Designing an information architecture for data management technologies: Introducing the DIAMANT model

Blask, Katarina; Förster, André

Article

Journal of Librarianship and Information Science
2020, Vol. 52(2) 592–600

DOI: 10.1177/0961000619841419
journals.sagepub.com/home/lis
SAGE

Designing an information architecture for data management technologies: Introducing the DIAMANT model

Katarina Blask
Service Centre eSciences, Trier University, Germany

André Förster
GESIS – Leibniz Institute for the Social Sciences, Germany

Abstract

Although research institutions take on increased responsibility for providing infrastructures and services around the proper handling of research data, there is no comprehensive framework addressing the ideal conditions of this implementation process. To overcome this gap, we present the DIAMANT model, a reference model aimed at providing an orientation framework for the implementation of research data management guided by the research process itself. It builds upon a central research data management information unit controlling the information flow between all other organizational units involved in research data management. Due to the possibility of outsourcing organizational units, the implementation process is maximally flexible and efficient.

Keywords

Information architecture, research data management, reference model, research process service and infrastructure landscape

Introduction

Universities and research institutions are increasingly responsible for providing researchers with institutional structures and services designed to maintain the sustainable handling of research data. When it comes to the boundary conditions of this process, research institutions often have to face insufficient data literacy, unclear responsibilities or the lacking integration of research data management (RDM) into study or advanced training curricula. In sum, knowledge of RDM and the adequate handling of research data often seems to be limited to the process itself rather than its boundary conditions.[1]

To the best of our knowledge, there is currently no comprehensive *theoretical* framework regarding the institutional conditions, under which the RDM process reaches maximum effectiveness and efficiency. This is quite surprising, since research institutions have found many different paths to incorporate RDM *in practice* over the last years. However, this may be because the goals associated with RDM have mainly been considered from an infrastructural perspective. Focusing on technical infrastructures is the result of treating RDM as a tool to fulfil specific quality criteria stated by the open science movement. In this context, Wilkinson et al. (2016) have excellently summarized that, for example, data curation and long-term archiving of research data should be guided by the FAIR principles. According to these principles, guidelines regarding the documentation of research data as well as infrastructures for their long-term preservation should ensure that research data are **f**indable, **a**ccessible, **i**nteroperable and **r**eusable. Besides the FAIR principles, lots of research funding agencies (for instance, the German Research Foundation DFG, 2013) have established similar principles, for example by declaring long-term preservation of research data mandatory. However, both the FAIR principles as well as funding agency policies and other guidelines in the context of RDM and open science rather describe general quality criteria that should be fulfilled by research data and the technologies entrusted with their curation and preservation. Thus, the quality of the RDM process is mainly determined by technical infrastructure or more specifically by the quality of archives and repositories.

Corresponding author:
André Förster, GESIS – Leibniz Institute for the Social Sciences, CESSDA Metadata Office, Unter Sachsenhausen 6–8, D-50667 Köln, Germany.
Email: Andre.Foerster@gesis.org

This *technical* view disregards the fact that RDM is not a purely *technical* process. Instead, we understand RDM as an *information* process, its implementation demanding the integration of different process perspectives. This means that the realization of RDM functions and activities does not only require technological solutions, but also an organizational structure that describes the organizational units involved in RDM as well as the communication and direction relations between them (Scheer, 2000).[2] This organizational perspective, which considers organizational units, positions and persons as well as their specific tasks, is essential for the implementation of an optimized RDM process. In particular, an organizational perspective increases the quality and functionality of the RDM process because it allows for a synchronization of the information processes and related workflows involved in RDM. The importance of a synchronization on an organizational level cannot be overstated notably in that it lays the foundations for the important technical synchronization. The remainder of this article is dedicated to the description of an RDM reference model for the implementation of RDM infrastructures and services at universities and research institutions. The aim of the model is to help research institutions fully incorporate RDM into their service and infrastructure landscape. To stretch the framework of our reference model, we begin with an introduction of its basic concepts, defining the research process itself as the *proximal* reference frame of the RDM process and the 'Architecture of Integrated Information Systems' (Scheer, 2000) as the *distal* reference frame. In this context, the dichotomy of proximal and distal – usually used as anatomical terms of location – describes the extent to which the circumstances of RDM have a direct (proximal) or indirect (distal) impact on the implementation of RDM functions. By integrating these two reference frames we then outline our RDM reference model, which helps to cut the 'rough diamond' of research data and is therefore named after the German term for diamond: **D**esigning an **I**nformation **A**rchitecture for **D**ata **Man**agement **T**echnologies (DIAMANT).

