

Deskripsi Metadata dalam Manajemen Data Penelitian pada Sistem Repositori Ilmiah Nasional

¹Seno Yudhanto, ²Nina Mayesti

¹Departemen Ilmu Perpustakaan dan Informasi, Fakultas Ilmu Pengetahuan Budaya, Universitas Indonesia

²Pusat Data dan Dokumentasi Ilmiah, Lembaga Ilmu Pengetahuan Indonesia

E-mail: seno.yudhanto91@ui.ac.id

Abstract

Organizing research data is very important for data and information managers through a research data management mechanism (research data management/MDP) in a repository system. In this mechanism, research data must be organized and described as an effort to provide access. One important aspect of organizing is the availability of metadata. This Study was supported by the Institute of Sciences of Indonesia (LIPI) and the SAINTEK Scholarship from the Ministry of Research and Technology/National Research and Innovation Agency of the Republic of Indonesia (KEMENRISTEK/BRIN) in 2020 and it's purpose is to identify and describe metadata standards and metadata elements used in research data management in the National Scientific Repository (RIN) system. This study uses a qualitative approach with a case study method. Sources of data come from literature / document studies and direct observation. The results of the study show that the RIN system adopts descriptive metadata from three main standards, they are DublinCore, DataCite, and DDI. As a medium for describing research data in general, the metadata sections provided by the RIN system in the dataset folder are quite large and complete. Of the 35 metadata fields available in the dataset folder in this system, the three metadata standards complement each other with an adaptation of the dominant DDI standard with 32 metadata fields. However, the fields that are available can also be found in other standards, such as the title, subject, or keyword fields that are also found in the DublinCore and DataCite standards. Thus, the metadata fields provided in the RIN system is good enough and sufficient for research data management needs.

Keywords: Metadata; Repository; Research Data; Dataverse

Abstrak

Pengorganisasian data penelitian sangat penting dilakukan oleh para pengelola data dan informasi melalui mekanisme research data management

Tik Ilmeu : Jurnal Ilmu Perpustakaan dan Informasi

IAIN Curup | p-issn: 2580-3654; e-issn:2580-3662

DOI: 10.29240/tik.v5i1.2486

(manajemen data penelitian/MDP) dalam suatu sistem repositori. Dalam mekanisme tersebut, data penelitian harus diorganisir dan dideskripsikan sebagai upaya dalam penyediaan akses. Salah satu aspek penting dalam pengorganisasian adalah tersedianya metadata. Kajian ini didukung oleh Lembaga Ilmu Pengetahuan Indonesia (LIPI) dan Beasiswa SAINTEK dari Kementerian Riset dan Teknologi/Badan Riset dan Inovasi Nasional Republik Indonesia (KEMENRISTEK/BRIN) tahun 2020 dan bertujuan untuk mengidentifikasi dan mendeskripsikan standar metadata dan elemen metadata yang digunakan pada manajemen data penelitian di sistem Repositori Ilmiah Nasional (RIN). Kajian ini menggunakan pendekatan kualitatif dengan metode studi kasus. Sumber data berasal dari studi literatur/dokumen dan observasi langsung. Hasil kajian menunjukkan bahwa Sistem RIN mengadopsi metadata deskriptif dari tiga standar utama, yaitu DublinCore, DataCite, dan DDI. Sebagai media pendeskripsian data penelitian secara umum, ruas metadata yang disediakan sistem RIN pada folder dataset sudah cukup banyak dan lengkap. Dari 35 ruas metadata yang tersedia dalam folder dataset di sistem ini, ketiga standar metadata tersebut saling melengkapi dengan adaptasi dari standar DDI yang dominan dengan 32 ruas metadata. Meskipun demikian, ruas-ruas yang tersedia juga terdapat di standar yang lain, seperti ruas judul, subject, atau keyword yang terdapat pula di standar DublinCore dan DataCite. Dengan demikian, ruas metadata yang disediakan pada sistem RIN sudah cukup baik dan mencukupi kebutuhan pengelolaan data penelitian.

