Göttingen eResearch Alliance

Report for External Advisory Board

2018/2019

Reporting Period:	03/2018 – 05/2019
Authors:	Jan Brase, Claudia Engelhardt, Timo Gnadt, Péter Király, Lena Steilen, Philipp Wieder
Date of Publication:	June 16, 2019

Summary

The Göttingen eResearch Alliance (eRA) is a service unit of the University of Göttingen with a focus on research data management and related topics. The eRA started 2014 to serve the Göttingen Campus offering consulting, training, networking and IT services. The report at hand summarizes the results and events between March 2018 and May 2019 for the External Advisory Board of the Göttingen eResearch Alliance.

Overview

The purpose of this document is to summarize the results of the work of the Göttingen eResearch Alliance (eRA) between March 2018 and May 2019. Furthermore, this document provides an outline of the plans of the eRA for the coming months.

Structure of the Report

Following the overview, we provide general feedback on the recommendations provided by the External Advisory Board for the period of 2017/2018. After that, we report on the results of the four main pillars of the eRA portfolio: (i) training, (ii) consulting, (ii) networking, and (iv) [digital] services. Each section contains selected results for each of the pillars. This part is followed by a description of the key performance indicators (KPIs) that have been defined to better judge and adjust the work and outcome of the Göttingen eResearch Alliance. Next, we describe a number of outreach activities and the results of 3rd party funded projects aquired by the eRA. After a brief description of the eRA team retreat the report is completed by an outlook on the plans for the next reporting period.

eRA Team and Governance

The core staff of the Göttingen eResearch Alliance consists of:

- Coordination & Management
    - Jan Brase (funded fully through the SUB)
    - Timo Gnadt (funded fully through the eRA)
    - Philipp Wieder (funded fully through the GWDG)
- Team and Associated Members
    - Sven Bingert (funded through GWDG)
    - Maximilian Brodhun (funded partially through eRA)
    - Claudia Engelhardt (funded through 3rd party funding: project GRAcE)
    - Péter Király (funded fully through the eRA)
    - Harald Kusch (funded fully through the Medical Informatics Department of the UMG)
    - Claudio Leone (funded partially through eRA)
    - Markus Matoni (funded through 3rd party funding: CRC INF)
    - Jens Nieschulze (funded fully though University of Göttingen)
    - Jan Rohden (funded fully through SUB)
    - Lena Steilen (funded partially through eRA)
    - Ubbo Veentjer (funded partially through eRA)

Depending on the actual request, other people from Göttingen Campus partners are involved.

Fatih Berber left the Göttingen eResearch Alliance End of October 2018.

Extended Steering Group

The steering group has grown within the last year. Especially more people from the SUB joined it. Now there are representatives from the following departments and groups involved: Electronic Publishing (SUB), Informations- und Literaturversorgung Zentrale Erwerbung und Erschließung (SUB), Research and Development Department (SUB), Stabsstelle Wissen als Gemeingut (SUB). The GWDG is still represented by members of the AG eScience (GWDG). All Steering Group members can be found in "Appendix A – Members of the eRA Steering Group".

Addressing the Feedback from the Previous Report of the eRA External Advisory Board

We would like to acknowledge the effort and the time invested by the External Advisory Board to analyse and assess the directions and deliverables of the Göttingen eResearch Alliance. As a result of the second meeting of the Board, we received a feedback report, which is attached to this document as "Appendix B – Feedback from the Göttingen eRA Advisory Board 2018".

The feedback report addressed three major opportunities for improvement

1. Redressing the gender balance in the team
2. More user research into the requirements and use cases for faculty/users
3. Establish key performance indicators to measure success and achievements.

Re. 1: There was no active hiring for eRA staff in the reporting period, therefore the gender balance is currently unchanged. Nevertheless, we tried to take the gender balance into account in all workshops, trainings and conference sessions that were organized by the eRA. It is worth noting that the gender balance in the Research and Development department of SUB Göttingen, one of the two administrative structures supporting the eRA has increased. Currently 11 out of 28 employees are female, making the proportion of women in the department 39%.

Re. 2: The bi-annual eRA council meetings have proven to be a valid instrument to gather feedback from the service users on campus. In addition to these meetings, we have started to hold extensive test sessions with potential users before launching new services. During the reporting period, we hold two test sessions related to the campus repository service "GRO.data" inviting customers from Göttingen Campus One meeting was organised with future scientific users from the campus and the other one with IT specialists. The sessions were used to identify gaps and receive further user requirements to improve the service. The same approach was applied to "GRO.publications" including multiple discussions with the university administration to realise an interface between the publication service and the local CRIS system. Last, but not least, participatory design was applied to the re-design of the eRA website putting the "service first" approach into the centre of this endeavour. Feedback was gathered from power users, colleagues, and the eRA Council.

Re 3: In the last period, the eRA has started to define and collect KPIs to measure the number of consultations, the number of consulted projects, objects in the data repository, and further key performance indicators. Some of the respective metrics and results are part of this report. Please refer to Section "Key Performance Indicators" for more details.

Consulting

The eRA offers consulting on a large variety of topics to individual researchers, groups, or projects. Consulting is not limited to a certain range of topics, i.e. the eRA does in general not refuse any consulting request. Requests that cannot be fulfilled be the eRA itself are normally brokered to other players on the Göttingen Campus or other partners from the extended network of the eRA.

Main consultation topics during the reporting period concerned research data management issues such as data storage and data exchange, data policies and best practices. In a number of cases, technical support through the provision of specific services such as a data platform was provided.

Summary

Based on the KPIs related to Consulting (see Section "Key Performance Indicators") the eRA started to evaluate all consultation activities since 2014. As of today, a number of 166 consulting activities (KPI C-01) have been collected from records, tickets, and mail communication. Please be aware, that be further from earlier reporting periods might be activities added as the collection process has not yet been finished.

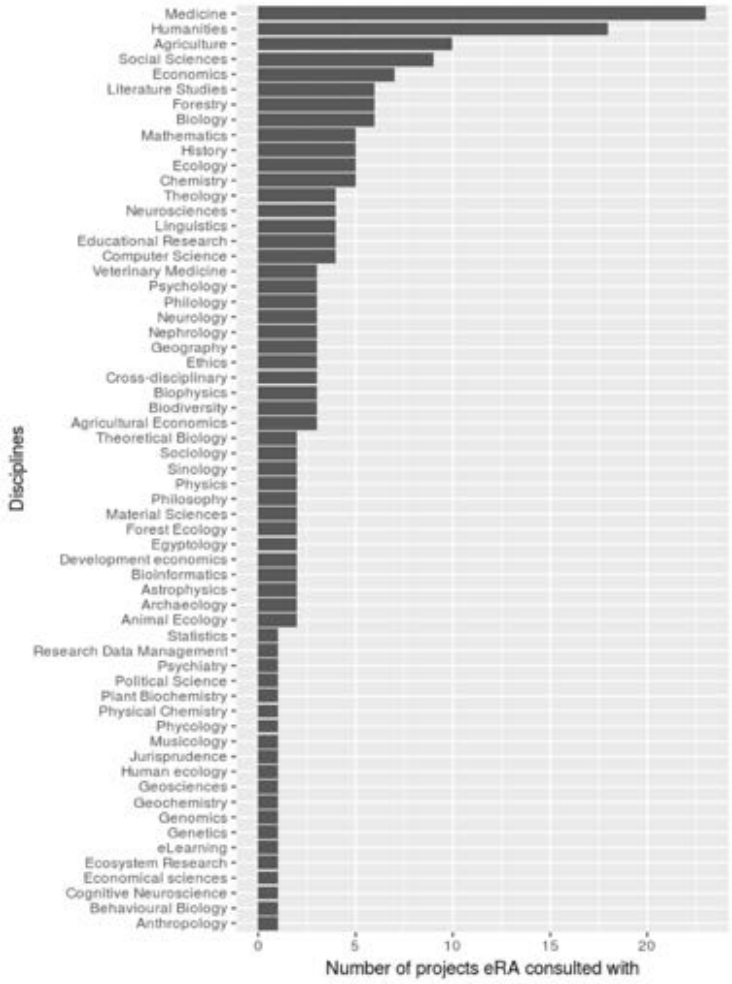

Figure 1: Consultation activities per discipline (KPI C-02).

Figure 1 shows the distribution of the various consultation activities per scientific discipline and Figure 2 depicts the projects requesting consultation (and training) in relation to the funding agency or program. This data is the baseline for further analysis and strategic decisions on the required expertise within the eRA, faculties to target, networking activities to be planned, or information material to be added to the web site.

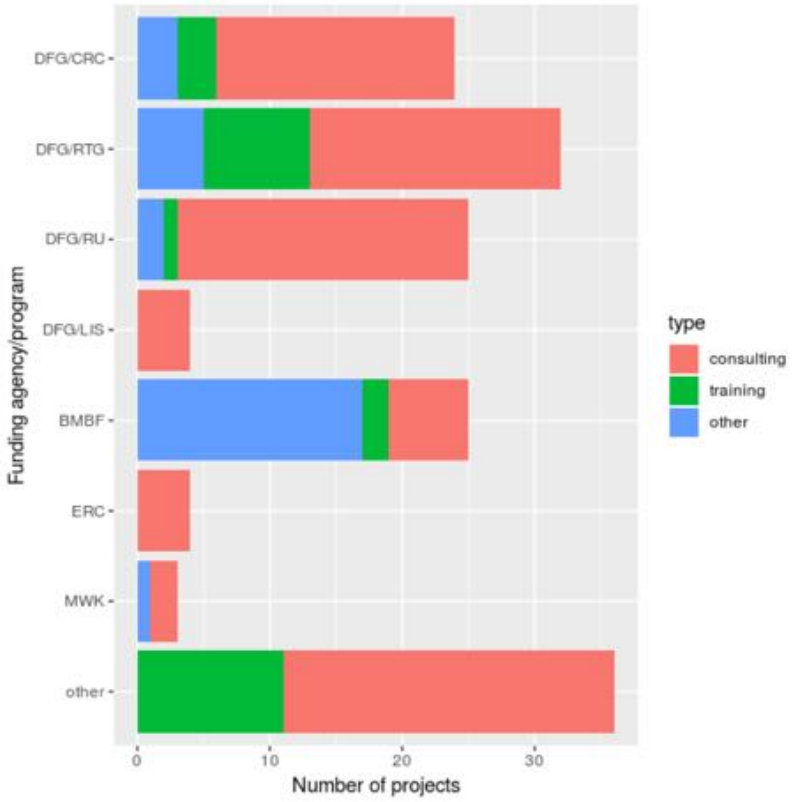

Figure 2: Consulting and training activities per funding agency and program (KPI C-03).

Consulting of Large-Scale DFG Projects

In the reporting period, a number of DFG coordinated projects have been consulted regarding their project proposals, but also during their run time. You can find an overview of consulted projects and the consultation topics in "Appendix C – Consulting: Large-Scale DFG Projects".

On-Demand Service Development

One specific offer within the portfolio of the eRA is related to service development. As projects and faculties not only struggle with the sustainability of their services, but also with the integration of their proprietary solutions into campus solutions, or the fact that recruitment of IT staff is getting more and more difficult, the eRA offers the following:

- Connection to campus services
- Customizing existing services to meet the requirements of specific projects
- Software development of complete services

Requests in these areas span from often rather simple questions like "How can I migrate my data into the campus repository GRO.data?" to requirements that cannot be fulfilled by any of the existing services on the campus and therefore require a completely new service development.

In the following sections we describe three examples of projects resulting from such requests. They all fall into the second, more complex category where complete services have been or are developed.

Example 1: CRC 990

The "Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems" project (EFForTS) is a large-scale, DFG-funded collaborative research centre (CRC). The project creates a large amount of research data in collaboration with Indonesian researchers.

