



Enhancing Findability and Searchability of Research Data: Metadata Conversion and Registration in Institutional Repositories

PRACTICE PAPER

MASAHITO NOSÉ

ATSUKI SHINBORI

YOSHIZUMI MIYOSHI

TOMOAKI HORI

TSUKASA OHIRA

JUNKO HASHIBA

CHIZUKO NAOE

RUI GAKIYA

MAIKO OKAMOTO

TAKESHI SAGARA

TAKAAKI AOKI

SHIGEKI MATSUBARA

ICHIRO TAKAHASHI

HIDEKAZU HAYASHI

KAZUNARI YAMADA

YASUYUKI
MINAMIYAMA

YOSHIMASA TANAKA

SHUJI ABE

SATORU UENO

SHUN IMAJO

YASUO SAITO

TAKUYA ASHIKITA

YUKO HORI

TOSHIYUKI SHIMIZU

NANAKO OKAMURA

KAORU HIRANO

LEE BARGATZE

ubiquity press

CORRESPONDING AUTHOR:

Masahito Nosé

School of Data Science,
Nagoya City University,
Nagoya, Japan

nose@ds.nagoya-cu.ac.jp

*Author affiliations can be found in the back matter of this article

ABSTRACT

This paper outlines our practice to enhance the findability and searchability of research data through metadata conversion from the Space Physics Archive Search and Extract (SPASE) schema to a more generic schema, the Japan Consortium for Open Access Repository (JPCOAR) schema, and registration of converted metadata in institutional repositories. Traditionally, earth and space science research data have been organized using the SPASE schema. Although the SPASE schema is comprehensive, its usage has been restricted to highly professional databases, limiting broader accessibility and impeding cross-disciplinary research. We discuss a case study at Nagoya University where 284 metadata records were converted from SPASE to JPCOAR, and illustrate the process and benefits of this conversion. This approach significantly improved the visibility and usability of metadata across various platforms like the Institutional Repositories DataBase (IRDB), the Data Catalog Cross-Search System, and Google's Dataset Search, extending access to a wider range of users beyond the highly professional scientific community. This approach also aligns with national policies in Japan on research data management and simplifies metadata handling for researchers. Future direction includes expanding this conversion and registration model to other universities and institutions by leveraging the ubiquity of the SPASE schema in the earth and space science fields. Our practice may be useful in other research fields. This initiative aims to improve the overall findability of research data, to foster cross-disciplinary collaboration, and to enhance the value of research data itself and of creators and managers of the data.

KEYWORDS:

Metadata conversion;
Findability and searchability;
Institutional repositories;
SPASE metadata schema;
JPCOAR metadata schema;
Research data management

TO CITE THIS ARTICLE:

Nosé, M, Shinbori, A, Miyoshi, Y, Hori, T, Ohira, T, Hashiba, J, Naoe, C, Gakiya, R, Okamoto, M, Sagara, T, Aoki, T, Matsubara, S, Takahashi, I, Hayashi, H, Yamada, K, Minamiyama, Y, Tanaka, Y, Abe, S, UeNo, S, Imajo, S, Saito, Y, Ashikita, T, Hori, Y, Shimizu, T, Okamura, N, Hirano, K and Bargatze, L 2024 Enhancing Findability and Searchability of Research Data: Metadata Conversion and Registration in Institutional Repositories. *Data Science Journal*, 23: 40, pp. 1–9. DOI: <https://doi.org/10.5334/dsj-2024-040>

In the field of earth and space science, metadata to describe the content of research data is generally created following the Space Physics Archive Search and Extract (SPASE) metadata schema (Roberts et al., 2018). The SPASE metadata schema is established by an international consortium including the National Aeronautics and Space Administration (NASA) and the Japan Aerospace eXploration Agency (JAXA). The consortium updates the schema periodically with the latest version being 2.6.0 as of April 2024. The SPASE metadata schema has been developed primarily for describing satellite observation data, but the latest version, 2.6.0 can also describe data obtained from ground-based observations, computer simulation, and modeling. The definition of the SPASE metadata schema is available on the SPASE website (<https://spase-group.org/>). Figure 1 shows the highest layer of the namespace diagram of the SPASE metadata schema. It features a multi-layered hierarchical structure designed to describe detailed information on research data generated through various methods. The hierarchical structure includes approximately 2,200 descriptive elements.

