

Data Management Policy Plan

M.N. van't Hoog

Version March 2016

1 Introduction

In 2013, following publication of the KNAW's Schuyt committee recommendation entitled 'Managing academic research data with care and integrity' (*Zorgvuldig en integer omgaan met wetenschappelijke onderzoeksgegevens*¹), the University of Groningen started thinking about policy on the storage and management of research data. In 2014, the faculty boards and the directors of the SEP institutes were asked to develop data management plans for their research institutes, setting out how data must be stored, managed and shared. In February 2015, after the faculties indicated that there were still many uncertainties regarding the responsibility for and basic principles of data management and storage, the University of Groningen published a memorandum² that set out the basic principles for the management of research and other data. The Faculty of Theology and Religious Studies' Data Management Policy Plan is based on this memorandum.

1.1 National and international developments in data management

Research data management is a relatively new field of expertise and subject to continuous development. The basic principles set out by the University of Groningen are in part based on a number of important national and international practices and developments:³

- Some academic journals have data access policies that aim to ensure that the researcher takes responsibility for ensuring access to the relevant research data when they publish an article.
- Research funders such as NWO require that good data management practices are put in place at an early stage of a project in the form of a data management plan. In relation to the control of databases, NWO states in its grant regulations (2014) that the research funder and knowledge institute are jointly regarded as the 'producer of the database'. The knowledge institute can only exercise this control effectively if it has set out clear internal agreements to this end.
- There is a growing international call for more open access to data (e.g. the EU pilot project on research data in Horizon2020).
- The generally applicable principles of good data management in the latest version of the VSNU Code of Conduct for Academic Practice (2014). Among other things, the Code of Conduct recommends a retention period of ten years in order to assure verifiability. ALLEA's European Code of Conduct for Research Integrity sets out similar standards.
- The new national Standard Evaluation Protocol (2015-2019) requires information on how 'raw and processed data' is managed under the SEP format heading 'research integrity'. Given the significant differences between the various disciplines, this information will need to be provided in the research data management plans of the individual SEP institutes and/or faculties.
- The new PhD regulations of the University of Groningen stipulate that research data must always be accessible for the purposes of further research.

¹ <https://www.knaw.nl/nl/actueel/publicaties/zorgvuldig-en-integer-omgaan-met-wetenschappelijke-onderzoeksgegevens>, June 2015.

² <http://www.rug.nl/research/search/research-data-office/policy>, August 2015.

³ More information about these developments is available at <http://www.rug.nl/research/search/research-data-office/policy>.

2 Current situation

2.1 Current situation analysis

The current situation in regard to the management of and access to data at the Faculty of Theology and Religious Studies was analysed by means of a questionnaire (response rate 27%) among the academic staff members. The results of this questionnaire were then discussed with the departments and colleagues were also interviewed on this subject individually.

The policy plan was developed in consultation and following an exchange of information with the Research Data Management Committee of the Faculty of Arts, the Research Data Office and the recently installed Ethics Committee of the Faculty of Arts.

2.2 Types of data

The Faculty of Theology and Religious Studies collects and uses a wide diversity of research data. In order to define the necessary measures for data management, storage and access, a distinction needs to be made between unique data that is generated by the researcher and existing data that is freely accessible through public archives. In the departments of Jewish, Christian and Islamic Origins and Christianity and the History of Ideas, many researchers make use of historical and contemporary data that is available in public archives and libraries in digital and/or physical forms. This data is freely accessible to the public and therefore verifiable: if the references in the publications are precise and correct then the analyses and conclusions can be verified by anybody.

Not all archives and databases are freely accessible. Some archives require payment for the provision of access and/or limitations apply due to the presence of confidential information. The archive managers have primary responsibility for the management of the data in their archives.