The eRA is partner of the Infrastructure (INF) project and has developed jointly with other campus partners a research data platform based on a system called BEXIS[1]. The platform and a WebGIS system are in operation for a few years now.

Figure 3: The WebGIS platform of the CRC 990

As the project will not be able to sustain the development without funding, the eRA currently evaluates a migration path into the campus repository GRO.data (see Section "GRO.data"). Depending on whether the projects ends in 2019 or will be funded for another three years, the migration will be implemented in 2019 or later.

Example 2: Research Unit 2432

The Research Unit 2432 "Social-Ecological Systems in the Indian Rural-Urban Interface: Functions, Scales, and Dynamics of Transition" observes social and ecological sources from macro- to microscopical scale. The project creates and shares a large variety of data sources between Indian and German partners. This setting requires an elaborated data management solution, which has been

[1] http://bexis2.uni-jena.de

implemented by the eRA. It is based on a standard service for sharing working data (GWDG own-Cloud[2]) and a service implemented by the eRA to store and share published data.

Within 2019, an integration of a WebGIS solution is planned.

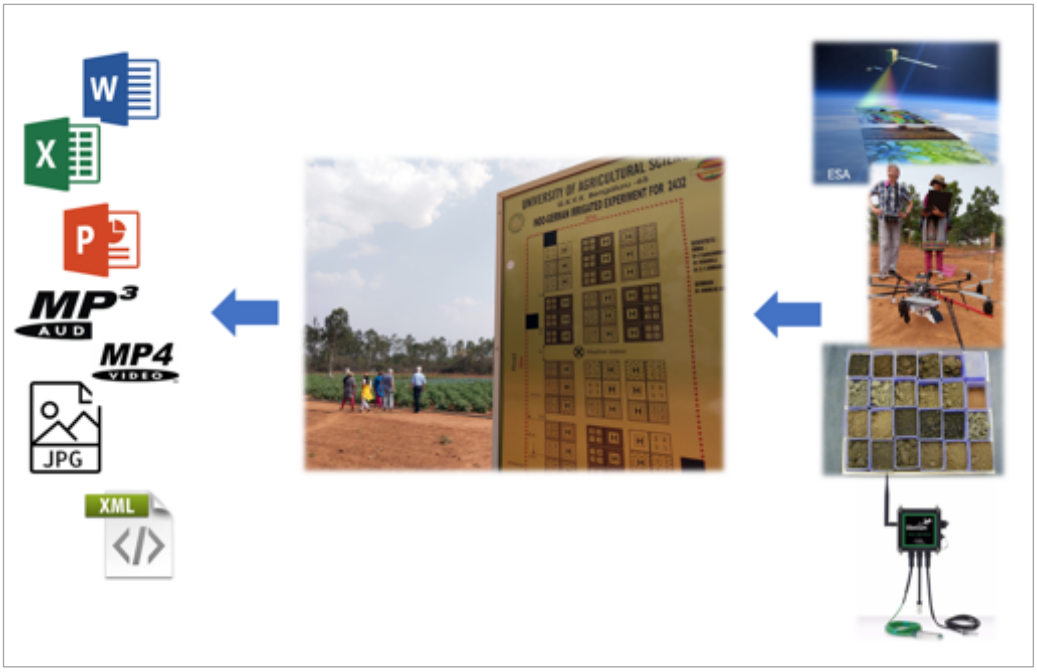

Figure 4:The research unit gathers data from a large variety of sources and creates a large variety of digital data objects.

Example 3: Research Unit 2705

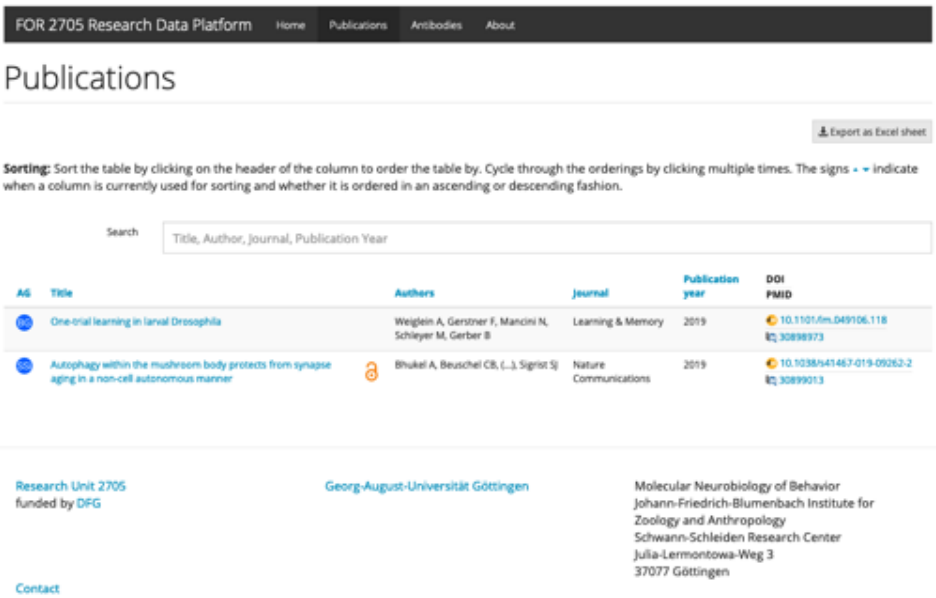

Figure 5: The initial research data platform for the RU 2705 showing the view on publications.

The Research Unit 2705 "Dissection of a Brain Circuit: Structure, Plasticity and Behavioral Function of the Drosophila Mushroom Body" has just started in 2019 and integrated the development of a

[2] https://www.gwdg.de/de/storage-services/gwdg-owncloud

research data platform into its project from the beginning. Of particular importance for the researchers is an integrated view of publications and research data, as well as the integration of external data sources and catalogues of antibodies and drosophilae. The eRA evaluated the requirements of the project and suggested to use the life sciences platform already used in e.g. CRC 1002 and CRC 1190 as a foundation for the Research Unit 2705.

The basic service has been made available and will be further developed within 2019.

Training

Second International Göttingen Summer School on Data Science 2018

From August 2 – 16, 2018, the Second International Göttingen Summer School on Data Science (DS3GOE) took place. 32 students (mainly Master and PhD students) participated in the school. They came from research institutions of nine different countries with diverse disciplinary backgrounds. Together they learned about the opportunities and challenges of data science in this intense two weeks Summer School. The teaching focus was on providing an overview to the multi-layered topics and methods:

- Dealing with the data lifecycle
- Best practices in data management
- Methods for data analysis (modeling, statistics, mining, and more)
- Infrastructures and platforms
- Application examples
- Ethical, legal, and social aspects

The keynote emphasized the importance of the FAIR data principles and the challenges of data quality assessment.

It was in particular valuable for the students to bring their own data and use them in hands-on sessions as example data. During these sessions they e.g. verified their own data quality, proofed their methodological approach, or tested new tools.

The multidisciplinary input was a challenge and an opportunity for the participants at the same time. The students' ability to meet this challenge was confirmed by their presentations on the last day: they showed their lessons learned, own ideas, new insights and also memorable moments of the Summer School.

Figure 6: Impressions from the Second International Göttingen Summer School on Data Science

The organisation team received a lot of valuable feedback from the participants and incorporated this into the planning of the summer school 2019.

Third International Göttingen Summer School on Data Science 2019

Preparation Phase

From August 5 - 16, 2019, the Third International Göttingen Summer School on Data Science will be held. As the University of Göttingen had only approved funding for the event for two years (i.e. 2017 and 2019) as part of the so-called Short-Term Program, the organisation tea, had to apply for external funding this time. The proposal to the DAAD was successful and resulted in receiving the full required amount (25.000 €). Additional funding (10.000€) is provided by University of Göttingen.

Regarding the overall organization, the successful cooperation between the Institute of Computer Science, Göttingen International, the Göttingen eResearch Alliance, and the University Medical Centre is being continued. We are also much obliged that lecturers from the German-Japanese university cooperation HeKKSaGOn will contribute to the summer school again as they already did in 2017 and 2018.

35 participants have been chosen from about 300 applications. The programmatic focus will shift more to practical learning and hands-on sessions this year based on the feedback from the previous summer school. Therefore, longer sessions are planned with a smaller group of lecturers.

Further Training Events

The eRA offers a large variety of training topics to individuals, groups, as well as projects. Here, we report on the events that a) target larger groups and b) consume a lot of time to plan and implement. During the reporting period, this includes:

- *Training event for MPG PostDocs:* The eRA organized a PostDoc training event for the Max-Planck-Society (MPG) in May 2018, where participants from all national Max-Planck Institutes were invited. While in the end only 12 researchers participated, the feedback was very positive, and the following contacts and requests encourage the eRA to repeat this kind of event.
- *Training for the Neurosciences:* In June 2018, the Göttingen Graduate Center for Neurosciences, Biophysics, and Molecular Biosciences received a training on research data management for around 15 PhD and PostDoc researchers.

- *DataCite Fabrica Workshop:* In December 2018, the eRA organized an info event for the SUB clients of its DataCite DOI registration service with DataCite representative Martin Fenner. Although only around half of the registered participants were able to join the event due to a strike in public transport, it was a successful exchange of questions and information on the new DataCite web frontend Fabrica, and upcoming new services and features.
- *Dataverse - How to use it:* The CRC 1136 received a training on the use of the Dataverse-based university data repository GRO.data (see also Section "Services") in May 2019, which brought forward some new challenges on how to manage the case of consortia where a large fraction of researchers is leaving or have already left the institution.
- *Research Data Management and Electronic Lab Notebooks:* The eRA organized a training event for the German Primate Center (DPZ) in May 2019 on the topics of Research Data Management and Electronic Lab Notebooks, which received a very positive feedback and will be mentioned as success story on the new eRA website. Over 30 staff members from different group of the DPZ joined and received a certificate for training.
- *Info Events at UMG:* The regular info events at UMG were continued on their half-yearly schedule. While the number of participants dropped in 2018, the eRA adopted a new approach in advertising the events in cooperation with the UMG academy for education, which resulted in a significant rise of participants and more focussed interest in May 2019. The feedback was largely positive. The events will be continued and extended to other faculties, starting with the institute of Geosciences, in late 2019.

eRA Lab

The eRA is developing a plan to provide training on research data-related topics complementing the existing offers at the Göttingen Campus. The so-called eRA Lab will become an open space for information exchange on eResearch topics. The lab is a joint service of different actors of the Göttingen Campus, hosted by eRA. The lab is primary an easily accessible space where the customers find domain experts. It provides regular programmes, such as lectures, workshops, consultations, trainings, hackathons, meetups, or a helpdesk.

Since the lab hosts a diverse set of activities, the space, and its internal organization is crucial. It should be flexible to be adaptable to the type of event it hosts. It means that great part of the space should not have fixed furniture, like standard office desks with a full set of wired equipment. The eRA has therefore ongoing discussions with the people responsible for the "Digital Creative Space", a large area that will be opened end of 2019 within the central library. The Digital Creative Space will feature multiple areas, including a relaxed working scenario, a business lounge, and digital, collaborative working environment. It is planned to integrate the eRA Lab as much as possible into this space and offer the respective trainings demand-driven. The eRA will test this as early as possible and adapt the Lab offers to the demands.

When the Lab is open, it implies that the customers find someone there from the eRA who is ready to talk to her or him. The eRA is the host: it organizes programmes on its own, but also let others use the space for their programmes - if it fits to the eResearch domain.

