Academic Libraries Meet Research Data Management: A Case Study of the University of Ghent’s Central Library’s Practices

Emilie Hermans
National Open Access Desk at Ghent University
Antwerp Area, Belgium
Emilie.Hermans@UGent.be

Abstract
The emerging of research data management (RDM) as a mature part of the research lifecycle, provides a major challenge for research organizations. It gives rise to the need for solid research data services (RDS) to be developed. In the university-wide efforts being deployed in the realization of these services, an academic library can appeal to its expertise in information management to provide a valuable contribution to the discussions. As the context academic libraries operate in varies, the roles it can take upon itself in the process might differ, as working collaboratively with other stakeholders is an essential part in the complex matter of RDM.

This paper analyses the contribution of the academic library of Ghent University (Belgium) to research data management services in the wider institutional context. It explores the hybrid RDM services model employed to encompassing both university-wide and decentralized initiatives.

Current emphases in the RDM program at Ghent University is on the development of consultative RDS.

The starting point for the University of Ghent’s academic library to commit to the development of RDS lay not, in the first place in the new demands from funders in regards to data management planning but stems from a longstanding tradition of providing access to information which reflects in their mission statement: “Facilitating open knowledge creation”. In this paper we hope to provide an example of what the role of a research library can be, reveal its possible strengths (involving a diverse range of stakeholders and expertise, achieving goals through cooperation) and how it tried avoiding pitfalls (a pragmatic view rather than a top-down approach to facilitate researchers need and combining autonomous domain-specific initiatives and central services, reflecting the institutional context).

Keywords: academic libraries, research data management, research data services, case study, Ghent University, OpenAIRE

The role of an academic library in RDM - a conversation starter

Open Access to research data and good research data management (RDM) is increasingly regarded as an element central to all parts of the research lifecycle.1

Good RDM allows to insure transparency and accountability of scientific research, aids the preservation of data now and for the future, enables reuse and reanalysis of data and data sets and can accelerate the advancement of scientific knowledge. It adds further to the paradigm shift towards greater openness of the research process.

In addition, the demand for Data Management Plans (DMPs) and research data to be available for reuse and accountability has been pushed forward by government funding agencies and publishers.

1 This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 643410.
I would like to thank Myriam Mertens, Research Data Officer at Ghent University Library who’s insight and expertise greatly improved this paper.
This evolution has raised the interest in clear and concise information on how to develop RDM services for institutions. The amount of information which is to be found to help guide institutions to organize services for researchers and staff on RDM is increasingly abundant.\(^2\)

But as RDM is more and more recognized as an essential part of the scholarly record, the question where these services should be developed and who can lead the task often remains.

Research Data Services (RDS)\(^3\) are not a single service and partnership is deemed essential to the success of the deployment of data initiatives.\(^4\) With their expertise in information management, it is not surprising that there is a growing consensus that academic libraries can play an essential role in the task of developing research data management services.

Academic libraries have a traditional role in preserving and providing access to the scholarly record in many forms. With its expertise in metadata, selection, collections, institutional repositories, preservation, curation and access, the library can be a trusted steward for RDM.\(^5\) By contributing its expertise, an academic library can be at the forefront of good RDM in a university-wide conversation.

As cooperation is key, a good balance should be found between the library, central administration, faculties and other administration such as the research office and ICT, to develop a spectrum of data services.

We explore the steps taken by Ghent University’s Central Library to start and engage in supporting activities, hereby taken well into account that starting new services is a complex task, not to be completed by a single entity.

To fully support and further expand the ambition towards greater openness to the field of research data, a research data officer has been appointed to assist the development of RDS since June 2015. As collaboration proofed crucial to gain insight in, not only the needs, but also the existing solutions and expertise already (partly) in place in the development of new RDS, the succeeding discussion of the position of Ghent University’s academic library must be viewed in a hybrid model of RDM service development, comprising both decentralized initiatives at faculty level as central, university-wide endeavors both involving library and non-library staff.