Most researchers of the Department of Comparative Study of Religion and a number of researchers of the Department of Christianity and the History of Ideas generate unique data in the form of field notes of observations (including participant observations), transcripts, audio and visual recordings of interviews, questionnaires, audio and visual recordings of ritual practices, etc. This data is unique to the researcher who collected it and therefore is not freely available as a matter of course. In most cases this concerns confidential data. A number of researchers collect and analyse publicly accessible photographic and video material such as films and photos of church interiors. The Digital Palaeography research project, a collaboration with the Faculty of Mathematics and Natural Sciences, uses the storage capacity of this faculty and is hence subject to its data management policy.

2.3 Storage during the research

The Faculty of Theology and Religious Studies does not have any separate guidelines for data storage, management and access in addition to the University-wide guidelines (e.g. use of the X and Y drives and registration of publications in PURE). The questionnaire and the interviews revealed that, during the actual research, data is mainly stored on the hard drive of a private laptop and/or external media (e.g. USB stick, external hard drive, Dropbox), and sometimes also on a University PC (X or Y drive). UWP2 is used at the University workstations, but this is not standard practice at external locations. The University Y drive is not used as standard. Non-digital data is stored and archived both at home and at work.

2.4 Archiving after the research

Data obtained from public archives and databases is not archived by the researchers because this data is managed by third parties. Researchers who generate new data archive this data in much the same way as they store it during the research: on the hard drive of their private laptop and/or external media and usually on the University's X drive as well. This research data (both digital and physical) is archived in perpetuity. There are no agreements on data retention

periods at the Faculty or department level. Publications are registered in PURE, but this system does not contain data sets nor references to data sets.

2.5 Data reuse and verifiability

The data generated by the departments of Comparative Study of Religion and Christianity and the History of Ideas is reused by the researchers themselves but not by others. Anthropologists with the Comparative Study of Religion department may request permission from information providers to use their data for specific research, but not for reuse. One anthropologist does request permission to view interview recordings or transcripts in order to verify data. Another applies the data only on request principle.

2.6 Responsibility for storage, archiving, security and access

Currently, the researcher has full responsibility for data storage, archiving, security and access. There are no agreements or guidelines on data storage, archiving and access. Nor are there arrangements on what to do with the data of researchers who have left the employment of the Faculty of Theology and Religious Studies. In practice, if a researcher leaves the employment of the Faculty then only they can make the data available for reuse or verification purposes. New staff members (PhD candidates, research assistants, postdocs, assistant professors, associate professors and full professors) are not provided with data management guidelines or instructions upon appointment.

2.7 Teaching

No specific attention is paid to data management in the teaching programmes for Master's students and PhD candidates. One or two staff members provide explicit instruction to Master's students working on a thesis on the storage of and access to archive material using coding or a database.

3 Faculty of Theology and Religious Studies Data Management Policy

3.1 University data management policy⁴

Our Faculty data management policy is bound by the general data management policy of the University of Groningen.

This policy stipulates that data generated during the course of research must:

- be accurate, comprehensive, reliable, authentic and accompanied by metadata
- be securely stored with a minimum risk of loss
- be registered in a CRIS (e.g. PURE)
- be traceable
- be accessible and citable
- comply with the law, the requirements of ethically responsible research, the agreements in the various contracts and the conditions of the research funders
- be available for verification and further research upon completion of the research and/or termination of the researcher's employment
- be retained for a period of at least 10 years (in principle)

3.2 Data management plan

As of 1 September 2016, a data management plan must be drawn up for all new and current research within the Faculty. Any research project completed before this date does not require such a plan.

3.2.1 Procedure for academic staff

- 1 All researchers must draw up a data management plan based on the standard format (see appendix 4.2).
- 2 They will discuss the data management plan with their supervisor during their Results and Development (R&O) interview. They must send the signed data management plan to their supervisor prior to the interview.
- 3 Once the data management plan has been discussed and approved in the R&O interview, the researcher must submit a digital copy to Willeke van de Pol who will store it.
- 4 Data management is discussed once a year by the CRS based on a comparative analysis of all data management plans.
- 5 A new data management plan must be drawn up in the following cases:
 - new project (e.g. NWO/ERC/etc.)
 - newly collected or generated data not covered in an existing data management plan

The R&O form contains a standard query as to whether a new data management plan is required.

