ZBORNÍK FILOZOFICKEJ FAKULTY UNIVERZITY KOMENSKÉHO
Ročník XXVIII KNIŽNIČNÁ A INFORMAČNÁ VEDA Bratislava 2019

RESEARCH DATA MANAGEMENT AND DATA LITERACY AS WE SEE THEM TODAY

Tibor Koltay

1 Introduction

Research data is the raw material for scholarly research. In the same time, it is the output of research activities (Pryor 2012). It consists of "heterogeneous objects and items used and contextualized, depending on the academic discipline of origin" (Semeler, Pinto and Rozados 2017, p. 3). To serve as research data, little data can be just as valuable as big data and – in many cases – there is no data, because relevant data does not exist, cannot be found, or is not available (Borgman 2015).

As succinctly described by Cox (2018), library and information professionals (librarians) usually have been "thinking of data as sitting beneath information and knowledge in a pyramid of value. But the concept of data has grown significantly in importance in the last five years, driven by public speculation about the power and risks of big data. In parallel, data related roles are becoming increasingly important in professional practice". He adds that "academic librarianship is becoming more of a data profession and outlines a spectrum of data roles from the familiar to the new".

Considering the above development, this paper focuses on research data management (RDM), which builds a complex of services that libraries can or should offer to a wide community of researchers, working in the natural sciences, the social sciences and the humanities. Such services require fulfilling new professional roles, the question of which will also be addressed. Data literacy in its own virtue and being closely connected to RDM cannot be left out. Neither should we forget about the similarities and differences between data-librarians and data scientists.

2 Methods

Making use of varied current awareness services, and recommendations of academic social networking services, mainly Mendeley and ResearchGate, a number of writings, including review articles have been selected to reflect on varied selected issues of research data management and data literacy and the attributes of data-librarians and data scientists.

3 Discussion

Scholarly research does not need more data. Rather, it requires having the right data (Borgman 2015). In other words, researchers require high quality, curated data, to work with. On the other end of the spectre, we see librarians' traditional willingness and service mindset for answering these needs, which require additional knowledge and skills, not known before, as well as confidence to work in this area (Corrall, Kennan and Afzal 2013). One of the main requirements of acquiring all related competencies is that librarians are given opportunities to learn new skills (Tenopir et al. 2017).

3.1 Needs and expectations

The fact that data is of no use if not analysed, is a challenge for librarians, because many of them do not possess analytical skills and lack the abilities needed for working confidently in this new field of the changing information environment (Kirkwood 2016). As a result – just as it was said above – re-skilling librarians is necessary, as it allows them to be prepared to the changing circumstances and requirements of the data-intensive research environment. Such education must be directed towards changing librarians' self-identification. Accordingly, they should be able to handle data of all kinds, making use of their traditional skills of dealing with textual information. This involves a conceptual understanding of data, and "the abilities to find, extract, collect, clean, organise, analyse, and present data", even if they do not serve users, who deal with data (Robinson and Bawden 2017, p. 314). We can say that similarity with traditional tasks and the obvious and less known differences between old and new responsibilities are key to going forward.

If we look at the researchers' side, we see that the level of acceptance and the perception of Open Data, or the perceived need for RDM is far from uniform by discipline and by countries, showing a wide variety, much broader than that we could experience it in the case of Open Access to publications (Cox and Corrall 2013). We will see that neither is the preparedness of academic libraries to take responsibility on the same level in every country and institution (Tenopir et al. 2016; Tenopir et al. 2017).

To provide successful services under these circumstances, a basic appreciation of issues, related to research data is indispensable for librarians. Such appreciation should include technical and socio-ethical competences, the latter containing knowledge of legal and ethical issues. As already mentioned, many of these have been well-known to librarians in the paper-based world (Robinson and Bawden 2017).

Special attention needs to be given to services, offered to digital humanists, as the digital humanities is not only one of the sciences that are interested in investigating information (Koltay 2016a), but relies heavily on the interpretation of data (Gibbs 2011) thanks to the ability of computers to handle large corpuses of textual material (ACRL 2013). Distinguished attention is in place here, because we tend to easily accept the prevalence of data in such brands of "big science" as astronomy or physics and easily admit that new disciplines such as astroniformatics, or bioinformatics are data-driven, while this does not seem to be self-explanatory in the case of the humanities (Borgman 2012).