Kata Kunci: Metadata; Repositori; Data Penelitian; Datavers

A. PENDAHULUAN

Pada era digital seperti saat ini, tidak menampik bahwa “tsunami data” terjadi dalam kegiatan penelitian. Fokus pada pusat-pusat data dan informasi seperti perpustakaan saat ini tidak hanya pada koleksi konvensional, melainkan juga pada koleksi digital, dimana salah satunya adalah data primer penelitian. Data penelitian merupakan salah satu koleksi yang menjadi perhatian utama dalam pengelolaannya di semua pusat data dan informasi (Tang & Hu, 2019). Dalam dokumen Peraturan Kepala LIPI Nomor 12 Tahun 2016 menyebutkan bahwa data penelitian (data primer) adalah kumpulan fakta objektif yang dapat berbentuk teks, gambar, rekaman suara atau video, bilangan atau simbol dan dapat dijadikan sebagai dasar empiris untuk penelitian. Data penelitian yang bersifat primer membutuhkan upaya yang lebih dalam proses produksinya di kegiatan penelitian, sehingga

keberadaannya sangat fundamental dan membutuhkan pengorganisasian yang baik dan berkualitas.

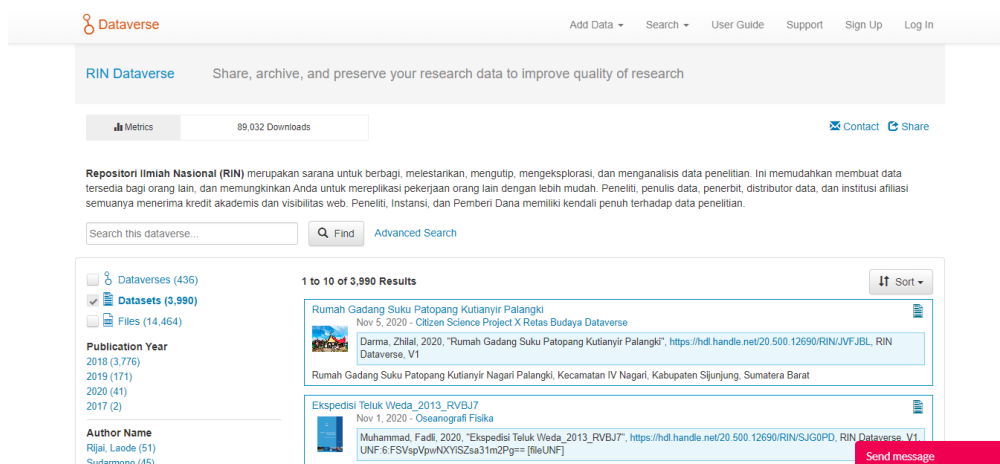
Pengorganisasian koleksi ini sangat penting dilakukan oleh para pengelola data dan informasi melalui mekanisme *research data management* (manajemen data penelitian/MDP) dalam suatu sistem repositori. MDP dianggap dan diakui secara internasional sebagai faktor kunci dalam perkembangan ilmu pengetahuan (Redkina, 2019). MDP adalah konsep luas yang mencakup proses yang dilakukan untuk membuat data penelitian berkualitas yang terorganisir, terdokumentasikan, dapat diakses, dan dapat digunakan kembali (Miller et al., 2018).

Dalam proses MDP tersebut terdapat subproses pengorganisasian, dimana data penelitian deskripsikan untuk dapat diakses. Salah satu aspek penting dalam pengorganisasian adalah metadata. Metadata, yang terdiri dari metadata struktural, administratif, dan deskriptif, dapat mendukung pengguna dalam menemukan dataset dan memungkinkan pengguna mengetahui apa yang ada di dalamnya (Allen, 2020). Dalam proses kurasi dan penggunaan kembali data, metadata berfungsi untuk memastikan bahwa data penelitian dapat ditemukan, diakses, digunakan, disimpan, disebarluaskan, dan dilacak dengan baik untuk waktu jangka panjang (Lee & Jeng, 2019; Rousidis et al., 2014). Pada skema metadata umum, suatu sistem dapat memeriksa dan mengindeks objek data yang disimpan dan membuat katalog konten yang dikumpulkan di berbagai repositori terkait. Hal ini akan memungkinkan objek data tersedia secara detail bagi calon pengguna, sehingga mereka dapat memutuskan apakah mereka dapat mengaksesnya (Canham & Ohmann, 2016).

