

Machine-actionable Data Management Plans Model Analysis and Improvement Direction

Suntae Kim*

Department of Library and Information Science,
Jeonbuk National University, Jeonju, Korea
E-mail: kim.suntae@jbnu.ac.kr

ABSTRACT

In this study, the RDA DMP Common Standard (RDCS), a data model for implementing a machine actionable Data Management Plan (maDMP), was analyzed in four aspects. First, the twelve class models proposed by RDCS were analyzed. Second, whether the DMP attribute was included in the class attribute was analyzed. Third, we analyzed the namespace used for RDCS properties. Fourth, the values and identifiers used in RDCS properties were analyzed. As a result of the analysis, four directions for improvement were derived. First, it is necessary to add an academic record class to describe information such as papers and reports, which are representative academic documents. Second, the primary research institution, responsibility, resources, option attribute, and additional attributes are needed to describe the researcher's affiliation information. Third, it is necessary to additionally use a namespace such as Friend of a Friend that can be used universally. Fourth, the use of digital object identifier should be considered to identify academic literature.

Keywords: machine-actionable DMP, data management plans, DMP, research data, research data management

Open Access

Received date: May 26, 2020
Accepted date: December 10, 2020

*Corresponding Author: Suntae Kim
Assistant Professor
Department of Library and Information Science, Jeonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju 54896, Korea
E-mail: kim.suntae@jbnu.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

Data are fundamental to the research lifecycle in all subject areas, both as an essential raw material and as a reusable output of research (Hodson & Molloy, 2014). So, more research funding agencies, institutions, journals, and publishers are introducing policies that encourage or require the sharing of research data that support publications (Colavizza, Hrynaszkiewicz, Staden, Whitaker, & McGillivray, 2020). With the spread of awareness that research data is a national asset, the above activities are continuously expanding. Data Management Plans (DMPs) were born in this context. DMPs are generally short, high-level descriptive plans that prescribe the data to be generated by a research project, how that data will be stored (securely, as required), who will have access, what documentation and metadata will be created with the data, and preservation intentions if the data are to be preserved long-term (Burnette, Williams, & Imker, 2016). According to the Regulation on the Management of National Research and Development Projects in Korea, DMP refers to the plan for production, preservation, management, and joint use of research data. The method defines the research data as data necessary for the verification of research results as fact data calculated through various experiments, observations, investigations, and analysis conducted in the course of conducting research and development tasks.

DMP is a tool that helps researchers manage data and increase the quality, accessibility, and reusability of data after the project is over. DMPs are created by researchers using checklists or online tools. However, static DMP created before the start of the project cannot be effectively used for actual data management activities. It is also common to recognize that DMP have only the function of a document for a document. In recognition of these problems, much attention is being paid to machine actionable DMPs (maDMPs) that can be recognized by machines and humans. Researchers, funders, repository managers, research managers, and data librarians are particularly interested in maDMPs as various stakeholders related to research data. Various information related to research data can be automatically generated through maDMPs. The generated information can be automatically shared systematically, reducing various management costs. It also has the advantage of improving the quality of DMP information.

In addition, various effects can be expected when maDMPs are common. If maDMPs are shared in public and under an open license, anyone can aggregate them, re-slice the corpora, use, and re-share the resulting information. Such front-ends to maDMP collections could be generic—which would help

with the standardization and spread of good data management practices across domains—or be tailored for specific audiences, e.g., to facilitate discovery in a given area or education about research in the domain, including associated data management practices (Miksa, Simms, Mietchen, & Jones, 2019). maDMPs can also provide a variety of benefits to funding institutions. According to Noh, Kwon, and Moon (2018), it is important for funding institutions to understand the impact of research and investment support programs in funding strategies, program design, and mission coordination. Funding institutions continue to challenge the performance of R&D investment programs and measure their impact.

To enjoy the various expected effects described above, a data model commonly used by various stakeholders surrounding DMP is required. Therefore, this study analyzes the RDA DMP Common Standard (RDCCS) presented in the DMP Common Standard WG of the Research Data Alliance (RDA) held in March 2019, from the aspect of the common data model, and aims to propose an improvement direction.