Research data management and its different reference frames

The research process as the proximal framework of RDM

The relevance of RDM mainly results from two goals formulated by the open science movement, namely improving research *economy* and research *integrity*. Research economy addresses the circumstances under which scientific exchange is promoted. This includes establishing conditions that enable the quick and easy access to research data, facilitate a better return of research investments and contribute to more sustainable research data. While research economy is mainly directed to improving the efficiency of the research process (Hinrichs-Krapels and Grant, 2016), research integrity focuses on improving the quality of scientific work. According to this concept, researchers should ensure reliability, honesty, respect and responsibility regarding all activities and results associated with the research process (ALLEA – All European Academies, 2017).

As the research process and the quality criteria directed towards it determine the conditions directly affecting RDM, we define it as the proximal reference frame for the implementation of RDM infrastructures and services. Thus, the primary function of RDM is implementing both research economy and research integrity within the research process. RDM reaches this functionality by telling a tale about research data (Surkis and Read, 2015). In contrast to publishing a journal article, which mainly reports the results of a study, RDM tells the whole story according to the research process. This means that RDM already starts with developing a research concept, and then goes on during data collection, analysis, dissemination and archiving (Vardigan et al., 2008). Conducting and documenting all these activities not only ensures research integrity, but also fosters efficiency of the research process, since additional work is averted, errors are avoided and resources are saved due to re-use of the data (ICPSR, 2012; UK Data Archive, 2011).

The Architecture of Integrated Information Systems as the distal reference frame of RDM

Whereas the research process provides a rather clear proximal reference frame of RDM, the information architecture, in which RDM is conducted, has been rather poorly conceptualized so far. As RDM is mainly performed in research institutions (i.e. universities and other research institutions), its information architecture should reside within this framework. The Architecture of Integrated Information Systems (ARIS) (Scheer, 2000) is especially suitable for application in this context, since the integration of IT and management – a crucial issue in RDM – was the origin of the concept. More specifically, ARIS relies on a 'common language for IT and management' (Schwickert et al., 2011), and thus allows modelling business processes accounting for the needs of all stakeholders involved in RDM. Furthermore, ARIS helps reducing the perceived complexity of RDM by allowing for an independent view on organizational, functional or performance issues of the business process. For this purpose, the ARIS information architecture is split into five perspectives, from which three are relevant to the application of ARIS on RDM.[3] While the functional and the organizational perspective enable the review of a specific level, the governance perspective represents the relationships between the two. Each of these perspectives is represented by three description levels (subject, data processing and implementation).

On the subject level, ARIS requires that users model actual and target situations, while on the data processing level the description of these situations is translated into IT requirements and tasks. Finally, software and hardware solutions carry out these tasks on the implementation level. As our reference model aims at providing a general basis for establishing business processes in the context of RDM, we restrict the model to the subject level. Thus, research institutions remain flexible when modelling the concrete specifications on the data processing and implementation level according to their specific needs.

The DIAMANT model

The aim of our reference model is to enable research institutions to **D**esign an **I**nformation **A**rchitecture for Data **MAN**agement **T**echnologies (DIAMANT) that allows for the ideal use of these technologies. This means, by implementing an appropriate information architecture, research institutions support researchers in cutting their 'rough diamonds' of research data into 'high-carat' ones. While the quality of the research data 'diamonds' depends on their compliance with open science criteria (e.g. FAIR), the quality of their manufacturing conditions depends on a proper subject level concept for effective and efficient RDM. Therefore, the following sections elaborate on the subject level concept of the different ARIS perspectives within the DIAMANT model (for a summary of the reference model see Figure 1).