Temporary project researchers are also required to draw up a data management plan, and they are included in the R&O cycle. In research projects that involve several Faculty researchers (e.g. NWO, ERC), the project leader has final responsibility for the data management plan(s).

⁴ <http://www.rug.nl/research/search/research-data-office/policy>

3.2.2 Procedure for PhD candidates

1. The PhD candidate draws up a data management plan at the start of the PhD programme, discusses this with the main supervisor and submits it together with the TSP. The plan must be approved by the Director of the Graduate School.
2. The PhD candidate then submits the approved data management plan and the TSP to Willeke van de Pol, who will store the documents in Hora Finita.
3. The data management plan will be discussed during the interim evaluation and updated if required. The new version will be stored in Hora Finita by Willeke van de Pol.
4. At the end of the PhD programme, the candidate must declare that the data have been stored on the Y drive.
5. External PhD students submit a data management plan as part of the admissions procedure for the Graduate School. The plan must be approved by the Director of the Graduate School.

3.3 Data storage instructions

3.3.1 During the research

All data generated during the research must be retained (e.g. audio and visual recordings, interviews, questionnaires, drawings, photographs, field notes, memos, codes, interview transcripts, SPSS data files). Digital data must be stored on the University X or Y drive to ensure automatic and secure backups. The formats used must be future-proof (see appendix 4.1).

Request extra storage capacity if there is more than 8 GB (X drive) or 100 GB (Y drive) of data by contacting the Faculty Demand Manager, Dries Gankema (a.j.gankema@rug.nl). If more than 3 TB of capacity is required then a charge will apply. Make sure to include the cost item for this extra data storage capacity in the research budget. The Research Data Office (researchdata@rug.nl) can help to calculate the data management costs as part of a grant application.

During the research project, researchers can share data with other University staff and students using the Y drive storage option or Unishare (100 GB storage capacity). Unishare can also be used to share data with colleagues from other universities. Arrange access to Unishare at <https://unishare.nl>, <http://uwp.rug.nl> or through the UWP. Contact the Research Data Office (researchdata@rug.nl) to discuss other situations involving data sharing with external parties (such as businesses). They can provide advice on capacity, security and legal requirements.

The Research Data Office has a data protection specialist whom researchers can approach with questions about storing, archiving and reusing confidential and other data.

3.3.2 After the research

Once the research has been completed, the researcher is required to archive it. This applies to:

- data on which the publication(s) is/are based, unless this data is fully accessible in public archives or databases
- data that is required to be archived by the funders of the research
- data that may be required for reuse by other parties (with the researcher's permission)

It is also recommended to archive unique data that can only be generated once and data that would be extremely costly (in money and/or time) to regenerate.

3.3.2.1 Archiving procedure

Publications are registered in PURE. The data from each publication must be archived on the Y drive. The Secretariat of the Faculty Board reserves capacity on the Y drive for each researcher.

Only the individual researcher (and the members of the research team, if they have authorization) has complete access to this capacity on the Y drive. The Director of the Graduate School (for PhD candidates) and the Faculty Dean (for academic staff) are authorized to access the data if there is sufficient reason to do so (only with read rights).

Data files up to 2 GB containing data that may be made public are stored in the data repository of DataverseNL. Arrange this through the Research Data Office (researchdata@rug.nl).

Data must be archived such that:

- it is traceable and verifiable, i.e. such that an academic staff member from the same discipline can understand and verify that data
- it is accessible in the long term, i.e. stored in a future-proof format (see appendix 4.1)

Non-digital data must be duplicated in digital or non-digital form, so that both the researcher and the Faculty can keep a copy. If it is not possible to make a duplicate, for example in the case of visual recordings, then the data must be stored at the Faculty whereby the researcher will have unrestricted access, including after they leave the employment of the University.