2. PREVIOUS STUDIES

Koo and Kim (2019) proposed a research records management plan that applied DMP in recognition of the importance of managing relational information between R&D development projects and research records. Kim (2020) derived the functional requirements of the research data repository by analyzing the items required in DMP and CoreTrustSeal. For reference, the ICSU World Data System and Data Seal of Approval create a CoreTrustSeal organization that certifies data repositories. Deposit, Ethics, License, Discovery, Identification, Reuse, Security, Preservation, Accessibility, Availability, and (Meta) Data Quality, commonly required by DMP and CoreTrustSeal, were derived as functional requirements that should be implemented first in implementing data repositories (Kim, 2020).

Meanwhile in Korea, various research activities of public institutions are being conducted in relation to identifiers. The National Library of Korea (2016) aims to upgrade its KOLIS-NET system to solve the problem that author identification metadata used by domestic organizations are different from each other. To manage and ultimately increase the efficiency of public information services, a model was developed for the interconnection of International Standard Name Identifier (ISNI) and domestic and international identifiers for integrated management of standard identification systems. The National Assembly Library is also promoting ISNI issuance of researchers

or creators as a domestic ISNI registrar by issuing individual author authority IDs (Byeon & Oh, 2018). The National Research Foundation of Korea (NRF) issues Researcher Registration Numbers to researchers, and the NTIS system operated by the Korea Institute of Science and Technology Information (KISTI) issues Researcher Numbers. The Researcher Registration Number and Researcher Number are identical identifiers that differ only in their names. In addition, KISTI is currently conducting research to identify researchers, articles, reports, and research institutions belonging to articles and reports written by Korean researchers.

The above previous studies are limited to suggesting concepts for maDMPs and principles for becoming maDMPs. However, practical studies have been conducted to implement maDMPs, and are specifically discussed through the RDA WG. At the RDA held in March 2019 there was the activity of DMP Common Standard WG, a working group on maDMP(RDA DMP Common Standard WG, 2020a, 2020b). A review of the metadata structure in JSON format based on the model being released and reviewed by the Lucidchart platform was conducted. According to working group members Miksa, Walk, and Neish (2019), this model (application profile) does not prescribe how information must be presented to the end user and does not enforce any specific logic on how this information must be collected or used. The application profile is an information carrier and full machine-actionability can only be achieved when systems using the application profile implement appropriate logic. The results of this study are practical examples of the common data model proposed by Miksa, Simms, et al. (2019) as the sixth of the ten principles for maDMPs.

3. RESEARCH METHODS AND LIMITATIONS

The purpose of this study is to analyze a practical data model that can implement maDMPs and to propose a model improvement direction. The RDCS proposed by RDA's DMP Common Standard WG was selected as the data model selected for analysis. RDA conducts various academic activities related to research data, and major funders, research institutes, and researchers from each country participate. Therefore, it was judged that the probability of spreading RDCS could be greater than any other data model.

Fig. 1 shows the research method used in this study. In this study, the RDCS data model was analyzed in four aspects. First, twelve class models suggested by RDCS were analyzed. Second, the comprehensiveness of RDCS properties was analyzed using DMP properties which were derived from previous studies.

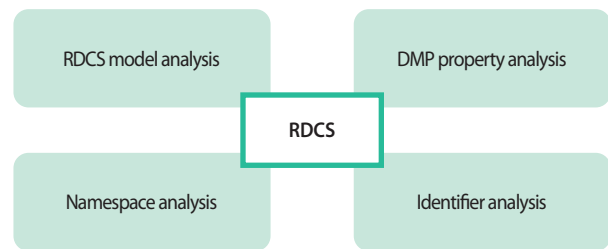


Fig. 1. Concept diagram of research method. RDCS, RDA DMP Common Standard; DMP, Data Management Plan.

Third, the namespace used in RDCS properties was analyzed. Fourth, the values and identifiers used in RDCS attributes were analyzed.

