. It allows for the superimposition of large amounts of Geological Thematic Maps and GIS data held by users. The platform also provides various information services, including Multiplatform Geoscience Information, Korea Marine Geological Geophysical Map Data, Geoenvironmental Information Verification System, Infocenter for Environmental Geology, Mineral Commodity Information, Earthquake Research Information, and the Dokdo Drone Platform.
- Integrated search with DataON (<https://dataon.kisti.re.kr/>): The Geo Big Data Open Platform facilitates integrated searching of research data linked to DataON, a national research data platform that offers a centralized location for domestic and foreign research data information.
- Data-based cooperative research ecosystem: The platform registers research data held by domestic and foreign researchers, fostering a data-driven connection, sharing, and convergence platform.
- IGSN issuance: We are the first in Asia to provide systematic management of geological sample data and related information by issuing International Generic Sample Numbers (IGSN).

* International Generic Sample Number (IGSN): The only globally identifiable generic sample identification code.

The Geo Big Data Open Platform has the capability to archive, publish, and distribute data from users, including projects, organizations, and individuals in the field of geoscience. It guarantees the long-term availability of the data for a period exceeding 10 years (<https://data.kigam.re.kr/about/info>). Article 16 of Chapter 5 of the KIGAM Data Management Rules stipulates that the retention period for research data is permanent (<https://data.kigam.re.kr/manage/data/rule#item-5> & <https://data.kigam.re.kr/manage/data/preserve#item-2>).

Links:

- [Data Management Rules #1](#)
- [Data Management Rules #5](#)
- [Data Preservation Guideline #2](#)
- [Geo Big Data Open Platform](#)
- [DataON](#)
- [Information on Geo Big Data Open Platform](#)

R2 Rights Management (R02)

R02. The repository maintains all applicable rights and monitors compliance.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geo Big Data Open Platform currently utilizes the Creative Commons (CC) license (<https://creativecommons.org/licenses/by/4.0/deed.en>) and prominently displays the license chosen by the depositor for each data item. In addition to the CC license, other options are available to data depositors, enabling them to specify their own licenses, including copyright notices, citations, and usage caveats for the materials. There are six types of licenses, each offering four terms and conditions. Once you have selected a Creative Commons (CC) license with the desired terms and conditions, you can refer to the CC License Markup Guide to properly apply the selected CC license to your work (<https://creativecommons.org/>).

- CC BY: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for
- CC BY-NC: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator.
- CC BY-ND: This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, and only so long as attribution is given to the creator. The license allows for commercial use.
- CC BY-SA: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use. If you remix, adapt, or build upon the material, you must license the modified material under identical terms.

Geo Big Data Open Platform

● CC BY-NC-SA: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator. If you remix, adapt, or build upon the material, you must license the modified material under identical terms.

● CC BY-NC-ND: This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, for noncommercial purposes only, and only so long as attribution is given to the creator.

The Geo Big Data Open Platform complies with KIGAM Data Management Rules, Chapter 6, The Conditions for disclosure and Use of research data (<https://data.kigam.re.kr/manage/data/rule#item-6>). Users of research data should be informed about the contribution of the person(s) who produced the data. If the research data license does not specify this information, the use of the research data should be properly acknowledged using appropriate citation or quotation methods. When research data is licensed, it is essential for the user to review and comply with the terms and conditions of the license, which may include requirements such as specifying the author and source of the data. The terms and condition specify the obligations of users, and failure to comply with these obligations may result in restrictions on the use of the Service (<https://data.kigam.re.kr/auth/join?lang=en>). The contents of use and service use restrictions are outlined in Chapter 3, Article 10 (Member's Obligations), Chapter 5, Article 14 (Service Use Restrictions), and Chapter 6 (Damages and Other Matters) of the terms and conditions. During the registration process, users can be categorized as internal or external, allowing the administrator to effectively manage and enforce use restrictions based on the user type.

According to the user guide (<https://data.kigam.re.kr/guide/dataset>), users who have obtained dataset registration rights can register datasets in the Geo Big Data Open Platform. The registered contents undergo an internal review and approval process before being officially registered as datasets. Users can track the progress of their data registration through the 'Register data' in 'My data' section. Notifications regarding dataset additions will be sent via email. Please note that the review and approval process may take approximately 3-5 days. Incomplete or inappropriate registrations may be rejected during this process.

All personal and sensitive information handled by the Geo Big Data Open Platform is collected, retained, and processed in accordance with relevant laws and policies.

● PERSONAL INFORMATION PROTECTION ACT (<https://law.go.kr/LSW/lsInfoP.do?lsiSeq=195062&viewCls=engLsInfoR&urlMode=engLsInfoR#0000>)

● Privacy Policy (<https://www.kigam.re.kr/menu.es?mid=a10906020000>)

● Rejection of unauthorized e-mail collection (<https://data.kigam.re.kr/about/email?lang=en>)

● Terms and Conditions (<https://data.kigam.re.kr/auth/join?lang=en>)

● Agreement for collecting private user information (<https://data.kigam.re.kr/auth/join?lang=en>)

Links:

- [CC license](#)
- [Creative Commons](#)
- [Agreement for collecting private user information](#)
- [Rejection of unauthorized e-mail collection of KIGAM](#)
- [Data Management Rules #6](#)
- [User guide of Geo Big Data Open Platform](#)
- [Privacy policy](#)
- [Terms and Conditions](#)
- [Personal Information Protection Act](#)

R3 Continuity of Service (R03)

R03. The Repository has a plan to ensure ongoing access to and preservation of its data and metadata.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Act on the Establishment, Operation, and Development of Government-funded Research Institutions in the Field of Science and Technology was enacted for establishing a national science and technology innovation system and promoting the development of government-funded research institutions in the field of science and technology. KIGAM is one of 22 research institutes established under Article 8 of the Act on the Establishment, Operation, and Development of Government-funded Research Institutions in the Field of Science and Technology (https://elaw.klri.re.kr/kor_service/lawView.do?hseq=59752&lang=ENG), ensuring its long-term existence and stability.

KIGAM traces its origins back to the establishment of the 'Geological Survey' in 1918. In 2001, KIGAM was rebranded and currently operates as the sole government-appointed organization dedicated to geoscience & mineral resources in Korea. Over its more than 100-year history, KIGAM has been at the forefront of addressing public safety and social challenges through its leadership in geoscience-related endeavors, a crucial role that the private sector cannot fulfill (<https://www.kigam.re.kr/menu.es?mid=a20503000000>). KIGAM's role, as stated in Article 5 of Chapter 1 of its Research Data Management Rules (<https://data.kigam.re.kr/manage/data/rule#item-1>), is to manage and preserve research data generated through projects and facilitate its reuse through free access. KIGAM operates the Geo Big Data Open Platform to fulfill this role, ensuring continued access to and preservation of budgets and

Geo Big Data Open Platform

data.

The Geo Big Data Open Platform receives funding from KIGAM's basic business budget, which is renewed annually. Because KIGAM operates under the framework of 'the Act on the Establishment, Operation, and Development of Government-funded Research Institutions in the Field of Science and Technology' the likelihood of funding interruptions is minimal. The platform operates with an annual budget exceeding 2,600,000,000 (KRW) (\$1,980,000).

The 'Geoscience Data Center,' which operates the Geo Big Data Open Platform, is an integral part of KIGAM's Geology and Space Division. The center comprises 30 researchers, including the center's head, 26 research and technical staff members, and four research students (<https://www.kigam.re.kr/menu.es?mid=a2050400000>). The Geoscience Data Center is responsible for various tasks, including business and budget management, data linkage, data quality, metadata, information systems and security, development, and operation of the data analysis environment, repository certification management, research data repository hosting service management, expert council operation, function development and operation, user support and service, and community activities. Staff members are highly qualified, hold doctoral degrees, possess specialized knowledge, and have high-level research capabilities in the fields of geoscience and IT (see Section R5).

