

The Re3data.org: Reservoir of Open Access to Research Data

Rushmansaab Gurikar*, Gururaj S. Hadagali**

Abstract

Open data and openness is a phenomenal paradigm shift in the knowledge society. The purpose of the present paper is to identify the open data repositories especially re3data.org. The authors used re3data.org registry as the source for the present study. Re3data.org is a registry which indexes the data repositories, i.e. mainly open data repositories. There were a total of 2327 repositories found during the study. The results of the study indicate that the data repositories consist the research data largely on Science and Technology discipline; the European and North American countries are the major contributors to the data repositories; the repositories indexed in re3data use variety of software; the data sets are mainly available in English language and the Dublin core metadata standard is widely used by the research data repositories.

Keywords: Open Data, Open Access, Data Repository, Research Data, Data Management, Re3data

Introduction

The researchers and knowledge seekers are increasingly more aware of the need for more reliable and proficient data access and sharing across all fields (Borgman, 2012). Open data and openness is a phenomenal paradigm shift in the knowledge society. Research data plays an imperative role in scientific progress of any nation. Further, the research is also generating and adding the data to the pool of existing research data. Hence, managing the research data has become a major challenge to the research organisations in order to provide the data to the right

clientele (Pinfield et al., 2014). Library and Information Centres are so involved in recent past in managing the research data (Cox & Pinfield, 2013). Many Research Institutions have established data centres in order to collect, compile, analyse, organise, store and disseminate the research data. However, much of the research data generated by these institutions are not accessible to every individual unless they subscribe it.

The scientific progress always suffers due to the limitation and/or restriction on access to research data. The renowned scientist C.N.R. Roa (India) opines that 'Data is like Viswaroopadarshan in the Universe'. It explains the importance of data in research. The International Council of Scientific Unions (ICSU) realised the creation of an International Committee to mobilise and share the scientific and technical research data among the scientific community at the international level. As a result, the 'Committee on Data for Science and Technology (CODATA)' was created. The scientific community acclaims that the success of research depends on how you present and use the relevant research data. The open access to research data helps every individual to access and use data (www.codata.org). R3data.org is such a repository which allows individuals to use the research data.

Open Access to Research Data

The Research data became a vital part of scientific documentation (Schopf et al., 2014). The Research data are valuable and unique for science progress.

* Research Scholar, Department of Library and Information Science, Karnatak University, Dharwad, Karnataka, India.
Email: rrgurikar@gmail.com

** Assistant Professor, Department of Library and Information Science Karnatak University, Dharwad, Karnataka, India.
Email: gururajhadagali123@gmail.com

The everlasting access to research statistics is a major challenge for all the stakeholders in the scholarly communication. The long term preservation and open access to research data opens a wide range of opportunities to the research community. The open repositories are of a great help to the scientific community in this endeavour. The Public Library of Science (PLOS) states that “PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception.” The European Commission (Kroes, 2012) also recommends that “research data that results from publicly funded research become publicly accessible, usable and re-usable through digital e-infrastructures.”

Finding the right data is a major challenge to the scientific community. Therefore, evaluation of such data is crucial in the open science and open data environment (Schopfel et al., 2017). The quality and thorough data helps for the assessment and good strategy (Alshawi et al., 2003). The CODATA promotes the open data and open science practices without any barriers and improves the data skills and capacity building to utilise the open data. The re3data.org provides open access to research data to the scientific community.

Re3data.org

The re3data (The Registry of Research Data Repositories) is a global registry which is listing the Research Data Repositories (RDR), mainly the Open Data Repositories (ODR). The re3data indexed 2327 RDR (as on 13 May, 2019). The re3data was launched in 2012, funded by the German Research Foundation and various institutions working as partners in Germany. It permanently stores the data set in repositories which is useful to the researchers, funding agencies, publishers, academic institutions and scholarly community. The registry covers research data from different academic fields. It aims to promote the research data for sharing, access and better visibility. It is a tool to identify the research data for publishers and research institutions. The registry helps the researchers, scientists and other scholarly community to deposit and access the research data. It publishes the data under the

license of Creative Commons Attribution 4.0 (www.re3data.org).