The Lab's two main goals are knowledge sharing about and community building around eResearch. The Lab encourage faculties and departments to use it for programmes organized in interdisciplinary and transdisciplinary domains for a more mixed audience. The eRA will cover the following topics itself:

- eResearch 101
    - Data documentation
    - Data storage
    - Data sharing

        - Licenses
        - Solutions available at the Göttingen Campus: GRO, CDSTAR, PIDs, and more
- Software development 101 (similar or identical to the Carpentries)
- Open science
- Data science and analytics
        Data analysis
        - Visualisation
        - Data modelling
- General IT topics
        - Methods (e.g. agile development, documentation, or TDD)
        - Tools (databases, IDE-s, or CI tools)
        - Languages (Java, Python, or R)
- Special IT topics
        - IT needs of specific disciplines, such as library IT related topics (OAI-PMH, MARC, Signposting) for librarians, chemistry-related ones for chemists, etc.
        - Open Source development. The eRA Lab could e.g. host the meetings of different open-source developer communities - in collaboration with existing facilities, such as the Chaos Computer Club
        - Supercomputing/HPC
        - "Making your software robust": it would be a special service for programming researchers. They could reach a point where their software is working in their own environment, but the next step, making the code robust and sustainable, requires some extra help and extra tasks, which are outside of their research interest.
        - "Bring your own data" sessions (see also the summer schools)
- Digital humanities related topics
        - NLP
        - Text mining
        - Stylometrics

The eRA Lab is complement existing offers from Göttingen Campus partners and central units. We will not double existing offers. The eRA Lab programme has to be assessed and evaluated regularly and specific KPIs have to be defined to make sure that the programme is built for the customer and evolves depending on their needs.

Networking

CODATA Symposium and Follow-up Events

From Sunday 18th to Tuesday 20th March a symposium took place in Göttingen to exchange views on the role of universities and libraries worldwide in supporting research data management. The aim of the event was to give an overview of the challenges that exist in this area and to see successful examples from practice. The symposium entitled "'The critical role of university RDM infrastructure in transforming data to knowledge" was organized by the eRA together with the Committee on Data of the International Council for Science (CODATA).

Figure 7: The Pauliner Church served as an adequate location for the workshop.

The symposium was deliberately intended as a pre-event for the 11th Plenary of the Research Data Alliance (RDA) in Berlin from 21st to 24th March. In fact, with more than 160 participants the expectation was more than exceeded, as many participants of the RDA conference took the opportunity to spend 2 days in Göttingen on their way to Berlin. With almost 90 presentations the number of presentations was much higher than expected, so that we were forced to offer 3 parallel tracks with presentations at the event.

Due to the support of the University of Göttingen, it was also possible to get top-class keynote speakers from the USA to give an overview of the state of research data management there. In detail they were:

- Amy Nurnberger, *MIT*, Libraries tense: future continuous
- Peter Fox, *Rensselaer Polytechnic Institute*, Setting the scene from a researcher's point of view
- Michael Witt, *Purdue University*, Research Data Services at Purdue University: An Institutional Approach
- David Minor, *UCSD*, Filters and firehoses: reflections on a decade of research data curation

- Lisa Johnston, *University of Minnesota Libraries*, Rethinking Institutional RDM Strategy: Applying Collaboration to Affect Change in the Research Landscape
- Bob Freeman, *Harvard*, Dataverse and RDM Opportunities at Harvard
- Tom Cramer, *Stanford*, Actors, Factors, and Gaps: Managing Data Holistically at Stanford
- Liz Lyon, *University of Pittsburgh*, Stewardship and Science: Reflections on Data-driven Opportunities for Libraries

Figure 8: David Minor giving his talk.

The response to the event was extremely positive. Already during the symposium there was so much feedback on Twitter that the Hashtag #RDMGoettingen on March 19 was listed in the German Trends on Twitter[3].

With the symposium, Göttingen and the eResearch Alliance were able to further position themselves as leading players worldwide in the field of solutions for research data management at universities. As a result of the symposium, it was decided that further similar events would follow. Two of those took already place during the reporting period of this report:

In October 2018 at the International Data Week, the joint event of 12th RDA plenary and the Conference on Scientific Data (SCiDataCon), Jan Brase organized and moderated a session on research data on campus (https://www.scidatacon.org/IDW2018/sessions/183/) which was well attended and especially gave interesting insights on activities at African universities.

[3] https://twitter.com/NichtTomJones/status/975700377885605888

Figure 9: The auditorium at the International Data Week (October 2018)

Prior to the 13th RDA plenary in April 2019, a follow-up workshop of the Göttingen symposium took place at Drexel University in Philadelphia. Jan Brase and Wolfram Horstmann represented the eRA in the organisation committee of the workshop and both moderated a session at the event. Claudia Engelhardt made a presentation of the eRA and the GRACE project. During the 13th plenary a bird of feather session took place to discuss the possibility of establishing a RDA interest group on the topic of research data services at campus and eRA was mentioned at the session and various members of the eRA took part in the discussion. For the 14th RDA plenary in October 2019 in Helsinki, again a pre-RDA workshop is planned in cooperation with CODATA and there will likely be a first meeting of such an RDA interest group. Both with involvement and support of eRA team members.

Toolbox Series

The Göttingen eResearch Toolbox Series (GeTS) was continued with a workshop in May 2018 in the topic of Workflow management software. The workshop featured short presentations and intensive discussions about key aspects and different flavours of workflow management and documentation in academic research. Some of the software solutions presented and discussed were: Taverna, Common Workflow Language, KNIME, Galaxy, and Orange 3.

The series will be continued in late 2019, the topic is to be determined.

Other eRA Networking Events

Further noteworthy networking events during the reporting period have been:

- *Exchange with the National Library of the Czech Republic:* On April 4th and 5th 2018 a team from the National Library of the Czech Republic visited the SUB Göttingen and Jan Brase gave an overview on the services of the eRA.
- *Exchange with Kazakh University*: In June 2018, Jan Brase and Philipp Wieder presented the eRA to an interested group from the Kazach L.N. Gumilyov Eurasian National University via a web-conference, and answered subsequent questions regarding the establishment of research data infrastructure and networking as well as curricular activities in this area at the Göttingen Campus.
- *Göttingen Campus cooperation:* Following its strategy to intensify collaboration with other groups at the Göttingen Campus, the eRA met with the Research Department of the University in August 2018 and with subject specialists (so called "Fachreferate") of the SUB Göttingen in March 2019 in order to discuss and align cooperation strategies and possible synergies.
- *DINI/nestor participation:* The eRA was also participated with different contributions in three workshops of the DINI/nestor network, which covered the following topics:
    - Requirement analyses - a basis for custom-fit infrastructures? (January 2019)
    - Developing structures - organisation and governance for local RDM services (April 2019)
    - Who is going to pay? - Cost and business models for sustainable research infrastructures and RDM services (June 2019)
- *PODMAN workshop:* With a presentation of Göttingen Research Online, the eRA contributed to the final workshop[4] of the PODMAN project[5] addressing the topic: "Research + Data management = RDM: When does it add up?"
- *Exchange with University of Maribor:* In May 2019, Timo Gnadt presented the eResearch Alliance to interested library staff from the University of Maribor, Slovenia, and discussed with them the organizational and practical challenges in supporting researchers with data management planning.
- *Dataverse network meeting:* On June 3 2019, the eResearch Alliance organized a meeting of German and Austrian research institutions who are using or building up a Dataverse instance as data repository. The discussion covered topics such as "Developing the Dataverse code base", "Curation workflows", "Handling large amounts of data", Translations" and "Communication channels" and was enhanced by the intermittent online participation of Dataverse developers from Harvard University and a DataverseNL representative. As result, the participants agreed on closer collaboration across different channels and pursue the idea of building up a German-speaking Dataverse community. (->PK: anything to add?)

GRAcE Networking Events

The GRAcE project (see also Section "GRAcE") is one of the 3rd part funded projects that has been aquired by the eRA and which is collaborating closely on a day-to-day basis. Events are in general co-organized by projects partners and the eRA.

Workshop on RDM and infrastructure in DFG CRCs

On 26/27 November 2018, the GRAcE project organized a nationwide workshop on the topic of Research Data Management and Infrastructure in the large, interdisciplinary DFG funded Collaborative Research Centres (CRCs). The workshop was specifically targeted at persons involved in INF-subprojects or similar activities of CRSs. It was well received with almost 60 participants including

[4] https://fdm.uni-trier.de/einladung-zur-podman-abschlusstagung/

[5] https://fdm.uni-trier.de/

representatives of 30 CRS and the DFG. The agenda included a presentation of GRAcE project results, a poster session and a World Café. The slides, posters and other workshop materials are openly available via GRO.data[6]. Detailed workshop results along with CRC case studies will be published in summer 2019 in the Open Access Online Journal Bausteine Forschungsdatenmanagement[7].

Workshop on RDM costs and efforts

On 28 May, GRAcE presented relevant project results and invited representatives of partner institutions (e.g. the U4 network) to talk about how the topic of RDM costing and staffing is dealt with at their institutions and to discuss different aspects with interested members of the Göttingen Campus. In total, 19 persons participated in the event. The slides will be published via GRO.data in the middle of June 2019.

[6] https://data.gro.uni-goettingen.de/dataverse/GRAcE: https://doi.org/10.25625/97D6TL and https://doi.org/10.25625/22Y1WC.

[7] https://bausteine-fdm.de/

Services

One central task of the Göttingen eResearch Alliance is the support for and the development of services for research data management. As support is mostly covered by consulting and training (except the actual technical support of the service operation), the Services pillar mainly reflects the portfolio of services developed by the eRA.

The overall goal of the provision of this service portfolio is to establish a core set of essential service that is required by researchers and to integrate existing an upcoming service in a best possible manner. This portfolio will be marketed under the label "Göttingen Research Online", or in short GRO.

Overall Architecture

The general idea of GRO is to have a central access portal to all services. This portal has a particular recognition value and offers centrally needed functions like single-sign-on (SSO). The various services are integrated into this portal and accessed through their respective APIs. Depending on the service, the portal can also support the controlled input of metadata. All services are integrated through APIs, which makes it easier to replace a particular product or software in case of it coming end-of-life or unbearable license cost.

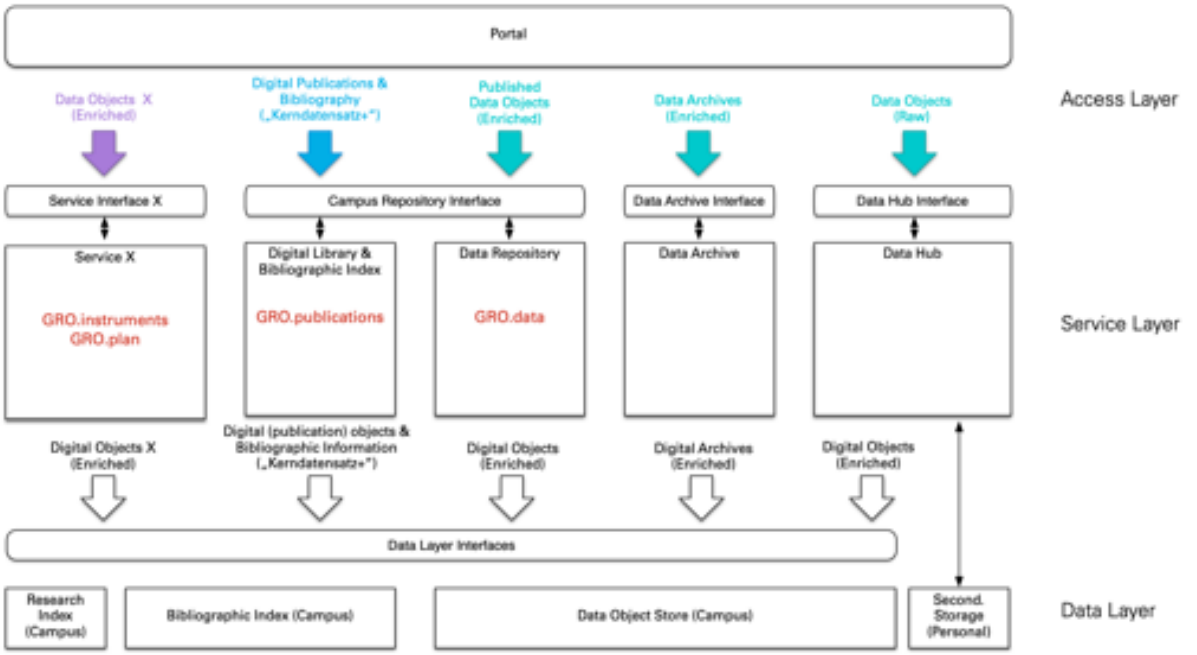

Figure 10: Overall architecture of Göttingen Research Online.