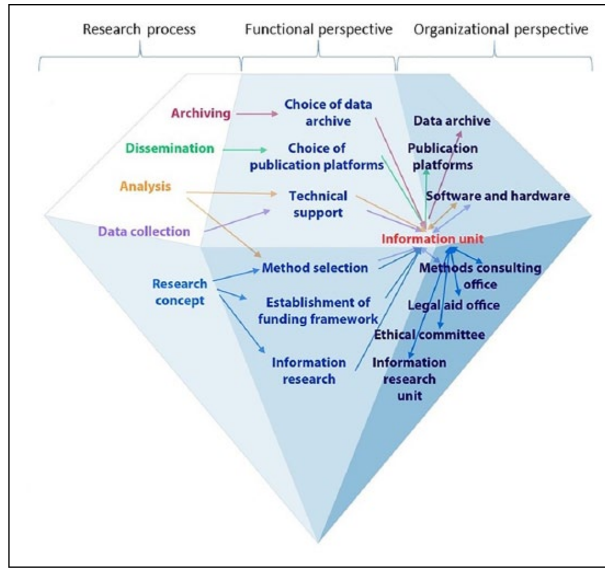

Figure 1. Schematic depiction of the DIAMANT model.
Note: This includes the envisaged relations between the functional and the organizational perspective within the governance perspective and following the different stages of the research process. For reasons of clarity, we omit the depiction of the direct relations between RDM functions and executing organizational units.

The ideal RDM process from the functional perspective

The functional perspective of the ARIS concept applied to the RDM process defines and describes all functions required from universities or research institutions in order to maintain research integrity and research economy. Therefore, we present these functions and their respective location within the research process. To help fulfilling these ideal RDM functions, we additionally provide to-do lists that define the conditions under which the respective activities have to be performed (King, 1967).

Typically, the RDM process starts with conceptualizing the research process. In developing a new research concept, the primary RDM functions are information search (Table 1), method selection (Table 2) and the establishment of a funding framework (Table 3). Whereas information search is designed to validate and conceptualize the research question, the method selection is associated with its operationalization. In other words, theoretical concepts and hypotheses are made measurable by defining observable events. In contrast, finding and establishing an adequate funding framework determines the effective and efficient conduct of the research project. It will be successful only when there are enough resources for staff and equipment. Additionally, the funding framework specifies strategies to fulfil the relevant RDM guidelines and determines which ethical and legal requirements have to be met.

In fact, fulfilling these three RDM functions is mostly the basis for every research project seeking external funds. Thus, it is an essential part of the value chain of universities and research institutions in general. Following the classification of Porter (1985), we can define these functions as

Table 1. Decision table: Information research.

Decision table: Information research		Rules			
		R1	R2	R3	R4
Conditions	Access to up to date information search services	Y	Y	N	N
	Information search competence	Y	N	Y	N
Activities	Establish access to up to date information search services	–	–	x	x
	Offer trainings/workshops	–	x	–	x

Table 2. Decision table: Method selection.

Decision table: Method selection		**Rules**							
		R1	R2	R3	R4	R5	R6	R7	R8
Conditions	Methodological competence	Y	Y	Y	N	Y	N	N	N
	Access to relevant software and hardware	Y	Y	N	Y	N	Y	N	N
	Access to discipline-specific methodological consulting	Y	N	Y	Y	N	N	Y	N
Activities	Provide methodological trainings/ workshops	–	–	–	x	–	x	x	x
	Provide software and hardware components	–	–	x	–	x	–	x	x
	Provide access to discipline-specific methodological consulting	–	x	–	–	x	x	–	x

Table 3. Decision table: Establishment of funding framework.

Decision table: Establishment of funding framework		**Rules**							
		R1	R2	R3	R4	R5	R6	R7	R8
Conditions	Knowledge of relevant legal and ethical parameters	Y	Y	Y	N	Y	N	N	N
	Knowledge of relevant discipline-specific and interdisciplinary RDM guidelines	Y	Y	N	Y	N	Y	N	N
	Knowledge of relevant scientific network structures	Y	N	Y	Y	N	N	Y	N
Activities	Establishment of a legal aid office	–	–	–	x	–	x	x	x
	Establishment of an ethical committee	–	–	–	x	–	x	x	x
	Establishment of an information webpage regarding relevant RDM guidelines	–	–	x	–	x	–	x	x
	Establishment of an information webpage regarding relevant funding agencies	–	x	–	–	x	x	–	x

primary functions or primary activities. Thus, research institutions should ensure support regarding these functions in order to compete efficiently.