The VSNU Codes of Conduct stipulate a retention period of at least ten years.

3.3.3 Confidential data

If confidential data is collected, the permission of the providers of this data is required in advance for:

1. The use of the data for the research in question.
2. The verification of the data by faculty staff members with the appropriate authorization.
3. The use of the data by the researcher, the research team or other researchers for the purposes of reuse or further research.

These agreements must be set out in the data management plan.

3.3.4 Data ownership

In principle, the data is jointly owned by the researcher and the institute (Faculty). If a researcher leaves the employment of the institute, they retain their right of use. All data that must be retained will remain at the institute; the researcher is permitted to make copies. All agreements on confidential data in the data management plan remain intact.

3.3.5 Responsibility

Data management is the joint responsibility of the researcher, the Faculty and the University of Groningen.

- The researcher has primary responsibility for the data management plan.
- The researcher discusses the data management plan with their supervisor, who must approve it.
- Once a year, the Faculty Board member for Research will verify that data management plans have been drawn up and approved by means of a random sample.
- The supervisor will instruct new staff members regarding the data management agreements and will ensure adequate data access if a staff member leaves the employment of the institute.
- Staff members who leave the employment of the institute are required to sign a standard statement that their data have been stored or archived at the institute in the appropriate manner.

3.3.6 Teaching

Data management must be made an integral part of teaching and instruction, for example in the course unit 'Methods and techniques of research', during thesis seminars and methodology clinics and as part of the instruction on academic integrity during the Bachelor's and Master's degree programmes. This is the responsibility of the Programme Director.

PhD candidates can ask the Research Data Office to help them draw up their data management plans (researchdata@rug.nl). E-learning modules on academic integrity and data management will be developed for the use of external PhD candidates.

3.3.7 Funding

The digitization of non-digital data can be included in the budget for a grant application. In other cases, researchers can discuss the options with their supervisor and the Managing Director of the Faculty.

3.3.8 Concluding remarks

Research data management is a relatively new field of expertise. During the coming years, the existing policies will need to be reviewed regularly and changed where necessary. At the same time, new and better infrastructures and resources to support data management will become available. This policy memorandum will therefore also have to be regularly reviewed and updated. The first review is planned in the autumn of 2018.

The following items will in any case be on the agenda of this review:

- data management procedure for scholarship PhD students
- review of the non-open access option
- the functionality of the data management plan format
- data management quality
- decision-making after the ten-year retention period (how do you decide which data must be retained and which must be deleted after ten years?)

4 Appendices

4.1 Preferred Data Formats⁵

Data type	Preferred format	Acceptable format
Documents	PDF/A (.pdf) Unicode TXT (.txt)	OpenDocument Text (.odt) MS Word (.doc, .docx) Rich Text File (.rtf) PDF (.pdf) Non-unicode TXT (.txt)
Spreadsheets	PDF/A (.pdf) Comma separated values (.csv)	OpenDocument Spreadsheet (.ods) MS Excel (.xls, .xlsx)
Databases	ANSI SQL (.sql) Comma separated values (.csv)	MS Access (.mdb, .accdb) dBase III or IV (.dbf)
Statistical data	R	SPSS Portable (.por) SAS transport (.sas)
Video	MPEG-2 (.mpg, .mpeg, ...) MPEG-4 H264 (.mp4) Lossless AVI (.avi)	QuickTime (.mov)
Pictures (vector)	PDF/A (.pdf) Scalable Vector Graphics (.svg)	Adobe Illustrator (.ai) PostScript (.eps) PDF (.pdf)
Pictures (raster)	JPEG (.jpg, .jpeg) TIFF (.tif, .tiff)	
Audio	WAVE (.wav)	MP3 AAC (.mp3)

⁵ http://www.rug.nl/research/institute-evolutionary-life-sciences/data-management/repository/procedure_gelifes_data_repository_v3.pdf