On the service layer, a variety of services have already been developed or are still under development. As of today, these are:

- A data repository (service name: GRO.data)
- A publication service (service name: GRO.publications)
- A service to manage large-scale instruments (service name: GRO.instruments)
- A service to develop data management plans (service name: GRO.plan)

The respective services are described in detail in the following sections.

The data layer is accessed through a variety of interfaces. As of now, there are different data and metadata sources, but the long-term goal is to have a single point of truth (SPOT) for as many services as possible.

GRO.data

Status: Production (not officially announced yet)
Launch date: June 2019
Service manager: Peter Kiraly
Software: Dataverse[8]

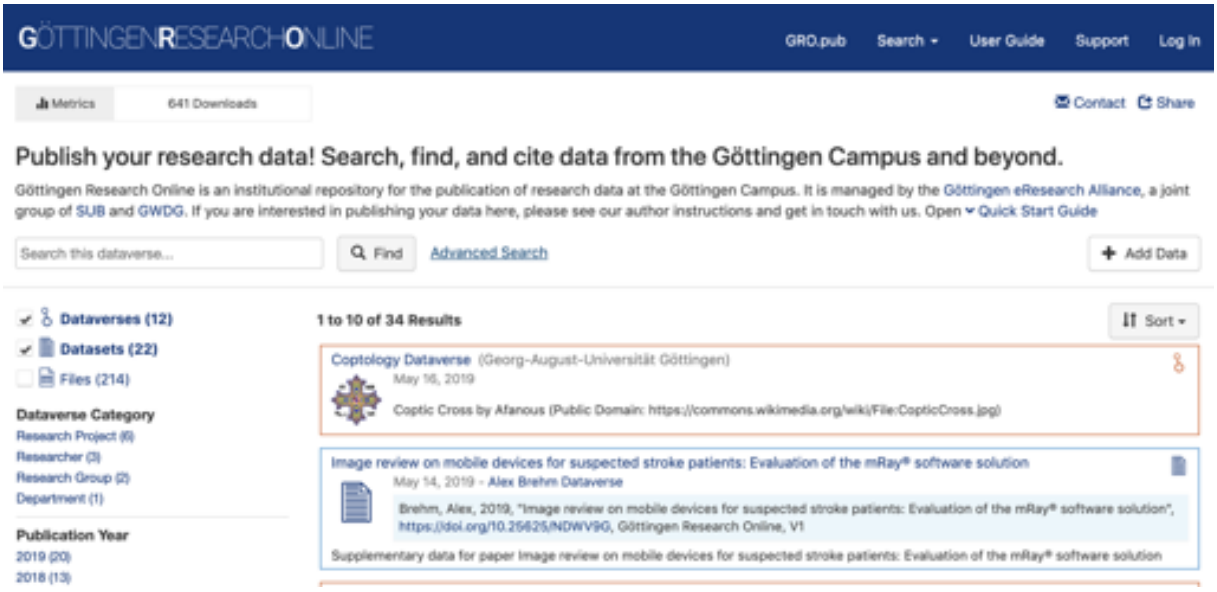

Figure 11: GRO.data home page.

GRO.data is a data repository with the primary function to publish research data and make it citable. The service is based on the software Dataverse, which is provided openly and for free by Harvard University. Citations are realised through the DOI service offered by the eRA (see below).

In GRO.data individual researchers, groups, projects, or departments can create their own space, which is called a dataverse, and can work collaboratively until the data can be published. Researchers outside Göttingen can be invited; the respective function to self-register is offered centrally through the SSO. Once a data object is created it immediately gets a DOI assigned. This allows to cite it within publications already during the writing process. The final DOI registration is done once the object is published. The respective KPIs can be found here: https://data.gro.uni-goettingen.de/metrics/. Please note that the numbers still represent a pre-launch status with only very few data is published by specific user requests.

A closer integration with GRO.publications is currently planned to link between publications and related data objects, and to make both entities equal peers in the long run. Other requirements are gathered through user engagement. The respective list of feature requests is maintained in Gitlab and also discussed with the Dataverse developers.

The eRA is member of the Global Dataverse Community Consortium[9] to be closely linked to the development process and it is doing community building within Germany (through a Dataverse interest group) and in Europe (through the SSHOC project (see below)).

[8] https://dataverse.org
[9] http://dataversecommunity.global/members

GRO.publications

Status: Test
Launch date: June 2019 (soft launch)
Service manager: Daniel Beucke
Software: DSpace CRIS[10] (heavily modified version)

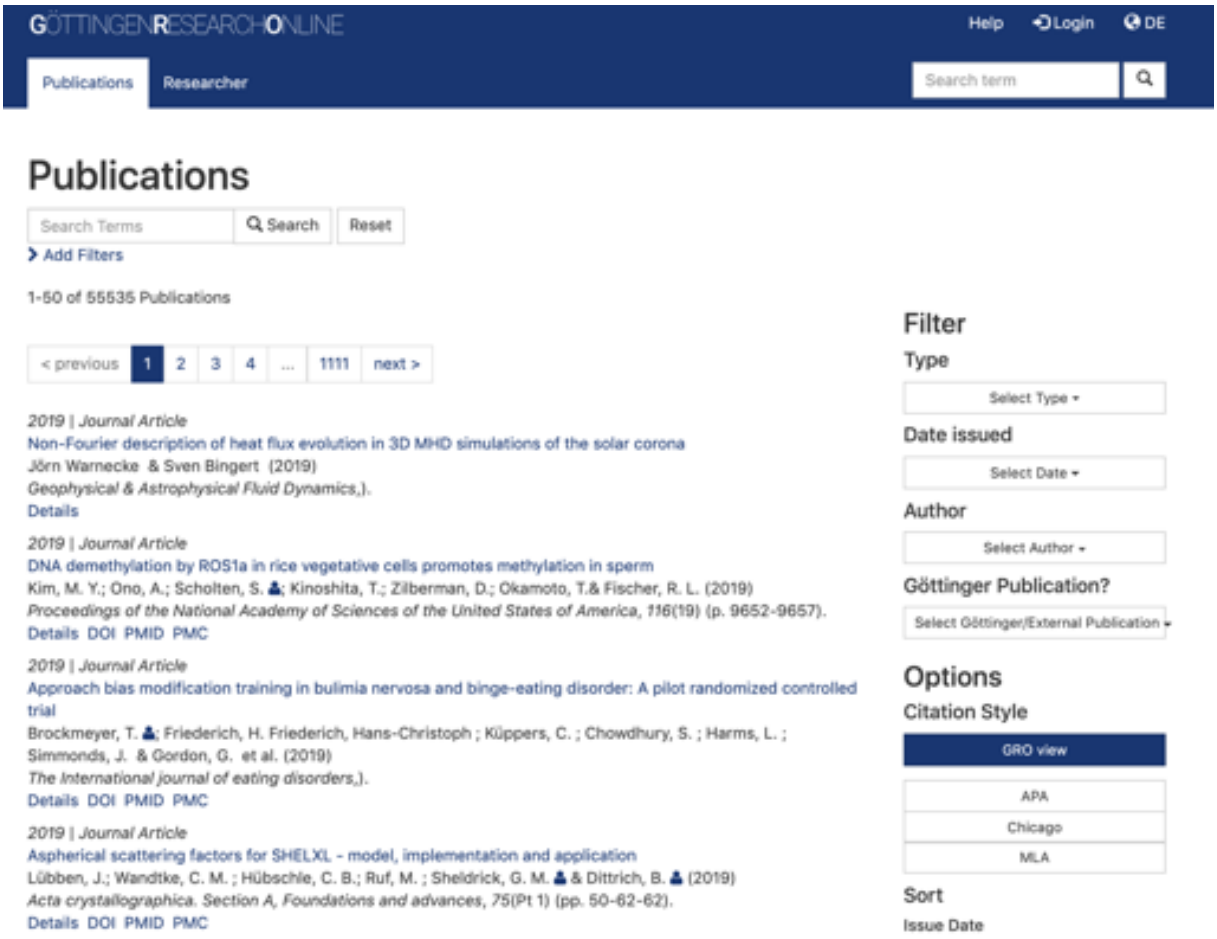

Figure 12: GRO.publications home page.

GRO.publications is a service that allows researchers to claim and maintain their own publications. As a foundation, the SUB has ingested and curated over 55.000 publications of researchers from the Göttingen Campus from sources like Web of Science and Scopus.

GRO.publications offers in addition to the mere publication management a variety of other functions including:

- Curation support
- Full ORCHID integration
- Export functions
- Different citation styles
- Creation of individual publication lists and integration into websites (through automatically generated Java Script snippets)

More functions, including am integration with GRO.data, are on the feature request list and will be implemented in the future.

[10] https://wiki.duraspace.org/display/DSPACECRIS/DSpace-CRIS+Home

GRO.instruments

Status: Production (not officially announced yet)
Launch date: To be decided
Service manager: Claudio Leone
Software: OpenIRIS

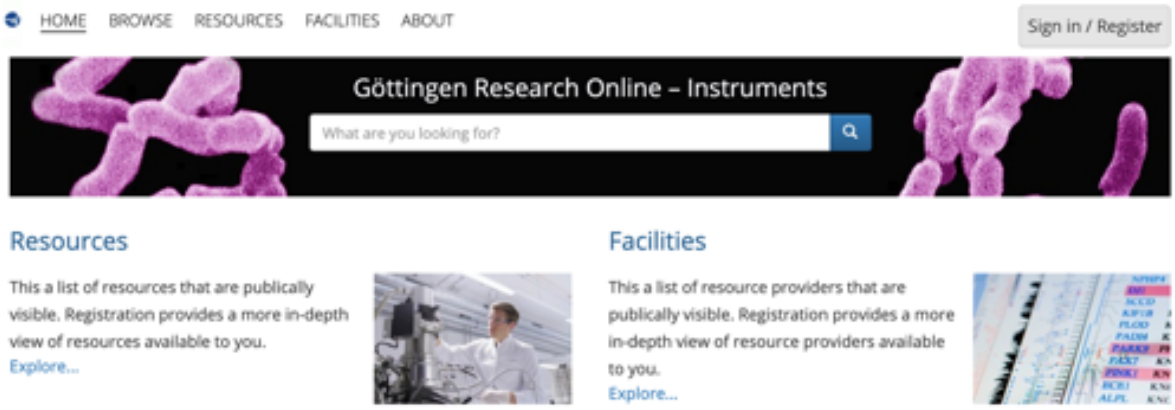

Figure 13: GRO.instruments home page.

GRO.instruments is the service to manage large-scale instruments like microscopes, MRTs, or other. It allows to register facilities and their instruments, assign user groups, schedule the usage of the instruments, and provide further information.

The service is currently used by a number of Max-Planck-Institutes, but has not yet been generally announced as a campus service.

GRO.plan

Status: Development
Launch date: To be decided
Service manager: Ubbo Veentjer
Software: RDMO[11]

As data management plans are becoming more and more instruments of planning and controlling data generation, management, and usage, the eRA plans to offer a respective service. GRO.plan uses the Research Data Manager Online (RDMO) as the basic software, but in particular the UI is not yet what is expected by the eRA. Therefore, the service development is currently planned to have a concise roadmap and be able to announce the launch of the service.