Related Literature

Managing research data is a big challenge to the research institutions. In the recent past, the data storage and data curation was fairly an ad hoc and altruistic. Recently, libraries are stepping into the data management services for their parent organisation (Fox, 2013) and play a crucial role in the research data management (Medeiros, 2013). The librarian certainly supports the research process through extending the library data services (Wang, 2013) and also the research data management leads to the development of research data services in libraries (Whitmir et al., 2015). The library always adopts new technology and methodologies to provide value added services to the user community. The librarians are collaborating with scientists and researchers to address the research data management (Medina-Smith et al., 2016). The study conducted by Diekema et al. (2014) reveals that the researchers and teaching faculty would like to share their research data through an open access mode. However, they often lack with the expertise to share the research data. Though libraries are involved in management of research data they offer limited services due to certain challenges like skill gap, cultural change etc. (Cox & Pinfield, 2013). However, the librarians need to acquire data management skills through various means (Tenopir et al., 2014). The researchers of Chemistry discipline are confident that the librarians help the research community through data management services including the data generation, collection, recording, processing, preservation and backup, publication and sharing (Chen & Wu, 2017). Pinfield et al. (2014) analyse the association between academic libraries and other organisational stakeholders involved in research data management activities to evaluate its drivers. To improve the impact of library on learning and research, Laitinen and Saarti (2012) opine that there is a need to develop the data system.

The Elsevier team for research data management finds out that 60% of the respondents agreed that the research

data management specialists play an important role in research data sharing among the researchers without any restrictions (Waard & Cousjin, 2017). However, certain hurdles obstruct to access the research data. Childs et al. (2014) argue that there are three key issues i.e. which research data is appropriate to make open access to share and reuse, how ethical the research data is openly accessible for sharing and reuse and what are the practicability to make open access to research data for sharing and reuse? Patel (2016) proposes a conceptual framework for research data management which certainly helps the library professionals to build a data repositories. Parinov (2014) discusses the data model (CRIS-CERIF approach) for how the scientist uses the research data to produce the explicit knowledge. Hellberge and Hedstrom (2015) explain that public wants to reuse the data is the myth behind open government and open data. Pampel et al. (2013) discusses the different kinds of data repositories, their landscapes, typology and features of research data repositories listed in the re3data.org. Re3data.org allows finding the relevant research data to the scholarly community. Hence, re3data is the best source of information to access research data.

Research Objectives

The following are the specific objectives set for the study to:

- Find out the country wise distribution of research data repositories in re3data.org;
- Trace the software used for the creation and maintenance of research data repositories listed in re3data;
- Identify the subject domain research data repositories;
- Find out the language wise distribution of research data repositories;
- Know the metadata standards used in the research data repositories.

Scope and Methodology

The scope of the study covers research data repositories indexed in the Registry of Research Data Repositories (re3data.org). The authors accessed the data from re3data.org. The authors have collected the data on research data

repositories as on 13 May, 2019. The re3data was harvested and indexed a total of 2327 research data repositories (as on 13 May, 2019). The data then collected from www.re3data.org was classified under various categories. The MS Excel was used for processing and tabulation of data.

Analysis and Interpretation

Subject Wise Distribution of Data Repositories

Table 1: Subject Wise Distribution of Data Repositories

<i>Subjects</i>	<i>No. of Data Repositories</i>	<i>Percentage</i>
Geosciences	694	29.82
Botany	671	28.84
Medicine	567	24.37
Engineering Sciences	365	15.69
Physics	275	11.82
Social Sciences	247	10.61
Zoology	220	9.45
Humanities	231	9.93
Chemistry	194	8.34
Economics	174	7.48
Agriculture, Forestry, Horticulture & Veterinary	177	7.61
Education Sciences	46	1.98
Law	38	1.63
Psychology	29	1.25
Mathematics	27	1.16

The Table 1 shows the subject wise distribution of data repositories listed in re3data. The re3data has shown 2327 repositories. However, a total of 3955 entries were found in the subject category. This indicates that repositories consist the research data on more than one subject category. It is found from the study that the re3data.org organises repositories under few disciplines with many facets. However, the authors categorised the subjects into 16 major headings. It was observed from the table that the majority of the subjects are from science discipline, which means that the data repositories consist the research data largely on science and its allied subjects. The study shows that the research

data was mainly generated and used by the scientists, Technologists, Doctors and others from the science discipline. The study also reveals that the research data on Geosciences subject includes the highest number (694) of repositories which account 29.82% of the total listed repositories. Geosciences cover the research data

from Geography, Geology, Climatology and other allied subjects, followed by, Plant Sciences (671, 28.84%), Medicine (567, 24.37%). Mathematics (27, 1.16%), Psychology (29, 1.25%) and Law (38, 1.63%) are some of the subjects which consist the less number of repositories on research data.