As Cox and Corrall (2013) pointed it out, the literature of data-centred academic library work can be characterized by both optimism and pessimism. The example of the former is the attention, given to the potential roles in research data management, based on the hope that the library's role in supporting scholarship can be extended by librarians, who become embedded in academic departments and research teams. However, the authors remind us that

– while this aspiration is natural and could lead to highly desirable results – we should remember the disappointment that was brought by the attempts of acquiring a role in knowledge management. For that reason, it is better to be cautious with exaggerated expectations.

3.2 Research data management as a critical service

In this part of our paper, we will offer a selection from the issues related to the multitude of RDM activities. Therefore, it is time to provide a simple general definition of RDM, so we can say that it is a comprehensive set of activities for the organization, storage, access, and preservation of data (Semeler, Pinto and Rozados 2017).

Even if not simply reinventing or extending what librarians already do today, RDM is not exception from the strong continuities with the tasks that they are already expected to do (Cox 2018).

In many cases, there is willingness to offer RDM. For instance, the majority of library directors, participating in a survey of European LIBER member libraries from 22 countries[1] "recognize the growing importance of research data and are looking for solutions that fit their institutional needs and priorities" (Tenopir et al. 2017, p. 38). The survey team, led by Carol Tenopir found that – similarly to previous surveys, conducted with libraries in North America (Tenopir et al. 2014) – European academic libraries are more likely to offer informational services than technological ones.

Informational services, such as consulting with staff or students on Data Management Plans (DMPs), on metadata standards and providing reference support for finding and citing data (datasets) occur with relatively high frequency. By involving a personal client-librarian relationship these services align with long-established, traditional reference services, supplied by libraries.

Technological services appear with a lower frequency. They include creating or transforming metadata for datasets, identifying datasets that could be candidates for repositories, selecting and preparing datasets for deposit, deaccessioning, or deselection of these datasets.

This survey also shows that consulting with staff or students on *data-management plan*s (DMPs) happens with a relatively high frequency (Tenopir et al. 2017). Compliance with the varied data-related requirements of research funders has begun with DMPs, and assistance of preparing them is a basic level informational service (Molloy and Snow 2012) that helps researchers to plan, articulate, and execute data management (Penfield, Cox and Smith 2014).

DMPs formalize all technical and non-technical details, and should be an integral part of any project that requires some kind of data management (Solodovnik and Budroni 2015). Let us add that – while the skills needed to assist in producing DMPs effectively are the same as the ones for any informational service – such support requires librarians to be familiar with funder requirements, relevant standards, and local data management processes (Cox and Verbaan 2018).

Similarly to traditional reference services, *data reference interviews* build on both the researcher and the librarian to bring their own unique expertise into the interaction. This is in conformity with the *RUSA Professional Competencies for Reference and User Services Librarians* documents, which says that [The reference and user services librarian] engages the

[1] Data does not include responses from Croatia, Cyprus, France, Hungary, Luxembourg, Malta, Portugal, Romania, Serbia, Slovakia, and Turkey.

user in the process and in making decisions (Huling et al. 2017). The earlier version of this document expresses this rule more outspokenly, when stating that the librarian "acknowledges the knowledge brought by the user to the interaction" (Whitlatch et al. 2003).

Despite similarities, data reference interviews may consist of more questions than in traditional reference interactions, and it might happen that some useful starting points can be found, which however are far from being the perfect source (Rice and Southall 2016).

Besides assessing researcher's data needs, consultations may require the librarian to assist the researcher in finding an existing dataset. Traditional skills play a role here again, as – similarly to the situation of reference librarians, who are aware of the key reference works and authoritative texts within their fields – data librarians should be familiar with major repositories and other sources of datasets (Federer 2014). Reviewing datasets for their potential inclusion into repositories, already mentioned above, is also familiar to librarians, as it shows similarities to activities, rooted in collection management principles (Cox 2018).