In this paper, we present the activities associated with the RDM functions via decision tables. These tables define the criteria that have to be met in order to fulfil a specific function. Additionally, they define those activities that should be done in case the requirements are not met yet. Whether a specific activity has to be carried out ('x') or not ('-') is defined by the specific decision rules (R1, R2, …, Rn). These rules then in turn describe which criteria are met by the institution ('Y' = yes), and which are not met yet ('N' = no).[4]

After conceptualizing a research project, data collection and analysis are usually the next steps. In these stages, RDM has to provide technical support and consulting regarding the method selection. The technical support of the RDM process mainly accounts for safe data storage, data cleaning and documentation within the data lifecycle and the persistent findability of research data. Technical support activities (Table 4) usually do not have a direct impact on the creation of research data and thus can be defined as secondary activities (Schwickert et al., 2011). In general, secondary activities can be outsourced without impairing the value chain.[5]

The last two stages of the research process are publication and archiving. In these stages, researchers publish results, archive their data and disseminate them if appropriate. The main RDM function required in the stage of publication (Table 5) is the choice of a publication platform, while in the archiving stage (Table 6) the main RDM function is choosing a suitable archiving infrastructure. Again, these functions can be defined as secondary, so that outsourcing is possible regarding the related activities.

In total, there are six RDM functions related to the research process, while half of them are defined as primary. The other three functions are secondary or supporting functions not directly involved in the creation of research data. Thus, we conclude that these functions can be outsourced to external infrastructure and service partners. However, in case universities or research institutions would like to outsource these functions, they should make sure that there is sufficient information about these partners at hand (e.g. on the institution's website). This way, researchers have a quick access to relevant infrastructure and service providers.

Table 4. Decision table: Technical support.

Decision table: Technical support		Rules							
		R1	R2	R3	R4	R5	R6	R7	R8
Conditions	Safe data storage	Y	Y	Y	N	Y	N	N	N
	Data cleaning and documentation within the data lifecycle	Y	Y	N	Y	N	Y	N	N
	Persistent findability of research data	Y	N	Y	Y	N	N	Y	N
Activities	Provide necessary data storage capacities	–	–	–	x	–	x	x	x
	Provide access to relevant support tools for the (discipline-specific) documentation of data	–	–	x	–	x	–	x	x
	Provide information about the use of data documentation software and/or provide access to respective trainings/ workshops	–	–	x	–	x	–	x	x
	Provide infrastructures ensuring the persistent findability of data (e.g. bitstream preservation, persistent identifiers)	–	x	–	–	x	x	–	x

Table 5. Decision table: Choice of publication platforms.

Decision table: Choice of publication platforms		Rules			
		R1	R2	R3	R4
Conditions	Knowledge of suitable (open access) publication platforms	Y	Y	N	N
	Access to (open access) publication platforms	Y	N	Y	N
Activities	Provide information material regarding relevant (discipline-specific) publication platforms for research data and research data-related publications	–	–	x	x
	Provide access to (open access) publication platforms (e.g. open access publication fund)	–	x	–	x

Table 6. Decision table: Choice of data archive.

Decision table: Choice of data archive		Rules							
		R1	R2	R3	R4	R5	R6	R7	R8
Conditions	Knowledge of and access to existing data centres and repositories	Y	Y	Y	N	Y	N	N	N
	Warranty of persistent findability of data	Y	Y	N	Y	N	Y	N	N
	Matching between terms of use on the one hand, and project-specific, discipline-specific and legal requirements of data and texts on the other	Y	N	Y	Y	N	N	Y	N
Activities	Provide information about relevant existing data centres and repositories	–	–	–	x	–	x	x	x
	Support mechanisms guaranteeing the persistent findability of data (e.g. labelling with persistent identifiers)	–	–	x	–	x	–	x	x
	Provide support for assessing the matching between terms of use on the one hand, and project-specific, discipline-specific and legal requirements of data and texts on the other	–	x	–	–	x	x	–	x

The ideal RDM process from the organizational perspective

The organizational perspective of ARIS offers a proper representation of the organizational units and key players of a research institution as well as their relations and structures. Thus, it also presents an overview of the ways of communicating and directing between the different organizational units occupied with RDM. As organizational units do not only structure human resources, but also physical ones, services performed by machines can also be included in organizational units (e.g., in an IT centre).[6]