Country Wise Distribution of Data Repositories

Table 2: Geographic Distribution of Data Repositories

<i>Sr. No.</i>	<i>Name of Country</i>	<i>No. of Data Repositories</i>	<i>Sr. No.</i>	<i>Name of Country</i>	<i>No. of Data Repositories</i>
1	Azerbaijan	1	38	International	227
2	China	40	39	Austria	36
3	Cyprus	1	40	Belgium	22
4	French Polynesia	1	41	Bosnia & Herzegovina	1
5	Hong Kong	1	42	Croatia	1
6	India	49	43	Czech Republic	8
7	Indonesia	3	44	Denmark	22
8	Israel	10	45	Estonia	6
9	Japan	57	46	European Union	241
10	Korea, North	1	47	Finland	10
11	Korea, South	9	48	France	101
12	Pakistan	1	49	Germany	357
13	Philippines	2	50	Greece	10
14	Singapore	4	51	Greenland	3
15	Sri Lanka	1	52	Hungary	5
16	Taiwan	9	53	Iceland	1
17	Thailand	2	54	Ireland	9
18	Turkey	1	55	Italy	36
Asia		193	56	Lithuania	5
19	Benin	2	57	Luxembourg	3
20	Burkina Faso	2	58	Netherlands	52
21	Cameroon	1	59	Norway	25
22	Costa Rica	1	60	Poland	5
23	Cote d'Ivoire	1	61	Portugal	5
24	Egypt	1	62	Romania	1
25	Ethiopia	1	63	Russian Federation	22
26	Ghana	2	64	Slovakia	2
27	Kenya	4	65	Slovenia	4
28	Senegal	1	66	Spain	27
29	South Africa	8	67	Sweden	21

Sr. No	Name of Country	No. of Data Repositories	Sr. No.	Name of Country	No. of Data Repositories
30	Tunisia	1	68	Switzerland	68
Africa		25	69	Ukraine	2
31	Australia	88	70	United Kingdom	279
32	New Zealand	9	Europe		1390
Australia		97	71	Canada	219
33	Argentina	4	72	El Salvador	1
34	Brazil	8	73	Mexico	11
35	Chile	1	74	Panama	2
36	Colombia	2	75	United States	1039
37	Peru	2	North America		1272
South America		17	Total		3221

The Table 2 reveals the geographic distribution of data repositories. There were 75 countries including the European Union and International bodies contributing the research data to the 2327 repositories. However, the total entries found in the table were 3221, because many of the repositories were receiving the data from more than one country. Therefore, the data repository was added in the respective contributing countries. The USA (1039), Germany (357) and the UK (279) are the major contributors. There are 32 countries adding research data from the European countries, Germany (357), UK (279) and European Union (241) are the major contributors from Europe. There were 18 countries contributing to the re3data.org from Asia viz. Japan (57), India (49) and China (40) are the key contributors. The European countries contributing 1390, (43.15%) of the repositories listed in re3data.org, followed by, North America (1272, 39.49%), International organizations (227, 07.05%), Asia (193, 05.99%), Australia (97, 03.01%), Africa (25, 00.78%) and South America (17, 0.53%). The table also reveals that the countries from Europe and North America are dominating in contributing to the research data.

Type of Contents

The Table 3 gives an insight on the type of contents included in the re3data. The re3data.org identified and classified into 15 types of contents that was added in repositories. The repositories consists variety of contents viz. images, statistical data, scientific data, audiovisual data, office documents etc. However, a single repository

does not consist all types of data or only one type of data. It helps the users who specifically look for a specific type of content. The scientific and statistical data formats are the highest i.e. 1442 repositories, followed by, Standard Office Documents (1397), Images (1165), Plain Text (1052) and so on. The table reveals that the scientific and statistical data and office documents are the type of research data included in most of the repositories.