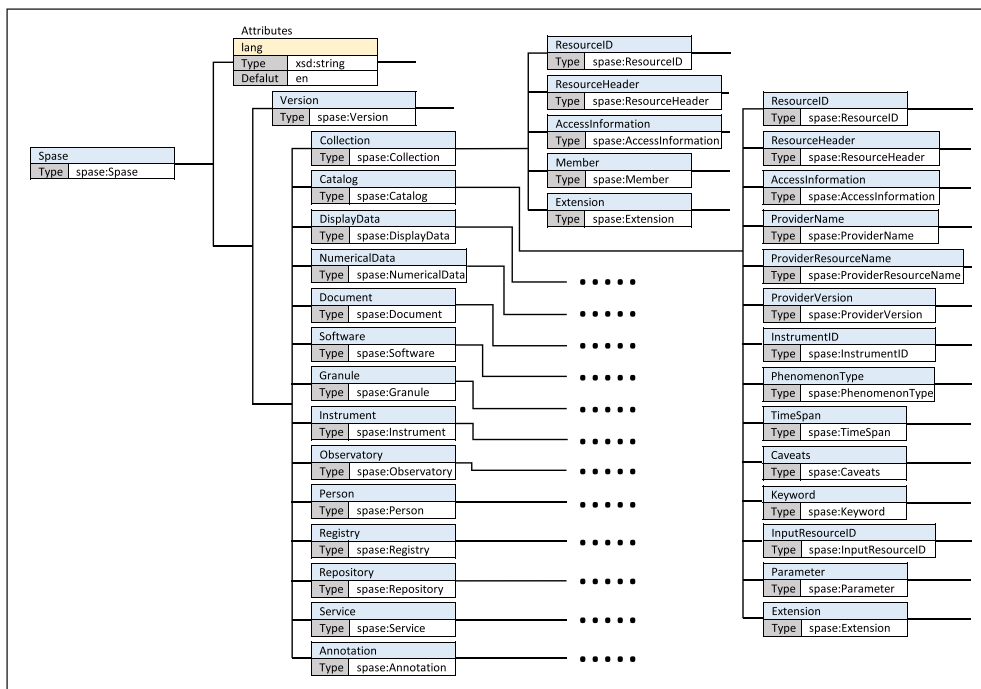


Figure 1 The highest layer of the namespace diagram of the SPASE metadata schema version 2.6.0.

One example of a metadata database created with the SPASE metadata schema is the Heliophysics Data Portal (<https://heliophysicsdata.gsfc.nasa.gov/>). This database contains about 3,000 entries of metadata, mainly for satellite observations. Researchers can use the search function on this website to obtain information regarding types of research data, periods of availability, data creators, and procedures for accessing the data.

Another example of a metadata database is that provided by the Inter-university Upper atmosphere Global Observation NETwork (IUGONET) project (<https://www.iugonet.org/>). This project was initiated in 2009 by seven institutions from five universities in Japan that promote the research of earth and space science based on ground-based observations (Hayashi et al., 2013; Tanaka et al., 2022). To date, metadata have been created according to the SPASE metadata schema for approximately 1,200 different types of observations relating to research data acquired and stored by each institution. This metadata database is publicly available from the IUGONET website (<http://search.iugonet.org/>) in the same manner as the Heliophysics Data Portal, allowing researchers to obtain information on the research data they need to conduct their studies.

As described above, in the field of earth and space science, metadata databases have been created, and research infrastructure has been developed so that users can easily obtain information on research data through the search of such databases. In this sense, the “Findability” aspect of the FAIR Principles (Wilkinson et al., 2016) has been greatly improved within this research field. However, the search for these metadata databases is still available

only through the highly professional websites of Heliophysics Data Portal and IUGONET. Thus, for the broader research community outside of the field and the general public, the ‘Findability’ has not increased. To promote interdisciplinary research and citizen science, the findability and searchability for research data need to be improved.

We attempted to solve this problem by (1) converting metadata described in the SPASE metadata schema to those described in a widely-used generic metadata schema and (2) registering the converted metadata in an institutional repository. Section 2 describes in detail the above two steps to improve the findability and searchability of research data. We also describe the results of this practice performed at Nagoya University for 284 metadata files. In Section 3, we report that, as a derivative effect, the periodic harvesting among metadata databases has broadened the findability and searchability of research data to a greater degree than expected. We also explain that this practice fulfills demands for research metadata management from policy-making authorities. Section 4 describes future prospects for further improvement of findability and searchability of research data.