The Geo Big Data Open Platform provides a set of compliance guidelines that serve as a reference for the operation of the research data platform. These guidelines are disclosed on the platform's webpage to ensure transparency and accessibility.

- Data Management Rules #1 (<https://data.kigam.re.kr/manage/data/rule#item-1>)
- Data Collection Guideline (<https://data.kigam.re.kr/manage/data/collect>)
- Data Management Guideline (<https://data.kigam.re.kr/manage/data/guide>)
- Data Preservation Guideline (<https://data.kigam.re.kr/manage/data/preserve>)
- Ethics, Copyright, and License Guideline (<https://data.kigam.re.kr/manage/data/license>)

Additionally, the Geo Big Data Open Platform adheres to the following policies to ensure the preservation of data and metadata, and to provide permanent services.

- Geo Big Data Open Platform terms and conditions (<https://data.kigam.re.kr/about/terms>).
- Data citation and publication (see section R4)
- Digital Object Identifier (DOI) access (search) and identification through IGSN issuance (see section R7)
- Software Management Guidelines, Operating Manual of the Informatization Deliberative Committee, and Guidelines for the Use of Computer Facilities
- Guidelines for Security Management of National Geographic Information, Regulations on Security Work, Operating Manual of Security Assessment Committee, Secure Transfer and Destruction Plan

The metadata of the Geo Big Data Open Platform was automatically deposited into DataON (<https://dataon.kisti.re.kr/>), a national research data platform operated by the Korea Institute of Science and Technology Information (KISTI, <https://www.kisti.re.kr/eng/>). This ensures that even in situations where The Geo Big Data Open Platform may face risks, such as disasters or funding suspension, users can still access the data deposited by KIGAM through DataON (see section R6). In the event of a disaster or catastrophe, such as destruction of infrastructure, interruption of power supply, or loss of life, we will attempt to restore the system through backup data by following the 'Crisis Response Manual for Personal Information Processing System in Preparation for Accidents and Disasters' of the parent organization, KIGAM. If an existing server room cannot be used, the Information Technology Department & Library, Earthquake Research Center, Korea Seismological Research Station server room in Wonju, KIGAM Pohang Branch, etc. A disaster recovery and data linkage policy was established that uses a cloud center (KoreaV-G Cloud).

Links:

- [Agreement for collecting private user information](#)
- [Data Management Rules #1](#)
- [History of KIGAM](#)
- [Organization of KIGAM](#)
- [Data Collection Guideline](#)
- [Data Management Guideline](#)
- [Data Preservation Guideline](#)
- [Ethics, Copyright, and License Guideline](#)

R4 Legal & Ethical (R04)

R04. The repository ensures to the extent possible that data and metadata are created, curated, preserved, accessed and used in compliance with legal and ethical norms.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

Geo Big Data Open Platform

- [Information on Geo Big Data Open Platform](#)
- [Agreement for collecting private user information](#)
- [Rejection of unauthorized e-mail collection of KIGAM](#)
- [Data Management Rules #6](#)
- [Data Management Rules #2](#)
- [Data Management Guideline](#)
- [Ethics, Copyright, and License Guideline #4](#)
- [Ethics, Copyright, and License Guideline](#)
- [Ministry of Science and ICT](#)
- [National Research Council of Science & Technology \(NST\)](#)
- [Privacy policy](#)
- [Data Management Guideline #7](#)
- [Terms and Conditions](#)
- [Creative Commons Attribution 4.0 International Public License](#)

R5 Governance & Resources (R05)

R05. The repository has adequate funding and sufficient numbers of staff managed through a clear system of governance to effectively carry out the mission.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

'The Act on the Establishment, Operation, and Development of Government-funded Research Institutions in the Field of Science and Technology (https://elaw.klri.re.kr/kor_service/lawView.do?hseq=59752&lang=ENG)' was enacted with the aim of establishing a national science and technology innovation system and fostering government-funded research institutions in the field of science and technology. KIGAM is one of the 22 research institutes established under Article 8 of this act and operates the Geo Big Data Open Platform. KIGAM, the parent organization of the Geo Big Data Open Platform, is an agency under the Ministry of Science and ICT* and a member of the National Research Council of Science & Technology (NST)*. The Geo Big Data Open Platform serves as a centralized data hub for the management and preservation of geoscience research data at the national level. It is funded through the government budget under the Act on the Establishment, Operation, and Development of Government-funded Research Institutions in the Field of Science and Technology, which ensures the stability and long-term sustainability of the repository.

● Ministry of Science and ICT (<https://www.msit.go.kr/eng/index.do>)

● National Research Council of Science & Technology(NST) (<https://www.nst.re.kr/eng/index.do>)

The Geo Big Data Open Platform is operated by the Research Data Management Division, also known as the Geoscience Data Center, which is a part of KIGAM. The division is responsible for various tasks related to research data, including collection, management, preservation, sharing, and utilization. KIGAM, the parent organization of the Geo Big Data Open Platform, has a total of 529 employees, including 469 regular employees and 33 contract employees. It is comprised of nine departments, namely the Geology&Space Division, Mineral Resources Division, Marine Geology&Energy Division, and Geologic Hazards Division, among others (<https://www.kigam.re.kr/menu.es?mid=a20504000000>). Within KIGAM, the Geoscience Data Center, which operates the Geo Big Data Open Platform, is a part of the Geology&Space Division. It consists of a team of 8 researchers, including 8 research and technical staff, as well as 1 research students. The center is led by the head of the center. The staff members are highly qualified, holding doctoral degrees and possessing specialized knowledge in the field of geoscience & IT. They demonstrate a strong research capacity in related disciplines.

The Geo Big Data Open Platform is operated with an annual budget of more than ■2,600,000,000 (\$1,980,000) through the basic project "Establishment of National Geoscience Data Center (NGDC) foundation through the Geo Big Data Open Platform development" of KIGAM. The total research cost for the development of the Geo Big Data Open Platform is ■4,941,076,000 (\$3,780,000), which corresponds to the second year (as of 2021) of the total research period (2020~2024). Of the total research expenses, about ■200,000,000 (\$152,000) is devoted to data refinement, processing, and construction, and about ■200,000,000 (\$152,000) is devoted to system operation and maintenance.

Links:

- [Ministry of Science and ICT](#)
- [National Research Council of Science & Technology \(NST\)](#)
- [Organization of KIGAM](#)

R6 Expertise & Guidance (R06)

R06. The repository adopts mechanisms to secure ongoing expertise, guidance and feedback-either in-house, or external.

Geo Big Data Open Platform

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

Internal consultation within the Geo Big Data Open Platform is facilitated by KIGAM's Research Data Management Committee. This committee serves as an advisory body that discusses issues related to data management plans and the execution of research data-related activities. It comprises a maximum of 25 members, including department heads from various departments (<https://data.kigam.re.kr/manage/data/rule#item-2>).

To support the continuous growth and professional development of our employees, the Geo Big Data Open Platform offers a range of training programs tailored to their needs. These programs include specialized training in geoscience, regular IT skills training, and compliance with legal, ethical, and regulatory frameworks such as tenure rights. Our employees have access to a training portal with diverse courses, and they are required to participate in mandatory training sessions at least three times a year. We are committed to enhancing their expertise through training and activities that promote a culture of data sharing and utilization.