Penelitian terdahulu mengenai metadata untuk data penelitian adalah (Farnel & Shiri, 2014) dengan tujuan menguji standar dan format metadata yang digunakan oleh sejumlah layanan data penelitian, yaitu Datacite, Dataverse Network, Dryad, dan FigShare. Temuan studi ini adalah bahwa ada berbagai elemen metadata yang digunakan oleh layanan data penelitian dan penggunaan kosakata terkontrol adalah hal yang umum di seluruh layanan. Adapun (Rousidis et al., 2014) mengidentifikasi masalah kualitas data yang terkait dengan metadata yang digunakan dalam repositori Dryad. Temuan penelitian adalah banyaknya elemen metadata yang tidak sesuai dan tidak konsisten dalam pengisiannya, sehingga berpengaruh terhadap analisis kuantitatif terhadap metadata, seperti untuk keperluan *data mining*. (Gregg et al., 2019) mengidentifikasi tantangan, peluang, dan kesenjangan pengetahuan yang berkaitan dengan penggunaan metadata dalam komunikasi ilmiah. Mereka mengumpulkan dan menafsirkan literatur dalam beberapa bagian

berdasarkan kelompok profesional, atau pemangku kepentingan, dalam metadata komunikasi ilmiah: peneliti, penyandang dana, penerbit, pustakawan, penyedia layanan, dan kurator data.

Di Indonesia, beberapa studi telah mengkaji mengenai metadata sistem repositori (Darmawiguna & Purnamawan, 2016; Pertiwi et al., 2017; Rachmat C., 2014), namun belum ada kajian yang meneliti mengenai analisis standar metadata pada sistem repositori khusus data penelitian. Untuk itu, diperlukan suatu studi analisis deskriptif mengenai metadata untuk repositori data penelitian. Berdasarkan pemaparan di atas, kajian ini bertujuan untuk mengidentifikasi dan mendeskripsikan standar metadata dan elemen metadata yang digunakan pada manajemen data penelitian di sistem Repositori Ilmiah Nasional (RIN). Ruang lingkup kajian ini membatasi pada elemen metadata deskriptif pada sistem RIN. Penelitian ini adalah penelitian kualitatif dengan metode studi kasus. Sumber data berasal dari studi literatur/dokumen dan observasi langsung. Studi literatur digunakan dengan mengidentifikasi literatur/dokumen yang terkait dengan metadata untuk data penelitian dan dokumen teknis dari sistem RIN dan penyedia standar metadata, sedangkan observasi langsung dengan mengidentifikasi metadata yang digunakan pada sistem repositori ilmiah nasional. Daftar ruas metadata yang digunakan kemudian disajikan, diidentifikasi dan dianalisis secara deskriptif.



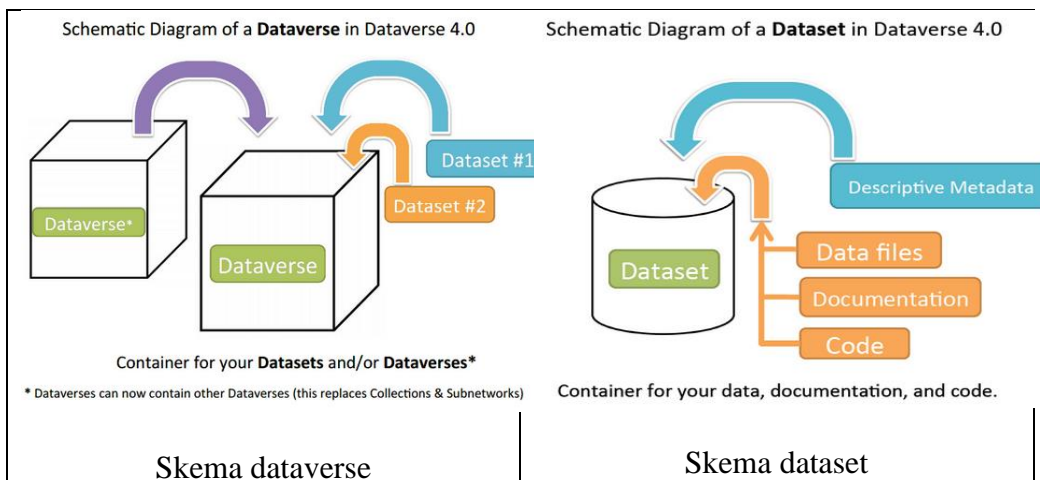
Gambar 1. Sistem Repositori Ilmiah Nasional (data.lipi.go.id)