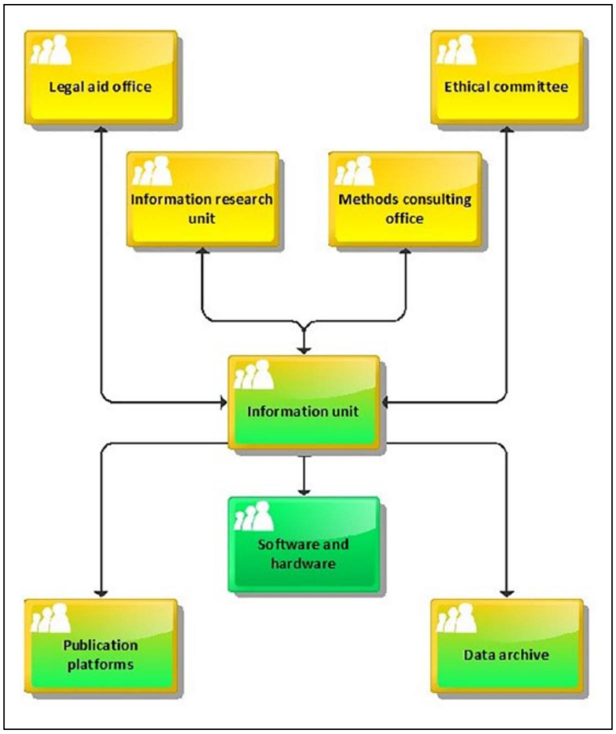

Figure 2. The organizational structure as the basis of the RDM process.

From the organizational perspective, eight general organizational units can be defined within the RDM process. Their relevance stems from their key roles in fulfilling the respective RDM functions. These relevant units are an information research unit, an ethical committee, a legal aid office, a methods consulting office, software and hardware providers, an information unit, publication platforms and a data archive. It is important to note that naming these units in our model should be understood as naming general types of organizational units. This means that we do not make any suggestions regarding the structure and resources of these units, but rather think that research institutions should find their own ways of designing them based on their needs.

To illustrate the relations between the different organizational units within RDM, we assume a process-oriented view of the organizational structure depicted in Figure 2. According to this model, there are four organizational units with human service providers only (highlighted in yellow). Together with the information unit, they are grouped on a higher hierarchical level. Thus, they form the functional units already needed in the early stages of the research process. The lower hierarchical levels, whose functionality is dependent on the interaction between human and information processing service providers (highlighted in green), correspond to the later stages of the research process.

The ideal RDM process from the governance perspective

While the functional and the organizational perspective reduce the complexity of RDM by focusing on a specific aspect of the process, the governance perspective accounts for the proper implementation of RDM infrastructures and services by modelling the interactions and relationships between the different perspectives. For reasons of keeping the DIAMANT model generalizable, we restrict it to modelling the structural relations between RDM functions and organizational units.[7] Thus, it offers structural relations between RDM functions and organizational units that allow for the efficient and effective performance of the different RDM activities. By relying on already existing RDM literature as well as expressed needs of researchers (Blask and Förster, 2018), we define how strong each organizational unit is involved in fulfilling RDM functions within the model.

The best way to depict the structural relations between RDM functions and organizational units is a matrix (Table 7). Within this matrix, we define clear areas of accountability and responsibility allowing research institutions to identify

Table 7. Type of involvement of the different organizational units regarding the initiation, advancement, provision and conduction of support mechanisms fulfilling the RDM functions.

RDM functions \ Organizational units	Information research service	Methods consulting office	Ethical committee	Legal aid office	**Information unit**	Software and hardware	Publication platforms	Data archive
Information research	p				r x			
Method selection		p			r x	p		
Establishment of funding framework			p	p	r x			
Technical support					r x	p		x
Choice of publication platform					r x		p	
Choice of data archive					r x			p

Functional relations are marked as follows: r = responsible; p = provides/executes; x = receives results.

the ways of communicating and information processing that have to be implemented for realizing the respective RDM functions. For this purpose, we suggest three types of relationships between RDM functions and organizational units within the DIAMANT model. The first type describes which organizational unit is responsible for initiating support mechanisms regarding the fulfilment of a specific RDM function. Specifically, they are authorized to issue directives toward the organizational units of the second type. These second type units are involved in the provision and execution of support mechanisms. In order to ensure effective, accurate and complete execution of RDM functions the organizational units of the first type supervise these executive units. The third type of relationship is the link between the various function-specific organizational units. Units of this type have to be informed about relevant results stemming from the execution of function-specific event process chains. In doing so, they help to advance support activities, reduce additional work and create a more transparent and flexible RDM process.