4.2 Data management plan format, Faculty of Theology and Religious Studies

1. General

1.1 Name	
1.2 Department	<input type="radio"/> Jewish, Christian and Islamic Origins
	<input type="radio"/> Comparative Study of Religion
	<input type="radio"/> Christianity and the History of Ideas
1.3 Name and (if applicable) number of research project*	
1.4 Data management plan document name (year-dmp-name)	
1.5 Date	

* Complete a separate form for each project

2. Data types

2.1 What types of data will be collected and/or generated?	<input type="radio"/> <i>Solely</i> data from public archives and databases? If yes, the form is now complete. Save and sign the form and send it to your supervisor or head of department.
	<input type="radio"/> If no, continue to 2.2
2.2 Describe the data that will be collected and/or generated For example: visual and audio recordings, interviews, questionnaires, drawings, photographs, field notes, memos, codes, call transcripts, SPSS data files	

3. Data storage

3.1 Where are digital data stored during the research?	<input type="radio"/> On the X and/or Y drive
	<input type="radio"/> Not on the X and/or Y drive

	Describe: <ul style="list-style-type: none"> - where the data will be stored - how backups and security have been assured - why the X and/or Y drive are not going to be used
3.2 How much storage capacity is required during the research?	<ul style="list-style-type: none"> ○ Storage on X drive: less than 8 GB ○ Storage on Y drive: less than 100 GB
	<ul style="list-style-type: none"> ○ Storage on X drive: more than 8 GB ○ Storage on Y drive: more than 100 GB <p>You can request extra storage capacity from the Demand Manager, Dries Gankema (a.i.gankema@rug.nl)</p>
3.3 Will you be sharing the data with others during the research?	<ul style="list-style-type: none"> ○ No
	<ul style="list-style-type: none"> ○ Yes, with: <ul style="list-style-type: none"> - University of Groningen colleagues: use Y drive (contact Willeke van de Pol) - Students and colleagues of other universities: use Unishare Go to https://unishare.nl, http://uwp.rug.nl or UWP - Others: ask the Research Data Office researchdata@rug.nl
3.4 Is the data confidential?	<ul style="list-style-type: none"> ○ No
	<ul style="list-style-type: none"> ○ Yes Describe the agreements reached with the providers of the data and where the documents recording these agreements have been stored <p>Any questions? Contact the data protection specialist of the Research Data Office researchdata@rug.nl</p>

4. Archiving after the research

4.1. Is the data suitable for reuse?	<ul style="list-style-type: none"> ○ No
---	--

	<p>Explain why not</p> <p>Continue to 4.3</p>
	<p><input type="radio"/> Yes</p> <p>Continue to 4.2</p>
<p>4.2. In which data repository will the data be stored?</p>	<p><input type="radio"/> DataverseNL (up to 2 GB). You can arrange this through the Research Data Office (researchdata@rug.nl)</p> <p><input type="radio"/> Other (please specify):</p>
<p>4.3. Will non-digital data be used (photos, videos, <i>field notes</i>, diaries, etc.) that has to be digitized for archiving?</p>	<p><input type="radio"/> No</p>
	<p><input type="radio"/> Yes</p> <p>Discuss with your supervisor</p>
<p>4.4. Will non-digital data be stored at the Faculty?</p>	<p><input type="radio"/> Yes</p>
	<p><input type="radio"/> No</p> <p>Explain why not</p>
<p>4.5. Will the digital data be stored in a preferred format (see appendix 4.1 of the Data management policy plan)?</p>	<p><input type="radio"/> Yes</p>
	<p><input type="radio"/> No</p> <p>Explain why not</p>
<p>4.6. How and where can the data be accessed for the purposes of verification?</p>	
<p>4.7. How and where can the data be accessed for the purposes of verification if the researcher leaves the employment of the University?</p>	

5. Conclusion

Date agreed with supervisor	
Signature of researcher	
Signature of supervisor	