Persistent Identifier Services

Status: Production
Launch date: ./.
Service manager: Claudio Leone (DOI Service) and Sven Bingert (ePIC PID Service)
Software: Datacite[12] (DOI Service) and ePIC PID Service[13]

The eRA offers two different services for persistent identifiers: the DOI Service and the ePIC PID Service. While the former mainly aims at persistent and long-term identification of published elements

[11] https://rdmorganiser.github.io
[12] https://datacite.org
[13] https://www.pidconsortium.eu

(whether this are papers or data objects), the latter is mainly used to integrate its API into added value services (like repositories or archives) to persistently identify large numbers of data objects.

Both services are based on the handle.net system[14]. Parts of the core global infrastructure of the ePIC PID Service are operating in Göttingen.

DOI Service

The provision of DOIs for the Humanities to German research institutions via the DataCite membership of the SUB Göttingen has continued and grown throughout the past year. With 13 new client institutions in the reporting period, we now have 35 clients and currently three contracts in negotiation. Six of the new clients were Humanities-related clients transferred from GESIS, another German DataCite member who is now focusing on serving clients from its primary target disciplines in the Social Sciences. Another eight clients from GESIS are expected to become eRA clients within 2019.

We have a total number of around 11,000 findable DOIs registered (KPI S-07) as of June 1, 2019, with additional numbers of around 1,000 in draft and 360 in registered state. In December 2018, we organized a training session for our current and prospective clients on the new DataCite web interface Fabrica (see the respective Section "Training").

ePIC PID Service

Also, the ePIC PID Service is increasingly used by the community. As of May 28, GWDG hosts in total prefixes for 44 customers (KPI C-08) totalling to a number of 42.819.389 PIDs registered (KPI C-09) and another 30.824.329 replicated PIDs.

[14] http://handle.net

Key Performance Indicators

The purpose of key performance indicators is to measure the performance of actions or processes against predefined target values. Based on the assessment of these measurements further actions can be defined to improve the performance based on the overall strategic goals of a company or institution. A common KPI example from IT service management is the availability of a service, which is mostly set to be 99,9…% (the famous 3/4/5/… nines). Closely linked to KPIs is the Deming- or PDCA Cycle that defines a plan – do – check – act cycle most often based on KPIs: (i) a target value for a certain process is defined (plan), (ii) the process is executed (do), (iii) the outcome is measured and assessed (check), and further actions are defined to improve the performance (act).

In case of the eRA, it makes sense to define general KPIs for the four different pillars (consulting, training, networking, and services) and add specific KPIs for individual services depending on their function.

There is no generally valid way to specify a KPI. This should be done depending on the strategy of the respective institution and the processes to be assessed. With respect to the eRA, KPIs are defined and measured as input to the strategic process, to evaluate the success of services and actions, as well as to adjust priorities and resource usage. KPIs help, for example, to understand which services are used to which extend and where the eRA has to increase marketing or to check whether the service actually fulfils the needs of the users. The KPIs themselves are subject to change: the eRA will regularly check the feasibility and applicability of its KPIs, adjust them or add further KPIs. This will be done with the help of the External Advisory Board, the eRA Council, the Steering Group, and the Quality Management Team of the GWDG, which is responsible for the ISO 9001 certified quality management system.

Basic Set of KPIs for the Göttingen eResearch Alliance

Based on the recommendation from the External Advisory Board, the eRA has specified a basic set of KPIs. For each KPI the following information is provided:

- Identifier and Name
- Metric
- Target value
- Source of data
- Means of acquisition
- Objective
- Related KPIs

Consulting

KPI C-01: "Consultations"

- Metric: Number of consultations per year
- Target value: > 40 consultations per year (defined based on the mean value for the previous years)
- Source of data: List of consultations, trainings, and service developments
- Means of acquisition: Manual based on support tickets and other support requests
- Objective: Understand whether the mean number of consultations per year (combined with the time consumed) corresponds to
- Related KPIs: C-02, C-04, N-01

KPI C-02: "Consultations per Discipline"

- Metric: Amount of consultations per discipline
- Target value: Distribution should be in line with the expertise of the eRA team members
- Source of data: List of consultations, trainings, and service developments
- Means of acquisition: Manual based on support tickets and other support requests
- Objective: Understand whether the eRA team has the right expertise with respect to the requirements of different scientific disciplines on the campus.
- Related KPIs: C-01, C-04, N-01

KPI C-03: "Consultations per Funder"

- Metric: Amount of consultations per funder
- Target value: Distribution should be in line with the expertise of the eRA team members
- Source of data: List of consultations, trainings, and service developments
- Means of acquisition: Manual based on support tickets and other support requests
- Objective: Understand whether the eRA team has the right expertise with respect to the requirements of particular funders.
- Related KPIs: C-01

KPI C-04: "Time Consumed per Consultation"

- Metric: Amount of time required per consultation
- Target value: To be defined
- Source of data: List of consultations, trainings, and service developments
- Means of acquisition: Manual based on support tickets and other support requests
- Objective:
- Related KPIs: C-02, C-04, N-01

Training

KPI T-01: "Trainings"

- Metric: Number of trainings per year
- Target value: All training requests should be fulfilled
- Source of data: List of consultations, trainings, and service developments
- Means of acquisition: Manual based on support tickets and other support requests
- Objective: Understand whether actual demand is fulfilled and whether the eRA has enough ressources to offer appropriate training.
- Related KPIs: N-01

Networking

KPI N-01: "Stakeholder Coverage"

- Metric: Percentage of stakeholders covered (Göttingen Campus partners, faculties of the University of Göttingen, selected individuals)
- Target value: 100%
- Source of data: List of consultations, trainings, and service developments
- Means of acquisition: Manual based on support tickets and other support requests
- Objective: The eRA and its management need to understand whether relevant stakeholders have been contacted.
- Related KPIs: C-01, C-02

Services

KPI S-01: "Service Usage"

- Services: Every service launched through Göttingen Research Online

- Metric: Number of new users per year per service
- Target value: To be defined based on initial experience with the services
- Source of data: Service logs
- Means of acquisition: Automatic per service
- Objective: Understand the overall usage per service to assess the success and to decide upon potential actions to increase the number, market services better, re-design the service portfolio, or alike.
- Related KPIs: S-02 – S-09, N-01

KPI S-02: "Published Data Objects"

- Services: GRO.data
- Metric: Number of data objects added per year
- Target value: To be defined depending on initial experience with the service
- Source of data: https://data.gro.uni-goettingen.de/metrics/
- Means of acquisition: Automatic
- Objective: Understand the growth of the service and (re-)design the service offer accordingly.
- Related KPIs: S-01

KPI S-03: "Publications"

- Services: GRO.publications
- Metric: Number of publications added and curated per year
- Target value: To be defined depending on initial experience with the service
- Source of data: https://data.gro.uni-goettingen.de/metrics/
- Means of acquisition: Automatic
- Objective: Understand the growth of the service and (re-)design the service offer accordingly.
- Related KPIs: S-01

KPI S-04: "Registered Facilities"

- Services: GRO.instruments
- Metric: Number of facilities registered per year
- Target value: To be defined depending on initial experience with the service
- Source of data: Service logs
- Means of acquisition: Automatic
- Objective: Understand the growth of the service and (re-)design the service offer accordingly
- Related KPIs: S-01

KPI S-05: "Registered Ressources"

- Services: GRO.instruments
- Metric: Number of ressources registered per year
- Target value: To be defined depending on initial experience with the service
- Source of data: Service logs
- Means of acquisition: Automatic
- Objective: Understand the growth of the service and (re-)design the service offer accordingly.
- Related KPIs: S-01, S-04

KPI S-06: "Data Management Plans"

- Services: GRO.plan
- Metric: Number of data management plans created per year
- Target value: To be defined depending on initial experience with the service

- Source of data: Service logs
- Means of acquisition: Automatic
- Objective: Understand the usage of data management plans and (re-)design the service offer accordingly.
- Related KPIs: S-01

KPI S-07: "DOIs"

- Services: DOI service
- Metric: Increase in DOIs registered through the DOI Service
- Target value: To be defined depending on initial experience with the KPI
- Source of data: Service logs
- Means of acquisition: Automatic
- Objective: Understand the growth of DOIs and (re-)design the service offer accordingly.
- Related KPIs: S-01

KPI S-08: "ePIC Prefixes"

- Services: ePIC PID Service
- Metric: Increase in Prefixes registered through the ePIC PID Service
- Target value: To be defined depending on initial experience with the KPI
- Source of data: Service logs
- Means of acquisition: Automatic
- Objective: Understand the growth of ePIC prefixes and (re-)design the service offer accordingly.
- Related KPIs: S-01, S-09

KPI S-09: "ePIC PIDs"

- Services: ePIC PID Service
- Metric: Increase in PIDs registered through the ePIC PID Service
- Target value: To be defined depending on initial experience with the KPI
- Source of data: Service logs
- Means of acquisition: Automatic
- Objective: Understand the growth of ePIC PIDs and (re-)design the service offer accordingly.
- Related KPIs: S-01, S-08

Miscellaneous

KPI M-01: "3rd party Funding"

- Metric: Amount of 3rd party funding raised per year
- Target value: Amount large enough to fulfil existing contracts
- Source of data: Grant agreements
- Means of acquisition: Manual through the administrations of participating partners
- Objective: This KPI helps to a) make sure that enough money is raised to fulfil existing contracts of staff from the extended eRA team who is not permanently funded and b) to execute mid-term planning according to the overall strategy of the eRA and its tasks.
- Related KPIs: ./.

Outreach

Relaunch of the Web Site

The current eRA website (eresearch.uni-goettingen.de) was an internal effort that started in early 2015. It was designed, managed and curated solely by members of the eRA team. Over the years, not only the eRA has evolved, but so have website design options as well as requirements and expectations from prospective users. Information on our website became hard to find for our target groups, resulting in frequent feedback from various sides over the past years that the eRA website needed a general overhaul.

As a result, SUB and GWDG decided to invest into a complete redesign of the webpage (see also Figure 14) and in parts also of the corporate design of the eResearch Alliance. After selecting from several offers for presenting our current and future content, an external developer started the design process in December 2018. Since March 2019 the team is busy with restructuring information and (re-)writing texts. Since May 2019 the content is in place inside the new website instance. The new design has been approved by the steering group and the eResearch Council. After some more feedback from both boards has been considered, the launch of the new website is expected to launch in July 2019. Subsequent modifications and extensions are already in the pipeline, and will be covered by a follow-up order to the same developer.