B. HASIL DAN PEMBAHASAN

Sistem RIN merupakan sarana manajemen data penelitian nasional yang dikelola oleh Pusat Data dan Dokumentasi Ilmiah – Lembaga Ilmu

Pengetahuan Indonesia (PDDI LIPI). Sistem RIN menggunakan *platform open source* Dataverse. *Platform* ini dikembangkan oleh The Institute for Quantitative Social Science at Harvard, suatu lembaga penelitian di bidang ilmu sosial di Universitas Harvard. Dataverse dapat terdiri dari beberapa folder dataverse, dimana setiap folder dataverse adalah arsip virtual yang menyimpan data yang dihasilkan oleh seorang peneliti, replikasi data untuk jurnal, arsip dari keseluruhan proyek penelitian, atau semua data yang dibuat oleh disertasi suatu institusi. Setiap dataverse berisi studi, yang dapat diatur dalam kumpulan studi yang terkait, dan setiap studi berisi bidang metadata atau katalog serta file data dan file pendukung, seperti dokumentasi, *codebook*, kode, dan file tambahan (Crosas, 2012).

(Riyanto et al., 2019) menganalisis kelebihan platform Dataverse sebagai sistem pengelola data penelitian dibandingkan *platform* sejenis, seperti Dryad, Zenodo, CKAN, Galaxy, dan Geonode. Menurut mereka, *platform* Dataverse memiliki pengaturan hak akses terhadap data dan mekanisme kurasi data. Hal ini bertujuan agar data penelitian yang didepositkan dalam sistem memiliki kualitas yang baik dan dapat dipertanggungjawabkan.



Gambar 2. Skema folder pada platform Dataverse

Sumber: Dataverse Userguide: <https://guides.dataverse.org/en/4.17/user/index.html>

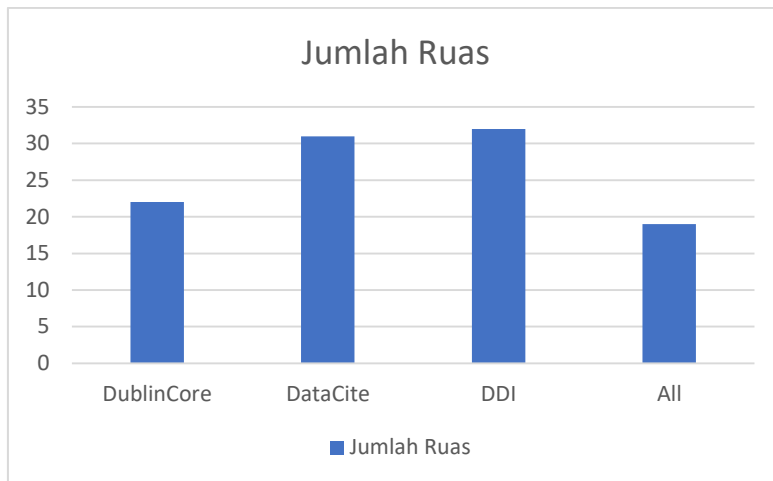
Pada sistem RIN yang berbasis Dataverse, hirarki penyimpanan terbagi menjadi dua skema, yaitu folder dataverse dan folder dataset. Folder Dataverse adalah wadah kumpulan dari beberapa Dataset dan juga dapat berisi kumpulan Dataverse (sub-dataverse), sedangkan folder Dataset adalah

wadah untuk menyimpan dan mengelola data, dokumentasi, kode, dan deskripsi metadata yang menjelaskan Dataset tersebut (Nashihuddin et al., 2020). Melalui sistem ini, skema pengorganisasian metadata deskriptif terdapat pada bagian folder dataset.