- Korea Institute of Human Resources Development in Science and Technology (KIRD): AI integrated education, research ethics, and research security programs (<https://alpha-campus.kr/en>)

- International School for Geoscience Resources (IS-Geo): Python education, geological modeling practice, etc. (<https://www.kigam.re.kr/educationList.es?mid=b40402000000>)

- Geo big data contest: Research data analysis and use case development, AI model development (<https://www.geodata-con.kr/2022/index.php>)

The Geo Big Data Open Platform relies on an external monitoring organization to provide technical and data advice. This organization conducts regular checks on data errors, suggests feature enhancements, and provides recommendations to improve the platform (<https://data.kigam.re.kr/monitoring/board/?lang=en>). The external monitoring organization welcomes participation from university (graduate) students, experts in the field of geoscience, and individuals interested in the subject. Their activities are carried out once a year for a duration of five months per term. The 2nd Geo Big Data Open Platform Monitoring Team (2022) comprised 17 members. The main responsibility of the monitoring team is to submit monthly reports on data errors and propose measures to enhance the platform's performance. Through the activities of the monitoring team in the 1st period, a total of 168 cases of data errors and system improvements were reported (2021, aggregated from the 1st period).

KIGAM has signed MOUs with various domestic and foreign professional organizations to facilitate collaborations and exchanges.

- Domestic

- GEO DATA: <https://geodata.kr/>

- DataOn: <https://dataon.kisti.re.kr/>

- The Korean Earth Science Society: <https://www.kess64.net/>

- The Korean Society of Economic and Environmental Geology: <https://www.kseeg.or.kr/>

- The Geological Society of Korea: <https://www.gskorea.or.kr/html/?pmode=english>

- Memorandum of Understanding with Overseas Organizations (Total 44 institutional MOUs, 24 countries, 2 international organizations. <https://www.kigam.re.kr/menu.es?mid=a20302000000>)

- Europe: NORWAY(NGU-NORCE), GERMANY(AWI-IFM-GEOMAR), RUSSIA(GS RAS), ITALY(IRPI), AUSTRIA(GBA), UK(BGS), FINLAND(GTK)

- Asia: UZBEKISTAN(GOSCOMGEOLOGY), KAZAKHSTAN(KAZGEOLOGY), MONGOLIA(MRPAM-MUST-IAG), CHINA(CGS-CEA-IGGCAS), JAPAN(AIST-GSJ-JAMSTEC-NIED), TIMOR-LESTE(IPG), MYANMAR(DGSE-Dagon Univ.), INDONESIA(DGGMR-LEMIGAS), MALAYSIA(DMG), VIETNAM(HCMUT-VIGMR), THAILAND(DMR), IRAN(RIPI)

- North America: USA(USGS-AFTAC-NSF-Planetary Science Institute-Univ. of Arizona-Virginia Polytechnic Institute), CANADA(NRCAN GSC-CANMET)

- South America: CHILE(SERNAGEOMIN)

- Oceania: AUSTRALIA (CSIRO)

- Africa: ETHIOPIA (AAiT)

- International Organizations: CCOP, UNESCO

Staff of both KIGAM and the Geo Big Data Open Platform actively engage in knowledge exchange with a wide range of organizations. (similar institutions globally, international organizations, and programs: <https://www.kigam.re.kr/menu.es?mid=a20301000000>)

- ASEAN (Association of Southeast Asian Nations)

- CCOP (Coordinating Committee for Geoscience Programmes in East and Southeast Asia)

- CTBTO (Comprehensive Nuclear Test Ban Treaty Organization)

- ICDP (International Continental Scientific Drilling Programme)

- ICL (International Consortium Landslides)

- IEA-GIA (International Continental Scientific Drilling Programme)

- IODP (International Ocean Discovery Program)

- SCUFN-IOC (Sub-Committee on Undersea Feature Names)

- OneGeology

- IGSN e.V. (International Generic Sample Number)

- DataCite

Recently, KIGAM signed a 'Critical Minerals Cooperation between MRIWA and KIGAM

(https://www.kigam.re.kr/board.es?mid=a20401000000&bid=0032&act=view&list_no=56201)' with the Minerals Research Institute of Western Australia (MRIWA).

Geo Big Data Open Platform

The Geo Big Data Open Platform has been certified by the Korea Data Agency as Platinum Class (https://data.kigam.re.kr/img/ci/data_certificated.png) in Database Quality Certification (DQC), a system that examines and certifies data quality, data management system, and data security elements to ensure data quality. (See section R10)

KIGAM, the parent organization of the Geo Big Data Open Platform, has obtained the distinction of being the first institution in Asia to be issued an IGSN (International Generic Sample Number). The Geo Big Data Open Platform utilizes this unique identifier to enable systematic management of geological sample data and associated information (<https://data.kigam.re.kr/about/info?lang=en>).

The Geo Big Data Open Platform is registered with re3data, enabling access for users in relevant fields and communities (<https://www.re3data.org/repository/r3d100014025>).

Links:

- [Korea Institute of Human Resources Development in Science and Technology \(KIRD\) alpha campus](#)
- [International School for Geoscience Resources \(IS-Geo\)](#)
- [Monitoring organization of Geo Big Data Open Platform](#)
- [GEO DATA](#)
- [Korean Earth Science Society](#)
- [Geological Society of Korea](#)
- [DataON](#)
- [Memorandum of Understanding with International Organizations](#)
- [Similar global institutions, international organizations, and programs](#)
- [MOU Signing Ceremony between KIGAM and MRIWA](#)
- [Data Quality Certification from Korea Data Industry Promotion Agency](#)
- [Information on Geo Big Data Open Platform](#)
- [Data Management Rules #2](#)
- [Re3data Identifier](#)

Digital Object Management

R7 Provenance and authenticity (R07)

R07. The repository guarantees the authenticity of the digital objects and provides provenance information.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

KIGAM has a guide that encompasses Data Collection, Management, Preservation, Ethics, Copyright, and Licenses. The Geo Big Data Open Platform operates in accordance with these guidelines.

The Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect>) focuses on the collection and deposition of research data, as well as the creation of a Data Management Plan (DMP).

The Data Management Guide (<https://data.kigam.re.kr/manage/data/guide>) provides information on maintaining the authenticity of research data and metadata through versioning and quality assurance.

The Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve>) includes backup and recovery policies, as well as preservation strategies for research data.

Lastly, the Ethics, Copyright, and License Guide (<https://data.kigam.re.kr/manage/data/license>) covers ethical considerations, copyright concepts, and licensing standards for research data.

To submit data to the Geo Big Data Open Platform, depositors are required to complete a membership registration and login process for identity verification. The identity information collected during the membership registration is outlined in the 'Agreement for collecting private user information' (<https://data.kigam.re.kr/auth/join>).

The Geo Big Data Open Platform specifies the procedure for research data registration (<https://data.kigam.re.kr/guide/dataset>), which can also be found in the KIGAM Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect#item-4>). Researchers should check the quality and abnormality of the research data within 30 days after acquiring the research data and submit the research data to the data repository. The principal investigator or the person in charge of research data management at the institution or department will review and approve the submitted research data. The main aspects considered during the review include the data format, adherence to metadata technical requirements, compliance with research data disclosure and licensing standards, and the presence of any sensitive information.

When depositing data into the Geo Big Data Open Platform, researchers are required to provide metadata that describes the contents of the data file.

This metadata should follow the Dublin Core format, which consists of 15 elements: title, creator, type, contributor, publisher, date, language, format,

Geo Big Data Open Platform

description, subject, related resources, identifier, copyright, source, and coverage. For more information on metadata techniques, please refer to the KIGAM Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect#item-3>) and KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-6>).