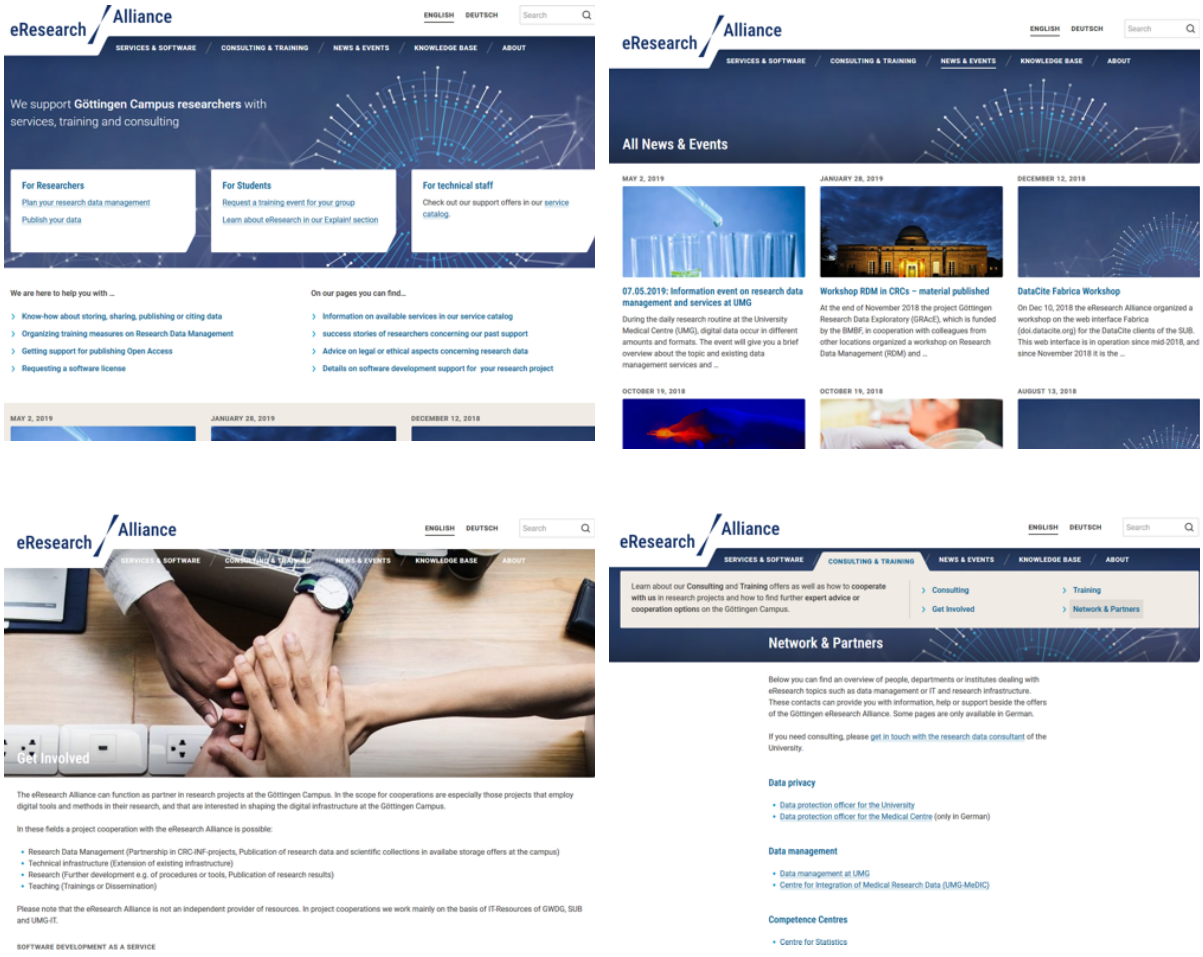

Figure 14: Screenshots from the new eRA web site. Home page, news and events, user involvement, and partner network (clockwise from the top left).

Intensified Coordination with Göttingen Campus

The Göttingen eResearch Alliance serves the whole Göttingen Campus and all of its partners:

- University of Göttingen
- University Medical Center Göttingen
- Göttingen Academy of Sciences and Humanities
- German Aerospace Center
- German Primate Center
- MPI for Biophysical Chemistry
- MPI for Dynamics and Self-Organization
- MPI of Experimental Medicine
- MPI for Solar System Research
- MPI for the Study of Religious and Ethnic Diversity

This fact is not properly communicated (neither by the eRA nor by the campus partners) and therefore not known everywhere on the campus. To improve this situation, we intensified our communication and coordination efforts with the team from Campus Göttingen and Public Relations of the University of Göttingen. This effort has the following major objectives:

- Getting feedback on the web site before its relaunch.
- Discussing means to improve the linkage between the eRA web site and those from campus partners.
- Creating information material for the eRA services and integrating it properly into the offers from other campus partners.
- Establish means to communicate offers and results from the eRA through channels of the Göttingen Campus (like web site, Twitter, blog, event calendar.

The respective feedback concerning the web site was already incorporated, further action plans regarding the other objectives have to be generated. In general, actions of mutual benefit have already been identified (like offering eRA service information and training to the young researchers on the campus through the Göttingen Campus media channels) and the implementation will be planned accordingly.

EFI Report

In February 2019, the German Federal Expert Commission on Research and Innovation (EFI) published their report on research, innovation and technological capability in Germany[15]. In this report, in the section on Digitization of Universities (page 95), the eResearch Alliance was featured as a particular example of good practice in research. The EFI is installed by the German government and reports directly to it. This high profile and widely visible report acknowledge the offers of the eRA as well as its approach and work

[15] https://www.e-fi.de/fileadmin/Gutachten_2019/EFI_Gutachten_2019.pdf (p 95)

Figure 15: Front page of the EFI report, the respective section on the eRA (in German), and the handover to Chancellor Merkel).

3rd Party Funded Projects

In this section we describe the results of those projects, which are funded 3rd parties like the German ministry for education and research (BMBF), the DFG, or the European Commission.

GRAcE

In GRAcE, a variety of activities were carried out in 2018. Regarding the costs, a series of expert interviews in the pilot area of Medical Sciences / University Medical Center (UMG) was undertaken in March and April 2018. The aim was to gather information about the research process in different areas of medical research and the nature and extent of data management tasks in the different phases of this process. The results of these interviews informed the work on the questionnaire for an online survey that was carried out in October 2018. It was directed at all UMG staff involved in medical research and the invitation sent to about 2000 employees of the UMG. The main aim of the survey was the estimation of the average working time spent on different activities and tasks in the research process, which was categorised in the three big areas of Documents, Data and Publication. 46 persons participated in the survey. Another strand of activities concentrated on roles and competences in the area of research data management. Based on a quantitative literature analysis, including for example PubMed and Google Scholar, a dictionary of relevant terms was created[16].

In terms of the generalisability of RDM services that were developed for specific environments, one of the first steps was to investigate the needs regarding research data management infrastructures and related services among the researchers on Göttingen Campus. Therefore, interviews with both researchers of different faculties, representatives of the German Primate Center as well as with IT coordinators of different faculties were carried out. The results were matched with information about existing RDM solutions which had been developed for specific contexts. The research data platform of the CRC 1002 on Heart Insufficiency was identified as a promising candidate and is currently being adapted for being used by the research group FOR 2705 studying the drosophila brain. Another use case will be CRC/TRR 274 on Checkpoints of Central Nervous System Recovery, which is currently being applied for.

In addition to the two workshops organised in the reporting period (see Section "GRAcE Networking Events") GRAcE organised a training event at the German Primate Center on RDM in general and Electronic Lab Notebooks in particular in May 2019. Originally, GRAcE was to end by the end of May 2019, but applied for a cost-neutral extension of three months until the end of August 2019, which was granted by the funding organisation. The rest of the project lifetime will be dedicated primarily to wrapping-up open activities, presenting results and documenting the project results in the form of several reports.

OLA-HD

The OLA-HD project - an OCR-D long-term archive for historical prints - is part of the large OCR-D[17] collaborative project. OCR-D focuses on the development of Optical Character Recognition (OCR) methods for historical prints. OLA-HD was selected as one of the eight model projects out of 20 submitted project proposals. The eRA, represented through SUB and GWDG, is executing the project.

A preferably unrestricted access to sources is indispensable in the field of historically working sciences. In numerous research projects, historical works from the 15th century to the present day have been digitized in masses. OCR enables the creation of full texts. The number of such full texts is growing rapidly. Intensive work is being done to make them directly usable for high-quality full text indexing.

[16] https://www.forschungsdaten.org/index.php/Rollen

[17] http://ocr-d.de/eng

Historical texts show great orthographic differences, which makes it almost impossible to specifically search and reuse digitized works. This requires a standardised concept. In addition, the availability and citation capability of the OCR texts is an important prerequisite for the verifiability of scientific results. This means that the existing archiving of an object (as a rule, the objects are catalogued with structure and metadata as well as images according to national bibliographic standards) must be supplemented by OCR texts.

According to these requirements, the main objectives of the project are the development of an integrated concept for long-term archiving, persistent identification of OCR objects and the implementation of this concept. It should also be borne in mind that different versions of a historical source may exist due to research, improvements or technical improvements of the OCR procedures or the use of different OCR techniques. The respective architecture is depicted in Figure 16.

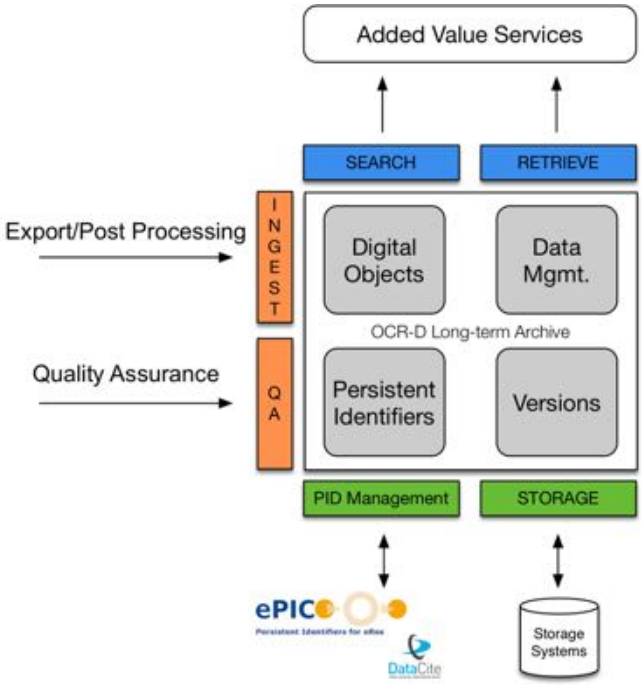

Figure 16: The architecture of the long-term archive.

The eRA is taking care of a tight integration of already existing campus solutions and, wherever possible, designs the long-term archive in a flexible manner to also fulfil use cases from the Göttingen Campus. Furthermore, the project applies an agile development approach integrating members from different eRA-related organisations into one team. The experience from this endeavour will be used for future service developments (see also Section "On-Demand Service Development").

SSHOC

The Social Sciences & Humanities Open Cloud[18] (SSHOC) is a Horizon 2020 project that "*will realise the social sciences and humanities part of the European Open Science Cloud (EOSC) by offering a scalable and flexible access to research data and related services adapted to the needs of the SSH community.*"

SSHOC plans to integrate existing and new infrastructures from the Social Sciences & Humanities ERICs to foster synergies over disciplines and foster interdisciplinary research and collaboration. To achieve this, SSHOC will integrate services and infrastructures from the EOSC-hub project.

[18] https://sshopencloud.eu/

The project started in January, 2019. The Göttingen eResearch Alliance is sub-contractor of the DARIAH ERIC and is responsible for a number of tasks. In particular, this includes the development of a marketplace of the Social Sciences and Humanities and the development of an extended service offer for repositories based on Dataverse, the software behind the Göttingen Campus repository GRO.data. Currently, most of the work is of conceptual nature, but the building of a Dataverse community in Germany has already been initiated by the eRA. Further results will be reported in the next report.

Team Retreat

On March 13th and 14th, the core eRA and GRAcE team members attended a self-organized team retreat at Burg Ludwigstein near Witzenhausen, which was supervised by an external coach from Göttingen[19]. The team discussed the current status, self-understanding and external perception of the eResearch Alliance as well as team roles and current and future tasks. Several team-building activities stimulated the discussions and brought the team closer together, both on a working level and also on a personal basis.

As results, several activities were identified and prioritized with respect to developing the future role and tasks of the eResearch Alliance. The next steps to be taken by the team were identified as the following:

1. Being more proactive
    a. pursue the establishing of an eRA lab, inspired by the Berkeley Institute for Data Science (see subsequent section)
    b. establish regular workshops and courses
    c. make eRA more visible
    d. work towards establishing data stewards at faculties
2. Create new checklist for consulting
3. Being more innovative
    a. define responsibilities for Tech- and Community-Watch
    b. establish a monthly exchange on current eResearch topics
4. Create a modular toolbox for trainings and presentations

The results were presented to the eRA coordinators in April and are currently being worked on by the team. Some impressions from the event can be found in the figure below.

