Research Data Infrastructures – A Perspective for the State of North Rhine-Westphalia in Germany

Ania López[1], Raimund Vogl[2], Sabine Roller[3]

[1]University of Duisburg-Essen, Universitätsstr. 9 - 11, 45141 Essen, ania.lopez@uni-due.de
[2]University of Münster, Röntgenstr. 7-13, 48149 Münster, rvogl@uni-muenster.de
[3]University of Siegen, Hölderlinstr. 3, 57068 Siegen, sabine.roller@uni-siegen.de

Keywords
Research Data Management, Cooperative Storage Infrastructures, Storage of Research Data

1. ABSTRACT

With the digital transformation progressing fast, scientific research is massively affected by the ever-increasing amount of digitally generated research data. New ways have to be established to properly manage this digital research data to ensure that the standards of good scientific practice, first of all the reproducibility of scientific results, can be fulfilled. Sustainable management of digital research data requires standards for metadata annotation and a future-proof infrastructure for handling and storing these data. In this paper, we describe the current state of affairs in Germany with respect to research data management (RDM) and the establishment of a national research data infrastructure. The special focus is on the currently ongoing activities in the state of North Rhine-Westphalia to start implementing research data infrastructures for universities through several cooperative multi-university projects, tailored to the specific needs and approaches of the participating universities.

2. RESEARCH DATA MANAGEMENT IN GERMANY

During the last decades, an enormous quantitative and qualitative growth of research data, i.e. data used or generated during a research project, can be observed. Depending on the respective scientific discipline, research data exist in many formats including numerical data, text, transcripts, images, video, audio recordings, etc.

The relevance for research data itself has drastically increased, as technology is making it feasible to ensure reliable verification of results and permits new and innovative research built upon existing information. Research data has shifted from being considered a residue of the scientific process, to an important resource for future research.

Worldwide, several research funding institutions as well as various publishers are more and more requiring good data management practices from scientists, such as the use of data management plans or the publication of research data associated to classical paper publications. On a national level, several countries have built technical infrastructures for the handling of research data.

In Germany, the importance that has been given to the topic of RDM is reflected and has been promoted by several organisations and institutions, such as the German Research Foundation[1] (DFG, 2013), the German Rectors' Conference[2] (HRK, 2015), the Alliance of Science Organisations in Ger-

[1] The German Research Foundation (DFG) is the largest independent research funding organisation in Germany, see also www.dfg.de/en/ (retrieved March 3, 2017).
[2] The German Rectors' Conference (HRK) is the voluntary association of public and government-recognised universities and other higher education institutions in Germany, having currently 268 member institutions in which around 94 per cent of all students in Germany are enrolled, see also https://www.hrk.de/home/ (retrieved March 3, 2017).

many[3] (Alliance of German Science Organisations, 2010) and the Council for Scientific Information Infrastructures[4] (RfII, 2016) among others. These statements come along with similar international activities, just to mention the vision of the European Commission where "a scientific e-infrastructure that supports seamless access, use, re-use and trust of data" (European Union 2010, p. 4) is described.

Besides this, the Council for Scientific Information Infrastructures has stated the need for national coordinated activities, and described a National Infrastructure for Research Data (known under the German acronym NFDI) which is expected to be built within the next 15 years. The NFDI on the one hand is regarded as a net of data centers with discipline-specific focus, which will offer solutions for RDM. On the other hand, the NFDI will have to focus on the needs for standardization and basic general services, addressing the explicit wish of the Council to facilitate crossing disciplinary borders and therefore allowing universities or groups of universities to play an important role providing basic data services for researchers and therefore being part of the NFDI (RfII 2016, p. 38).

As a consequence, universities have to face the challenge to build and develop (infra-)structures for a professional research data management (RDM) to support their researchers, such as technical (storage) infrastructures, training and support, research data policies, etc.

Being a collective task, cooperation between universities and research institutions seems to be the best road to go.