On the other hand, it can be pointed out that only RDCS was considered in this study. However, since various countries and stakeholder groups participate in RDA activities, it is considered that the results of this study can be used as meaningful basic data for subsequent studies related to maDMPs.

4. THEORETICAL BACKGROUND

4.1. maDMPs

The purpose of DMP is to consider various aspects such as data management, metadata creation, data preservation, and data analysis before the project starts. With DMP, current data can be well-managed and be ready for data retention in the future. Preparing a data management plan before data are collected ensures that data are in the correct format, organized well, and better annotated. There is now widespread recognition that the DMP can have more thematic, machine-actionable richness with added value for all stakeholders: researchers, funders, repository managers, research administrators, data librarians, and others (Miksa, Simms, et al., 2019). Further, maDMPs are DMPs written in a machine-readable format as well as for humans.

4.2. Persistent Identifiers

A persistent identifier (PID) is an ongoing, long-lasting digital reference to a resource. An identifier is a label which gives a unique name to an entity: a person, place, or thing. It is designed to always point towards the intended object and does so more reliably compared to URLs. Digital object identifiers (DOIs) are PIDs for entities such as journal articles, books, and datasets (CrossRef, 2020). maDMPs have emerged due to the need for free-form DMPs to be machine-readable to address the needs of various stakeholder groups related to research

data. Therefore, the use of standardized properties that describe the values of maDMPs is recommended, and the use of PIDs should be actively considered for describing related resources. Target resources that need to be identified include researchers, institutions, and content. At the time of this study, studies on global identifiers have been actively conducted, and the results are used in the academic ecosystem. Typical global identifiers include ISNI identifiers that identify people and institutions, Open Researcher and Contributor IDs (ORCID) that identify researchers, and DOIs that identify content.

ISNI is a 16-digit international standard name identifier that is assigned to identify individuals and organizations related to research and creative activities, such as writers, researchers, performers, and video producers. ISNI was established as an international standard in March 2012 by the International Organization for Standardization (ISO) TC46 (literature information technology committee) SC9 in March 2012. In Korea, it was established as the KS X ISO 27729 national standard in May 2018 (National Library of Korea, 2020). ORCID issues an identifier that is identified by 16 digits for the researcher. In addition, ORCIDs are issued to organizations, and they are supported to register their research results. ISNI and ORCID operate differently, and they operate a separate identification system. However, cooperation between the two organizations began in 2013. According to ORCID (2013), ORCID and ISNI have already taken a first collaborative step in defining system interoperability. The ORCID ID is compatible in format with the ISNI ISO Standard (ISO 27729). The ORCID registry randomly assigns ORCID IDs from a block of numbers set aside for them by the ISNI International Agency, which

avoids having the same number assigned to different people.

The DOI® system provides an infrastructure for persistent unique identification of objects of any type. DOI is an acronym for “digital object identifier,” meaning a “digital identifier of an object” rather than an “identifier of a digital object” (International DOI Foundation [IDF], 2019). DataCite, one of the IDF agencies that support the academic ecosystem to assign DOI to digital data objects, is undertaking significant activity in the data age where data is used as a key tool in research. DataCite is a leading global non-profit organization that provides PIDs (DOIs) for research data and other research outputs.

5. maDMPs PROPERTY ANALYSIS

In this study, twelve classes of RDCS examined in previous studies were analyzed. In addition, the comprehensiveness of RDCS properties was analyzed by using the DMP property derived from Kim’s (2020) study. Next, the namespace used in the RDCS attribute was analyzed, and finally, the property’s value and identifier used in the RDCS attribute were analyzed.