2. PRACTICES FOR IMPROVING FINDABILITY AND SEARCHABILITY

2.1 CONVERSION FROM SPASE METADATA SCHEMA TO GENERIC METADATA SCHEMA

As a first step towards increasing discoverability, we used the schema developed by the Japan Consortium for Open Access Repository (JPCOAR) (<https://jpcoar.repo.nii.ac.jp/>) as a target generic metadata schema. As of April 2024, the JPCOAR metadata schema is version 2.0 and consists of 44 elements as shown in Figure 2 (<https://schema.irdb.nii.ac.jp/>). Because it is general-purpose, the JPCOAR metadata schema cannot contain detailed and professional information that can be described in the SPASE metadata schema, but it does include the minimum information necessary to access the data, such as title, creator, contributor, access rights, description, and identifier. We created a table of mapping items from the SPASE metadata schema to the JPCOAR metadata schema using an Excel table. Figure 3a is a part of the mapping table, showing which items in the JPCOAR metadata schema (green boxes) correspond to items within the SPASE metadata schema (blue boxes). The creation of this mapping table required discussions between researchers in the field of earth and space science who understand the contents of the SPASE metadata schema and library staff who understand the contents of the JPCOAR metadata schema. The completed mapping table was implemented with XSL Transformations (XSLT) with assistance from an information science engineer. Figure 3b shows the initial part of the XSLT file. The important thing was to have face-to-face meetings among the researchers in the field of earth and space science, the library staff, and the information science engineer. This practice could not be carried out if any one of them was missing.

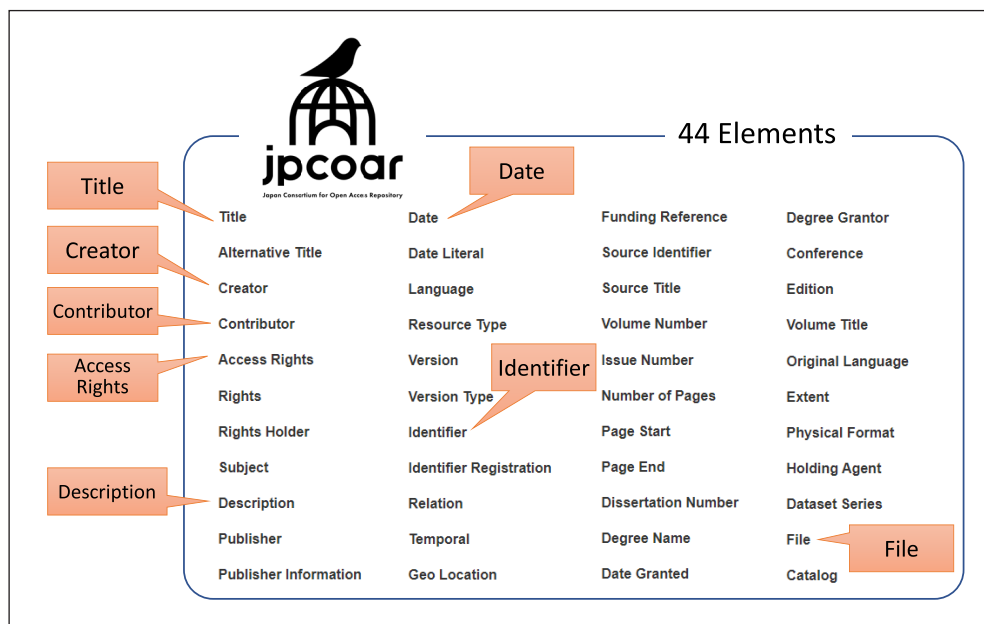


Figure 2 Forty-four elements for JPCOAR metadata schema version 2.0.

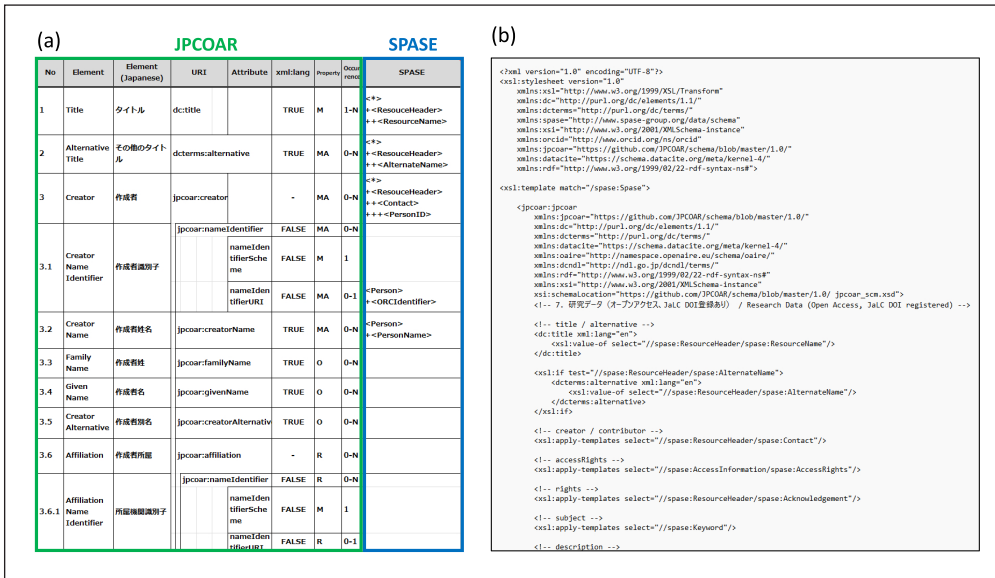


Figure 3 (a) A part of the mapping table between items in the JPCOAR metadata schema (green boxes) and items in the SPASE metadata schema (blue boxes). (b) An initial part of XSLT program implementing the mapping table of Figure 3a.

This XSLT file and metadata files based on the SPASE metadata schema were processed into xsltproc, an XSLT processor, to convert metadata files from SPASE to JPCOAR. We targeted metadata files that have been created and managed at Nagoya University, which amount to 284 out of approximately 1,200 metadata files within the IUGONET project.

2.2 REGISTRATION OF JPCOAR METADATA IN INSTITUTIONAL REPOSITORY

The 284 JPCOAR metadata files were registered in Nagoya University's institutional repository (NAGOYA Repository, <https://nagoya.repo.nii.ac.jp/>). The JPCOAR metadata schema is widely used in institutional repositories in Japan, and therefore, the registration process was simple. Figure 4 displays a web page showing a list of the registered metadata (upper-left) and a web page showing general information on a specific database that is accessed from the list (lower-right).

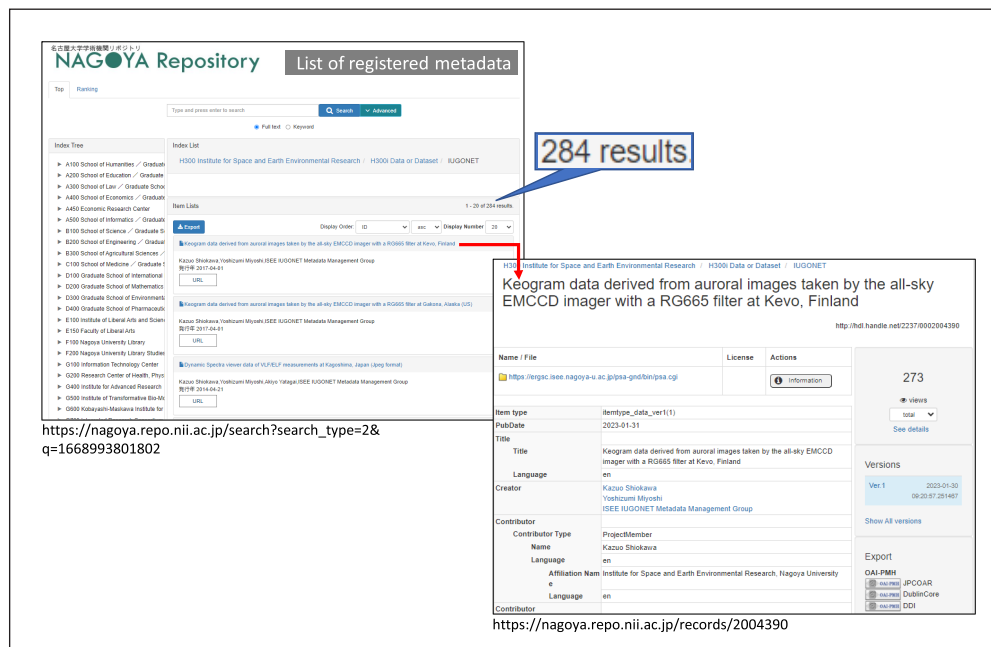


Figure 4 Screenshots of a NAGOYA Repository site showing a list of the 284 registered metadata (upper-left) and a NAGOYA Repository site showing general information on a specific dataset that is accessed from the list (lower-right).

Institutional repositories are familiar to researchers and general users for accessing scholarly information such as research bulletin papers, journal articles, dissertations, research reports, and various teaching materials. Making the information discoverable by a broader audience through the institutional repository greatly improved the findability and searchability of research data, versus being accessible only through the domain-specific IUGONET website.

3. DERIVED EFFECTS

3.1. IMPROVEMENT OF SEARCHABILITY IN WIDER RANGE OF FIELDS BY METADATA HARVESTING

Metadata registered in academic institutional repositories such as the NAGOYA Repository are automatically harvested on a regular basis by the Institutional Repositories DataBase (IRDB, <https://irdb.nii.ac.jp/>) operated by the National Institute of Informatics (NII). Then, the IRDB metadata database has been incorporated into the Data Catalog Cross-Search System, which collects information on data held and released by governmental agencies or local governments in response to the government's promotion of open data. The metadata can also be searched at the Data Catalog Cross-Search website interface (<https://search.ckan.jp/>). The metadata are also available through Google's Dataset Search (<https://datasetsearch.research.google.com/>).

Figure 5 shows an example of how information on research data registered in the NAGOYA Repository is found by searching on the other metadata databases described above, owing to the metadata schema conversion conducted in this practice. Figure 5a shows the information on aurora image data from Finland managed by the IUGONET project, which was converted from the SPASE metadata schema to the JPCOAR schema and registered in the NAGOYA Repository as described in Section 2. This information is regularly harvested among metadata databases such as IRDB (Figure 5b), the Data Catalog Cross Search System (Figure 5c), and Google's Dataset Search (Figure 5d), so that information on the Finnish aurora image data became available more broadly.

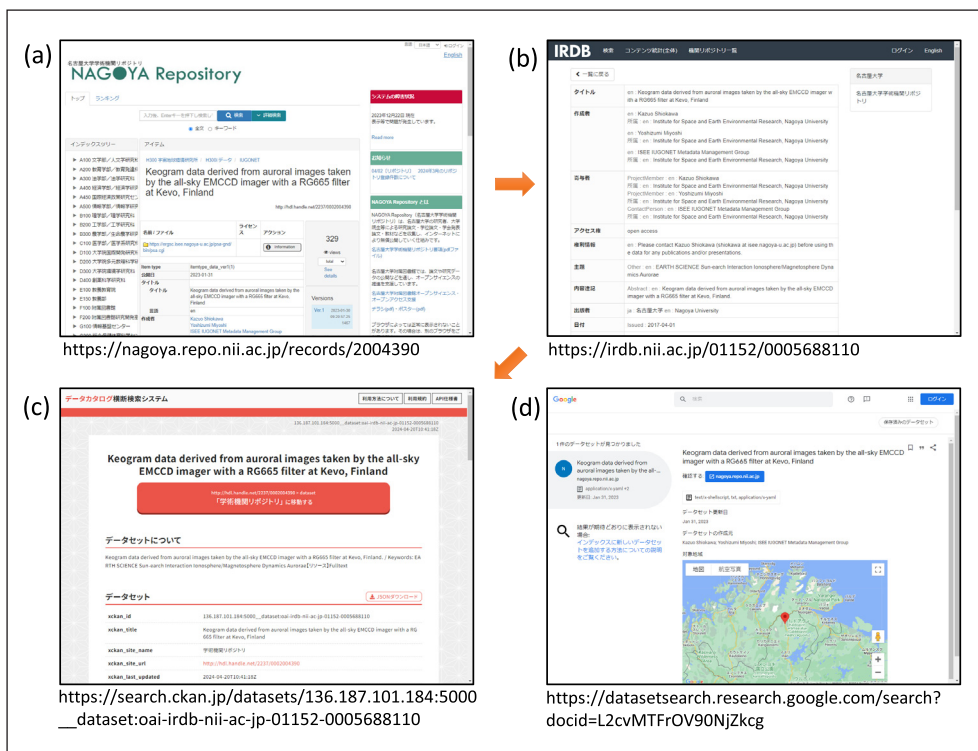


Figure 5 Examples of metadata information found in various metadata databases that are regularly and automatically collecting metadata from each other. (a) A web page of the NAGOYA Repository showing the information on aurora image data in Finland managed by the IUGONET project. (b) Metadata information on the aurora image data in Finland that is available from IRDB. (c) The same as Figure 5b, but available from the Data Catalog Cross-Search System. (d) The same as Figure 5b, but available from Google's Dataset Search.

Converting the SPASE metadata schema to a generic metadata schema and registering it in an institutional repository not only enables searches of the Nagoya University institutional repository, but also enables data to be found more easily over a wider area due to auto-harvesting among metadata databases.

3.2. HANDLING REQUESTS FROM POLICY-MAKING AGENCIES

In Japan, the *Basic Policy on the Management and Utilization of Publicly Funded Research Data* was published by the Council for Science, Technology and Innovation of the Cabinet Office in April 2021 (<https://www8.cao.go.jp/cstp/kenkyudx.html>). According to this policy, researchers are required to define the scope of research data to be managed based on the data policy which is defined by their institution and funding agencies for publicly solicited research funds,

and to register metadata of the research data so that they can be searched on a research data infrastructure system. The research data infrastructure system referred to here is the information infrastructure called ‘NII Research Data Cloud’ (<https://rcos.nii.ac.jp/service/>), which NII started full-scale operation in 2020. The metadata schema used in this infrastructure system is called ‘Common Metadata Items’ and consists of 15 items, including title, research field, creator name, and so on, as shown in Figure 6.

Figure 6 “Common Metadata Items” metadata schema that is defined by the information infrastructure called “NII Research Data Cloud”.

(As of 31 March 2023)		
Elements	Mandatory/Optional	Remarks
1 Funder	Mandatory*	English abbreviation of funding agencies (including ministries) providing public research funds. No entry required except for public research funds.
2 Funding Stream Code in Japan Grant Number	Optional	In the case of public research funds, the code for “Agency Code” and “Program Specific Code” in the Japan Grant Number. No entry required except for public research funds.
Program Name	Optional	The name of the competitive research funding program
3 Japan Grant Number	Mandatory*	In the case of public research funds, the Japan Grant Number assigned to each research grant. No entry required except for public research funds.
Project Name	Mandatory*	The name of the scope of R&D under the PM (e-Rad Project Name, etc.). If there is no official name, enter this according to the rules of the researcher’s affiliated institution.
4 Data No.	Mandatory	The number to uniquely identify managed data. Assigned by the funding agency providing public research funds. Assigned by the Hosting Institution (element 14) except for public research funds.
5 Title	Mandatory	Avoid ambiguous titles such as ‘Conference Materials’, ‘Report Materials’, ‘Measurement Results’, etc.
6 Date (Issued/Updated)	Mandatory	The date on which metadata was issued/updated
7 Description	Mandatory	Describe the contents in a simple and clear manner
8 Research Field	Mandatory*	e-Rad research field (main field). This will be automatically entered in cooperation with e-Rad.
9 Data Type	Mandatory	On the research data infrastructure system, the default is usually ‘dataset’; however, a type other than ‘dataset’ can be selected according to the characteristics of the data.
10 File Size	Optional	Describe by classifications such as less than 1GB, more than 1GB and less than 10GB, more than 10GB and less than 100GB, 100GB or more
11 Data Utilization and Provision Policy	Mandatory	Describe free/not free, license information, other conditions (how to cite, etc.)
Access Rights	Mandatory	Select from open access/ restricted access/ metadata only access/ embargoed access
Available Date	Mandatory	If the embargoed access is selected, enter the Available Date.
12 Repository Information	Mandatory	Current repository information or repository information after the project
Repository URL/DOI Link	Optional	Enter this if there is information
13 Creator Name	Optional	The name of the researcher who created the managed data.
Creator Name Identifier (e-Rad)	Optional	e-Rad Researcher Number of the creator of the managed data
14 Hosting Institution	Mandatory	The legal name of the R&D institution that manages each data
Hosting Institution Code	Optional	The code of the Hosting Institution
Data Manager	Mandatory	The name of the person who manages each managed data in the Hosting Institution
Data Manager Identifier (e-Rad)	Optional	e-Rad Researcher Number of the manager. No need to fill this in if the Data Manager does not have an e-Rad Researcher Number; Mandatory if he/she has, unless she/he wishes to keep the number private.
Contact of Data Manager	Mandatory	The address, telephone number, email address, etc. of the Data Manager’s affiliated institution
15 Remarks	Optional	

*In the case of publicly funded research activities

Courtesy of Bureau of Science, Technology and Innovation, Cabinet Office, Government of Japan
 (Taken from https://www8.cao.go.jp/cstp/ms_metadatainstructions_en.pdf)

For researchers, in terms of distribution of research data and contribution to the research community, it is more important to create detailed metadata based on the metadata schema used as a standard in their field (in the case of space and earth science, the SPASE metadata schema). On the other hand, creating new metadata according to the Common Metadata Items described above may complicate research data management because of the doubling of metadata management.

The metadata schema conversion and registration in institutional repositories introduced in Section 2 satisfies both the requirement of the policy-making agencies and unification of metadata management, because institutional repositories are included in the NII Research Data Cloud. Consequently, this will reduce the burden on researchers.

4. FUTURE PLANS

Figure 7 shows an overview of this practice and future plans. We have converted the metadata schemas for the 284 metadata files that have been created and managed at Nagoya University, and registered them in institutional repositories (Figure 7, Phase 1). Similar procedures can be applied to other universities and research institutes if they have metadata created with the SPASE metadata schema and institutional repositories (Figure 7, Phase 2). In the IUGONET project, in particular, each of the seven participating institutions creates and manages metadata based on the SPASE metadata schema, which will assist in making rapid progress. Kyushu University has created 180 metadata files. We are currently in the process of converting these metadata and registering them in Kyushu University’s institutional repository to further improve the findability and searchability of research data.

In addition, the SPASE metadata schema is often used in the field of earth and space science, so the XSLT file we have created (Figure 3b) is available on GitHub (<https://github.com/iugonet/SPASE-JPCOAR>) for wider use.

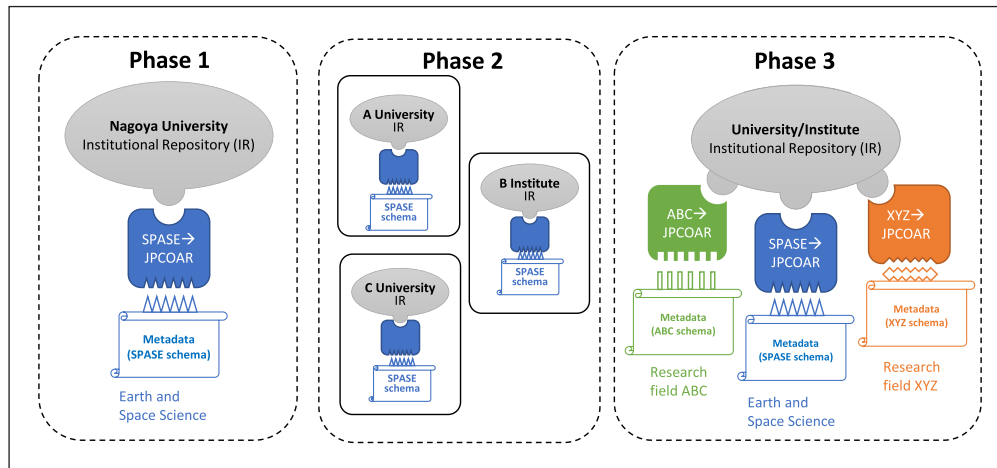


Figure 7 A schematic diagram of this practice and future plans. At the present time, we have converted the metadata schemas for the 284 metadata files that have been created and managed at Nagoya University, and registered them in institutional repositories (Phase 1). This conversion and registration procedure will be applied to other universities and research institutes where the SPASE metadata files and institutional repositories are available (Phase 2). The same procedure will be applied to the research fields other than earth and space physics, where metadata are created according to a domain-specific metadata schema (Phase 3).

Outside of the field of earth and space science, there are likely to be many metadata databases created on the basis of metadata schemas of specific fields. Such metadata databases may be difficult to notice and use by researchers outside of the fields. We can improve the overall findability of research data by converting a domain-specific metadata schema in each field to the JPCOAR metadata schema and registering the metadata in institutional repositories, as we did in this practice (Figure 7, Phase 3). We will contribute to the promotion of research data distribution and cross-disciplinary joint research by providing the knowledge and experience we have obtained in this practice. We also expect to contribute to an appreciation of the research data itself and of the creators and managers of the research data through the increased use of the research data.

ACKNOWLEDGEMENTS

We thank Ms. S. Tanaka for her cooperation in this work. We would also like to thank all those who have been involved with IUGONET in the past, other than the authors.

FUNDING INFORMATION

This study is supported by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (B) (21H01147), Grant-in-Aid for Fund for the Promotion of Joint International Research (Fostering Joint International Research(B)) (22KK0045), Grant-in-Aid for Scientific Research (A) (20H00099, 24H00164) as well as ROIS-DS-JOINT 2023 (043RP2023) and MEXT, 'Developing a Research Data Ecosystem for the Promotion of Data-Driven Science' (2024-01).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

MN is the primary author of the manuscript, and coordinated and participated extensively in the practice. He also contributed to drafting, reviewing, and editing the manuscript. AS, TO, RG, MO, TS participated practically in handling the SPASE metadata files, converting them into JPCOAR, and registering them in the NAGOYA repository. YM (Yoshizumi Miyoshi), TH, JH, CN, together with MN, AS, TO, RG, MO, TS, involved in creating the metadata mapping table and supported this practice. TA (Takaaki Aoki), SM, IT, HH, KY, YM (Yasuyuki Minamiyama) discussed

this practice at regular meetings. YT, SA, SU, SI, YS, LB provided information about the SPASE metadata schema and metadata database. TA (Takuya Ashikita), YH, TS, NO, KH, SA examined the possibility of implementing this practice at Kyushu University. All authors contributed to the manuscript by editing or commenting on the text and figures.

AUTHOR AFFILIATIONS

Masahito Nosé  orcid.org/0000-0002-2789-3588

School of Data Science, Nagoya City University, Nagoya, Japan

Atsuki Shinbori  orcid.org/0000-0002-6018-7431

Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan

Yoshizumi Miyoshi  orcid.org/0000-0001-7998-1240

Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan

Tomoaki Hori  orcid.org/0000-0001-8451-6941

Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan

Tsukasa Ohira  orcid.org/0000-0002-2514-1075

Nagoya University Library, Nagoya, Japan

Junko Hashiba

Nagoya University Library, Nagoya, Japan

Chizuko Naoe

Nagoya University Library, Nagoya, Japan

Rui Gakiya

Nagoya University Library, Nagoya, Japan

Maiko Okamoto

Nagoya University Library, Nagoya, Japan

Takeshi Sagara  orcid.org/0000-0002-2924-6739

Info Proto Co., Ltd., Tokyo, Japan

Takaaki Aoki

Information Strategy Office, Information and Communications, Nagoya University, Nagoya, Japan

Shigeki Matsubara  orcid.org/0000-0003-0416-3635

Information Technology Center, Information and Communications, Nagoya University, Nagoya, Japan

Ichiro Takahashi

Information Technology Center, Information and Communications, Nagoya University, Nagoya, Japan

Hidekazu Hayashi

Information Technology Center, Information and Communications, Nagoya University, Nagoya, Japan

Kazunari Yamada

Information Technology Center, Information and Communications, Nagoya University, Nagoya, Japan

Yasuyuki Minamiyama  orcid.org/0000-0002-7280-3342

Research Center for Open Science and Data Platform, National Institute of Informatics, Tokyo, Japan

Yoshimasa Tanaka  orcid.org/0000-0001-7276-4868

Polar Environment Data Science Center, Research Organization of Information and Systems, Tachikawa, Japan; National Institute of Polar Research, Tachikawa, Japan; Graduate Institute for Advanced Studies, SOKENDAI, Tachikawa, Japan

Shuji Abe  orcid.org/0000-0001-8270-0024

International Research Center for Space and Planetary Environmental Science, Kyushu University, Fukuoka, Japan

Satoru UeNo  orcid.org/0000-0002-2165-7998

Astronomical Observatory, Graduate School of Science, Kyoto University, Kyoto, Japan

Shun Imajo  orcid.org/0000-0002-9862-844X

Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science, Kyoto University, Kyoto, Japan

Yasuo Saito

Polar Environment Data Science Center, Research Organization of Information and Systems, Tachikawa, Japan

Takuya Ashikita  orcid.org/0000-0003-1853-7108

Kyushu University Library, Fukuoka, Japan

Yuko Hori  orcid.org/0000-0003-1469-9142

Kyushu University Library, Fukuoka, Japan

Toshiyuki Shimizu  orcid.org/0000-0003-1677-2366

Kyushu University Library, Fukuoka, Japan

REFERENCES

- Hayashi, H., Koyama, Y., Hori, T., Tanaka, Y., Abe, S., Shinbori, A., Kagitani, M., Kouno, T., Yoshida, D., UeNo, S., Kaneda, N., Yoneda, M., Umemura, N., Tadokoro, H., Motoba, T. and IUGONET project team.** (2013) 'Inter-university Upper Atmosphere Global Observation NETWORK (IUGONET)', *Data Science Journal* 12. DOI: <https://doi.org/10.2481/dsj.WDS-030>
- Roberts, D. A., Thieman, J., Génot, V., King, T., Gangloff, M., Perry, C., Wiegand, C., Zeeuw, D. D., Fung, S. F., Cecconi, N. and Hess, S.** (2018) 'The SPASE data model: A metadata standard for registering, finding, accessing, and using Heliophysics data obtained from observations and modeling', *Space Weather* 16(12), pp. 1899–1911. Available at: <http://doi.wiley.com/10.1029/2018SW002038>. DOI: <https://doi.org/10.1029/2018SW002038>
- Tanaka, Y., Umemura, N., Abe, S., Shinbori, A., and UeNo, S.** (2022) 'Advanced tools for guiding data-led research processes of upper-atmospheric phenomena', *Geoscience Data Journal* 10, pp. 130–141. DOI: <https://doi.org/10.1002/gdj3.170>
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al.,** (2016) 'The FAIR guiding principles for scientific data management and stewardship', *Scientific Data* 3, 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>

TO CITE THIS ARTICLE:

Nosé, M., Shinbori, A., Miyoshi, Y., Hori, T., Ohira, T., Hashiba, J., Naoe, C., Gakiya, R., Okamoto, M., Sagara, T., Aoki, T., Matsubara, S., Takahashi, I., Hayashi, H., Yamada, K., Minamiyama, Y., Tanaka, Y., Abe, S., UeNo, S., Imajo, S., Saito, Y., Ashikita, T., Hori, Y., Shimizu, T., Okamura, N., Hirano, K and Bargatze, L 2024 Enhancing Findability and Searchability of Research Data: Metadata Conversion and Registration in Institutional Repositories. *Data Science Journal*, 23: 40, pp. 1–9. DOI: <https://doi.org/10.5334/dsj-2024-040>

Submitted: 22 April 2024

Accepted: 11 June 2024

Published: 26 July 2024

COPYRIGHT:

© 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.