3. ACTIVITIES IN GERMAN FEDERAL STATES

Due to its federal organisation, several coordinated activities at German federal state level can be observed.

In Baden-Wuerttemberg, the Ministry of Science, Research and Arts (MWK) identified several fields of action in the context of digitalization of science, such as licensing of electronic information, digitalization, open access, research data management and virtual research environments. Within each of these topics, several projects have been funded. In the context of RDM, a starting project called bwFDM-Communities was funded during 2014 and 2015 and had the aim to contact researchers among all universities within Baden-Wuerttemberg as to understand the needs for research data management infrastructure. During the project, 627 interviews were carried out. In total, 779 persons were asked about their usage of data within their research activities, as to understand the needs of infrastructure. Description of methodology and the results can be found within a final report (Tristram 2015, German only), and the huge amount of resulting data is properly visualized on the web[5]. The outcome of the project spawned a wave of funding, where the ministry awarded grants with a total amount of 3 million Euros for 7 projects that are being carried out by different universities across Baden Wuerttemberg. The projects address different needs in the research data management cycle, first results are expected this year, and will be available for all universities within the state of Baden Wuerttemberg.

In Hessia, the universities and the state have settled a target and performance agreement for the future of scientific infrastructure of the universities within the state for the years 2016-2020. For research data management, the goal is to develop a joint RDM-infrastructure that can be used by the

---

[3] The Alliance of Science Organisations in Germany (Allianz der Wissenschaftsorganisationen) is a union of the most important German research organisations. It issues statements relating to research policy and funding and the structural development of the German research system. See also https://www.leopoldina.org/en/about-us/cooperations/alliance-of-science-organisations/ (retrieved March 3, 2017).

[4] The Council for Scientific Information Infrastructures (RfII) was initiated by the Joint Science Conference of the Federal States (GWK) and the Federal Government of Germany for a service period of four years in 2014. The 24 members represent a broad spectrum of scientific disciplines and institutions. The Council´s scope of work is the strategic development of a contemporary and sustainable infrastructure for access to scientific Information. See also http://www.rfii.de/en/ (retrieved March 3, 2017)

[5] See http://bwfdm.scc.kit.edu/101.php. (retrieved March 17, 2017).

universities within the state. The project is funded with a total of 3.2 million Euro and will be evaluated at the end of the third year by the German Research Foundation (DFG). For the enforcement of the project and to ensure communication, part of the funds have been distributed for staff being responsible for RDM at each of the five universities within the state of Hessia.

In the state of North Rhine-Westphalia (NRW), the Digitale Hochschule (DH-NRW) is a cooperation structure of the 40 public universities and the Ministry of Innovation, Science and Research (MIWF) of NRW to address the challenges of digitalization and to foster cooperation in information management at the universities within the state. It aims to develop strategies in the field of research, teaching, education and infrastructure such as to initiate projects and cooperation across universities.

Concerning RDM, on a call of the DH-NRW (formerly DV-ISA), a working group was formed in 2015, composed of representatives of libraries, IT-centers and researchers from 9 universities within the state with the aim to chart the activities regarding RDM found within all universities in NRW. Similar to known national and international survey results to RDM (part of the activities has been reported in (Thoring, 2016)), the biggest need for services for RDM can be found in the fields of information and technical infrastructure (DV-ISA, 2016). After the status quo analysis, in spring 2016 a follow-up expert group composed of 6 representatives of libraries, IT-centers and researchers from universities in NRW was nominated within DH-NRW. The aim for this group was to work out a proposal for sustainable RDM-services for universities in NRW, to ensure cooperation and coordination within the universities and to establish solutions to RDM, which can be incorporated into a developing national strategy. The so-far focus of the working group was on communication and information for ensuring the possibility of cooperative approaches within universities. Some of the results are the monthly exchange format "Jour Fixe RDM" (regularly visited by 20-30 different participants from the universities of NRW, coming from libraries, IT-centers and research support departments) and general information material as a set of slides for individual re-use[6]. Having started the discussion about RDM with different stakeholders, the dialogue and coordination of a planned NRW-wide storage-infrastructure for research data was initialized.