5.1. RDCS Model Analysis

The working group for maDMP, DMP Common Standard WG, presented twelve classes related to maDMP as shown in Fig. 1. Classes suggested in the model are composed of DMP, Project, Funding, Cost, Contact, Contributor, Dataset, Distribution, Host, License, Security and Privacy, Technical Resource, and Metadata as shown in Fig. 2. DMP information

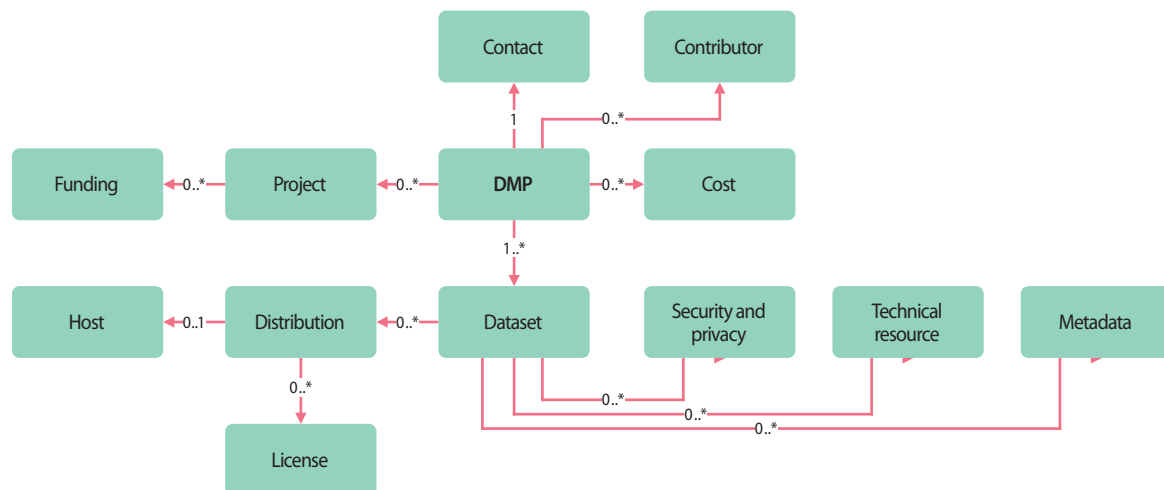


Fig. 2. Machine actionable Data Management Plan (maDMP)-diagram. Adapted from <https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/diagrams/maDMP-diagram.png>.

can relate to zero or more project information, and one project can relate to zero or more funding information. DMP can be related to zero or more contributor information and has one contact information. DMP can relate to zero or more cost information. DMP can be associated with more than one dataset. A dataset can be related to zero or more distribution information. Distribution information can be related to zero or more single host information and can have zero or more license information. A dataset can be related to zero or more Security & Privacy information, Technical Resource information, and Metadata information.

Colavizza et al. (2020) conducted a study to automatically tag Data Availability Statements to provide other researchers with data to verify the content of the published paper. Detailed

descriptions of the data are available through literature such as papers and reports. In other words, the most detailed description of the data is possible through relevant academic literature. Viewers can see examples of global publishers such as Elsevier (<https://www.sciencedirect.com>) and data repositories like Pangaea (<https://www.pangaea.de>) already linked to each other. Fig. 3 shows an example where academic and data services are intertwined. This means that the researcher obtains the paper of interest from Elsevier and additionally obtains research data related to the paper from a data service such as Pangaea. Of course, the reverse process is the same. Therefore, the scenario is possible to search for DMP, and acquire dataset described in DMP. Next, we obtain papers and reports, which are academic literature related to the dataset.



Fig. 3. Data repository (Pangaea) and academic thesis (Elsevier) link service case.