Librarians can offer help in *data citation.* Though much more complex than citing research publications, data citation is an unquestionable necessity (Silvello 2018). It also has the potential to become a source of motivation for researchers to share and publish their data, provided that it becomes a source of reward and acknowledgment for them (Candela et al. 2015). We know that supporting citation activities is familiar to librarians in all environments, including data-intensive ones. Besides of citing data, there is the possibility of *data publication*, even though there is uncertainty about its precise meaning, proper forms and extent of validating published datasets (Kratz and Strasser 2014).

Beyond assisting researchers in assigning metadata and citations to datasets, helping them in *data retrieval* is also a potential task. However, data retrieval is a less obvious task than retrieving research papers, which are available in a predominantly textual form. For instance, the average file size of retrieved datasets is several times larger than that of retrieved research publication files; and these in turn vary from one discipline to another. Besides of this, the retrieved datasets may be of different file types or formats and often must be downloaded before they can be read or used. Research datasets are almost always accompanied with separate documentation files and a single dataset item record may constitute several composite files (Bugaje and Chowdhury 2017).

3.3 The role of data literacy

Data literacy can be defined as a set of skills and abilities related to accessing research data, understanding, interpreting, managing, critically evaluating and ethically using it (Koltay 2015).

Research data management is inconceivable without data literacy education, because people, who will use research data, need education about how to understand, interpret, and apply what they find, and researchers are no exception from this rule. Besides of this, the fact that data is not being restricted to quantitative results drawn from surveys or scientific experiments, i. e. we have an expanded view on it, increases the importance of data literacy (ACRL 2013).

Data literacy competencies can be used in the following fields (Carlson et al. 2011):

- discovery and acquisition of data;
- data management;
- data conversion and interoperability;
- metadata;

- data curation;
- data re-use;
- data preservation;
- data analysis;
- data visualization;
- ethics;
- data citation.

Data literacy is not only cognate with information literacy in general, but is compatible with the information literacy focus of academic librarianship (Cox 2018). Presenting, evaluating, and interpreting qualitative and quantitative data is a learning outcome of IL. This statement of Andretta et al. (2008) shows not only the connection between IL and data literacy, but it is an early recognition of the growing emphasis on educational activities in libraries, explained in detail, for example by Grassian (2017). However, the strongest connection between data literacy and IL is made by their attention to critical assessment. In the case of data we must count with misleading or inappropriate use (Carlson et al. 2011). In consequence, attention has to be paid to the version of the given dataset and the person responsible for it (ACRL 2013). Data literacy is also connected to *media literacy* in pointing towards "the use of tools to use and reuse content in ways not imagined by the content creator are this literacy" (ACRL 2013, p. 11).

Data literacy also has an interface to *academic literacy* that involves the comprehension of the entire system of thinking, values, cultural identity and information flows of academia, which results in the ability to read, interpret, and produce texts, valued in academia (Elmborg 2006). While academic literacy's scope is much broader, it can be applied to data, especially if we adopt the idea of digital scholarship that relies – among others – on collections of information and data (Cox 2016).

Both academic literacy and data literacy are also loosely connected to *scientific literacy*, which is not tied to academia, but comprises methods, approaches, attitudes and skills, related to thinking scientifically and doing scientific research (National Academy of Sciences 1996), implying that everybody should be scientifically literate. *Statistical literacy* is often mentioned in connection with data literacy, due to the existence of a common set of problems and a similar level of approach (Schield 2004).

Data literacy is in close association with *data quality* that is one of the cornerstones of the data-intensive paradigm of scientific research (Koltay 2017). Therefore, teaching data literacy involves understanding the need for collecting data and the various ways to do it, and taking into consideration that – beyond basic literacy – there is a need for being *numerate*, i. e. understanding formulas, graphs, charts and tables.