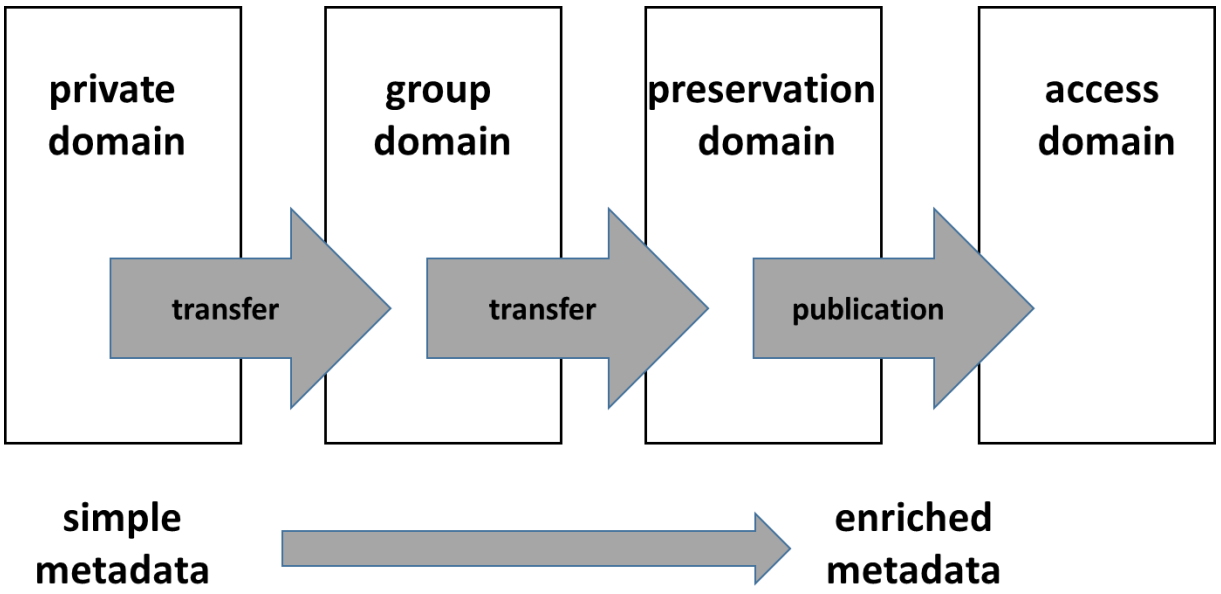
4. APPROACHES FOR STORAGE INFRASTRUCTURE IN THE STATE OF NORTH RHINE-WESTPHALIA

A widely-used model from a technical point of view for describing the management of data during the research process is the curation domain model (CDM; see Forschungsdaten.org, 2015; derived from Teloar, A. Groenewegen, D., Harboe-Ree, C., 2007). It describes 4 domains: The *private domain* describes the handling of data by a single researcher or a small team of researchers. On most large projects data is handled by bigger groups of researchers, a lot of times crossing institutional and disciplinary borders, this is designated as the *group domain*. The handling of data in a group domain requires explicit metadata to be understood not only by an individual researcher (as in the private domain). At some point within the research process, the data is to be archived. This is part of the principles of good scientific practice, to preserve research results, especially when publications have been derived from them. In this case, called the *preservation domain,* the data must be provided with more metadata, which can be descriptive but also technical, if for example the data is not publicly available, so it has to cover rights management, and also it has to be ensured that the data is persistently referenced, with a PID-Service (i.e. Handle, DOI, URN, etc.). The last described *accessibility domain* addresses the case, when research data is made publicly available, for example through a repository or a publisher.

The four domains in the curation domain model also give a clue that the physical systems used for generating, analyzing and storing research data will differ between these domains. While researcher will often use their personal computers and laptops (often with external USB disks, as (Thoring, 2016) shows) in the private domain, collaborative platforms (like sync & share cloud services, shared network drives or Sharepoint) might be used in the group domain. The preservation domain is often realized through a repository usually operated by university information or library services. The access domain will often also build on repository systems, but also community or project specific web based solutions can be found.

---

[6] https://doi.org/10.5281/zenodo.165126 (retrieved March 17, 2017).



The big issue here of course is the sustainability and long-term operation of these various systems to ensure access to research data for verification and reuse.

To create a sustainable infrastructure for research data management, especially for storage and long term availability of data, the board of IT directors of the universities in North Rhine-Westphalia (ARNW) in conjunction with DH-NRW has discussed possible solutions in a September 2016 workshop.

Out of the discussion of different user needs and requirements created through pre-established IT-environments at various participating universities, this workshop resulted in the formation of three multi-university consortia setting out to create joined grant proposals for solutions that address their shared research data management infrastructure needs.

5. THREE APPROACHES, ONE FOCUS

The common aim of all three consortia is the creation of sustainable infrastructure to support all or parts of the domains in the CDM. The individual approaches differ widely:

1. One consortium focusses mainly on a common storage systems solution with sub-systems established at multiple data center locations of the participating universities, but providing a unified common name space for data access, thus facilitating cross university research collaborations. At each participating university, a 2-tiered configuration is planned, addressing the group and preservation domain. Besides technical cooperation, the joint project will foster the creation of expertise in running such complex systems in a way to fit and benefit the specific demands on RDM processes within different user communities. An open standards interoperability strategy for later expansion with potentially different vendor storage products is planned.
2. A second consortium of universities focusses on the tamper- and audit-proof archival of research data. For data security and efficient re-use of these data, distribution of multiple data copies to the decentralized storage systems located at the data centers of the three participating universities is a key concern. Coordinated data life cycle management of these data instances (like implementation of retention and deletion periods for all data copies) and unified data access control (with identical rule sets for all data copies) are essential parts of this project.
3. In a third project, the creation of a cooperatively operated sustainable storage and service infrastructure for all stages of the data curation domain model, i.e. for creation, analysis, collaborative exchange, storage, archival and accessibility is planned. This project will be exclusively based on open source freeware software solutions: Ceph as a future-proof Software Defined Storage solution and OpenStack as an on-premise cloud platform for the flexi-

ble, resource efficient provision of virtual system resources for research groups along the stages of the data curation continuum from scientific computation and analysis down to the hosting of web based community focused accessibility for research data.

With the later project being coordinated by one of the authors, this will be the focused in more detail in the next section with an exposition on the design motivations for this innovative hardware-only approach to a research data infrastructure.

All three projects are meant to be incorporated into the established cross-university cooperative IT solutions landscape in NRW and align with other projects focused on different curation aspects of digital research data:

1. Long term archiving and backup of data, mainly in tape archives: a consortium of 13 universities in NRW are joined in a common project for the IBM Tivoli Storage Manager software product for the management of tape archive systems. This is seen as common basis for long term archiving for all above the research data infrastructure projects.
2. Offering for a common repository service: the university library service center (hbz) of NRW is pursuing a project to create a SaaS offering for a research data repository utilizing the Rosetta software product by ExLibris.
3. Research data services orchestrating the interconnection of data management utilities along the stages of the curation domain model: based on the well-established NRW-wide university operated cloud storage sync & share platform sciebo (Vogl, 2016), which is already supporting some 800 research project groups as a collaboration and storage resource, a set of research data services is planned to support the process of enriching the meta data and creation of submission and archival information packages (according to the OAIS model).

6. A SUSTAINABLE OPEN SOURCE FREEWARE APPROACH TO A RESEARCH DATA INFRASTRUCTURE FOR MULTIPLE UNIVERSITIES

The ongoing digital transformation has led to an exponential increase in digital data – and this is more than just a lip service to the currently unavoidable digitalization buzzword battle. It is e.g. very manifest at University of Münster and can be visualized with the development of storage capacity provisioned by the IT center and the backup data volume consumed in the central backup system. From 2010 onward, the disk capacity of storage systems has surpassed the backup data volume – the growth of data has led to a situation, where tape backup/restore and archival are no longer practicable – data has to remain available online and users rely on the data integrity features of the online storage systems to ensure the safe preservation of their data. Besides an ever-growing demand for user accessible highly performing online storage systems (tier 1 storage), a clear trend also leads to background storage systems, which are (at least currently) disk based but not so much performance oriented and not directly user accessible - they serve as online data pools for backup and archive systems (tier 2 storage), since tape archives are not suitable to provide the required restore times for large backup data sets. With disk capacities still growing (faster than tape) and a shift to affordable high volume nonvolatile solid state storage imminent, tape storage seems to be headed for a rather small niche of applications where offline media are of special importance, if not becoming completely obsolete. Still, tape used to be the medium that was accredited with very long term stability (30 years estimated archive life e.g. for IBM 3592 tape cartridges).

The preservation and accessibility of digital data for long periods of time has thus to be well considered for an all online storage future: 10 years at the minimum, this being the usual data retention requirement imposed by funding agencies, but for particularly valuable research data, unlimited preservation as part of the human heritage is called for. In such a context, a static offline media approach is doomed to fail, due to unavoidable media obsolescence. For online storage systems with a typical life cycle of 5-7 years, a passive approach to data preservation is out of the question, anyways.

The only way to deal with this, is constant migration of data to fight media and hardware obsolescence and silent bit flip data corruption. Sustainable storage systems thus must have instruments for data migration build in. Experience has shown how painful, and not unlikely even lossy, the migration from one file system based storage system to the next generation can be.

Object storage systems with standardized metadata hold the promise of a smooth migration process, especially when new storage nodes can be added to such a storage cluster on demand during operation, and obsolete hardware can be vacated by targeted data migration and then be taken offline. Data redundancy in the storage layer safeguards data loss due to media failure or corruption, and data replication between data center sites secures against catastrophic events.

A further possible step for obsolescence prevention is the reliance on open source software defined storage solutions, which eliminate the risk of end of life products – an active developer community takes care of the code-stewardship and in the worst case, skilled staff at universities can provide basic support to keep the system operational, even when fixes in the source code are necessary.

To create an infrastructure not only for research data storage, but also one that supports the research processes in data generation, analysis, and accessibility, a common approach is to provide an on-premise cloud environment for self-provisioned, scalable science & engineering resources through virtualized systems.

This can also be an infrastructure or platform as a service (IaaS, PaaS) for joint cloud-like collaborative solutions - such as sciebo, for instance.

For software defined storage, Ceph (Weil, 2006) currently is the top ranking open source freeware solution that holds the promise to deliver on all the above stated objectives for sustainable research data storage. With respect to cloud stacks, OpenStack (OpenStack, 2017) has established itself as the main open source product in the last few years.

These two products thus have been selected as the software part for this project – with confidence among the consortium partners that an open source freeware approach is the right way to go. It is further planned to leverage synergies, minimize operational cost and establish self-supported maintenance structures with minimum reliance on commercial support by pooling the skill and experience of the storage systems and virtualization specialists from the IT centers of the consortium universities for mutual assistance and stand in, for joint work like release testing and rollout concept, as well as coordination of overlapping administration tasks such as cross site data replication planning. This of course will be a virtual structure with a focus on periodical meetings for exchange and planning, jointly organized trainings and online collaboration resources for coordination of administrative activities. But some 20 years of experience with cooperation in the context of the IBM TSM backup and archive software has shown, that such cooperation is possible, effective and sustainable. So, there is reason to expect that it will be possible to bring this kind of multi-university IT operations cooperation to an even higher level in the context of open source software, where the field of activities for the IT specialists is much wider, beyond administration of a commercially supported proprietary product.

Open source freeware has grown in importance over the last 25 years – symbolized by the now tantamount presence of Linux in server computing not only in the academic and research domains. Similar developments are imminent in other fields of IT infrastructure – most of all software defined storage solutions and cloud stacks – and a major change of paradigm seems to be currently taking place there. The open source software packages that are part of the current planning may change in popularity in the next 15 years, but basic concepts of software defined infrastructures, open source and freeware, and cooperatively provided on premise IT services for higher education and research are to stay.

7. DISCUSSION

With good reason, the RfII recommendations for research data management in Germany have been subtitled “Performance through Diversity”. NRW as Germany’s largest state with some 40 universities, 14 of these research universities, offers a whole lot of diversity, even in the approaches and strategies towards sustainable research data management. The three approaches towards research data infrastructure described above will create offers that support both the phase of the acquisition and processing of data in the course of scientific work as well as the long-term archiving considerably better and more comprehensively than is currently possible. They aim at supporting consistent and reliable data storage in all domains of the research process. All three projects thus provide a discipline-agnostic storage infrastructure that offers the capacity and performance to cope with the research data management requirements of the respective universities, with the additional benefit of multi-location replicated storage for protection of sensitive data and data lifecycle support.

The timeline for the establishment of the national research data infrastructure (NFDI) laid out in the RfII recommendations is a realistic 15 years. With the life cycle of IT system platforms being in the 5-7-year timeframe, these three approaches will long have given way to follow ups by then. It is an iterative approach, and as it is common with research (and IT) projects, objectives and concepts have to be progressively adjusted to the ongoing developments. The projects described here will provide urgently required infrastructures for research now – tailored to the needs of the research groups within the universities in the respective consortia – and will help foster even deeper cooperation between universities in jointly operating IT services. Thus they will provide a valuable and lasting contribution to the long-term aim of a national research data infrastructure.

8. REFERENCES

Alliance of German Science Organisations (2010). *Principles for the Handling of Research Data*. Retrieved March 3, 2017, from: http://www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Principles_Research_Data_2010.pdf.

DFG Deutsche Forschungsgemeinschaft (2013). *Memorandum "Safeguarding Good Scientific Practice"*. Retrieved March 3, from: http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf.

DV-ISA Arbeitskreis DV-Infrastruktur der Hochschulen in NRW (2016). *Umgang mit digitalen Daten in der Wissenschaft: Forschungsdatenmanagement in NRW - Eine erste Bestandsaufnahme*. Retrieved March 3, 2017, from: https://doi.org/10.5281/zenodo.200429.

European Union (2010). *Riding the wave. How Europe can gain from the rising tide of scientific data*. Retrieved March 3, 2017, from: https://ec.europa.eu/eurostat/cros/content/riding-wave_en

Forschungsdaten.org (2015). Curation Domain Model, from: http://www.forschungsdaten.org/index.php/Curation_Domain_Model

HRK Hochschulrektorenkonferenz (2015). *Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien, Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel*. Retrieved March 3, 2017, from: https://www.hrk.de/uploads/tx_szconvention/Empfehlung_Forschungsdatenmanagement__final_Stand_11.11.2015.pdf.

OpenStack (2017): https://www.openstack.org/ (visited 24.3.2017).

RfII German Council for Scientific Information Infrastructures (2016). *Enhancing Research Data Management: Performance through Diversity. Recommendations regarding structures, processes and financing for research data management in Germany. Retrieved March 17, 2017 from* http://www.rfii.de/en/category/documents/

Thoring, A., Rudolph, D., Vogl, R., 2016. *How Do Researchers Manage their Data? An Empirical Comparison between Policies and Practice*. EUNIS 2016 Congress, Thessaloniki, Greece. http://www.eunis.org/eunis2016/wp-content/uploads/sites/8/2016/07/EUNIS2016_Book_of_Proceedings_v6.pdf

Treloar, A. Groenewegen, D., Harboe-Ree, C., 2007. *The Data Curation Continuum - Managing Data Objects in Institutional Repositories*. D-Lib Magazine, 13(9/10), S. 13. doi:10.1045/september2007-treloar.

Treloar, A., Wilkinson, R., 2008. Access to Data for eResearch: Designing the Australian National Data Service Discovery Services, International Journal of Digital Curation, 3(2), 151–158. doi:doi:10.2218/ijdc.v3i2.66.

Tristram, F. et al. (2015). *Öffentlicher Abschlussbericht von bwFDM-Communities. Wissenschaftliches Datenmanagement an den Universitäten Baden-Württembergs*. Retrieved March 17, 2017, from: https://bwfdm.scc.kit.edu/downloads/Abschlussbericht.pdf

Vogl, R., Rudolph, D., Thoring, A., Angenent, H., Stieglitz, S. & Meske, C. (2016). How to Build a Cloud Storage Service for Half a Million Users in Higher Education: Challenges Met and Solutions

Found, Proceedings of the 49th Annual Hawaii International Conference on System Sciences (HICSS 2016), 10.1109/HICSS.2016.658, https://www.computer.org/csdl/proceedings/hicss/2016/5670/00/index.html.

Weil, S. A., et al, 2006. *Ceph: a scalable, high-performance distributed file system*, OSDI '06 Proceedings of the 7th symposium on Operating systems design and implementation, 307-320. http://dl.acm.org/citation.cfm?id=1298485. See also http://ceph.com/.

Whyte, A., Tedds, J. (2011). *Making the Case for Research Data Management*. DCC Briefing Papers. Edinburgh: Digital Curation Centre. Retrieved March 3, from: http://www.dcc.ac.uk/resources/briefing-papers/.

9. AUTHORS' BIOGRAPHIES

Dr. Ania López is head of the group research and innovation at the University Library of Duisburg-Essen (Germany) since 2015. Besides this, she coordinates the expert group of research data management among Universities of North Rhine-Westphalia within the Digitale Hochschule NRW since 2016. She holds a PhD in mathematics from University of Düsseldorf (Germany). After completing her PhD studies in 2008, she joined the University Library of Duisburg-Essen as a subject librarian for mathematics, electrical and civil engineering and has been responsible for several IT-projects within the University Library. She is a member of the executive team of the DINI-nestor-group of RDM in Germany and a member of the european LIBER Metrics Working Group. Her current research interest lies in the management of research data and bibliometrics and metrics in general. Contact info: https://www.uni-due.de/zim/services/suchdienste/mitarbeiter.php?id=50724 .

Dr. Raimund Vogl is director of IT for the University of Münster (Germany) since 2007. He holds a PhD in elementary particle physics from University of Innsbruck (Austria). After completing his PhD studies in 1995, he joined Innsbruck University Hospital as IT manager for medical image data solutions and moved on to be deputy head of IT. He is active as a member of the executive team of the IT strategy board of the Universities in North Rhine-Westfalia and is board member and treasurer for EUNIS (European University Information Systems Organisation). He is a member of GMDS, EuSoMII and AIS and is representing Münster University in EUNIS, DFN, ZKI, DINI and ARNW. His current research interest in the field of Information Systems and Information Management focuses on the management of complex information infrastructures. More Info: http://www.uni-muenster.de/forschungaz/person/10774?lang=en .

Prof. Dr.-Ing. Sabine Roller is professor for Simulation Technics and Scientific Computing and director of the Center for Information and Media Technologies (ZIMT) at the University of Siegen (Germany) since 2013. She holds a PhD in aerospace engineering from University of Stuttgart (Germany). After completing her PhD studies in 2004, she joined the national supercomputing center HLRS in Stuttgart as head of the "Scalable Computing and Coupled Systems" group. She became professor for Applied Supercomputing in Engineering at RWTH Aachen University (Germany), before following the call to Siegen. Her current research interest is in procedures and infrastructures for research data management, with a special focus on media science as well as large scale simulation data. She is member of ACM, IEE, AIAA, and is representing the University of Siegen in DFN, ZKI, DINI and ARNW. More Info: http://www.zimt.uni-siegen.de/ and http://www.mb.uni-siegen.de/sts/