Data deposited in the Geo Big Data Open Platform follows the Data LifeCycle process. (See section R11) The Data LifeCycle process is thoroughly documented in the Data Management Plan (DMP), encompassing metadata, the data to be collected and managed, data formats and standards, as well as data sharing practices. Detailed information regarding these aspects can be found in the KIGAM Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect#item-5>) and the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-2>).

It is the responsibility of researchers and repositories to ensure regular and automatic backup of data to multiple locations, aiming to prevent and protect against data loss and corruption. Specific guidelines regarding data backup and recovery policies, as well as archiving practices within the Geo Big Data Open Platform, are outlined in the KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve#item-4>). When utilizing the administrator function to verify data files, it is recommended to check for corruption by comparing the current checksum with the checksum recorded when the digital object was initially saved.

The backup system of the Geo Big Data Open Platform comprises a dual backup approach utilizing the InnoStor Appliance (ISA-2000) and Quantum Scalar i500. Periodically, data from the service storage is backed up and stored in the backup system. The backup targets encompass data, databases, and user data files within the Geo Big Data Open Platform. Daily backups are performed for user data files, databases, and system data. Additionally, full backups of user data files, databases, and research data (files) are conducted on a weekly basis.

The Geo Big Data Open Platform has comprehensive preservation plans in place, which encompass format migration and emulation strategies (<https://data.kigam.re.kr/manage/data/preserve#item-4>). These preservation plans are designed to effectively manage changes in data, metadata, technology, and user requirements, ensuring reliable and timely handling of such changes. (See section R9, R13)

The Geo Big Data Open Platform ensures data integrity and authenticity through the implementation of versioning for data changes. Versioning is applied to both data and metadata, and the specific changes made to each version are displayed on the right side of the corresponding data. Version changes occur when a dataset is modified, and the decision to assign either a minor or major version lies with the research data depositor.

- Minor versions: changes to data basics and metadata (ex. 1.1, 1.2...)
- Major version: metadata and file data changes (ex. 2.0, 3.0...)

The Geo Big Data Open Platform utilizes a consistent version numbering scheme to effectively track the presence of new versions and changes made to the data. This numbering scheme serves to clearly differentiate between previously used versions and the current version being worked on. When a dataset undergoes modifications, a changelog record is stored, and previous versions of the dataset are retained. The dataset's changelog contains essential information such as the time of change and the identity of the individual who made the changes, aiding in identification purposes. The changelog can be accessed within the dataset's details. While direct links to previous versions are not currently provided, there are plans to develop a service in the future that will allow users to access previous versions. For more comprehensive details on data and metadata versioning, please refer to the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-4>).

The Geo Big Data Open Platform offers users access to data change history logs, which are accessible within the dataset details. Moreover, upon the completion of the final approval process for a dataset, a DOI or IGSN is issued via the registration API. These identifiers serve to uniquely identify, track utilization, and facilitate the management of the dataset.

IGSNs are assigned to sample datasets, specifically for geological samples such as rocks and drill cores. On the other hand, DOIs are issued for other types of datasets. In 2015, KIGAM became the first IGSN registrar in Asia (<https://igsn.github.io/>), enabling the assignment of internationally recognized unique identification numbers to geological samples. This ensures a higher level of reliability in the analysis of research data. Additional information regarding this topic can be found in the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-5>).

Links:

- [Data Collection Guideline #4](#)
- [User guide of Geo Big Data Open Platform](#)
- [Data Collection Guideline #3](#)
- [Agreement for collecting private user information](#)
- [Data Collection Guideline](#)
- [Data Management Guideline](#)
- [Data Preservation Guideline](#)
- [Ethics, Copyright, and License Guideline](#)
- [International Generic Sample Number \(IGSN\)](#)
- [Data Management Guideline #6](#)
- [Data Collection Guideline #5](#)
- [Data Management Guideline #2](#)
- [Data Management Guideline #4](#)
- [Data Management Guideline #5](#)
- [Data Preservation Guideline #4](#)

R8 Deposit & Appraisal (R08)

Geo Big Data Open Platform

R08. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for users.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geoscience Data Center, which serves as the operating department of the Geo Big Data Open Platform, has formulated guidelines for the standards of geoscience research data. Additionally, they have implemented a quality management and evaluation system to effectively manage the data (<https://www.kigam.re.kr/menu.es?mid=a20101040000>).

The Geo Big Data Open Platform provides a comprehensive procedure for dataset registration, wherein the deposit process is carried out through a review and approval system (<https://data.kigam.re.kr/guide/dataset?lang=en>). During data registration, specific requirements are defined within the metadata schema, indicating whether certain items are mandatory. If the required items are not registered, the data cannot be successfully submitted. Additionally, to ensure data quality and completeness, depositors are expected to provide thorough and well-documented data. Incomplete or insufficiently documented data may not be eligible for publication on the Geo Big Data Open Platform. Administrators are assigned the permissions to create, edit, and delete collections, enabling effective data management. Depositors select collections based on their designated permissions, allowing them to submit datasets within the appropriate scope.

The Geo Big Data Open Platform is built upon the Geoscience Data Repository System (GDR system)*, which facilitates the electronic opening and utilization of research data. A data depositor is a user who has been granted the privilege to register a dataset. The content submitted by the depositor undergoes an internal review and approval process before being officially registered in the dataset. Depositors can track the progress of their registration on the platform's homepage and receive email notifications regarding the registration process. The review and approval process typically takes 3-5 days to complete. In cases where the registration is deemed incomplete or inappropriate, it may be rejected.

* Geoscience Data Repository System (GDR) : A system that electronically performs tasks related to the collection, management, and preservation of research data.

During the dataset registration process, the depositor is responsible for interpreting the digital object and reviewing the status of the metadata inputs within the review and approval process. As data submission is only permitted when all the required metadata items have been entered, the depositor ensures the inclusion of the necessary items before submitting the data. Subsequently, the administrator examines the entered metadata during the review and approval process. The review points for data registration in the Geo Big Data Open Platform are as follows:

- checklist
- Follow naming conventions
- Check for typos
- File size (maximum single file size limit: 100 GB)
- DMP taxonomy code check

The Geo Big Data Open Platform operates in accordance with the KIGAM Data Management Rules, as outlined in Chapter 5, Article 15 of the rules (<https://data.kigam.re.kr/manage/data/rule?lang=en#item-5>). The main contents of the KIGAM Data Management Rules are as follows:

In general, it is the responsibility of the researcher to ensure the quality and integrity of the acquired research data before registering it in the GDR system. The registration process requires approval from the research director. The research data is registered in the GDR system in a format that allows software to verify the individual contents, internal structure, and perform operations such as modification, conversion, and extraction. The data file format adheres to the standards specified by the Geo Big Data Open Platform. Additionally, researchers are encouraged to register relevant documents, such as readme files, technical documentation, and auxiliary materials, that are essential for comprehending the research data. These additional documents contribute to a comprehensive understanding of the registered research data. Data that is registered in accordance with the KIGAM Data Management Rules undergoes a review process and, upon approval, is published on the Geo Big Data Open Platform.

The following data formats are registered in the Geo Big Data Open Platform. (<https://data.kigam.re.kr/advanced-search?q=&lang=en>)

- Compressed files: zip, gz, egg
- Text formats: xlsx, xls, hwp, docx
- Image formats: png, jpg
- 3D models: obj, mtl
- GIS formats: shp
- Other formats: dat, pptx, pdf

The Geo Big Data Open Platform provides research data management plan (DMP) guidelines and forms, which can be found in the KIGAM Data Management Rules (<https://data.kigam.re.kr/manage/data/rule#item-4>).