**Point of departure – policy and guidelines**

A RDM policy can be a good foundation to start from as it will make the aspirations and goals set for the curation of data clear. It can also give an indication of where the university stands with the RDS and provide the necessary grip to start developing services for the curation of data. As data management is a topic with many facets, getting the right people around the table will help to prevent from missing out important issues such as: data storage facilities, copyright and privacy issues, patenting and ownership of the data.

A first step towards developing a framework for data management within the university was to assemble a Research Data Management Group. The working group was set up in the broader context of the development of a Research Integrity Policy Plan. The working group was set up in the broader context of the development of a Research Integrity Policy Plan.

---

\(^2\) Hybrid structures of centralized and decentralized efforts are one among four approaches to developing e-science support systems identified by, Catherine Soehner et al., “e-Science and Data Support Services: a Survey of ARL Members”, (paper presented at the International Association of Scientific and Technological University Libraries, 31st Annual Conference, West Lafayette, USA, June, 2010), http://docs.lib.purdue.edu/iatul2010/conf/day3/1 (accessed January 2017).


The policy stipulates the basic principles for dealing with research data at Ghent University. It is a general framework set out as a way of enhancing the quality and credibility of research with a commitment from the university to support faculties and researchers via training, tools and infrastructure. And although it aims to develop a university-wide RDM system, autonomy of faculties and researchers to develop their own RDM initiatives is safeguarded.

It should enable researchers and students to carefully manage their research data with domain-specific procedures and guidance on faculty level and university-wide services and infrastructure on a central level. The implementation of the policy will be gradual and takes into account the national and international developments relating to RDM.

It deals with responsibilities, intellectual property rights, privacy, Data Management Plans (DMPs), storage of research data, retention and destruction of data, access to research data and so on.

Rather than posing obligations, the policy aims at giving advice and encouraging the use of DMPs, data documentation, and provides information on secure storage and processing of data during and after research. It recommends to preserve relevant data for a period of 5 years. The guidelines promote openness as it recommends that research data gathered and generated by Ghent University researchers is, wherever possible, made openly available for use by others in a manner consistent with relevant legal and ethical legislation. Although open research data is fostered, the policy recognizes that access may need to be carefully managed in order to maintain confidentiality, guard against unreasonable cost, protect individuals’ privacy, respect consent terms, as well as managing security or other risks.

**Staff skill development for DMP review – DMP analytic rubrics**

The successful development and implementation of RDM depends on a wide sets of skills. A university-wide survey of (doctoral) students and researchers on research skills, showed a lack in knowledge about the newest insights concerning data management and a shortcoming in experience with drafting a Data Management Plans (DMPs). To give more guidance, as a service to researchers or as a part of the evaluation process, some staff members started to review or pre-review DMPs.

Data management planning requires subject specific knowledge. To best judge the soundness of a DMP, peers and staff with the right knowledge of the research field are best suited to validate choices made in the DMP. DMP review is therefore best provided on faculty or department level. At Ghent University two faculties already provide subject specific consultation of DMPs: the faculty of Psychology and Educational Science and the faculty of Arts and Philosophy. However, the library can play a role in pre-reviewing DMPs, as it can give general advice, point out inconsistencies and function as a contact point to denote more specific services needed. For example, Ghent university’s central library provides informal pre-review for DMPs prepared in the context of Horizon 2020 projects.

---

6 UGent TechTransfer manages the intellectual property of the university, drives the use of that intellectual property through the creation of spin-offs and licensing, and supports the research community of Ghent University in creating research collaborations with industry.

7 Both faculties have run pilot projects implementing the use and development of DMPs. These pilot projects proved very useful as user-cases to judge the needs of researchers and to train staff in evaluating the quality of DMPs.
To deploy existing skills, combine efforts and learn from each other’s practices, a small group of DMP reviewers was brought together to share experience on reviewing DMPs. There are resources for reviewing DMP’s available online and one method for reviewing that came to the attention of the DMP-reviewing group was the use of an analytic DMP reviewing rubric.8

An analytic rubric is a standardized form to evaluate data management plans and can be used as a metric for compliance.9 It can be both a tool for reviewers to evaluate a DMP as for researchers to assess and progress the maturity of their data management plans.

To provide support and enable an objective and standardized evaluation in reviewing DMPs, the group decided on the development of their own DMP reviewing rubrics.

An analytic rubric gives an outline of what is expected for every question in a DMP template and describes the levels of detail for every subcomponent.10

For the time being the DMP rubrics drafted at Ghent University are not available for self-assessment by researchers. There are currently two DMP rubrics in development: one based on a generic template and one based on the H2020 template provided in the EC’s guidelines on FAIR data management.11 We decided to match the fixed measurement scale to the level of detail of the description for each characteristics of the level of performance. The scale was set to: “Complete/detailed”, “Addressed issue, but incomplete” and “Did not address issue”.

After developing a list of performance criteria based on the templates, a description of what constitutes each level of performance for each criterion was established. After reviewing, feedback on the descriptions was then gathered from the working group and incorporated. The rubric was then tested on relevant example DMPs and, if necessary, refined.

We hope that this method of reviewing can provide a starting point for other faculties to review DMPs and can bring staff quicker up to speed with the expectations a DMP should meet. In the future it may also prove useful as a self-assessment tool for researchers.

Support for drafting DMPs: DMPonline

DMPonline is a tool developed by the UK’s Digital Curation Centre that provides DMP templates and guidance based on specific requirements of different funders and institutions.12 The web-based tool enables researchers to respond to DMP requirements and with direct guidance and links for more information on specific topics, it provides instructions to deliver a mature DMP.13 Researchers can create and export customizable DMP’s depending on their needs and the requirements of the funder.14 Since the source code is available under a GNU Affero General Public License, you can redistribute or modify it to meet the requirements of your institution.15 Ghent University made use

---

8 Initially the DART project developed a methodology for an evaluation rubric for DMPs. It was presented by the Digital Curation Centre in a practical exercise on reviewing DMPs at the international event: Supporting Data Management Plans: examples from national pilots and institutional projects (25 February 2016). http://www.dcc.ac.uk/events/workshops/supporting-and-reviewing-data-management-plans (accessed January 2017).
12 The original online DMP-tool is available DMPonline, https://dmponline.dcc.ac.uk (accessed February 2017).
13 For more information on the dmponline tool DMPonline, About DMPonline, https://dmponline.be/about_us#Background (accessed February 2017).
of this possibility and developed a localized version. This shared platform for Belgian universities and research organizations: dmponline.be, has currently seven partaking members from different institutions. A shared DMPonline platform has the advantage that common DMP templates (e.g. for external funders) can be added, new releases can be processed for multiple institutions at once, and new features can be co-developed. At the same time institution-specific guidance can be provided, customizing the DMP templates to the liking and requirements of any given party. At the moment the DMPtool contains three templates: a Generic DMP template, a template for Horizon 2020 funded projects and a template based on the requirements of the faculty of Psychology and Education at the university of Ghent. The partaking institutions are in the process of forming a consortium that can manage and support the further development of the tool.

Outreach and advocacy services

Creating the right kind of environment for collaboration, sharing of knowledge and expertise rather than a one-sided top-down approach is a key part of the deployment of RDS for our institution. If data management is purely approached as a funder requirement, RDM in its own right as part of good research practices will be overlooked and researchers might thereby fail to adequately engage with the task at hand. It is therefore important to strike a balance between offering users appropriate levels of support from staff and peers and a broader university-wide vision of RDM.

Several initiatives to advocate for RDM were taken at Ghent University to present this new endeavor. An outline of the RDM guidelines and general information was combined on specific RDM webpages on the online platform of Ghent university. Although expansion of these informational webpages is planned, they give the researcher an overview of the requirement and a first lead in the direction of available services.

Workshops and sessions were set up both for interested parties from a research perspective as well as sessions for staff. The sessions topics and aims were dependent on the level of literacy on the topic; from introductory-level sessions to raise awareness of good data practices and encouraging the use of DMPs to field-specific sessions exploring useful tools. The central academic library organized various training and information sessions around data management. This included e.g. a half-day informational overview of the policy and services at Ghent University for employees to more specific presentations on the requirements for data management for Horizon 2020 projects for researchers. The need for advocacy is a contentious action point for all parties involved in RDM.

International cooperation

What is OpenAIRE

OpenAIRE17 stands for: Open Access Infrastructure for Research in Europe and is a European funded project aiming at facilitating Open Science.

It is a European community and service-driven framework with the overarching goal of providing an infrastructure and support network for enabling Open Science. It consist of an infrastructure: the research information system and a human network of supporting agencies for open sciences.

---

16 DMPonline.be, https://dmponline.be/ (Accessed March 2017) is online in a testing phase at the moment and can be accessed with the login credentials of Ghent University. The access to the tool will be broadened in the near future.

The Research Information System provides services for discovering, storing, linking, analyzing and monitoring all research output across all scientific disciplines. Behind OpenAIRE is a community of Open Access advocates working together to disseminate and provide information and services for better access to research results.18

OpenAIRE supports the European Commission’s open access policies by offering a platform, help desk and technological infrastructure to researchers, research coordinators and institutions. It assist in monitoring H2020 research outputs and will be a key infrastructure for reporting H2020’s scientific publications as it will be loosely coupled to the EC’s IT backend systems. The EC’s extended Research Data Pilot will be supported through European-wide outreach for best research data management practices and Zenodo,19 a data- and publication repository free for all to use, which provides long-tail data storage.

OpenAIRE’s support for Open Data

Ghent University’s central library is a long-standing partner in the OpenAIRE-network. It both functions as a means of aligning efforts as well as strengthening (pan-)European cooperation on supporting Open Access to research output. Horizon 2020 has been at the forefront as a major funder implementing Open Data and Data Management policies.20 An Open Research Data Pilot was piloted at the start of Horizon 2020 in 2015 and has been extended to cover all Horizon 2020 projects from 2017 onwards. For all new Horizon 2020 projects starting in 2017 Open Research Data will be the default option. The focus of the EC is on good data management, which means all parts of Horizon 2020 will be required to create a DMP.21 The policy has operated as an accelerator for the interest in data supporting services.

For this reason, OpenAIRE has developed supporting materials to aid projects to comply with the new guidelines on FAIR data management. The academic library in Ghent herby shared their experience in RDM and aided in the development of supporting materials.

OpenAIRE.eu provides information on the Open Data guidelines and an overview of the requirements and tools available to meet these. On the website you can also find an extensive briefing paper22 and fact sheets on the requirements of the EC for data management23 and handling personal data.24

A data anonymization tool, that will help data owners to safely share their data with experts and the scientific community, is being developed: AMNESIA. It is a web or standalone service that can be loosely integrated and readily fit in the repository infrastructure. Furthermore the infrastructure behind OpenAIRE makes it possible that available datasets are automatically linked with research publication based on those datasets.

OpenAIRE facilitates Open Science training for European research organizations and their stakeholders, and therefore organizes webinars with various presenters on a regular basis.25

---


19 OpenAIRE’s catch-all repository Zenodo was launched in May 2013 in partnership with CERN, pioneer in open source, open access and open data: Zenodo, https://zenodo.org/ (accessed February 2017).


Ghent University Library hosted one of the webinars on open research data and data management for Horizon 2020 projects. It also organized a workshop “Sharing research data and open access to publications in Horizon 2020” focusing on concepts and good practices of managing research data and open access to scientific publications in Horizon 2020.

By participating in the efforts of OpenAIRE the library has a relative advantage in obtaining information on the latest developments of RDM and accompanying initiatives as well as access to the expertise of a broad network of Open Access experts and international connection with the Open Science community.

Challenges and future developments

The developments of a RDM policy framework and guidelines was a first step to outline the different roles within the university and the aspirations towards stewardship and access to research data. Surveying research skills to indicate the most pressing needs, an essential need for support in composing data management plans (DMPs) was revealed. In response to the DMP challenge, initiatives have emanated from both the central administration and specific faculties, with two faculties piloting DMP practices. A systematic reviewing tool in the form of an analytic rubric resulted from sharing existing expertise between faculties and library staff, providing a standardized aid in (pre-)viewing DMPs. Furthermore a localized version of the online tool DMPonline, proofed not only a mode to guide researchers in drafting DMPs but fostered intra-university collaboration. Training of RDS-related skills to cater to the needs of scientific researchers, can foster a collaborative environment that will enable the support for the RDM mission of the university.

The long-term commitment to open access and open science reflects in the national and international involvement of the central library in Open Science initiatives. On a national level Ghent University advocates for Open Access to research results in various committees and the inter-university council, on an international front the library is a longstanding partner in the European project OpenAIRE which facilitates and promotes the European commission’s Open Science policy.

International cooperation, such as the case between Ghent University’s central library and the European project OpenAIRE, may provide the opportunity to weigh on international evolution in data management and open science as well as keep the library involved in a vibrant network of open science advocates.

Ghent University’s central library has invested in the development of RDS for the past 2 years. It is important to recognize that DMS are a shared responsibility. Various university players are relevant stakeholders to ensure efficient development of RDS. A further step in this process is examining how a more structured approach to the consolidation of the tasks needs to be undertaken.

In connection herewith, it is essential to continue the incremental approach towards effective data management. It seems to us, adopting a pragmatic and flexible approach that reflects the institutional realities is the best strategy to obtain researcher acceptance. Customized solutions as well as unified services provided by the library provide a flexible framework that can adapt to the wide variety of challenges presented.

In the future we hope a detailed institutional roadmap help consolidate existing expertise and establish a systematic RDM policy framework.

A more systematic and coherent approach to service development demands better coordination activities. Priority will be given to actions based on the RDM policy and requirement analysis.

---


27 A short summary of the workshop most important findings can be consulted: OpenAIRE, Sharing research data and open access to publications in Horizon 2020 (5th OpenAIRE workshop – November 2015), https://www.openaire.eu/ghent-workshop15 (accessed January 2017).
There is still a need for clearly defined roles on the one hand and guidelines for conducting cost-effectiveness and cost-benefit analyses and benchmarks for making decisions on the other.

The development of services is an ongoing task where consolidating existing efforts and expanding and stimulating new initiatives go hand in hand. DMPonline.be, the existing RDM web pages and the general training sessions will be at the center of the libraries outreach and development activities.

The complexity of RDM will remain a major challenge as well as the great variety of research cultures, which makes a one-size fits all approach inadvisable.

Through its related activities and expertise, Ghent University Library is a key RDM partner at the central level, but a university library also has its limits with specific domain expertise and research experience. In response, we hope to have proposed that adopting a mixed-approach of centralized and decentralized efforts towards a more data-centric academic culture with a pragmatic vision can present opportunities to library staff.

References:


Академске библиотеке и управљање истраживачким подацима: студија случаја централне библиотеке Универзитета у Генту

Сажетак
Појава управљања истраживачким подацима као зреле развојне фазе у истраживачком циклусу представља велики изазов за истраживачке организације. Она доводи до потребе за развојем сервиса истраживачких података. У оквиру сваког култури напора које читав университет улага у реализацију ових сервиса, академска библиотека се може позвати на своју стручност у управљању подацима. Како академске библиотеке делују у различитим контекстима, улоге које оне могу узети на себе у том процесу могу се разликовати, будући да је сарадња са осталим актерима есенцијални сегмент комплексног поступка управљања истраживачким подацима. У овом раду се анализира допринос академске библиотеке Универзитета у Генту (Белгија) услугама управљања истраживачким подацима у широм институционалном контексту. Истражује се модел хибридних услуга управљања истраживачким подацима који је конципиран тако да обухвата и универзитетске и децентрализоване иницијативе. Тренутни нагласак управљању истраживачким подацима која је на Универзитету у Генту је на развоју консултативних сервиса истраживачких података. Одлука академске библиотеке Универзитета у Генту да се посвети развоju сервиса истраживачких података није настала првенствено због нових захтева финансијера у вези са планирањем управљања подацима, већ је произилази из дугогодишње традиције пружања сервиса информацијама. Овај рад приказује пример онога што би могла да буде улога истраживачке библиотеке, открива њени могућности комбиновања аутономних иницијатива специфичних за одређене домене са централним сервисима, који одражавају институционални контекст. Кључне речи: академске библиотеке, управљање истраживачким подацима, сервиси истраживачких података, студија случаја, Универзитет у Генту, OpenAIRE