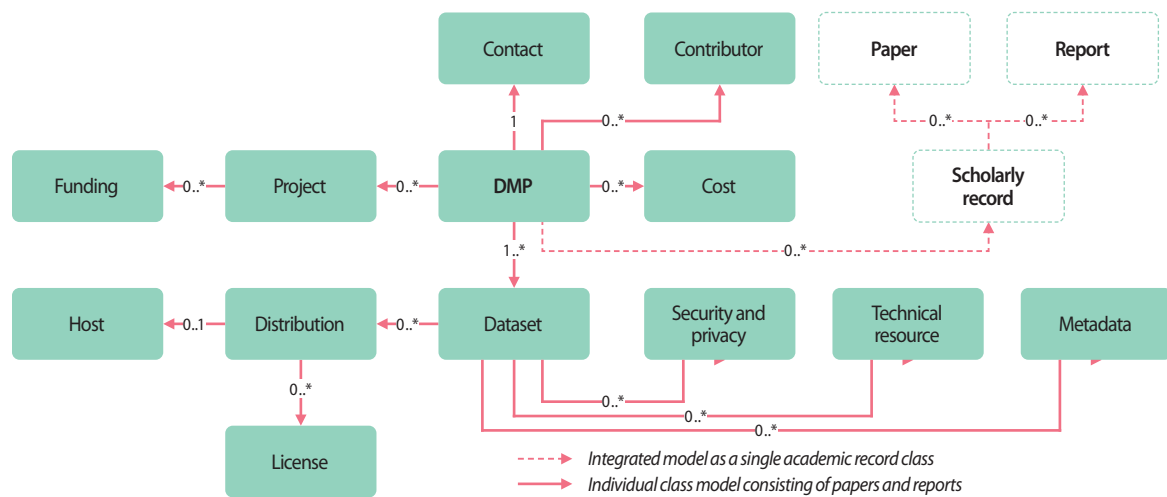


Fig. 4. Machine actionable Data Management Plan (maDMP)-diagram which is extended for the bibliography.

As described earlier, DMPs record various information about research data generated during the research process. maDMPs are developed by systems of various stakeholder groups related to research data in order to exchange understandable information with each other in a promised format. Therefore, it is necessary to exchange information on academic literature that contains abundant information on data among the information exchanged. Through this, the reuse of data described in DMPs can be ensured, and the availability of data can be ensured.

To enable the above scenario, it is necessary to add papers and reports, which are representative academic literature, to the maDMP diagram presented by RDCS.

Fig. 4 shows the addition of papers and reports to the maDMP diagram presented by RDCS. It is considered that the necessary sub-elements for the Paper and Report elements basically require identifier information to identify the resource.

5.2. DMP Property Analysis

Table 1 below shows the crosswalk mapping the RDCS attribute, the DMP attribute suggested in Kim (2020), and DC terms. DC terms refer to metadata terms defined in DCMI (Dublin Core Metadata Initiatives). The RDCS property column shows the properties of twelve classes suggested by DMP Common Standard WG, a working group for maDMP. The “Kim (2020) property” column shows DMP properties derived from Kim’s research process. The “DC terms” column shows terms suggested by DCMI (<https://www.dublincore.org/specifications/dublin-core/dcmi-terms>).

To derive the functional requirements of the research data repository, Kim (2020) analyzed the DMP of research fund support organizations registered in DMP Online, which is provided by the Digital Curation Centre, to derive the DMP attributes. As a result of comparing the derived attributes, most of the attributes were mapped to the RDCS attribute. However, the “Primary research organization” attribute in the project information did not exist to be mapped among the RDCS attributes. In addition, in the Responsibilities and Resources information, the Responsibilities attribute and the Resources attribute were not mapped among RDCS attributes. In addition, among the properties for describing Selection and Preservation information, there was no object to which the Selection property is mapped. Meanwhile, in the academic information distribution environment, identification of people, institutions, and contents is very important. It is possible to develop a service that links resources through identifiers for identifying each resource. In addition, it is possible to analyze the researcher’s relationship, the relationship between two distinct content

objects, and research trends by using precisely identified data. Through this work, more valued content can be produced, and new services can be provided to researchers. ORCID, ISNI, Author Identifier, Researcher ID, etc. may be used to identify the researcher. However, if there is no such identifier, the researcher’s email address or organization may be used as a key element to identify the researcher. Therefore, RDCS needs a property to describe the information of a researcher’s affiliated organization participating in the project. Researchers can belong to multiple institutions at any given time. Therefore, it is considered that the design of the organization entity connected to the contributor entity needs to be further reviewed. In addition, the organization entity needs an attribute to describe primary research organization information and an identifier attribute to identify the institution. The DMP attributes Responsibilities, Resources, and Selection derived from the Kim (2020) study are also required.