Links:

- [Data Management Rules #4](#)
- [Data Management Rules #5](#)
- [Geoscience Data Center of KIGAM](#)
- [User guide of Geo Big Data Open Platform](#)
- [Geo Big Data Open Platform Advanced search](#)

Geo Big Data Open Platform

R9 Preservation plan (R09)

R09. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geo Big Data Open Platform has documented preservation plans that encompass format migration and emulation, enabling effective management of changes to data, metadata, technology, and user requirements in a reliable and timely manner. The Geo Big Data Open Platform, as a national geoscience data center, stipulates the permanent preservation of research data in the KIGAM Data Management Rules (<https://data.kigam.re.kr/manage/data/rule>) and the KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve>).

The Geo Big Data Open Platform ensures the proper functioning of hardware and software systems and performs data backup and migration as necessary.

To prevent and protect data loss and corruption, researchers and repositories are responsible for regularly and automatically backing up their data to multiple locations, and details of the Geo Big Data Open Platform's data backup and recovery policies and archiving are set out in the Geo Big Data Open Platform's Research Data Retention Guidelines (<https://data.kigam.re.kr/manage/data/preserve#item-4>).

The backup system of the Geo Big Data Open Platform consists of a double backup of InnoStor Appliance (ISA-2000) and Quantum Scalar i500, and periodically backed-up data from the service storage is stored in the backup system. The backup targets are the Geo Big Data Open Platform's data, databases, and user data files. User data files, databases, and system data are backed up daily, and full backups of user data files, databases, and research data (files) are made once a week.

The following are the recovery policies and guidelines for the Geo Big Data Open Platform.

- System software and application software are recovered from the local GIT repository.
- Recovery of research data, databases, and user data files is performed with data stored on a backup device.
- In case of a problem with the backed-up data, restore from the tape backup at the source.

For archiving and preserving research data in the Geo Big Data Open Platform, the contents are as follows.

- Periodically archive (magnetic tape) and preserve research data for the purpose of preserving research data.
- Archiving and preserving backup tapes through tape backup as a method of archiving and preservation (yearly) by dissipating and archiving backup tapes to a remote location.
- The retention period for archiving tapes is at least 5 years.

The Geo Big Data Open Platform incorporates the Dublin Core (DC) metadata schema for effective utilization and provides specific schemas for each type of deposited data, such as survey, exploration, sample, analysis, and thematic maps. The Geo Big Data Open Platform follows the Leveled Digital Asset Preservation Framework, as outlined in the KIGAM Data Preservation Guide, which categorizes data preservation into four levels: Level 1 (Data Protection), Level 2 (Data Awareness), Level 3 (Data Monitoring), and Level 4 (Data Recovery) (<https://data.kigam.re.kr/manage/data/preserve#item-5>).

The Geo Big Data Open Platform does not require a formal contract for data deposition. Instead, it provides depositors with essential resources and documentation on its website, including research data collection, management, preservation, and licensing guides, as well as deposit forms (<https://data.kigam.re.kr/manage/data/rule>). Depositors are expected to adhere to these forms and guidelines when depositing data. The available documents include a description of data collection in the Geo Big Data Open Platform, a deposit form that outlines terms and conditions for data access and long-term storage, guidelines for good data deposit practices, and a metadata form. The KIGAM Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect>) establishes the repository's responsibilities concerning the long-term storage of data, providing guidelines and obligations for ensuring data preservation.

The metadata of the Geo Big Data Open Platform is automatically deposited in DataON (<https://dataon.kisti.re.kr/>), a national research data platform operated by KISTI. Even if the Geo Big Data Open Platform is unable to operate normally due to risky situations such as disasters or funding suspension, users can access the KIGAM data previously deposited in DataON through DataON (See section R6).

All research data in the Geo Big Data Open Platform is retained permanently as a general practice (<https://data.kigam.re.kr/manage/data/rule#item-5>). Additionally, users have the option to establish a retention policy for purged data, with available retention periods ranging from 3 months to 5 years, including options for 6 months, 1 year, 2 years, 3 years, and 5 years.

The Geo Big Data Open Platform follows a repository approach for deleting or removing data and metadata from collections. Administrators have the capability to discard datasets through the data management function. Moreover, the physical deletion of discarded data is carried out by configuring the data retention policy within the administrator preferences.

Links:

- [Data Preservation Guideline #5](#)
- [Data Management Rules](#)
- [Data Collection Guideline](#)
- [Data Preservation Guideline](#)
- [DataON](#)

Geo Big Data Open Platform

standards) checks, website openness checks, website accessibility (performance checks such as optimization) checks, e-government UI-UX (usability) checks, dead link (link error) and East Sea and Dokdo map API checks, and text content (efficiency, reliability) checks. To ensure the reliability of the website, we conduct checks on various aspects including the regular updating of materials, intellectual property protection, privacy policies, and identification of errors or inaccuracies in publicly available information.

The Geo Big Data Open Platform adopts DC's metadata elements as a standard. Notably, KIGAM has pioneered the issuance of the first IGSN in Asia, facilitating systematic management of geological sample data and associated information. Furthermore, the platform offers the flexibility to update and modify the metadata schema to align with user requirements. This capability is accessible through the admin page, allowing administrators to create, modify, and delete metadata schemas in response to evolving technical environments and user needs.

Furthermore, during the metadata authoring process for your dataset, you have the option to include external links to other relevant digital objects. Data users have the ability to raise concerns or issues regarding the dataset, such as data quality, availability, system-related matters, data format, missing or incomplete data, and other related issues. This can be done through the contact or email inquiry function provided by the Geoscience Data Center of KIGAM, the entity operating the Geo Big Data Open Platform. Additionally, data users have the opportunity to evaluate the data by accessing the dataset details.

Links:

- [Rule for Public data of KIGAM](#)
- [Korea Data Agency](#)
- [E-Government Website Quality Management, Ministry of the Interior and Safety Notice 2021-19](#)
- [User guide of Geo Big Data Open Platform](#)
- [Data Quality Certification from Korea Data Industry Promotion Agency](#)
- [Data Management Guideline #5](#)

R11 Workflows (R11)

R11. Digital object management takes place according to defined workflows from deposit to access.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

KIGAM has defined research data management standards to promote responsible research and data management practices among researchers, following the guidelines set forth in the KIGAM Data Management Rules (<https://data.kigam.re.kr/manage/data/rule#item-1>). Furthermore, KIGAM has developed a comprehensive Data LifeCycle policy and procedure for effective data management, which has been internally published within the organization. The Geo Big Data Open Platform follows a staged approach to research data management, aligning with the Data LifeCycle specified in the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-2>). This ensures that data management is carried out in a systematic and structured manner.

The Data LifeCycle process encompasses six distinct phases: Collect, Store, Manage, Access & Use, Preserve, and Share & Reuse.

The Collect phase involves defining the scope of research data collection and creating a Data Management Plan (DMP). KIGAM collects and manages various geoscience datasets, including Geology, Mineral Resources, Petroleum & Marine, Geologic Environment data. This includes research data generated by KIGAM researchers as well as data obtained from external sources. Specific guidelines and procedures for data collection are outlined in the KIGAM Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect#item-3>).

During the Store phase, the data obtained during the collection phase undergo a review process to ensure compliance with preferred file formats and naming conventions. Additionally, backup policies are implemented to guarantee data integrity and accessibility, particularly in the event of potential risks to the repository. More comprehensive information on these processes can be found in the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-4>).