The responsibilities and relations between RDM functions and organizational units depicted in Table 7 show that a researcher contacting the information unit always initiates the respective functions in the various stages of the research process. Thus, not only the research process itself but also the RDM process lies within the autonomy of each researcher. In contrast, the information unit provides researchers with RDM knowledge (e.g. via training, workshops or information material) and forwards researchers' requests to the appropriate executive unit, if necessary. For selected RDM activities, it is also possible for researchers to contact the relevant executive unit directly.[8] However, DIAMANT recommends the interposition of the information unit, especially if external units provide RDM support mechanisms. Restricting the main responsibilities to the information unit allows for boosting the efficiency of the RDM process. Furthermore, this reference model meets the demand of many researchers regarding the 'one-stop' solution of RDM service (Blask and Förster, 2018). Last but not least, feedback loops from the organizational units to the information unit help evaluating and refining RDM services and therefore increase both effectiveness and efficiency of the RDM process. This (self-) learning process shortens the event process chains within the whole process as growing RDM competencies of researchers render forwarding RDM needs to executive units more and more unnecessary.

Summary: The DIAMANT model and its implications

In this article, we have presented a reference model providing an orientation framework for the implementation of RDM infrastructures and services at universities and research institutions. The DIAMANT model hereby focuses on making the RDM process more effective and efficient. In other words, it provides the ideal framework for performing RDM tasks in a manner that minimizes the effort researchers have to put into RDM during their scientific work.

To this end, we have modelled the RDM process within the DIAMANT model on the basis of the ARIS concept (Scheer, 2000). In accordance with this concept, our modelling strategy covered multiple perspectives, namely the *functional* perspective, the *organizational* perspective and the *governance* perspective. Together, they form a model that enables the optimization of the RDM process for the benefit of a user-friendly implementation into the research process. Most importantly, this implementation process is practically independent of current resource allocation at research institutions and thus extremely flexible. Specifically, one might create the relevant boundary conditions for an optimized RDM just by implementing (and institutionalizing) the information unit and corresponding feedback loops with the remaining outsourced organizational units.

We believe that following the implementation strategy proposed in the DIAMANT model has several practical advantages regarding RDM. For one thing, it curbs *doing RDM things for the sake of doing RDM things* (Klump and Ludwig, 2013), since the DIAMANT model makes clear statements about which RDM functions are relevant for the value creation process (i.e. the research process) of research institutions, and which functions in contrast have a rather supporting character. For another thing, research institutions can use existing infrastructures and services on the basis of this classification without facing competitive disadvantages. Finally yet importantly, the DIAMANT model reduces the creation of isolated applications of RDM at research institutions as a whole, supporting the establishment of national as well as the formation of international research data infrastructures. This increased structuring and organization of existing infrastructures also promotes the implementation of quality criteria such as the FAIR principles or the development and establishment of community-specific standards.

In conclusion, the DIAMANT model demonstrates the conceptualization of an RDM information architecture that is guided by the requirements of the research process. Given that the requirements related to the RDM infrastructure and service landscape are oriented towards the needs of researchers this model essentially contributes to the implementation of an optimized RDM process. Optimized, because it is directed towards using RDM in its intended function, namely to cut the 'rough diamond' of research data.

Author Note

Katarina Blask is now affiliated to ZPID - Leibniz Institute for Psychology Information, Germany.

Acknowledgements

We would like to thank Marina Lemaire, Gisela Minn, Georg Müller-Fürstenberger and Erich Weichselgartner for their support in writing this paper within the PODMAN project.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is a result of the PODMAN project (www.fdm.uni-trier.de/en) that received a specific grant from the German Federal Ministry of Education and Research directed at the investigation of research data management and its life cycle at universities and research institutions (Grant Number 16FDM023).

Notes

1. One of the best known representations of the RDM process is, for example, the Curation Lifecycle Model of the Digital Curation Centre (Higgins, 2008).
2. In this context, the term 'organizational structure' does not mean to build up a specific institute, but rather to (re-)structure the design of an already existing organization or institution (e.g. a university) regarding its governance mechanisms of RDM.
3. Since the data perspective and the performance perspective are hardly generalizable in the context of RDM, we limit our model to the functional, organizational and governance perspective.
4. Example for Table 1: A research institution has established the access to up-to-date information search services (Y), but information search competence is insufficient (N). This situation is reflected in decision rule R2: appropriate training should be provided, whereas there is no call for action regarding the information search services.
5. However, the method selection is a primary function, since the activities associated with this choice have a direct impact on creating data. Thus, this function has to be assigned to the core business of any university or research institution.
6. Following this argument, organizational units mainly consist of human or technical service providers, while the latter can be divided into material machining and information processing (e.g. a computer) (Scheer, 2000). As the technical service providers involved in the RDM process are occupied with information processing only, a technical organization unit in our model always consists of information processing service providers.
7. Besides the *structural* relations, ARIS usually requires the additional description of system *dynamics* within the governance perspective via event-driven process chains (EPC). However, the design of these EPC is highly dependent on the structure of the specific institution. Thus, it would reduce the model's generalizability.
8. This should be possible in case these organizational units have already been established at the researcher's institution.