Table 3: Type of Content Available in Data Repositories

Content Types	No. of Data Repositories	Percentage
Archived data	505	21.70
Audiovisual data	436	18.74
Configuration data	59	2.54
Databases	460	19.77
Images	1165	50.06
Network based data	131	5.63
Plain text	1052	45.21
Raw data	1031	44.31
Scientific and statistical data formats	1442	61.97
Software applications	410	17.62
Source code	130	5.59
Standard office documents	1397	60.03
Structured graphics	822	35.32
Structured text	703	30.21
Others	828	35.58

Software Used in Data Repositories

The Table 4 depicts the software used for creation and maintenance of data repositories indexed in re3data. Majority of the repositories' software names are unknown (1172, 50.37%). On the other hand, 15.21% of the repositories (354) software names are not classified to identify the name of the software. It is clear from the table that 34.42% of the repositories' software names are clearly known and classified. Among these, 422 repositories' software names are classified into others category. The most popular known software are DataVerse (76, 03.27%), MySQL (73, 03.14%), Dspace (69, 02.97%), CKAN (65, 02.79%), Fedora (35, 01.50%), Eprints (33, 01.42%) and Nesstar (19, 0.82%). It is noticed from the study that the popular content management software like Dspace, Eprints and Fedora are also used for creation and maintenance of research data repositories. Along with content management software, some specialised software like DataVerse, CKAN, Nesstar are also used for maintenance of data repositories.

Table 4: Software Used for Data Repositories

Software Name	No. of Data Repositories	Percentage
CKAN	65	2.79
Dspace	69	2.97
DataVerse	76	3.27
DigitalCommons	03	0.13
EPrints	33	1.42
Fedora	35	1.50

Software Name	No. of Data Repositories	Percentage
MySQL	73	3.14
Nesstar	19	0.82
Opus	01	0.04
dLibra	02	0.09
eSciDoc	03	0.13
Other	422	18.13
Unknown	1172	50.37
Not Classified	354	15.21
Total	2327	100

Language Wise Distribution of Data Repositories

The Table 5 reveals the language wise distribution of data repositories. The re3data includes the data from a total of 59 languages from all over the world. Most of the repositories consist the data in multiple languages. The research data on English language can be found in 2270 data repositories. It means that 97.55% of the data repositories listed in re3data provides data in English language, followed by, French (169, 07.26%), German language (166, 07.13%) and Spanish (81, 03.48%) and so on. Most of the languages listed in the re3data are from the European countries. There are 13 data repositories providing the data from 08 Indian languages which includes Hindi (06) and each one from Bengali, Gujarati, Kannada, Malayalam, Marathi, Telugu and Tamil language respectively.

Table 5: Language Wise Distribution Data Repositories

Sr. No.	Language	No. of Data Repositories	Percentage	Sr. No.	Language	No. of Data Repositories	Percentage
1	Afrikaans	1	0.03	31	Korean	9	0.28
2	Albanian	1	0.03	32	Lao	1	0.03
3	Arabic	12	0.37	33	Latvian	4	0.12
4	Basque	2	0.06	34	Lithuanian	6	0.18
5	Bengali	1	0.03	35	Macedonian	1	0.03
6	Bulgarian	2	0.06	36	Malayalam	1	0.03
7	Catalan	9	0.28	37	Maltese	2	0.06
8	Chinese	45	1.38	38	Marathi	1	0.03

Sr. No.	Language	No. of Data Repositories	Percentage	Sr. No.	Language	No. of Data Repositories	Percentage
9	Croatian	2	0.06	39	Norwegian	15	0.46
10	Czech	10	0.31	40	Persian	3	0.09
11	Danish	25	0.76	41	Polish	11	0.34
12	Dutch	187	5.72	42	Portuguese	25	0.76
13	English	2270	69.44	43	Romanian	4	0.12
14	Estonian	7	0.21	44	Romansh	1	0.03
15	Finnish	11	0.34	45	Russian	35	1.07
16	French	169	5.17	46	Serbian	2	0.06
17	Galician	1	0.03	47	Slovak	5	0.15
18	German	166	5.08	48	Slovenian	5	0.15
19	Greek	9	0.28	49	Spanish	81	2.48
20	Greenlandic	1	0.03	50	Swedish	17	0.52
21	Gujarati	1	0.03	51	Tagalog	1	0.03
22	Hebrew	2	0.06	52	Tamil	1	0.03
23	Hindi	6	0.18	53	Telugu	1	0.03
24	Hungarian	10	0.31	54	Thai	4	0.12
25	Icelandic	2	0.06	55	Turkish	1	0.03
26	Indonesian	7	0.21	56	Ukrainian	4	0.12
27	Irish	2	0.06	57	Vietnamese	3	0.09
28	Italian	22	0.67	58	Welsh	2	0.06
29	Japanese	38	1.16	59	Yoruba	1	0.03
30	Kannada	1	0.03				

Keywords Used While Searching Data Repositories

Table 6: Keywords Used in Number of Data Repositories

Sr. No.	Keywords	No. of Data Repositories	Percentage
1	Multidisciplinary	184	9.05
2	Genomics	131	6.44
3	Bioinformatics	112	5.51
4	Genetics	108	5.31
5	Biology	92	4.53
6	Health	97	4.77
7	Biodiversity	86	4.23
8	DNA	77	3.79
9	Climate	75	3.69
10	Atmosphere	70	3.44
11	Meteorology	68	3.34
12	Agriculture	62	3.05
13	Statistics	59	2.90
14	Environment	54	2.66

<i>Sr. No.</i>	<i>Keywords</i>	<i>No. of Data Repositories</i>	<i>Percentage</i>
15	Cancer	52	2.56
16	Climate Change	50	2.46
17	Hydrology	50	2.46
18	Weather	50	2.46
19	Ecology	52	2.56
20	Oceanography	48	2.36
21	Economics	48	2.36
22	Molecular Biology	47	2.31
23	Climatology	46	2.26
24	Gene Expression	47	2.31
25	Proteins	44	2.16
26	Ecosystem	46	2.26
27	Social sciences	42	2.07
28	Remote Sensing	45	2.21
29	Education	47	2.31
30	Human	44	2.16
	Total	2033	100

Apart from browsing the repositories by subject, it is also possible to browse the repositories using some specific keywords in re3data.org. The re3data listed 30 broad keywords through which users can easily identify the specific repositories they needed. Browsing the repositories by subject will get more results which leads to recall than precision. However, if a user uses the keywords given by re3data then the users will get the precision results. The Table 6 shows that 184 repositories consist the data on multidisciplinary subject, followed by, Genomics (131), Bioinformatics (112), Genetics (108), Biology (92) etc. The keywords used in the re3data are mainly from Life Science, Earth Science and Medicine subject category.

Metadata Standards Used in the Re3data

The metadata standards were used for describing the Web resources as well as physical resources. The metadata standards help to access, share and exchange data. There are many metadata standards evolved for describing the resources. However, Dublin Core Metadata Standard was widely used and accepted metadata standard. The Table 7 indicates that a total of 1140 (48.99%) repositories used the metadata standards out of 2327 repositories. A total of 1140 repositories were from different metadata standards.

Among these repositories, Dublin Core (256) was widely used metadata standard, which accounted for 22.46% of the metadata used repositories, followed by, DDI-Data Document Initiative (152, 13.33%), ISO 19115 (140, 12.28%), DataCite Metadata Schema (125, 10.96%) are the major metadata standards indexed in re3data. It is clear from the study that there were some specialised subject metadata standards used for describing the resources like ABCD, AVM, EML, MIBBI and SDMX etc.

Table 7: Metadata Standards Used in Data Repositories

<i>Sr. No.</i>	<i>Metadata Standards</i>	<i>No. of Data Repositories</i>	<i>Percentage</i>
1	ABCD - Access to Biological Collection Data	14	1.23
2	AVM - Astronomy Visualization Metadata	01	0.09
3	CF (Climate and Forecast) Metadata Conventions	36	3.16
4	CIF - Crystallographic Information Framework	06	0.53
5	CIM - Common Information Model	03	0.26

Sr. No.	Metadata Standards	No. of Data Repositories	Percentage
6	CSMD-CCLRC Core Scientific Metadata Model	01	0.09
7	DCAT - Data Catalog Vocabulary	11	0.96
8	DDI - Data Documentation Initiative	152	13.33
9	DIF - Directory Interchange Format	37	3.25
10	Darwin Core	24	2.11
11	DataCite Metadata Schema	125	10.96
12	Dublin Core	256	22.46
13	EML - Ecological Metadata Language	31	2.72
14	FGDC/CSDGM - Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata	82	7.19
15	FITS - Flexible Image Transport System	10	0.88
16	Genome Metadata	01	0.09
17	ISA-Tab	11	0.96
18	ISO 19115	140	12.28
19	International Virtual Observatory Alliance Technical Specifications	04	0.35
20	MIBBI - Minimum Information for Biological and Biomedical Investigations	06	0.53
21	MIDAS-Heritage	01	0.09
22	OAI-ORE - Open Archives Initiative Object Reuse and Exchange	11	0.96
23	PROV	02	0.18
24	RDF Data Cube Vocabulary	25	2.19
25	Repository-Developed Metadata Schemas	108	9.47
26	SDMX - Statistical Data and Metadata Exchange	04	0.35
27	Other	38	3.33
	Total	1140	100

Discussions and Conclusion

The research data is crucial for scientific communication and evaluation (Schopf et al., 2017). The researchers generate a wide range of data sets and data types (Whitemire & Boock, 2015). The literature review reveals that the data management is a complex issue. The re3data is such a registry that it has brought many data repositories under a unified umbrella for easy access. The re3data.org registry organises, manages and enables the easy retrieval and sharing of data among the scholarly community. Pampel et al. (2013) describe and differentiate the registry into institutional, disciplinary, multidisciplinary, project specific data repositories. It provides prominent infrastructure to share the research data. However, the scholarly community posed with several difficulties in sharing the data. Waard and Cousijn (2007) identify the privacy concerns, ethical issues, intellectual property rights, and training in data sharing practices are the major problems faced by the scholarly community.

There are many efforts made to eradicate the problems regarding privacy concerns, ethical issues, intellectual property rights through creative commons and other such activities. The major hurdle was (Kuipers & Vander Hoeven, 2009; Tenopir et al., 2011) the lack of willingness to share the data among the scholarly community. The study conducted by Mueller-Langer and Andreoli-Versbach (2017) reveal that in order to exploit the collected data to the maximum extent and make use it for further research, sometime strategically delayed by the researchers. The libraries can overcome from data sharing problem by collaborating with researchers, organisations, data centres and the government departments etc. Cox and Pinfield (2013) notice that very limited extent of libraries offering the research data management and many of the libraries involved in placing the research data management policies. The data managers have an ocean of opportunities in data and information management, by using their experience and expertise in managing the open data (Childs et al., 2014). Therefore, re3data.org is the role model to the professionals to adopt and develop an altruistic nature to provide services to the research community.

The funding agencies must develop the mandates and policies towards making the mandatory open access to research data. Organisation for Economic Co-Operation

and Development (OECD) (2007) call for global guiding principles on the access to research data on January 2004 in Paris meeting. There were 30 OECD and other four countries adopted a 'Declaration on Access to Research Data from Public Funding'. The OECD countries recognised the importance of access to research data. Hence, they invited the OECD members to develop a set of guidelines to facilitate the optimal access to digital research data. The Research data management helps to develop the research data services in the libraries and fulfil the research data needs of scholarly community. The Leiden University's Centre for Science and Technology Studies (CWTS) and Elsevier undertook an assignment to explore open data practices at the workbench. The key findings of the report were: 73% of the respondents opined that having access to published research data would benefit their own research, 64% were willing to allow others to access their research data, 69% of the respondents said that the sharing research data is important for doing research in their field, 34% of the researchers do not publish their data at all. Open data practices facilitate collaboration, drive data analysis, and support transparency and reproducibility. Yet the research community has not consistently embraced open data or data sharing practices. Therefore, the library professionals need to practice in this regard and should help the scholarly community to share their research data to the repositories.