The Manage phase focuses on ensuring data authenticity through quality control measures and security maintenance. The KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-5>) provides detailed information on these processes. As part of this phase, KIGAM assigns IGSNs to sample datasets and DOIs to other types of datasets. These unique identifiers facilitate tracking, management, and utilization of the datasets, ensuring a high level of reliability. The Manage steps adhere to the FAIR principles (Findable, Accessible, Interoperable, Reusable) and employ a checklist to assess the level of compliance of research data with these principles.

The Access & Use phase covers the following points, which are detailed in the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-2>).

- Ensuring that both designated users and reusers can access data on a daily basis
- Applying access control and authentication procedures
- Supporting methods of discovery, analysis, repurposing, distribution, and representation.

The Preserve phase emphasizes the implementation of measures to ensure the long-term preservation and retention of research data. This includes preservation planning and comprehensive documentation. For more specific information regarding these processes, please refer to the KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve#item-4>).

Geo Big Data Open Platform

The Share & Reuse phase is designed to facilitate the sharing and reuse of data, adhering to the principles of the FAIR Principles. This phase is specifically structured to promote the accessibility, interoperability, and reusability of research data. (See section R13)

- Making research data available to the widest possible audience while protecting research confidentiality and within regulatory permissions.
- Allowing users to select private, internal, or external disclosure when submitting research data, as well as setting time limits for disclosure through licenses and embargoes.
- Providing rich contextual information through standards-enabled metadata to enable accurate interpretation and reuse of research data.
- Increase research transparency through citation of research data, improve searchability and reusability of research results, and measure research impact.

The KIGAM Data Collection Guide (<https://data.kigam.re.kr/manage/data/collect>) available on the Geo Big Data Open Platform provides an overview of the data types to be collected, the process of depositing and registering research data, and important notes to consider. Additionally, the webpage (<https://data.kigam.re.kr/guide/dataset?lang=en>) offers data registration guidelines for reference.

The Research Data Management Guidelines (<https://data.kigam.re.kr/manage/data/guide>) provide comprehensive guidance on data identification, publication, and citation practices.

The KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve>) covers essential concepts related to data preservation, including the selection and evaluation of preserved data, as well as information on data repositories.

Lastly, the KIGAM Ethics-Copyright-License Guide (<https://data.kigam.re.kr/manage/data/preserve>) addresses crucial aspects such as the disclosure and sharing of research data, as well as copyright and licensing considerations.

The Geo Big Data Open Platform is compliant with the Reference Model for Open Archival Information System (OAIS). The data integrity service of the OAIS provides a conceptual framework that considers the functional requirements of archive that is adequate for the long term management and preservation of digital data. The OAIS prevents data from any unpermitted change or destruction.

The Submission Information Package (SIP) refers to the additional data deposit required for documentation and metadata, ensuring the long-term preservation and access/reuse of data by consumers. The SIP serves as the foundation for creating the Archival Information Package (AIP) and Dissemination Information Package (DIP). If necessary, this process includes the creation of data versions for preservation and distribution. The data included in the SIP should be in a suitable preservation format with clear migration paths, accompanied by relevant documentation that is sufficient to support the creation of the AIP. A well-constructed SIP ensures a smooth process of creating the AIP. In other words, the AIP utilizes fixity or checksum values to verify the integrity of the data and enables ongoing management through batch processes such as migration. The data included in the SIP should be in a format suitable for distribution or have a migration pathway defined. The registered format for preservation and distribution can be the same. The SIP should encompass all necessary documents that facilitate easy reuse, including resource search, usability, access, transmission, and metadata related to utilization.

Links:

- [Data Collection Guideline #3](#)
- [Data Management Rules #1](#)
- [Data Collection Guideline](#)
- [Data Management Guideline](#)
- [Data Preservation Guideline](#)
- [Ethics, Copyright, and License Guideline](#)
- [User guide of Geo Big Data Open Platform](#)
- [Data Management Guideline #2](#)
- [Data Management Guideline #4](#)
- [Data Management Guideline #5](#)

R12 Discovery and Identification (R12)

R12. The repository enables users to discover the digital objects and refer to them in a persistent way through proper citation.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geo Big Data Open Platform offers a range of search functions, search filters (faceted search), and search utilization services to facilitate the convenient use of research data. Users can access various types of research data (such as data sets, reports and papers, statistical information, and file data like 3D models and images) provided by the platform through the search bar on the webpage, using various search methods.

The search filter service is available through the "Search in result" / "Title exact Match" functions located next to the search bar on the search result screen. Users can employ these filters to refine their search by specifying conditions such as keywords, locations, subjects, and types within the "Detailed Search" option. By clicking on the desired research data in the search results, users can access further information about that specific research data.

The Geo Big Data Open Platform offers a variety of search options to facilitate the discovery and access of research data. These include integrated

Geo Big Data Open Platform

search, detailed search, topic search, type search, and Geoscience Thematic Map (2D/3D Visualizing service) search. Members can check the registration status of research data by topic and type. Integrated search is a basic search that allows member to use various search functions, and member can select the exact data member want by utilizing various facet searches in the detailed search. In the subject search, member can check the search results categorized by Geology, Mineral Resources, Geologic Environment, and Petroleum & Marine. Type search allows member to check search results according to the type of research data.

The Geoscience Thematic Map 2D/3D Visualizing service offered by the Geo Big Data Open Platform enhances the presentation of geoscience topics by overlaying various geological information, such as Geological Map, Sea-floor Geological Map, and Geoenvironmental Information Map, onto the TerriaJS platform*. This integration creates a more immersive and realistic information service. In addition to location search and other search functions, this service enables users to overlay their own GIS data onto the platform, facilitating comparative analysis with other geological topics. Furthermore, the service provides a range of user-friendly features, including screen splitting, walking mode for virtual exploration, distance measurement, story storage, and sharing/printing functions.

* The TerriaJS Platform: An open-source framework for web-based geospatial catalogue explorers (<https://terria.io/>)

The Geo Big Data Open Platform follows standard practices for citing research data, treating it as any other scholarly output. The format and usage of data citations on the platform are as follows:

- Author. (Year of data production). Number of data (sets). Data publishing organization. DOI
- Example: Kangwon National University. (2019). Silicified wood. Korea Institute of Geoscience and Mineral Resources. <https://doi.org/10.22747/data.20201105.50>

The Geo Big Data Open Platform uses International Generic Sample Numbers (IGSNs) and digital object identifiers (DOIs) to identify data objects. It provides systematic management of geological sample and analysis data, along with related information, through the issuance of IGSNs. This facilitates the identification, sharing, and citation of each dataset using DOIs. Additionally, metadata is stored in DC format for the purpose of searching and identifying research data (<http://dublincore.org/documents/dces/>).

The Geo Big Data Open Platform is a registered repository with re3data. It serves as a gateway to freely access high-quality research data, allowing users to search and utilize the latest geological resource information. The platform provides open APIs and data, including a wide range of geoscience reports and search functions for both domestic and foreign papers, totaling over 130,000 resources. Additionally, it offers metadata-based data linkage and search capabilities in collaboration with the National Research Data Platform (DataON) (<https://www.re3data.org/repository/r3d100014025>).

Links:

- [Re3data Identifier](#)
- [Cesium Web 3D terrain platform: Basic open platform for creating 3D geospatial applications](#)
- [DunblinCore](#)

R13 Reuse (R13)

R13. The repository enables reuse of the digital objects over time, ensuring that appropriate information is available to support understanding and use.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The open science and open access movements are driving a shift towards the publication of data and publications resulting from publicly funded research projects in public repositories. Sharing research data offers numerous benefits, including increased utilization and recognition of data from other researchers or institutions, prevention of duplication of effort, and advancement of better research through the validation of data. Further information regarding data repositories for sharing, reusing, and managing research data can be found in the KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve>). This guide provides definitions, selection considerations, and examples of data repositories.