Figure 17: Various pictures from the team retreat.

[19] https://www.nevoteam.de/

Plans for the Next Reporting Period

Overall, the eRA is well established and its offers are demanded and used by a multitude of users. However, we experience that our self-conception not always matches the way others see the eRA. There are still user groups who do not know about the eRA and their services, others know about the eRA, but do not fully understand its purpose and/or service portfolio. Furthermore, the communication towards (and from) the different boards and the presidency of the University of Göttingen can be improved.

The relaunch and ongoing improvement of the web site will be a first step here. The direction is "service first" and it has to become clear that the eRA is a service unit for the Göttingen Campus. This includes links from all campus partners to the eRA web site with properly branded and consistent semantics. The site has been developed with user involvement from the beginning and this will be continued. This effort will be complemented by a closer collaboration with the Göttingen Campus team and by provision of information via their media channels.

Based on the assessment of the KPI "User Coverage" the eRA will contact the Deans of Research of the faculties, selected individual users, and other campus partners to first present the Göttingen eResearch Alliance and, second, to get a better understanding of their demand. This effort will be combined with a target-oriented service roadshow.

With respect to the organisation of work within the eRA and in collaboration with other groups from the eRA partners, we plan to do the following: First of all, it is essential for an improved team work to react to the outcome of the team retreat (see Section "Team Retreat"). To achieve this, the team will discuss the outcome, demand and issues with eRA Management who will derive appropriate actions from this discussion. The overall goal here is strengthen the team, balance work load, and resolve issues. Second, it became clear in 2018 that the self-conception of the Steering Group members differs when it comes to questions like "What is the offer of the eRA?" or "What is the decision-making scope of the Steering Group?". Following a number of meetings in 2018 and 2019, we plan to clarify the remaining issues in 2019 and strengthen the role of the Steering Group. And last, but not least, a discussion started on the development and operation of software and services. The Digital Infrastructure Group has been brought into being, including those players from eRA partners that deal with software and service development. The overall goal is to define DevOps processes across institutions, agree on a common platform, and include operations from day one. The Digital Infrastructure Group will intensify its planning within the next reporting period starting with a common project that will serve as an incubator.

Regarding the four pillars, the following particular plans will be implemented:

- Consulting. In 2019, we will put particular focus on consulting campus partners exploit with respect to the National Research Data Infrastructure (NFDI) initiative. This nationally funded effort wants to "*systematically manage scientific and research data, provide long-term data storage, backup and accessibility, and network the data both nationally and internationally. The NFDI will bring multiple stakeholders together in a coordinated network of consortia tasked with providing science-driven data services to research communities.*"[20] From Göttingen Campus multiple partners are involved in a number of these activities and, giving the nature of the programme, need support for research data-related topics and the respective institutional offers.
- Consulting is one of the most prominent and resource-intensive offers of the eRA. To better understand where resources are actually invested and whether this fits the current strategy,

[20] https://www.dfg.de/en/research_funding/programmes/nfdi/index.html

the eRA will extend the KPIs measurements beyond what is done so far (see also Section "Key Performance Indicators"). The respective information will be fed back into boards to assess and potentially adapt the overall strategy.

- Training: The current plan foresees the implementation of the eRA Lab (see Section "eRA Lab") within the next 12 months. Prior to the actually realisation, it is essential to consult the relevant players on the campus and incorporate the eRA Lab into their offers. Relevant stakeholders are, among others, the Digital Learning and Teaching team, the Campus Göttingen team, the project "Daten Lesen Lernen", the Institut of Informatics, the SUB, the "Digital Creative Space" team, and the GWDG.
- Training. After the Third International Göttingen Summer School on Data Science in August 2019, feedback will be gathered and the planning for the fourth summer school will begin directly.
- The plan regarding Networking mainly focuses on increasing marketing efforts through the collaboration with campus partners and in, in particular, the Göttingen Campus team as a multiplier. Furthermore, as already indicated above, the eRA will address particular faculties, groups, and people on the campus directly to understand their demand and, at the same time, market the offers of the eRA. The efforts will be complemented by
- Services. The continuation of the service portfolio development is the main plan in this area. With the official launch of GRO.data in June 2019 and the soft launch of GRO.publications in the same month, two of the core services will be available timely. For GRO.instruments, a service infrastructure will be set up and user manuals will be written. Although the system is already used in production, new users need a proper onboarding and testing infrastructure, which has to be built and trained. With respect to GRO.plan, it is planned to have a proper development and release roadmap in Q3.2019. Once the status and the timing of all these services is clear, the eRA will plan further integrations (e.g. the cross reference between publications and data objects from GRO.publications and GRO.data, respectively. Further services, like a long-term archive for the campus, will be discussed and planned based on the outcome of the other discussions and depending on the availability of resources.
- Services. The eRA will finally start the service roadshow, which has already been planned to start in 2018. The plan is to visit faculties and research groups to provide information about the service portfolio and its capacities.

Appendix A – Members of the eRA Steering Group

- Jan Brase - Head Research and Development, SUB
- Mustafa Dogan – Deputy Head Digital Library, SUB
- Wolfram Horstmann – Director, SUB Göttingen
- Frank Klaproth – Head of Digital Library, SUB
- Harald Kusch – Member of Medical Informatics Department, UMG
- Jens Nieschulze – Research data officer at University of Göttingen
- Birgit Schmidt – Head of "Wissen als Gemeingut", SUB
- Arnulf Timm – Co-Head of „Informations- und Literaturversorgung Zentrale Erwerbung und Erschließung", SUB
- Ulrich Schwardmann – Deputy Head eScience Group, GWDG
- Philipp Wieder – Head of eScience Group, GWDG, and Deputy Head, GWDG
- Ramin Yahyapour – Managing Director, GWDG

Appendix B – Feedback from the Göttingen eRA Advisory Board 2018

Feedback Report for the Göttingen eResearch Alliance

(Period: 01/2017-02/2018)

By the External Advisory Board

Dr ***Joanna McEntyre*** (Chair) (EMBL-EBI)
Dr ***Sünje Dallmeier-Tiessen*** (CERN-Library)
Dr ***Mark Hedges*** (King's College London)
Prof. Dr ***Stefan Liebig*** (Universität Bielefeld / Deutsches Institut für Wirtschaftsforschung Berlin)
Prof. Dr ***David De Roure*** (Oxford e-Research Centre)

General remarks

The team has done an excellent job this year in moving from start-up phase towards a consolidated suite of activities. The transition appears smooth, coordinated and well managed, as is evident from the maturity of the consultancy, training and networking activities in particular. This is particularly noteworthy given the changes in the eRA staffing and management over the past year which imposed additional challenges and reorganization.

The results delivered speak for themselves: We see a positive movement from consultancy through to a service role resulting in an impressive suite of services that seem to address institutional needs and exploit low hanging fruits (e.g. the existing instruments database); on this note however we would recommend some user research into the requirements and use cases for faculty/users.

It is encouraging to see the good support from the institutional management. This is particularly evident in the central role of eRA in the excellence initiatives and their preparations.

The gender balance in the team could use some redressing. While we acknowledge that this is not an easy task, it is certainly possible to foster more diversity. Perhaps lessons can be learnt from international approaches, e.g. Athena SWAN in UK. In addition, more generally, what is the scope for career progression for eRA staff?

Particular Comments

How has the feedback from last year been addressed?

The feedback from last year has been addressed in certain areas, for example the training activities; the team acknowledged that in some areas (e.g. KPIs), progress was as not as advanced as initially hoped. It is still unclear how to measure success and achievements, in particular in the evolving core area of business of consultancy and that is becoming services.

What is your opinion on the resource planning of eRA within the past year? Has it been reasonable and productive?

The use of resources has been reasonable, in particular considering the presented results. A wide range of activities have been started and consolidated during the past year, e.g. concerning the development of new services.

What is your opinion on the chosen eRA fields of activity? Were they chosen in a constructive and expedient way?
It seems the activities of the eRA have been chosen wisely and in a constructive and expedient way. Through stronger collaborations in Göttingen, within GWDG and the SUB, the eRA has certainly created collaborations to create services, training and trigger more impact than in the start-up phase.

Portal development activities
It is recommended to explore and consider the use cases for the research(er) portal (GRO). Given the effort and resources it takes to develop, operate and curate the service, it is not evident how it relates to the institutional repository for example and how it will be used by the research community in Göttingen.

The work on the Instruments Portal sounds interesting but seems somewhat of an outlier with respect to the other activities in the portfolio. A catalogue of large instruments might be of broad use (wider than the University itself), and in this case it would be useful to compare with initiatives from other institutions or research infrastructures. However there needs to be some thought around how to resource the curation of this catalogue/service effectively in the context of all the other activities.

Knowledge sharing
The networking activities as described seem impressive and wide-ranging. We wondered however how the knowledge outputs were shared across participating groups and across disciplines. Is there scope for "water fountain/tea room/coffee" informal discussions as a conduit for knowledge questions?

Perhaps the spectrum of data science, and the role of the eRA within that, needs more specificity/definition. For example, from research into new methods, vs application (which may use well established methods, but in new data contexts) - where does the eRA sit with its training and networking?

Do you think it is possible that eRA will achieve all its strategic goals within the next year?
Yes, with the current staffing.

As noted last year, there are large dependencies on the outcomes of the Federal Excellence Strategy, which will progress to the proposal stage by the end of 2018. If the funding is not forthcoming, what is the obligation of the eRA to execute the same strategy? The evaluation of impact of the various eRA activities will be critical for decision making.

Recommendations for 2018

It seems that the roadmap to the delivery of the identified services is clear and will be delivered throughout the year. On this path it is recommended, as said, to consider, incorporate and evaluate the use cases from the research communities.

Conclusion

The board congratulates the eRA team to a successful year which demonstrates a remarkable consolidation phase of the new initiative. Particularly, it is noted that a closer collaboration between the university's library (SUUB) and the data centre (GWDG) is visible and seem fruitful in terms of delivery of services.

Appendix C – Consulting: Large-Scale DFG Projects

Project Type	Name/number	Discipline(s)	Consultation topics
TRR	initiative Prof. Flügel	Neurobiology	Provision of data platform including support
CRC	755	Physics, Medicine	INF: Data management solutions
CRC	889	Neurobiology	Data and software management
CRC	937	Physics	Data management
CRC	990	Biology, Ecology, Social sciences	INF: Operation and support, repository, data management
CRC	1002	Medicine	INF: Data platform, electronic lab notebook, antibody registration
CRC	1136	Theology, Philology, Archaeology	Data management, Publication management
CRC	1190	Medicine	INF: Data platform, electronic lab notebook, antibody registration
CRC	1286	Neurosciences	INF: Requirements engineering, establishment of data exchange platform
CRC	initiative Prof. Hohage	Mathematics	Data management
RTG	1666	Social sciences	Data management
RTG	2455	Chemistry	Data management, consultation for data publication platform
RTG	initiative Prof. Blomer	Mathematics	Data management
RU	2064	Philology, History	Data management, data platform, High Performance Computing
RU	2083	Computer Science, Mathematics, Architecture	Data management
RU	2414	Quantum physics	Data and publication management (speaker now in Frankfurt)
RU	2432	Social sciences, Ecology	Data management, data exchange platform
RU	Initiative Prof. Vollmer	Economics	Data management, data policy
RU	Initiative Prof. Wiesemann	Medical ethics, Philosophy	Data management

Appendix D – Publications

Articles

Gnadt, Timo; Steilen, Lena: Beratungsangebote für Forschende. In Bausteine Forschungsdatenmanagement, Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managern. Issue 1/2018: pp. 30-38.
https://doi.org/10.17192/bfdm.2018.1.7816.