All literacies emphasise critical appraisal in general. Data literacy education in particular should accentuate checking the provenance and integrity of data. It is vital to understand factors that may impact the data, including bias, patterns, errors and omissions. It is also important to consider sampling techniques, sample selections and size, as well as survey design. We need to able to distinguish between correlation and causation, as well as understanding how variables influence each other. Having the skills of predicting, generalizing from available datasets, as well as understanding trends, and drawing inferences is indispensable (Mason, Khan and Smith 2016). Seen from a different angle, data quality analysis requires – among others – analysing the quality of the sources. The analysis of the quality of the data itself may consider accuracy, consistency, completeness, originality and the degree of timeliness of data (Daraio et al. 2016).

By stressing the need for critical approaches, data literacy can be subsumed within a core skillset of being able to *discriminate* or exercising discernment. Discernment equals to being able to make independent judgements on the validity of information and evaluating the information component of literacies (Mason, Khan and Smith 2016, Walton, Pickard and Dodd 2018). This substantiates arguments that the final goal of data literacy is fostering a culture of the semantic scholar, because – unless educated in data literacy and information literacy – scientific training alone does not make guarantee that researchers become aware that their own scholarly communications may constitute primary sources of data (Haendel et al. 2012). This type of data literacy should involve abilities to translate vast amounts of data into abstract concepts, to understand data-based reasoning that leads to seeing almost everything that we come into contact, through the lens of data, and will regard it as computational, programmable.

When teaching data literacy skills to students, the overall learning outcome should be becoming data literate in the context of their subject area. Students should identify relevant data for their field of study, understand how data is connected to the publication process and be able to explain what open data is and how open data influences science. Applying critical thinking strategies to data and understanding the challenges of reusing data, as well as awareness of the importance of metadata and citation are crucial (Duffner-Ylvestedt and Rayner 2016).

Data literacy education could be enhanced by applying data governance principles, partially related to data quality. These principles are beneficial in delineating decision domains and defining accountability for decision making. Data governance is advantageous, because it is based on standardised, repeatable processes, rules, policies, and standards in order to enable the transparency of data-related processes (Koltay 2016b).

In the next part of this paper, we will see that such thinking may serve as one of the few common denominators between data librarians and data scientists.

3.4 Data professionals

Data-related professional roles may be especially relevant for three specialist groups: data librarians, data scientists and records professionals. Nonetheless, the main protagonists in this field seem to be data librarians and data scientists.

As said above, academic librarianship is becoming more of a data profession that reflects both familiar and new roles (Cox 2018). However, it remains questionable whether there is a true similarity or convergence between the traditional, though changing profession of the librarian (in particular the *data librarian*) and the data scientist's evolving sphere of occupation. On the one hand, there are undeniable similarities between them due to the fact that these jobs appeared in a data-intensive era. On the other hand, their work environments, culture and – last, but not least – the scopes of duties are different. Neither are the required sills and competences fully identical with each other. In one sentence, even though there are similarities between their duties and skills, data librarians are not data scientists.

The job of *data scientists* is – at least today and in the near future – tied to business environments. When calling this nascent profession the sexiest job of the 21st century, Davenport and Patil (2012) did not forget about its close relationship with the business world.

In fact, a big number of data scientists are working at companies of different size and advancement. By merging techniques of computer science and statistics, combining data analysis with the development of new products and services based on the data (Cox 2018),

data scientists also must have a good understanding of the principles of design and information architecture (Voulgaris 2014).

Data librarians focus on research data. They create new library services, focusing on heterogeneous and contextualized research data and act as facilitators in all stages of scientific research (Semeler, Pinto and Rozados 2017). As Rice and Southall (2016, p.17) put it, "data librarianship is concerned with the representation, organization and dissemination of data, and the use of technologies to design research data management and data services". They add that such jobs may require detailed specialist knowledge about collections or giving advice on copyright.

Despite the existing differences, data science uses new methods and practices that may turn out to be applicable to data librarianship first of all, if they are willing to get more closely acquainted with programing languages and programming logic (Cox 2018) and data visualisation (Ogier and Stamper 2018).

Although currently not considered to be leaders in this field, we should not forget about *records professionals*, who by virtue of having comparable experience to data librarians and data scientists since the early twentieth century are equipped with skills and expertise that may be applicable to research data management (Grant 2017).

When approaching these issues, we should heed the warning of Robinson (2016), who states that being specialised in data, i. e. examining pattern and syntax and using quantitative methodology, or focusing on information, i. e. being interested in meaning, semantics and using qualitative methods, results in becoming entrenched in professions that are defined in an overly narrow way, thus leading to silos. This division is harmful for both data scientists and data librarians.

Conclusion

The prevalence of research data does not subvert academic librarianship, but deeply influences it, even though the changes induced by it appear with a different speed in different countries. This means that it remains to be seen when academic libraries will start to see data-related roles as standard tasks (Cox 2018). To reach this stage of development, requires answering questions about both the perception of teaching staff members about the supporting role of librarians in their research (Brydges and Clarke 2015) and the librarians' view on their own role in relation to the research activities of faculty and what and how they want to achieve (Cox 2018). Data librarians and data scientists may cooperate in the future, but more likely they will borrow methods form each other. Last but not least, data literacy will remain indispensable in academic settings, while attaining more recognition and use among the public sphere and people's private life.

The writing of this text was supported by the EFOP-3.6.1-16-2016-00001 project "Complex Development of Research Capacities and Services at Eszterházy Károly University".

List of References

ACRL, 2013. *Intersections of scholarly communication and information literacy: Creating strategic collaborations for a changing academic environment* [online]. Chicago, IL:

Association of College and Research Libraries [cit. 2018-06-24]. Available from: http://acrl.ala.org/intersections/

ANDRETTA, S., POPE, A. and G. WALTON, 2008. Information Literacy Education in the UK. In: *Communications in Information Literacy.* Vol. 2, pp. 36-51.

BORGMAN, Ch.L., 2012. The conundrum of sharing research data. In: *Journal of the American Society for Information Science and Technology*. Vol. 63, pp. 1059-1078.

BORGMAN, Ch.L., 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World.* Cambridge, MA: MIT Press.

BRYDGES, B. and K. CLARKE, 2015. Is it time to re-envision the role of academic librarians in faculty research? In: *Library Connect* [online]. Vol. 13 [cit. 2018-06-24]. Available from: https://libraryconnect.elsevier.com/articles/2015-07/it-time-re-envision-role-academic-librarians-faculty-research

BUCKLAND, M., 2012. What kind of science can information science be? In: *Journal of the American Society for Information Science and Technology*. Vol. 63, pp. 1-7.

BUGAJE, M. and G. CHOWDHURY, 2017. Is Data Retrieval Different from Text Retrieval?: An Exploratory Study. In: *International Conference on Asian Digital Libraries*. Springer, Cham, pp. 97-103.

CANDELA, L., CASTELLI, D., MANGHI, P. and A. TANI, 2015. Data journals: A survey. *Journal of the Association for Information Science and Technology*, Vol. 66, pp. 1747-1762.

CARLSON, J., FOSMIRE, M. and C.C. MILLER, 2011. Determining data information literacy needs: A study of students and research faculty. In: *Libraries and the Academy*. Vol. 11, pp. 629-657.

CORRALL, S., KENNAN, M.A. and W. AFZAL, 2013. Bibliometrics and research data management services: Emerging trends in library support for research. In: *Library Trends*. Vol. 61, pp. 636-674.

COX, J., 2016. Communicating new library roles to enable digital scholarship: A review article. *New Review of Academic Librarianship.* Vol. 22, pp. 132-147.

COX, A., 2018. Academic librarianship as a data profession. *Information Today Europe/ ILI365* [online]. 2 June 2018 [cit. 2018-06-24]. Available from: http://www.infotoday.eu/Articles/Editorial/Featured-Articles/Academic-librarianship-as-a -data-profession-125376.aspx?PageNum=2

COX, A.M. and S. CORRALL, 2013. Evolving academic library specialties. In: *Journal of the American Society for Information Science and Technology*. Vol. 64, pp.1526-1542.

COX, A. and E. VERBAAN, 2018. *Exploring Research Data Management*. London: Facet.

DARAIO, C., LENZERINI, M., LEPORELLI, C., NAGGAR, P., BONACCORSI, A. and A. BARTOLUCCI, 2016. The advantages of an Ontology-Based Data Management approach: openness, interoperability and data quality. *Scientometrics*, Vol. 108, pp. 441-455.

DAVENPORT, T. and D. PATIL, 2012. Data Scientist: The Sexiest Job of the 21st Century. In: *Harvard Business Review* [online]. Vol. 90, pp. 70-76 [cit. 2018-06-24]. Available from: https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/

Checklist for a Data Management Plan: Version 4.0. *DCC* [online], 2014. [cit. 2018-06-24]. Available from: http://www.dcc.ac.uk/resources/data-management-plans/checklist (accessed 3 August 2015).

DUFFNER-YLVESTEDT, N. and J. RAYNER, 2016. Hooking Up Data with Literacy: Creating an Educational Framework for Uppsala University Library. *Nordic Journal of Information Literacy in Higher Education*, Vol. 8, pp. 38-44. https://doi.org/10.15845/noril.v8i1.261

FEDERER, L., 2014. *Exploring New Roles for Librarians: The Research Informationist*. San Rafael, CA: Morgan & Claypool Publishers.

GIBBS, F., 2011. Critical discourse in digital humanities. In: *Journal of Digital Humanities* [online]. Vol. 1 [cit. 2018-06-24]. Available from: http://journalofdigitalhumanities.org/1-1/critical-discourse-in-digital-humanities-by-fred-gibbs/

GRANT, R., 2017. Recordkeeping and research data management: a review of perspectives. In: *Records Management Journal*. Vol. 27, pp. 159-174.

GRASSIAN, E., 2017. Information Literacy and Instruction: Teaching and Learning Alternatives: A Global Overview. In: *Reference and User Services Quarterly*. Vol. 56, pp. 232-239.

HAENDEL, M.A., VASILEVSKY, N.A., and J.A. WIRZ, 2012. Dealing with Data: A Case Checklist for a Data Management Plan 2014. *DCC* [online]. [cit. 2018-06-24]. Available from: http://www.dcc.ac.uk/resources/data-management-plans/checklist

Study on Information and Data Management Literacy. In: *PLoS Biology*. Vol. 10, e1001339.

HULING, N., DALLAS L.J., KINDER, R., WHITLATCH, J.B. and B. WOODARD, 2017. *Professional Competencies for Reference and User Services Librarians.* Chicago, IL: American Library Association [cit. 2018-06-24]. Available from: http://www.ala.org/rusa/resources/guidelines/professional

KIRKWOOD, R.J., 2016. Collection development or data-driven content curation? *Library Management*. Vol. 37, pp. 275-284.

KOLTAY, T., 2015. Data literacy: In search of a name and identity. In: *Journal of Documentation*. Vol. 71, pp. 401-415.

KOLTAY, T., 2016a. Library and information science and the digital humanities: perceived and real strengths and weaknesses. In: *Journal of Documentation*. Vol. 72, pp. 781-792.

KOLTAY, T., 2016b. Data governance, data literacy and the management of data quality. In: *IFLA Journal*. Vol. 42, pp. 303-312.

KOLTAY, T., 2017. Data literacy for researchers and data librarians. In: *Journal of Librarianship and Information Science.* Vol. 49, pp. 3-14.

KRATZ, J. and C. STRASSER, 2014. Data publication consensus and controversies. In: *F1000Research* [online]. Vol. 3 [cit. 2018-06-24]. Available from: https://doi.org/10.12688/f1000research.3979.3

MASON, J., KHAN, K. and S. SMITH, 2016. Literate, numerate, discriminate–realigning 21st century skills. In: *Proceedings of the 24th International Conference on Computers in Education.* Asia-Pacific Society for Computers in Education, pp. 609-614.

MOLLOY, L. and K. SNOW, 2012. The data management skills support initiative: Synthesising postgraduate training in research data management. In: *International Journal of Digital Curation*. Vol. 7, pp. 101-109.

OGIER, A.L. and M.J. STAMPER, 2018. Data Visualization as a Library Service: Embedding Visualization Services in the Library Research Lifecycle. In: *Journal of eScience Librarianship.* Vol. 7, e1126.

PINFIELD, S., COX, A.M., and J. SMITH, 2014. Research data management and libraries: Relationships, activities, drivers and influences. In: *PloS ONE*. Vol. 9, e114734.

PRYOR, G., 2012. *Managing research data*. London: Facet.

RICE, R. and J. SOUTHALL, 2016. *The data librarian's handbook*. London: Facet Publishing.

ROBINSON, L., 2016. Editorial. *Alexandria.* Vol. 26, pp. 73-76.

ROBINSON, L. and D. BAWDEN, 2017. "The story of data": A socio-technical approach to education for the data librarian role in the CityLIS library school at City, University of London. *Library Management*, Vol. 38, pp. 312-322.

ROBINSON, L., PRIEGO, E. and D. BAWDEN, 2015. Library and information science and digital humanities: two disciplines, joint future?: Re-inventing information science in the networked society. In: PEHAR, F., SCHLÖGL, C. and C. WOLFF, eds. *Re:inventing Information Science in the Networked Society: Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia, 19th-21st May 2015*. Glückstadt: Verlag Werner Hülsbusch, pp. 44-54.

SEMELER, A.R., PINTO, A.L., and H.B.F. ROZADOS, 2017. Data science in data librarianship: Core competencies of a data librarian. *Journal of Librarianship and Information Science* [online]. [cit. 2018-06-24]. Available from: http://journals.sagepub.com/doi/10.1177/0961000617742465

SI, L., ZHUANG, X., XING, W. and W. GUO, 2013. The cultivation of scientific data specialists: Development of LIS education oriented to e-science service requirements. In: *Library Hi Tech*. Vol. 31 No. 4, pp. 700-724.

SILVELLO, G., 2018. Theory and practice of data citation. In: *Journal of the Association for Information Science and Technology*. Vol. 69, pp. 6-20.

SOLODOVNIK, I. and P. BUDRONI, 2015. Preserving digital heritage: At the crossroads of Trust and Linked Open Data. In: *IFLA Journal*. Vol. 41, pp. 251-264.

TENOPIR, C., SANDUSKY, R.J., ALLARD, S. and B. BIRCH, 2014. Research data management services in academic research libraries and perceptions of librarians. In: *Library and Information Science Research*. Vol. 36, pp. 84-90.

TENOPIR, C., POLLOCK, D., ALLARD, S., and D. HUGHES, 2016. Research data services in European and North American libraries: Current offerings and plans for the future. *Proceedings of the Association for Information Science and Technology*. Vol. 53, pp. 1-6.

TENOPIR, C., TALJA, S., HORSTMANN, W., LATE, E. HUGHES, D., POLLOCK, D. SCHMIDT, B., BAIRD, L., SANDUSKY, R.J. and S. ALLARD, 2017. Research Data Services in European Academic Research Libraries. In: *LIBER Quarterly*. Vol. 27, pp. 23-44.

VOULGARIS, Z., 2014. *Data Scientist: The Definitive Guide to Becoming a Data Scientist*. Basking Ridge, NJ. Technics Publications.

WALTON, G., PICKARD, A. J., and L. DODD, 2018. Information discernment, misinformation and pro-active scepticism. In: *Journal of Librarianship and Information Science* [online]. Available from: https://doi.org/10.1177/0961000618769980.

WHITLATCH, J.B., BODNER, N.E., DIEFENTHAL, N.Z., HULING, N. and K.M. KLUEGEL, 2003. Professional competencies for reference and user services librarians. In: *Reference and User Services Quarterly*. Vol. 42. pp. 290-295.

Summary

RESEARCH DATA MANAGEMENT AND DATA LITERACY AS WE SEE THEM NOW

Tibor Koltay

This paper focuses on Research Data Management (RDM) and Data Literacy, which build a complex of services that libraries and librarians can or should offer to a wide community of researchers, working in the natural sciences, the social sciences and the humanities. To provide these services successfully, requires the appreciation of technical and socio-ethical issues related to research data. Such role can be played by data librarians, the knowledge, skills and mindset of whom show similarities to the competences required for not data-related tasks of academic librarianship. They also have communalities with the requirements set against data scientists, but the differences between the two professions are considerable.