KIGAM is committed to upholding the FAIR principles (Findable, Accessible, Interoperable, Reusable), which serve as a valuable framework for the sharing and reuse of research data. Specifically, KIGAM aims to enhance the reusability of research data deposited in the Geo Big Data Open Platform by utilizing a checklist of reusable items. Researchers, as data producers, are encouraged to make their data widely available by organizing it according to common standards. Proper licensing should be applied to ensure appropriate usage. It is important to prioritize research confidentiality while adhering to relevant regulations. During the data deposition process, researchers have the option to choose private, internal, or external disclosure based on their preferences. Additionally, they can specify a time limit for disclosure through an embargo period.

To protect privacy, private data, including personal information, should be transformed into a de-identified form before its release. Considering the ongoing access and potential reuse of research data, it is recommended that the data be made available in a machine-readable format to maximize its usability. The use of non-proprietary open formats is encouraged to ensure compatibility with different computing and technology environments. Research data should be in a format that can be easily converted to formats compatible with various statistical and analytical software, without compromising its research value. It is also advisable to use file formats that do not require additional software for data interpretation. For more comprehensive information on this policy, please refer to the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide>).

Geo Big Data Open Platform

When researchers deposit research data, they are required to include metadata that describes the contents of the data file. This metadata should contain information such as the project title, principal investigator's name, abstract, distributor, subject matter, geographic scope, temporal scope, and depositor. Additionally, researchers should provide any necessary documentation files along with the data files and metadata. These documentation files could include codebooks, data collection tools, summary statistics, project summaries, and lists of publications related to the data. For more specific guidelines regarding this policy, please refer to the KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve>).

Researchers have the option to specify a time period during which their dataset will not be accessible to the public. During this embargo period, the dataset's descriptions, including title, author, metadata, and abstract, will still be viewable, but the dataset itself will not be available for public or moderated access. Once the embargo period expires, the dataset will be made available for public or moderated access. Additional details about the Embargo policy can be found in the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-7>).

The publication and sharing of research data should be organized in a way that is easy to understand, adheres to common standards, and includes an appropriate license to facilitate its reuse by other researchers. The Geo Big Data Open Platform encourages the use of licenses that enable broad reuse of research data and defaults to the standard Creative Commons License (CCL). Licenses used by the Geo Big Data Open Platform should be clear, standardized licenses that align with the FAIR principles and are in a machine-readable format. More detailed information on licenses and copyrights for the Geo Big Data Open Platform can be found in the KIGAM Ethics-Copyright-License Guide (<https://data.kigam.re.kr/manage/data/license>).

The Geo Big Data Open Platform does not have a specific preferred file format for research data sharing and reuse. However, there is a list of registered file formats available on the Geo Big Data Open Platform, which is outlined in the KIGAM Data Management Guide (<https://data.kigam.re.kr/manage/data/guide#item-4>). (See section R8)

The Geo Big Data Open Platform requires specific metadata to be provided. Any significant changes made to the metadata will be versioned within the Geo Big Data Open Platform, and a changelog record will be made available to users. (See section R7) The metadata in the Geo Big Data Open Platform is managed to ensure searchability, accessibility, interoperability, and reusability, both internally and externally, in alignment with the FAIR principles (<https://data.kigam.re.kr/manage/data/guide#item-2>).

The Geo Big Data Open Platform offers technical support for preservation purposes, including migration, emulation, and normalization, to accommodate future format advancements. The Geo Big Data Open Platform provides technical assistance for the long-term preservation of research data, as outlined in the KIGAM Data Preservation Guide (<https://data.kigam.re.kr/manage/data/preserve#item-4>).

Designated communities in the Geo Big Data Open Platform continuously share data and information. KIGAM researchers and users who have obtained permission to register can register datasets. Users consist of government (affiliated) organizations, local governments, public institutions, research institutes, schools, industrial companies, and general organizations, and are used for research or thesis purposes, and business purposes of local governments, public institutions, and industrial companies.

The Geo Big Data Open Platform offers guides on its website to assist users in efficiently utilizing its services for data reuse. These guides provide instructions on data requests, use case access, and open API applications (<https://data.kigam.re.kr/guide/dataset?lang=en>).

Links:

- [Ethics, Copyright, and License Guideline](#)
- [Data Management Guideline](#)
- [Data Management Guideline #7](#)
- [Data Preservation Guideline](#)
- [User guide of Geo Big Data Open Platform](#)
- [Data Management Guideline #2](#)
- [Data Management Guideline #4](#)
- [Data Preservation Guideline #4](#)

Information Technology & Security

R14 Storage & Integrity (R14)

R14. The repository applies documented processes to ensure data and metadata storage and integrity.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geo Big Data Open Platform utilizes the OAI-PMH, which is a globally recognized standard protocol for collecting and connecting metadata. This protocol ensures the comprehensive documentation of data and metadata, indicating the completeness of the information. Additionally, the platform adheres to the OAIS framework, which provides a conceptual model for establishing archives capable of effectively managing and preserving digital data over the long term. As part of this framework, data integrity checks are implemented to safeguard against unauthorized alterations or destruction of data, ensuring its integrity and reliability.

Geo Big Data Open Platform

The Geo Big Data Open Platform has specific procedures (<https://data.kigam.re.kr/manage/data/preserve#item-4>) for backup and recovery to prevent and protect research data from loss and damage. (See section R16)

Ensure integrity and accessibility when backing up data, and suggest how many copies of research data to store and how to synchronize them. Backup systems and processes include.

- The backup system of the Geo Big Data Open Platform consists of InnoStor Appliance (ISA-2000) and Quantum Scalar i500. It provides dual backup functionality, storing periodic backups of data from the service storage.
- The backup targets are data, database, and user data files of the Geo Big Data Open Platform.
- The backup cycle is daily backup of user data files, database, and system data, and full backup of user data, database, and research data (files) take place every Saturday.

The system software and application software utilize the local GitHub for recovery purposes. In the event of data loss or damage, the recovery of research data, databases, and user data files is carried out using the data stored in a backup device. If any issues arise with the backed-up data, data recovery is performed through tape backups located at the Vaulting location/site.

The Geo Big Data Open Platform has implemented a risk management policy and technical measures, as outlined in the 'explanatory note of Cloud Service IaaS Security Certification Criteria,' specifically in section 3.3 on Risk Management. This policy aims to anticipate and manage risks proactively, and it undergoes annual evaluations to ensure its effectiveness. The platform's retention policy is designed to maintain the technical sustainability of all contents within the repository (<https://data.kigam.re.kr/manage/data/preserve#item-4>). Additionally, the admin feature includes a data file check that detects corruption by comparing the current checksum with the checksum recorded when the digital object was initially saved. (See section R7)

Links:

- [Data Preservation Guideline #4](#)

R15 Technical Infrastructure (R15)

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

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geo Big Data Open Platform utilizes a three-tiered client/server architecture, consisting of multiple client and middleware components that manage the flow and quality of information.

At the server level, an RDBMS (PostgreSQL) is employed to store data. To optimize performance, large and binary data is stored in a standardized format on the hard disk.

For efficient access to the compiled data, a data warehouse is implemented. This data warehouse replicates all metadata to the middleware/frontend systems, enabling fast and seamless access as well as robust search functionalities.