Gnadt, Timo: Not without my data. In GGNB Times, Newsletter of the Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences, Issue 3, 2018.
https://www.uni-goettingen.de/de/document/download/33d6ffd044986e98bdbb03629c5b39f6-en.pdf/GGNB%20Times2018Web.pdf

Király, Péter; Stiller, Juliane: Multilinguality of metadata. Measuring the multilingual degree of Europeana's metadata. In M. Gäde, V. Trkulja, and V. Petras, Eds.: Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017) (2017), Schriften zur Informationswissenschaft, Verlag Werner Hülsbusch, pp. 164-176. https://www.researchgate.net/publication/314879735_Multilinguality_of_Metadata_Measuring_the_Multilingual_Degree_of_Europeana's_Metadata

Király, Péter: Towards an extensible measurement of metadata quality. In Proceedings of the 2nd International Conference on Digital Access to Textual Cultural Heritage - DATeCH2017 (2017), ACM Press, pp. 111-115. https://doi.org/10.1145/3078081.3078109.

Király, Péter: Measuring completeness as metadata quality metric in Europeana. Extended abstract published in the Digital Humanities 2017 Conference Abstracts: https://dh2017.adho.org/abstracts/DH2017-abstracts.pdf

Király, Péter; Charles, Valentine; Stiller, Juliane; Bailer, Werner; Freire, Nuno: Evaluating data quality in Europeana: Metrics for multilinguality. In A. Caputo, N. Kanhabua, P. Basile, S. Lawless, D. Gavrilis, C. Papatheodorou, Gifu, and D. Trandabat, Eds.: Joint Proceedings of the 1st Workshop on Temporal Dynamics in Digital Libraries (TDDL 2017), the (Meta)-Data Quality Workshop (MDQual 2017) and the Workshop on Modeling Societal Future (Futurity 2017) (TDDL MDQual Futurity 2017) co-located with 21st International Conference on Theory and Practice of Digital Libraries (TPLD 2017), Thessaloniki, Greece, 21 September 2017, CEUR. http://ceur-ws.org/Vol-2038/paper6.pdf

Király, Péter; Stiller, Juliane; Charles, Valentine; Bailer, Werner and Freire, Nuno: Evaluating data quality in Europeana: Metrics for multilinguality. In E. Garoufallou, F. Sartori, R. Siatri, and M. Zervas, Eds.: Metadata and Semantic Research. Proceedings of the 12th Metadata and Semantic Research Conference - MTSR2018 (Cham, 2019), vol. 846 of Communications in Computer and Information Science, Springer International Publishing, pp. 199-211. http://doi.org/10.1007/978-3-030-14401-2_19.

Király, Péter; Büchler, Marco: Measuring completeness as metadata quality metric in Europeana. In 2018 IEEE International Conference on Big Data (Big Data) (2018), IEEE, pp. 2711-2720. https://ieeexplore.ieee.org/abstract/document/8622487.

Király, Péter: Adat a könyvtárban. In: Hagyomány és újitás a 21. századi könyvtárban (Erdélyi Évszázadok. A Koloszvári Magyar Történeti Intézet Évkönyve. III.) eds. Rüsz-Fogarasi Enikő, Monok István. Kolozsvár (Romania), 2018. pp. 49-74.

Kálmán, Tibor; Durco, Matej; Fischer, Frank; Larrousse, Nicolas; Leone, Claudio; Mörth, Karlheinz; Thiel, Carsten: A landscape of Data – working with digital resources within and beyond DARIAH. In: Int J Digit Humanities (2019) 1: 113. https://doi.org/10.1007/s42803-019-00008-6.

Leone, Claudio: DARIAH-DE – Forschen mit digitalen Diensten und Werkzeugen für alle Geistes- und Kulturwissenschaftler, DHdBlog, February 13, 2018, https://dhd-blog.org/?p=9082.

Steilen, Lena: International Summer School on Data Science 2018, in: GWDG News, 8-9 2018, p. 17-19.

Posters

Engelhardt, Claudia; Bingert, Sven; Kusch, Harald: Exploring the costs and scalability of research data management services: Göttingen Research Data Exploratory. RDA 12th Plenary Meeting, Berlin, March 21-23, 2018.

Presentations

Bingert, Sven; Savin, Valeria; Kraus, Inga; Engelhardt, Claudia; Kusch, Harald: GRAcE: Göttingen Research Data Exploratory. Management parameters derived from the Göttingen eResearch Alliance. VDI-Vernetzungstreffen, Berlin, October 18, 2018.

Engelhardt, Claudia; Gnadt, Timo; Beucke, Daniel; Bingert, Sven; Roertgen, Steffen; Wieder, Philipp; Kusch, Harald; Savin, Valeria: Building a campus-wide RDM structure at the University of Göttingen. Drexel CODATA FAIR-RRDM Workshop, Philadelphia, April 1, 2019.

Gnadt, Timo; Beucke, Daniel; Brase, Jan and Wieder, Philipp: Aufbau einer campusweiten FDM-Struktur an der Universität Göttingen, PODMAN Abschlusstagung „Forschung + Datenmanagement = Forschungsdatenmanagement – Wann geht die Gleichung auf?“, Trier, 27.-28.03.2019

Gnadt, Timo; Beucke, Daniel; Brase, Jan and Wieder, Philipp: Überinstitutionelle Zusammenarbeit in der Göttinger eResearch Alliance, nestor/DINI Workshop „Strukturen entwickeln: Organisation und Governance für lokale FDM-Services“, Siegen, 03.-04.04.2019

Király, Péter: Measuring Metadata Quality Workflow. LDCX at Stanford University, 2018 March. Slides: http://bit.ly/qa-workflow

Király, Péter: Measuring Completeness as Metadata Quality Metric in Europeana Computational Archival Science workshop (part of IEEE Big Data 2018), Seattle, 2018 December 10-13. Slides: http://bit.ly/qa-cas2018

Király, Péter; Stiller Juliane: Evaluating Data Quality in Europeana: Metrics for Multilinguality MTSR 2018. 12th International Conference on Metadata and Semantics Research, Limassol, 2018 Oct 23-26.

Király, Péter: Researching metadata quality, Open Research Knowledge Graph workshop, Hannover, 2018 November 22. Slides: http://bit.ly/qa-orkg2018

Király, Péter: Metadata quality in cultural heritage institutions, ReIReS (Research Infrastructure on Religious Studies) Workshop on FAIR Principle for Digital Research Data Management, Mainz, 2018 November 28. Slides: http://bit.ly/qa-relres-fair

Király, Péter: Measuring Metadata Quality in Europeana and Measuring library catalogs, ADOCHS meeting, Brussels, 2017 November 21. Slides: http://bit.ly/adochs-europeana, http://bit.ly/adochs-marc

Király, Péter: Measuring cultural heritage metadata quality. Linked Data Quality Workshop Semantics 2017, Amsterdam, September 14. Slides: http://bit.ly/mq-semantics17

Király, Péter: Measuring completeness as metadata quality metric in Europeana Digital Humanities 2017, Montréal, 2017 August 7-11. Slides: http://bit.ly/mq-dh2017

Király, Péter; Charles, Valentine: "Nothing is created, nothing is lost, everything changes". Measuring and visualizing data quality in Europeana, ELAG 2017, Athens, 2017. June 6-9. Slides: http://bit.ly/mq-elag2017

Király, Péter; Stiller, Juliane: Multilinguality of Metadata. Measuring the Multilingual Degree of Europeana's Metadata, International Symposium on Information Science, Berlin, 2017. March 13-15. Slides: http://doi.org/10.13140/RG.2.2.30771.43047

Kusch, Harald; Nußbeck, Sara; Király, Péter; Macneil, Rory; Croasas, Mercè: Pilot integration of an electronic lab notebook and an open source research data repository as part of a modular biomedical research data platform. Datavers Community Meeting, Camebridge (Mass., USA), 2018, June 15.

Kusch, Harald; Rörtgen, Steffen; Savin, Valeria; Wache, Stefanie; Kraus, Inga; Engelhardt, Claudia; Bingert, Sven: GRAcE. Anforderungen und Lösungsansätze für ein FDM-Konzept am Göttingen Campus. E-Science-Tage, Heidelberg, March 28, 2019. Slides: https://heibox.uni-heidelberg.de/d/31bb269467/files/?p=/Vortr%C3%A4ge/C1_Harald%20Kusch_Steffen%20Ro%CC%88rtgen.pptx

Kusch, Harald; Wache, Stefanie; Kraus, Inga; Engelhardt, Claudia; Bingert, Sven: What will sustainable research data management cost? Determination of factors to estimate the financial requirements of FAIR research data handling. Göttingen-CODATA RDM Symposium 2018, Göttingen, March 20, 2018. Slides: https://drive.google.com/file/d/1gYwJuUJFFapSZc2gDtCqnd3TEhogZft7/view?usp=sharing

Wieder, Philipp: Practical Research Data Management. Workshop at DPZ, Göttingen, 2019, May 15.

Appendix E – Acronyms and Abbreviations

AG	Arbeitsgruppe (Working group)
AIP Potsdam	Leibnitz-Institut für Astrophysik Potsdam
API	Application Programming Interface
BMBF	Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research)
CDSTAR	Common Data Storage ARchitecture
CODATA	Committee on Data of the International Council for Science
CRC	DFG Collaborative Research Center
DARIAH-DE	Digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften - German part of DARIAH-EU (pan-european infrastructure for arts and humanities)
DAAD	Deutscher Akademischer Austauschdienst
DB	Abteilung Digitale Bibliothek (Department Digital Library) at SUB
DFG	Deutsche Forschungsgemeinschaft (German Science Foundation)
DINI e.V.	Deutsche Initiative für Netzwerkinformation e.V. (German Initiative for Network Information)
DOI	Digital Object Identifier
DONA	Digital Object Numbering Authority
EOSC	European Open Science Cloud
ePIC	Persistent Identifiers for *e*Research
eRA	Göttingen eResearch Alliance
EUDAT	European Data Infrastructure
EPU	Abteilung Elektronisches Publizieren (Department Electronic Publishing) at SUB
FH Potsdam	Fachhochschule Potsdam (University of Applied Sciences Potsdam)
FOR	Forschergruppe, equivalent to RU - research unit
GC	Göttingen Campus is an alliance between the university and local non-university research institutions (http://grc.uni-goettingen.de)
GCDH	Göttingen Centre for Digital Humanities (http://www.gcdh.de/)
GRAcE	Göttingen Research Data Exploratory
GRO	Göttingen Research Online
GWDG	Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
HeKKSaGOn	Foundation of a German-Japanese University Consortium,

	http://www.hekksagon.net/.
ICTSI	International Council of Scientific and Technical Information
IDW 2018	International Data Week (5.-8.11.2018), Gaborone, Botswana (http://www.co-data.org/events/conferences/international-data-week-2018)
IFAP	UNESCO Information for All Programme
INF	Subproject for Infrastructure within joint research projects
ISO	International Organization for Standardization (https://www.iso.org/home.html)
KIT	Karlsruher Institut für Technologie (Karlsruhe Institute of Technology)
KPI	Key Performance Indicator
PI	Principle Investigator
PID	Persistent Identifier
RDA	Research Data Alliance
RDD	Research and Development Department at SUB
RDM	Research data management
RDMO	Research Data Management Organizer
RTG	DFG Research training group
RU	DFG Research unit, equivalent to FOR - Forschungsgruppe
SUB	Niedersächsische Staats- und Universitätsbibliothek (Göttingen State and University Library)
UMG	University Medical Center Göttingen
UNESCO	United Nations Educational, Scientific and Cultural Organization
uniVZ	Course catalogue and staff data base of University of Göttingen
VP	Vice President