5.3. Namespace Analysis

To become a maDMP, it is necessary to develop properties that can be understood by machines. However, it is advantageous to secure interoperability by using terms that are already developed and used universally, rather than developing new properties. In this study, the namespace used in the properties of twelve classes suggested in the RDCS model was investigated. Among the properties presented as properties of the Project class and the Host class, the <title> and <descript> properties are indicated using the ‘dct’ namespace. Also, the class containing the <identifier> attribute uses the ‘dct’ namespace. Also, among the RDCS properties, the <contact> property and the <DMStaff> property use <mail> and <name> properties as sub-properties. In this case, it was judged that the meaning of the property can be clarified by using the universally used FOAF (<http://www.foaf-project.org>) namespace. As a result of analysis, it was determined that the language, created, and modified attributes of the DMP class and the type, subject, and issued attributes of the Dataset class can use the ‘dct’ namespace. Also, among the RDCS properties, the <contact> property and the <DMStaff> property use <mail> and <name> properties as sub-properties. In this case, it was judged that the meaning of the property can be clarified by using the universally used Friend of a Friend namespace.

On the other hand, regarding metadata of academic documents such as papers and reports, the schema used by CrossRef (<http://crossref.org>) can be applied to include metadata in maDMPs. To publish DOI in CrossRef, it is considered appropriate to include necessary metadata elements as basic elements.

Table 1. RDCS property and Kim (2020)'s property & DC terms mapping

RDCS property	Kim (2020) property	DC terms	RDCS property	Kim (2020) property	DC terms
DMP > title	Plan name	title	Dataset > title	n/a	title
DMP > description	Description	description	Dataset > description	n/a	description
DMP > dmp_id > identifier	ID	identifier	Dataset > type	Data Type	type
DMP > language	n/a	language	Dataset > keyword	n/a	subject
DMP > created	n/a	created	Dataset > issued	n/a	issued
DMP > modified	n/a	modified	Dataset > Metadata > metadata_id > identifier	Metadata	identifier
Project > dct:title	Project title	title	Dataset > Metadata > description	Documentation	description
Project > dct:description	n/a	description	Dataset > Distribution	Data Sharing Method	-
Project > projectStart	n/a	date	Dataset > Distribution > title	n/a	title
Project > projectEnd	n/a	date	Dataset > Distribution > description	n/a	description
Project > funding > funder_id > identifier	Funding organisation	identifier	Dataset > Distribution > format	Data Format	format
Project > funding > grant_id > identifier	Grant number	identifier	Dataset > Distribution > byteSize	Data Volume	extent
Contact > contact_id > identifier	Principal Investigator/ Researcher	identifier	Dataset > Distribution > access_url	Access and Security	n/a
Contact > mail	Plan data contact	n/a	Dataset > Distribution > download_url	Access and Security	n/a
Contact > name	Plan data contact	n/a	Dataset > Distribution > data_access	Access and Security	accessRights
DMStaff (= contributor) > staff_id > identifier	n/a	identifier	Dataset > Distribution > available_till	Access and Security	n/a
ethical_issues_description	n/a	description	Dataset > Distribution > Host	Data Sharing Method	n/a
Cost > title	n/a	title	Dataset > Distribution > Host > title	n/a	title
Cost > description	n/a	description	Dataset > Distribution > Host > description	n/a	description
Dataset > dataset_id > identifier	n/a	identifier	Dataset > Distribution > Host > host_id	n/a	identifier
Dataset > language	n/a	language	Dataset > Distribution > Host > supports_versioning	Storage and Backup	n/a
Dataset > personal_data	protect the identity of participants	n/a	Dataset > Distribution > Host > backup_type	Storage and Backup	n/a
Dataset > sensitive_data	sensitive data	n/a	Dataset > Distribution > Host > backup_frequency	Storage and Backup	n/a
Dataset > preservation_statement	consent for data preservation and sharing / Preservation	n/a	Dataset > Distribution > Host > storage_type	Storage and Backup	n/a
Dataset > Security and Privacy > title	n/a	title	Dataset > Distribution > Host > availability	Access and Security	n/a
Dataset > Security and Privacy > description	protect the identity of participants	description	Dataset > Distribution > Host > geo_location	n/a	location
Dataset > TechnicalResource > technical_resource_id	n/a	identifier	Dataset > Distribution > License > license_ref	copyright/Intellectual Property Rights (IPR)	accessRights
Dataset > TechnicalResource > description	n/a	description	Dataset > Distribution > License > start_date	copyright/Intellectual Property Rights (IPR)	date
Dataset > Metadata	Metadata				

RDCS, RDA DMP Common Standard; DC, Dublin Core; n/a, not available.

Table 2. Identifiers used in RDCS

RDCS property	Entity type	RDCS property	Entity type
dmp_id > identifier	doc.	technical_resource_id	material
funder_id > identifier	org.	metadata_id > identifier	metadata
grant_id > identifier	doc.	host_id > identifier	org.
contact_id > identifier	person	Paper_id > identifier ^{a)}	paper
staff_id > identifier	person	Report_id > identifier ^{a)}	report
dataset_id > identifier	data		

RDCS, RDA DMP Common Standard.

^{a)}Thesis and report identifiers required to identify academic literature required for RDCS.

5.4. Identifier Analysis

PIDs used in RDCS are shown in Table 2. PIDs are used to identify people and organizations. In addition, PID properties are used to identify resources such as content, system, DMP, grant, metadata, dataset, technical resource, and hosting system. The dmp_id property is used to reference the DMP, and the funder_id property is used to reference the research fund support institution. The grant_id property is used to reference research fund information, and contact_id is used to reference DMP-related contact information. The staff_id property is used to reference the staff related to the research data, and the dataset_id is used to reference the dataset. The technical_resource_id property is used to reference the technical resource, and the metadata_id property is used to reference the metadata. The host_id property is used to identify the system storing the research data.

On the other hand, regarding metadata of academic documents such as papers and reports, the schema used by CrossRef can be applied to include metadata in maDMPs. To publish DOI in CrossRef, it is considered appropriate to include necessary metadata elements as basic elements.

6. CONCLUSIONS AND SUGGESTIONS

In this study, we analyzed RDCS being developed as a practical data model that can be used as a reference for implementing maDMPs. RDA conducts various academic activities related to research data, and major funders, research institutes, and researchers from each country participate in RDA activities. Therefore, it was judged that the probability of spread of RDCS is very high. In this study, the RDCS model was analyzed in four aspects. The first, to analyze twelve class models suggested by RDCS; the second, to analyze whether RDCS properties include DMP properties; the third, to analyze

namespaces used in RDCS properties; and the fourth, to analyze values and identifiers used in RDCS properties, were carried out.

As a result of analyzing the classes of the model suggested by RDCS, the classes presented in the model are composed of DMP, project, funding, cost, contact, contributor, dataset, distribution, host, license, security and privacy, technical resource, and metadata. The main purpose of the DMP is to systematically manage the research data collected and generated during the research process and to verify the research results. This is achieved through the reuse of research data. For reuse, various context information related to research data is required, and this information is described in detail in academic literature such as papers and reports. Therefore, in this study paper and report classes, which are representative academic literature, were added to the RDCS model. In actual implementation, one scholarly record class can also be implemented.

As a result of analyzing whether the RDCS properties include DMP properties, most of the properties suggested by Kim (2020) by examining the DMP registered in DMP Online were mapped to the RDCS properties. However, among the properties for describing the information of “Primary research organization,” Responsibilities, Resources, and Selection and Preservation, the Selection property was not mapped. Therefore, DMP attributes for describing unmapped information were proposed. In addition, since the researcher’s email address or affiliated organization can be used as a key element to identify the researcher, RDCS needs properties to describe the researcher’s affiliated organization information participating in the project.

As a result of analyzing the namespace used in the RDCS property, the properties of the class including the Project, Host class, and Identifier properties use the ‘dct’ namespace. The language, created, and modified attributes of the DMP class and the type, subject, and issued attributes of the Dataset class can also use the ‘dct’ namespace. Also, it was judged that mail and name, which are sub-attributes of Contact and DMStaff classes, can be described by applying the universally used FOAF namespace.