References

- Alshawi, S., Missi, F., & Eldabi, T. (2003). Healthcare information management: The integration of patients' data. *Logistics Information Management*, 16(3/4), 286-295.
- Borgman, C.L. (2012). The conundrum of sharing research data. *Journal of the Association for Information Science and Technology*, 63(6), 1059-1078.
- Chen, X., & Wu, M. (2017). Survey on the needs for chemistry research data management and sharing. *The Journal of Academic Librarianship*, 43(4), 346-353.
- Childs, S., McLeod, J., Lomas, E., & Cook, G. (2014). Opening research data: Issues and opportunities. *Records Management Journal*, 24(2), 142-162.
- Cox, A. M., & Pinfield, S. (2014). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, 46(4), 299-316.
- Diekema, A. R., Wesolek, A., & Walters, C. D. (2014). The NSF/NIH Effect: Surveying the effect of data management requirements on faculty, sponsored programs, and institutional repositories. *The Journal of Academic Librarianship*, 40(3-4), 322-331.
- Fox, R. (2013). The art and science of data curation. *OCLC Systems & Services: International Digital Library Perspectives*, 29(4), 195-199.
- Hellberg, A. S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. *Transforming Government: People, Process and Policy*, 9(1), 35-51.
- Higman, R., & Pinfield, S. (2015). Research data management and openness: The role of data sharing in developing institutional policies and practices. *Program: Electronic Library and Information Systems*, 49(4), 364-381.
- Kroes, N. (2012). Commission recommendation of 17.7. 2012 on access to and preservation of scientific information [No. SWD (2012) 221 final]. Retrieved on 30 May, 2018 from <http://register.consilium.europa.eu/doc/srv?l=EN&f=ST%2012846%202012%20INIT>
- Kuipers, T., & Van der Hoeven, J. (2013). Insight into digital preservation of research output in Europe. Survey Report. Retrieved from http://www.parse-insight.eu/downloads/PARSE-Insight_D3-4_SurveyReport_finalhq.pdf
- Laitinen, M., & Saarti, J. (2012). A model for a library-management toolbox: Data warehousing as a tool for filtering and analyzing statistical information from multiple sources. *Library Management*, 33(4/5), 253-260.
- Medeiros, N. (2013). A public trust: Libraries and data curation. *OCLC Systems & Services: International Digital Library Perspectives*, 29(4), 192-194.
- Medina-Smith, A., Tryka, K. A., Silcox, B., & Hanisch, R. J. (2016). Librarians and scientists partner to address data management: Taking collaboration to the next level. *Digital Library Perspectives*, 32(3), 142-152.
- Mueller-Langer, F., & Andreoli-Versbach, P. (2017). Open access to research data: Strategic delay and

- the ambiguous welfare effects of mandatory data disclosure. *Information Economics and Policy*.
- Organisation for Economic Co-operation and Development. (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding*. OECD Publishing. Retrieved June 6, 2018, from <https://www.oecd.org/sti/sci-tech/38500813.pdf>
- Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., & Dierolf, U. (2013). Making research data repositories visible: The re3data.org-registry. *PloS One*, 8(11), e78080.
- Parinov, S. (2014). Towards an open data on how the research data are used: CRIS-CERIF based approach. *Procedia Computer Science*, 33, 55-59.
- Patel, D. (2016). Research data management: A conceptual framework. *Library Review*, 65(4/5), 226-241.
- Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PloS one*, 9(12), e114734.
- PLoS. (2017). Data availability. Retrieved on December 11, 2017 from <http://journals.plos.org/plosone/s/data-availability>
- Schöpfel, J., Prost, H., & Rebouillat, V. (2017). Research data in current research information systems. *Procedia Computer Science*, 106, 305-320.
- Schöpfel, J., Chaudiron, S., Jacquemin, B., Prost, H., Severo, M., & Thiault, F. (2014). Open access to research data in electronic theses and dissertations: An overview. *Library Hi Tech*, 32(4), 612-627.
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PloS One*, 6(6), e21101.
- Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(2), 84-90.
- Waard, A., & Cousijn, H. (2017). Mapping the route to open data sharing. Elsevier Research Data Management Services. Retrieved from <https://libraryconnect.elsevier.com/articles/mapping-route-open-data-sharing>.
- Wang, M. (2013). Supporting the research process through expanded library data services. *Program*, 47(3), 282-303.
- Whitmire, A. L., Boock, M., & Sutton, S. C. (2015). Variability in academic research data management practices: Implications for data services development from a faculty survey. *Program: Electronic Library and Information Systems*, 49(4), 382-407.
- Wouters, P., & Haak, W. (2017). Open data: The researcher perspective. Retrieved June 6, 2009, from https://www.elsevier.com/data/assets/pdf_file/0004/281920/Open-data-report.pdf