Sistem RIN-Dataverse mengadopsi elemen metadata berdasarkan tiga standar metadata utama, yaitu DublinCore, DataCite, dan Data Documentation Initiative (DDI) (Dataverse Team, 2016). Hasil analisis perbandingan ketiga standar metadata dapat dilihat pada tabel 1. Standar metadata tersebut juga diadopsi oleh sistem repositori lainnya, seperti FigShare, Zenodo, dan Dryad.

- DublinCore, merupakan standar metadata yang paling banyak digunakan (Xie & Matusiak, 2016). Standar ini pertama kali dibuat pada tahun 1995 dan hingga saat ini memiliki 15 area deskripsi utama, antara lain *contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title*, dan *type* (The Dublin Core Metadata Initiative, n.d.).
- DataCite, merupakan daftar properti metadata inti yang dipilih untuk identifikasi sumber daya yang akurat dan konsisten untuk tujuan kutipan dan pengambilan, bersama dengan petunjuk penggunaan yang disarankan. Standar metadata DataCite memiliki 19 area deskripsi utama yang terdiri dari *identifier, creator, title, publisher, publication year, subject, contributor, date, language, resource type, alternate identifier, related identifier, size, format, version, rights, geolocation*, dan *funding reference* (Smaele et al., 2016).
- DDI, merupakan standar internasional untuk mendeskripsikan data yang dihasilkan oleh survei dan metode pengamatan lainnya dalam ilmu sosial, perilaku, ekonomi, dan kesehatan. DDI adalah standar gratis yang dapat mendokumentasikan dan mengelola berbagai tahapan dalam siklus hidup data penelitian, seperti konseptualisasi, pengumpulan, pemrosesan, distribusi, penemuan, dan pengarsipan. DDI dapat memfasilitasi pemahaman, interpretasi, dan penggunaan, baik oleh orang, sistem perangkat lunak, maupun jaringan computer (Data Documentation Initiative, 2012).

Berdasarkan identifikasi pada tabel 1, adopsi ruas metadata terbanyak berasal dari standar metadata DDI dengan 32 ruas (91,4%), diikuti dengan DataCite dengan 31 ruas (88,6%), dan terakhir DublinCore dengan 22 ruas (62,9%). Disamping itu, ruas metadata yang mencakup ketiga standar elemen ini berjumlah 19 ruas (52,2%).



Gambar 3. Grafik jumlah ruas yang diadopsi
Sumber: hasil olah penulis

Pada standar DDI, ruas-ruas yang terdapat pada RIN Dataverse adalah *Title, Sub Title, Alternative Title, Other ID, Author, Contact, Description, Subject, Keywords, Topic Classification, Related Publication, Notes, Producer, Production Date, Production Place, Grant Information, Distributor, Distribution Date, Depositor, Deposit Date, Time Period Covered, Date of Collection, Kind of Data, Series, Software, Related Material, Related Datasets, Other References, Data Sources, Origin of Sources, Characteristic of Sources Noted, dan Documentation and Access to Sources*. Ruas metadata yang diadopsi dari DataCite adalah *Title, Sub Title, Alternative Title, Alternative URL, Other ID, Author, Contact, Description, Subject, Keywords, Topic Classification, Related Publication, Language, Producer, Production Date, Production Place, Contributor, Grant Information, Distributor, Distribution Date, Depositor, Deposit Date, Date of Collection, Kind of Data, Series, Software, Related Material, Related Datasets, Other References, Data Sources, dan Origin of Sources*. Terakhir, berdasar elemen yang tersedia pada standar metadata DublinCore, ruas metadata yang digunakan dalam RIN Dataverse adalah *Title, Alternative Title, Other ID, Author, Description, Subject, Keywords, Topic Classification, Related Publication, Language, Production Date, Production Place, Contributor, Deposit Date, Time Period Covered, Date of Collection, Kind of Data, Software, Related Datasets, Other References, Data Sources, dan Origin of Sources*. Ruas-ruas metadata yang mencakup ketiga standar elemen ini terdiri atas *Title, Alternative Title, Other ID, Author, Description, Subject, Keywords, Topic Classification, Related Publication, Production Date, Production Place, Deposit Date, Date of Collection, Kind of Data,*

Software, Related Datasets, Other References, Data Sources, Origin of Sources.

Sebagai media pendeskripsian data penelitian secara umum, ruas metadata yang disediakan sistem RIN pada bagian dataset sudah cukup banyak dan lengkap. Metadata yang lebih banyak dapat mencerminkan keragaman, sehingga akan membuat sumber informasi lebih dapat ditemukan dan digunakan, dimana ini merupakan tujuan akhir dari metadata (Alemu, 2016). Meskipun demikian, sangat penting untuk memperhatikan kualitas dari setiap metadata yang digunakan.

C. KESIMPULAN

Sistem RIN yang dikelola PDDI LIPI merupakan sistem repositori pengelolaan data penelitian yang diperuntukan secara nasional. Sistem ini menggunakan platform Dataverse Project yang diinisiasi oleh The Institute for Quantitative Social Science, Universitas Harvard. Sistem RIN Dataverse mengadopsi dari tiga standar metadata utama, yaitu DublinCore, DataCite, dan DDI. Dari 35 ruas metadata yang tersedia dalam sistem repositori ini, adaptasi terbanyak berasal dari standar DDI dengan 32 ruas metadata. Meskipun demikian, ruas-ruas yang tersedia juga terdapat di standar yang lain, seperti ruas judul, subject, atau keyword yang terdapat pula di standar DublinCore dan DataCite. Dengan demikian, ruas metadata yang disediakan pada sistem RIN sudah cukup baik dan mencukupi kebutuhan pengelolaan data penelitian. Kajian ini adalah kajian awal dari mengenai metadata pada sistem repositori data penelitian yang ada di Indonesia. Penting untuk menindaklanjuti studi lain yang terkait dengan bidang ini, seperti kesiapan pengelola data penelitian (kurator data, pustakawan, atau arsiparis) dalam menyajikan metadata yang berkualitas.

DAFTAR RUJUKAN

- Alemu, G. (2016). A Theory of Metadata Enriching and Filtering: Challenges and Opportunities to Implementation. *Qualitative and Quantitative Methods in Libraries (QQML)*, 5(2), 311–334. <http://www.qqml-journal.net/index.php/qqml/article/view/343>
- Allen, R. B. (2020). Metadata for Social Science Datasets. In J. I. Lane, I. Mulvany, & P. Nathan (Eds.), *Rich Search and Discovery for Research Datasets: Building the Next Generation of Scholarly Infrastructure* (pp. 40–52). Sage Publishing.
- Canham, S., & Ohmann, C. (2016). A metadata schema for data objects in

- clinical research. *Trials*, 17(1), 1–11. <https://doi.org/10.1186/s13063-016-1686-5>
- Crosas, M. (2012). A Data Sharing Story. *Journal of ESscience Librarianship*, 1(3), 173–179. <https://doi.org/10.7191/jeslib.2012.1020>
- Darmawiguna, I. G. M., & Purnamawan, K. (2016). Rancang Bangun Sistem Informasi Repositori UNDIKSHA dengan Metadata Dublin Core Berbasis Web: Studi Kasus FTK UNDIKSHA. *Seminar Nasional Vokasi Dan Teknologi (SEMNASVOKTEK)*, 185–193.
- Data Documentation Initiative. (2012). *DDI-Codebook 2.5*. <https://ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/codebook.xsd>
- Dataverse Team. (2016). *Dataverse Documentation* (4.2.4).
- Farnel, S., & Shiri, A. (2014). Metadata for Research Data: Current Practices and Trends. *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 74–82.
- Gregg, W. J., Erdmann, C., Paglione, L. A. D., Schneider, J., & Dean, C. (2019). A Literature Review of Scholarly Communications Metadata. *Research Ideas and Outcomes*, 5, 1–23. <https://doi.org/10.3897/rio.5.e38698>
- Lee, J.-S., & Jeng, W. (2019). The Landscape of Archived Studies in A Social Science Data Infrastructure: Investigating the ICPSR Metadata Records. *Proceedings of the Association for Information Science and Technology*, 56(1), 147–156. <https://doi.org/10.1002/pra2.62>
- Miller, C. F., Miller, R. S., & Phillips, G. A. (2018). *Keeping Up With... Research Data Management*. http://www.ala.org/acrl/publications/keeping_up_with/rdm
- Nasihuddin, W., Yudhanto, S., Surapermana, A. S., & Rishadi, R. (2020). Manajemen Data Penelitian Dengan Dataverse: Best Practice Pustakawan Menggunakan Sistem Repositori Ilmiah Nasional LIPI. *LIBRARIA: Jurnal Perpustakaan*, 7(2), 331–362. <https://doi.org/10.21043/libraria.v7i2.6508>
- Pertiwi, A. N. F. I., Subroto, I. M. I., & Assegaf, B. (2017). Penerapan Standar Metadata Dublin Core (DC) dan Open Archive Iniatif (OAI) di Fakultas Teknologi Industri UNISSULA. *TRANSISTOR EI: Jurnal Elektro Dan Informatika*, 2(1), 21–30. <http://jurnal.unissula.ac.id/index.php/EI/article/view/1630>

- Rachmat C., A. (2014). Analisis Rancang Bangun Sistem Repositori Institusi Berbasis Metadata Dublin Core di UKDW Yogyakarta. *Ultima InfoSys : Jurnal Ilmu Sistem Informasi*, 5(2), 65–74. <https://doi.org/10.31937/si.v5i2.267>
- Redkina, N. S. (2019). Current Trends in Research Data Management. *Scientific and Technical Information Processing*, 46(2), 53–58. <https://doi.org/10.3103/S0147688219020035>
- Riyanto, S., Subagyo, H., Marlina, E., Yaniasih, Y., & Triasih, H. (2019). Bagaimana Menghubungkan Publikasi Ilmiah Dengan Data Penelitian? *Baca: Jurnal Dokumentasi Dan Informasi*, 40(1), 41–54. <https://doi.org/10.14203/j.baca.v40i1.485>
- Rousidis, D., Garoufallou, E., Balatsoukas, P., & Sicilia, M.-A. (2014). Metadata for Big Data: A preliminary Investigation of Metadata Quality Issues in Research Data Repositories. *Information Services and Use*, 34(3–4), 279–286. <https://doi.org/10.3233/ISU-140746>
- Smaele, M. de, Dasler, R., Ashton, J., Martínez, I. B., Burger, M., Fenner, M., Habermann, T., Ilik, V., Jacobson, M., Raugh, A., Roi, A. la, Roy, S., Yahia, M., & Zolly, L. (2016). *DataCite Metadata Schema for the Publication and Citation of Research Data (Version 4.3)* (4.3). <https://doi.org/https://doi.org/10.14454/7xq3-zf69>
- Tang, R., & Hu, Z. (2019). Providing Research Data Management (RDM) Services in Libraries: Preparedness, Roles, Challenges, and Training for RDM Practice. *Data and Information Management*, 3(2), 84–101. <https://doi.org/10.2478/dim-2019-0009>
- The Dublin Core Metadata Initiative. (n.d.). *DCMI Metadata Terms*. Retrieved January 4, 2021, from <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#http://purl.org/dc/elements/1.1/source>
- Xie, I., & Matusiak, K. K. (2016). Discover Digital Libraries: Theory and Practice. In *Discover Digital Libraries: Theory and Practice*. Elsevier.