As a result of analyzing the values and identifiers used in RDCS properties, RDCS uses PID properties to identify resources such as people and organizations, content, systems, DMP, grant, metadata, dataset, technical resources, and hosting system. As a result of analyzing the RDCS model in the first analysis step of this study, it was proposed to add information of academic literature related to the dataset to the model. Accordingly, it was proposed to use DOIs, which are used universally as an identifier for identifying academic documents such as papers and reports.

As a follow-up to this study, additional research on the organization entity connected to the contributor entity is required to contain the researcher's plural organization information in maDMPs. In addition, in this study namespaces were searched for terms suggested by DCMI, but in the future it is necessary to further investigate and analyze terms presented by Schema.org and FAOF.

This study analyzed the RDCS model for maDMPs to improve the convenience of various interest groups surrounding the study data and suggested an improvement direction. It is expected that the content proposed in this study will be of great help in the development of maDMPs to be implemented in the future.

ACKNOWLEDGMENTS

This paper was supported by research funds for newly appointed professors of Jeonbuk National University in 2018.

REFERENCES

- Burnette, M. H., Williams, S. C., & Imker, H. J. (2016). From plan to action: Successful data management plan implementation in a multidisciplinary project. *Journal of eScience Librarianship*, 5(1), e1101.
- Byeon, H. -K., & Oh, B. (2018). A study on integrated management system of researcher identifiers based on the ISNI (International Standard Name Identifier). *Journal of the Korean BIBLIA Society for Library and Information Science*, 29(3), 139-155.
- Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., & McGillivray, B. (2020). The citation advantage of linking publications to research data. *PLoS ONE*, 15(4), e0230416.
- Crossref. (2020). *Persistent identifiers*. Retrieved August 12, 2020 from <https://www.crossref.org/education/metadata/persistent-identifiers>.
- Hodson, S., & Molloy, L. (2014). *Current best practice for research data management policies*. Retrieved August 12, 2020 from <https://apo.org.au/node/58192>.
- International DOI Foundation. (2019). *DOI® handbook*. Retrieved August 12, 2020 from https://www.doi.org/doi_handbook/1_Introduction.html.
- Kim, S. -T. (2020). Functional requirements of data repository for DMP support and CoreTrustSeal authentication. *International Journal of Knowledge Content Development & Technology*, 10(1), 7-20.
- Koo, C. M., & Kim, S. -H. (2019). Research records management in scientific research institutes by applying DMP. *Journal of Korean Society of Archives and Records Management*, 19(1), 1-21.
- Miksa, T., Simms, S., Mietchen, D., & Jones, S. (2019). Ten principles for machine-actionable data management plans. *PLOS Computational Biology*, 15(3), e1006750.
- Miksa, T., Walk, P., & Neish, P. (2019). *RDA DMP Common standard for machine-actionable data management plans*. Retrieved August 12, 2020 from <https://www.rd-alliance.org/group/dmp-common-standards-wg/outcomes/rda-dmp-common-standard-machine-actionable-data-management>.
- National Library of Korea. (2016). *Information strategy planning for advancement of national resources union catalog: Final report*. Seoul: National Library of Korea.
- National Library of Korea. (2020). *About ISNI*. Retrieved August 12, 2020 from <https://www.nl.go.kr/isni/about/isniIntroduce>.
- Noh, K. -R., Kwon, O. -J., & Moon, Y. -H. (2018). Utilization of ORCID data to track the performance impact of national R&D investment. *Proceedings of the Korea Contents Association*, 2018, 405-406.
- ORCID. (2013). *ORCID and ISNI issue joint statement on interoperation*. Retrieved August 12, 2020 from <https://orcid.org/blog/2013/04/22/orcid-and-isni-issue-joint-statement-interoperation-april-2013>.
- RDA DMP Common Standard WG. (2020a). *ex3-dataset-finished.json*. Retrieved August 12, 2020 from <https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/examples/JSON/ex3-dataset-finished.json>.
- RDA DMP Common Standard WG. (2020b). *ex1-header-fundedProject.json*. Retrieved August 12, 2020 from <https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/examples/JSON/ex1-header-fundedProject.json>.