ORCID iDs

Katarina Blask https://orcid.org/0000-0003-2062-4059
André Förster https://orcid.org/0000-0002-5451-4598

References

ALLEA – All European Academies (2017) *The European Code of Conduct for Research Integrity*. Available at: https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf (accessed 8 November 2018).

Blask K and Förster A (2018) *Problembereiche der FDM-Integration (Daten der 1. Interviewwelle des BMBF-Projekts PODMAN; PODMAN W1): Version 1.0.0 [Data set]*. Available at: http://doi.org/10.5281/zenodo.1492182 (accessed 20 November 2018).

DFG (ed.) (2013) *Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Denkschrift Empfehlungen der Kommission 'Selbstkontrolle in der Wissenschaft' = Proposals for Safeguarding Good Scientific Practice Memorandum: Recommendations of the Commission on Professional Self Regulation in Science*. Weinheim: Wiley-VCH.

Higgins S (2008) The DCC Curation Lifecycle Model. *International Journal of Digital Curation* 3(1): 134–140.

Hinrichs-Krapels S and Grant J (2016) Exploring the effectiveness, efficiency and equity (3e's) of research and research impact assessment. *Palgrave Communications* 2. Available at: https://www.nature.com/articles/palcomms201690.pdf (accessed 26 November 2018).

ICPSR (2012) *Guide to Social Science Data Preparation and Archiving*. Available at: http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf (accessed 27 October 2017).

King PJH (1967) Decision Tables. *Computer Journal* 10(2): 135–142.

Klump J and Ludwig J (2013) Forschungsdaten-Management. In: Neuroth H, Lossau N, Rapp A, et al (eds) *Evolution der Informationsinfrastruktur: Kooperation zwischen Bibliothek und Wissenschaft*. Glückstadt: VWH Verlag Werner Hülsbusch, pp. 257–275.

Porter ME (1985) *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: Free Press.

Scheer A-W (2000) *ARIS – Business Process Modeling*. Berlin, Heidelberg: Springer.

Schwickert AC, Müller L, Bodenbender N, et al. (2011) *Geschäftsprozessmodellierung mit ARIS*. Available at: https://wiwi.uni-giessen.de/dl/down/open/Schwickert/f1dbb464f72be49391c38d2e5328678cd7653e95e908312925e82e84c835203f7020665f52efdafe51443bc769f387fb/Apap_WI_GI_2011_01_PDF_Download_gesch_tzt.pdf (accessed 8 November 2018).

Surkis A and Read K (2015) Research data management. *Journal of the Medical Library Association JMLA* 103(3): 154–156.

UK Data Archive (2011) *Managing and sharing data: Best practice for researchers*. Available at: http://www.data-archive.ac.uk/media/2894/managingsharing.pdf (accessed 27 September 2016).

Vardigan M, Heus P and Thomas W (2008) Data Documentation Initiative: Toward a standard for the social sciences. *International Journal of Digital Curation* 1(3): 107–113.

Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3: 160018.

Author biographies

Katarina Blask studied psychology at Trier University. From 2010 to 2014 she worked at the Chair of Social Psychology at Trier University and received her doctorate there. As a postdoc she worked for the project 'Understanding the basic processes in affective learning' (funded by the German Research Association DFG) and the PODMAN project (funded by the German Federal Ministry of Education and Research BMBF). In the future, she will focus her research on the optimization of research-supporting services in the PsyCuraDat project at the Leibniz Centre for Psychology Information (ZPID).

André Förster studied political science, sociology, and communications and media at the Universities of Düsseldorf and Cologne. From 2013 to 2017 he worked for the DFG-funded German Longitudinal Election Study (GLES) and earned his doctorate in Social Sciences at the University of Cologne. After working for the PODMAN project at Trier University from 2017 to 2019, he became project leader of the CESSDA Metadata Office at the GESIS – Leibniz Institute for the Social Sciences. His research interests are data management as well as political participation and election studies.