The interface of the Geo Big Data Open Platform adheres to recognized standards such as W3C, ISO, and OGC. It utilizes web services, primarily REST, for various functionalities including mapping (Google Maps) and search (Apache Solr).

Metadata is extracted from the RDBMS in a specific format and stored in Apache Solr. The frontend system then transforms this metadata into different content standards, such as JSON-LD based on schema.org and Dublin Core XML. And data distribution is facilitated through HTTP services, OpenAPI, and OAI-PMH, allowing users to access and retrieve data using these protocols.

The Geo Big Data Open Platform is developed using various software tools such as iRODS, Apache Tomcat, Redis, PostgreSQL, Solr, and Nginx. These software components enable the platform to align with international standards. Additionally, the platform incorporates integrity inspection based on the OAIS Reference Model (See section R7).

The research data repository is designed to integrate metadata from other repositories using the OAI-PMH protocols. It utilizes the iRODS Federation function, which enables the sharing of source data directly through a high-capacity data repository.

All the standards employed in the Geo Big Data Open Platform are well-documented, and as part of the governance process, regular examinations are conducted to ensure compliance and adherence to these standards.

The Geo Big Data Open Platform is built on NaRDA version 3.1.0. When changes are made to a dataset, the platform ensures that the previous version of the dataset is preserved and saved.

To ensure the prevention and protection of research data from loss and corruption, the Geo Big Data Open Platform has implemented procedures for rapid recovery and backup of essential services in case of a repository outage. Regular and automatic backups of the data are performed in multiple locations, ensuring its safety and integrity. (See section R16)

*NaRDA is an IDR, NaRDA (National Research Data Archive, hereinafter NaRDA, hereinafter NaRDA. <https://idr.kisti.re.kr/>) has been developed and distributed to a total of 25 institutions (as of October 2022), including government-funded research institutes, government-affiliated research institutes, and universities (<https://post.naver.com/viewer/postView.nhn?volumeNo=31229238&memberNo=2355052&navigationType=push>). NaRDA was specifically designed and implemented to cater to the data management needs of researchers throughout the entire research data lifecycle, from the initial planning stage to the final publication of their research results. Its purpose is to provide comprehensive support and facilitate the seamless

Geo Big Data Open Platform

management of research data, ensuring its integrity, accessibility, and usability throughout the entire research process.

The Geo Big Data Open Platform receives reports on performance issues through the public cloud (KOREAV) and can increase infrastructure performance by adding costs if necessary.

The Geo Big Data Open Platform conducts quarterly regular inspections, four times a year, to monitor and manage the need for technology changes. During these inspections, system updates and reboots are performed to ensure optimal system performance. The KIGAM Data Preservation Guide provides detailed information on backup methods, including backup procedures for databases and file systems. It emphasizes the importance of maintaining data integrity and accessibility during the backup process. The guide also provides recommendations on the number of copies of research data to store and how to synchronize them. Furthermore, the platform applies data security requirements, access levels, and legal conditions based on the level of data disclosure(<https://data.kigam.re.kr/manage/data/preserve#item-5> & R16). (See section R16)

Links:

- [Data Preservation Guideline #5](#)
- [IDR \(Institutional Data Repository\)](#)
- [NaRDA \(National Research Data Archive\)](#)

R16 Security (R16)

R16. The repository protects the facility and its data, metadata, products, services, and users.

Compliance level:

Implemented: the requirement has been fully implemented by the repository - 1

Response:

The Geo Big Data Open Platform has established procedures for backup and recovery to safeguard research data against potential loss or damage (<https://data.kigam.re.kr/manage/data/preserve?lang=en#item-4>). These measures prioritize maintaining the integrity and accessibility of the backed-up data. The platform specifies the appropriate number of copies to be stored and determines the synchronization method to ensure data consistency. Furthermore, the Geo Big Data Open Platform applies data security requirements, access levels, and legal conditions based on the level of data disclosure, thereby upholding the necessary protections for sensitive information.

The backup system of the Geo Big Data Open Platform utilizes a dual backup setup comprising InnoStor Appliance (ISA-2000) and Quantum Scalar i500. Regular backups are conducted by periodically transferring data from the service storage to the backup system. The backup process covers various components, including data, databases, and user data files associated with the platform. The specific backup cycles are outlined as follows:

- Backup of user data files, databases, and system data: Daily
- Full backup of user data, database, and research data (files): once a week

System software and application software facilitate recovery from the local GitHub repository. In the event of issues with the backed-up data, recovery of research data, databases, and user data files is carried out using the data stored on backup devices. If necessary, data recovery is performed using tape backups at the designated Vaulting location/site. (See section R14)

The security of the Geo Big Data Open Platform is ensured through access control (<https://www.koreav.kr/product/security/zone.html>) implemented via the public cloud (koreav) firewall and web firewall. Data backup and recovery support (<https://www.koreav.kr/product/management/backup.html>) follow the policies set by the Public Cloud (koreav). Additionally, a backup copy site is available for the transmission system data stored in the cloud-based service. Within the Geo Big Data Open Platform, a designated security officer is responsible for ensuring the security of the public cloud (koreav) environment.

This officer maintains a channel of communication to promptly report any detected risks to agency and company representatives.

The Geo Big Data Open Platform ensures data safety through various measures. It conducts backups in a physically separated location, around 30km away from the public cloud (koreav) environment. Additionally, it has a disaster recovery center in place to mitigate the impact of potential disasters. The platform operates IDC center computer facilities to enhance its infrastructure resilience.

To enhance physical security, the Geo Big Data Open Platform implements a robust system. It installs access control devices, chip intrusion detection and alarm devices, and image information processing devices specifically designed for non-relatives. The details of the video information processing equipment operation and management policy can be found on the KIGAM platform (<https://www.kigam.re.kr/menu.es?mid=a10913010000>).

The Geo Big Data Open Platform includes a change log feature that records the history of log changes, including the time and user who made the changes. The platform categorizes users into different roles such as 'user', 'depositor', 'review researcher', 'approved researcher', and 'administrator'. Repository administrators have the ability to manage user rights and user group rights through the user management function. If a user without access requests to download a file, the repository administrator can review the request. Additionally, there is a data approval process in place for such requests. Research data providers and administrators have the ability to review data disclosure restrictions and set appropriate restriction levels.

Research data security in the Geo Big Data Open Platform is crucial to prevent unauthorized access, disclosure, or modification. The principal investigator is responsible for ensuring data security, with the level of security determined based on the nature of the data. Sensitive data requiring higher security measures may include personally identifiable information or pose risks to commercial, intellectual, or national security. The platform aims to protect the integrity and confidentiality of research data throughout its lifecycle.

The Geo Big Data Open Platform provides the following data security options:

- Controlling access to digital files through password protection or encryption.

Geo Big Data Open Platform

- Installing firewall and antivirus protection on every computer you use
- Identifying and ensuring the location of stored data
- Controlling access to locations and equipment where digital or physical data is stored

Access control in the Geo Big Data Open Platform is implemented through user-specific access rights. To manage access, the platform establishes member types and permission policies. Depending on the access level of each member, functional permissions are granted to utilize analytics, repositories, and data linkage management systems. (See section R2)

Links:

- [Data Preservation Guideline #4](#)
- [Security of Public cloud \(Koreav\)](#)
- [KIGAM's image information processing equipment operation and management policy](#)
- [Data backup and recovery support of Public cloud \(Koreav\)](#)

Applicant feedback

R17 Applicant Feedback

We welcome feedback on the CoreTrustSeal Requirements and the Certification procedure.

Compliance level:

In Progress: the repository is in the implementation phase - 0

Response:

N.A.

Links: