Research traditions and emerging expectations: PhD students and their research data management

Kerstin Helbig
Humboldt-Universität zu Berlin

ETD 2016 “Data and Dissertations”
Session Uptake and awareness
July 12, 2016

Photo: Creative Commons CC0

Unless otherwise stated this work is licensed under Creative Commons Attribution 4.0 DE License
Research data management in Germany

1998/2013 German Research Foundation (DFG)
Safeguarding Good Scientific Practice

2009 German Research Foundation
Recommendations for Secure Storage and Availability of Digital Primary Research Data

2010 Alliance of German Science Organisations
Principles for the Handling of Research Data

2014 German Rectors’ Conference (HRK)
Management of research data - a key strategic challenge for university management

2015 German Research Foundation
DFG Guidelines on the Handling of Research Data
Research Data Management Initiative at Humboldt-Universität zu Berlin

- **2012**: foundation of the initiative through Computer and Media Service, Research Service Centre, University Library and Vice President for Research

- **2013**: survey among researchers of HU regarding their research data handling and desired support

- **2014**: 17 face-to-face semi-structured interviews

- **2014**: adoption of a research data management policy as well as supplementing guidelines

- Since January **2015** training courses
Key results of survey and interviews

„A central backup service and storage space for research data is required.“ [2, p. 30] (Ø 20-50 GB)

„A legal clarification of fundamental issues is [...] absolutely essential.“ [2, p. 30] (Advice)

Services are necessary to „integrate data with metadata, so that they are more discoverable and accessible“ [2, p. 20] (Advice)
Services of the initiative

• Advice on handling of research data
• Communicating with and transfer to internal and external services
• Support for data management planning
• Workshops on research data management
• Information events within teaching for PhDs and students
• Information material
• Legal advice regarding research data in cooperation with experts
Principles of Humboldt-Universität zu Berlin

Policy

Humboldt-Universität zu Berlin Research Data Management Policy

Preamble

Handling of research data responsibly is critical for confirming research results, for the advancement of science and for the dissemination of knowledge. This policy is addressed to all researchers at Humboldt-Universität zu Berlin both in their role as independent researchers, and as teachers and supervisors. Their obligation includes instructing students and doctoral candidates about handling of research data properly, as well as imparting subject-specific skills and standards.

What is research data?

Research data means all data that was created in the course of research or results from it. Since the generation or collection of research data depends on a variety of research questions and methods, followed by processing, analysis and publishing and/or archiving, the data may involve a range of media, aggregation levels and formats. Sharing and reusing research data requires documentation about the original context and about tools that were used.

Policy

To further the Proposals for Safeguarding Good Scientific Practice of the German Research Foundation, and the Rules of the Humboldt-Universität zu Berlin for Safeguarding Good Scientific Practice and Investigating Allegations of Scholarly Misconduct (Satzung der

Guidelines

A supplement to the Humboldt-Universität zu Berlin Research Data Management Policy

Preface

Humboldt-Universität zu Berlin acknowledges significance of research data as a key basis for scientific and scholarly knowledge. In light of the “Rules of the Humboldt-Universität zu Berlin for Safeguarding Good Scholarly Practice and Investigating Allegations of Scholarly Misconduct” (Satzung der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens, in German) the Academic Senate adopted the “Humboldt-Universität zu Berlin Research Data Management Policy” (Grundsätze zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin, in German). Following guidelines complement this document and provide practical advice for implementing this policy.

There is a need for the University to provide central support to implement the policy. Under the aegis of the Vice President for Research, central services including the Computer and Media Service, the University Library and the Research Service Centre will cooperate on a strategic roadmap for the development of research data management services. These services will be integrated into existing service infrastructure and will be coordinated with academic units. International cooperation with other higher education institutions has guided the development of these. The Research data management initiative, which started in the summer of 2012, has coordinated this issue.1

Research traditions and emerging expectations: PhD students and their research data management
Humboldt-Universität zu Berlin Research Data Management Policy

Policy

To further the Proposals for Safeguarding Good Scientific Practice of the German Research Foundation, and the Rules of the Humboldt-Universität zu Berlin for Safeguarding Good Scientific Practice and Investigating Allegations of Scholarly Misconduct (Satzung der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens, in German) the University adopts following policy:

1. HU researchers are committed to the secure storage, processing, documentation and preservation of their research data for the long term. Responsibility for ensuring these processes lies with HU staff in charge of research projects or programmes.
2. All HU researchers are encouraged to process research data resulting from their research activities according to the conventions and standards of their respective scientific community. They should document the complete research lifecycle including tools and procedures that they used.
3. HU researchers should take responsibility for deciding at what time and on what legal terms research data may be accessed. The Open Access Declaration of the HU (Open-Access-Erklärung der Humboldt-Universität zu Berlin, in German) recommends making research data as well as scholarly publications publicly available in a timely manner. Protection of personal data, copyrights and the legitimate interests of third parties must also be ensured.
4. Research data underlying scholarly publications should be archived for the long-term and/or published in an appropriate trustworthy data archive or repository. They count as scholarly research output at Humboldt-Universität zu Berlin.

Humboldt-Universität zu Berlin has committed itself to establishing the means to enable this policy to be fulfilled.

• Storage/Back-up
• Processing according to discipline-specific standards
• Documentation of the complete research lifecycle including tools and procedures used
• Long-term preservation (at least 10 years) in an appropriate data archive/repository
• Access to data in a timely manner, especially combined with publication – depending on data protection and copyright
Technical tools and services at HU Berlin for research data management

Backup
SAN
HU-Box (Seafile)

edoc Publication Server
Research Data Management
Persistent Identifier (DOI/handle/URN)
Database Service

Media-Repository

Photo: Creative Commons CC0
Target groups

Professors

Research associates and doctoral candidates

Students

Others
Online survey among doctoral candidates

Source: Research data management of doctoral candidates at Humboldt-Universität zu Berlin; own calculation
Training

• Discipline-specific workshops
e.g. Workshop “Research Data Know-How for Geographers”

• Topical workshops
e.g. Workshop "Data management plans for EU, DFG and BMBF"

• Audience-specific workshops
e.g. Workshop “Research data management for PhDs”

• Outreach and information sessions
e.g. “Research data management for librarians”

We also offer workshops for small groups on demand
Workshop „Research Data Management“

- Audience-specific workshop for PhDs
- Aspects of good scientific practice with regard to research data
- Introductory training (120 minutes)
- Group work
- Workshop is offered in English
Workshop agenda

• What is research data?

• Why research data management?

• How does good research data management look like?

In the groups:

• Brainstorming about own data management

• How to describe data and what needs to be kept in mind for storage/archiving?

• Which repositories/data archives are relevant and how to find and use them?
Answers made by doctoral candidates during group session

Source: Word cloud of answers made by doctoral candidates during the group session; own calculation
Specifics of PhD research data management

- Time pressure
- Funder requirements vs. discipline-specific expectations
- Publication pressure
- Missing credit
- Diverse knowledge base (depending on discipline)
- Unclear legal grounds
- Practical help and support
- Easy to use and straightforward tools
Recommendations for other higher education institutions

- Cooperate with strategic partners
- Offer technical solutions
- Include research data into electronic thesis consultation
- Start training as early as possible
- Continuously offer advice and information material
- Try to respond to discipline-specific needs
- Raise awareness
- Offer consultation for conflict situations
- Advocate recognition of a transparent research process
- Create guidelines and standards
New in 2016

• Workshop „Data Management Plans“ for PhDs

• Topical workshop, practice-oriented

• Structure of data management plans

• Varying funder requirements

• Practical work with DMPonline tool
Further activities

• German video tutorials

• Cooperation projects with researchers (third-party funding)

• Project eDissPlus (see presentation and paper by Niels Fromm)
Thank you very much for your interest!

Kerstin Helbig, research data management coordinator
kerstin.helbig@cms.hu-berlin.de

Further information about research data management as well as our workshops and training:
https://hu.berlin/dataman

Follow us on Twitter @DatawomenHUB
Bibliography


Bibliography


Bibliography