Data Management Protocol of the Nieuwenhuis Institute (NI)
Version PhD theses 1.0

Aim of data archiving
A mindful and deliberate method of archiving research data is vital to ensure:

- Transparency of conduct and methods used in research at our institute;
- Scientific progress: enabling replication of research (by providing data sets, methods used and other relevant materials);
- Quality control actions.

Documenting research according to the protocol is, besides reasons of transparency, scientific progress and quality control, useful for researchers themselves. It ensures that they can always, even many years later, retrieve details about their research.

Target group
This protocol relates to data collected by members and affiliate members of the NI and to statistical meta-analyses, qualitative research, data simulations and analyses/secondary analyses of data collected by third parties (i.e. persons or bodies that are not students or staff of the NI) or available in databases.

Members/affiliate members of the NI are also responsible for ensuring that there is a publication package for every PhD. PhD thesis members/affiliate members are responsible for ensuring that there is a publication package for their dissertation. These rules apply from 1 January 2016 and apply to everything published after that date and PhD theses defended after that date.

Data storage
Every member/affiliate member of the NI has access to storage space on the Y: drive (Y:\Staff\gmw\NI\your group\your p-number). The publication package should be saved on the Y: drive by the PhD student. PhD supervisors/co-supervisors working at the University of Groningen have the right of access to the packages in which they are involved (writing rights, i.e. the ability to save data). The PhD student should apply for these rights via the NI data storage folder access request-form (https://fd2.formdesk.com/rug/GMW_Request_access_NI_folder). At present it is not possible for supervisors/co-supervisors at other universities to have access to the packages, but this should be possible at some future date (once PURE or another solution is able to deal with persistent identifiers). Subscriptions will be safeguarded by Miranda Cazemier van den Berg (m.cazemier.van.den.berg@rug.nl).

Publication package
In the case of all PhD theses submitted after 1 April 2016 publication packages must be saved on the Y: drive by the PhD student. The PhD student should apply for writing rights to the publication package for all co-authors and supervisors/co-supervisors affiliated to the University of Groningen for these rights via the NI data storage folder access request-form [https://fd2.formdesk.com/rug/GMW_Request_access_NI_folder](https://fd2.formdesk.com/rug/GMW_Request_access_NI_folder).

**Access and verification**
All PhD thesis supervisors/co-supervisors who are working at the University of Groningen have writing rights to the packages saved by their PhD students. The Research Director and his/her Secretariat have reading rights to all publication packages stored on the Y: drive to enable them to carry out spot checks to see (a) whether there are publication packages for the PhD theses of NI members.

**Retention period**
In principle all data on the Y: drive is retained permanently, even after a member/affiliate member of the NI has left. 
*Note:* The American Psychological Association (APA) lays down the rule that data must be available for five years after the publication of the research. The Y: drive, however, is a backup system that cannot be deleted (everything stored on that drive can be retrieved, even if someone edits or deletes data after it has been saved). Publication packages can therefore always be retrieved, even after a member/affiliate member of the NI has left the University.

**Consequences for PhD’s**
Providing the complete publication package for a PhD thesis is a condition for manuscript approval. If no publication package is stored in the correct folder, or parts are missing the manuscript will not be sent to the reading committee.
Appendix: Publication package PhD
How to store data in the case of a PhD theses?

1. For every empirical PhD thesis submitted after 1 April 2016 there must be a folder on the Y: drive. This is the publication package. The name of the publication package should clearly comprise PP (Publication Package), the wording PhD, the name of the PhD student and the year of PhD submission (e.g. PPPhDVanderVeen2016). All supervisors/co-supervisors working at the University of Groningen must have writing rights to the publication package: application for these should be made via the NI data storage folder access request-form (https://fd2.formdesk.com/rug/GMW_Request_access_NI_folder).

2. The publication package should comprise:
   a. A data archiving form stating for each PhD thesis (or project within a PhD thesis):
      i. Whether an ethical review has been carried out and if so where, the protocol number of the review and the date of approval (if approval was required).
      ii. A description of the contribution made by each supervisor/co-supervisor, for example ‘devising and organizing the project’, ‘data collection’, ‘data analysis’, ‘critical reading of the text/feedback’.
      iii. Detailed information on who collected the data, where and at what location, and where the data is stored if not on Y: (e.g. separate data-bank or on a protected repository of an institution outside of the University). This also applies to data collected by third parties, insofar as this information is available.
      iv. In the case of data from existing databases and repositories the origin, version and date must be specified.

   b. (If applicable) The ethics committee application and certificate of approval.

   c. (For each project within the PhD) The ‘raw database’ as initially available to the member of the NI digitally or in digitized form.
      i. Relevant data on participants (guaranteeing their anonymity), including the date(s) on which they took part should be included in this file.
      ii. Data in the raw database must be anonymized in accordance with ethical guidelines.
      iii. The data package should not include information that identifies persons or consent forms (which must also be retained unless other arrangements have been made on account of ethical objections).
      iv. In the case of non-digital source material (physiological findings, video/audio recordings, paper questionnaires – which may also be scanned and saved on the Y: drive – consent forms, etc.) the data archiving form must state where this is stored and who has access to it.
      v. In cases where the PhD student is not allowed to have the data on account of ethical guidelines, insurmountable logistical problems, privacy-related reasons or formal property rights the data archiving form must clearly state the reasons for the absence of the data from the package. In such cases the publication package should if possible include a random sample from the original database or a screenshot of part of the raw data.
d. (For each project) all the digital research material (or materials that can be scanned) used in the project that is needed to replicate the research, including:
   i. Questionnaires, stimuli, instructional texts, experiment leader protocols, video/audio material, software for simulation studies, computer scripts, a description of equipment used, etc.
   ii. Any source references made to material published elsewhere must be accurate and sufficiently specific.

e. (For each project) syntax, computer scripts or statistical logbooks of the processing of the raw data and a properly documented and edited database to enable analyses such as those reported in the article to be replicated from the raw database.

f. The PhD thesis

3. The publication package must be available as soon as a PhD thesis is submitted to an assessment committee. Any supervisor(s) approving the PhD thesis thereby confirm(s) that a proper publication package is available.