



Publication Packages & Research Packages

Contents

- 1 Publication and research packages: Why?
- 2 A short introduction: Who, what, where, and when?
- 3 Publication packages
- 4 Example of a publication package
- 5 Research packages
- 6 Example of a research package
- 7 Appendix 1. Checklist publication packages.
- 8 Appendix 2. Sustainable data formats

1. Publication and research packages: Why?

Following the guidelines of the University of Groningen⁸ and the Association of Universities in the Netherlands (VSNU)⁹, the Department of Sociology implemented policy to stimulate researchers to handle their research data in a responsible way.

In short, researchers in the department are from September 2016 onwards requested to document important steps that are taken during their research and store their digital data in special folders or *packages*. In the Department of Sociology there are two types of packages:

- **Publication packages** contain all data, syntaxes, and other important information needed to replicate results that are published in a scientific journal or book (chapter)
- **Research packages** contain all data, syntaxes, and other important information belonging to a larger research project (i.e., projects where several publications are based on one dataset such as the KiVa project)

The most important reason to document your research and store your data in a safe place is that you and your co-authors/research colleagues/supervisors will still have access to the data and know which important steps were taken during the research in a few years (i.e., **safekeeping** of valuable information). In addition, if needed, others can **verify** and **replicate** your research.

The packages are in principle intended to verify and replicate research. They are not meant to share data for reuse. However, the policy of the Department of Sociology is that data become publicly available unless there are ethical, juridical, contractual reasons not to do so. Sharing data contributes to the scientific body of knowledge and saves effort, time, and costs. Grant suppliers (such as NWO, KNAW, and ZonMw) and journals increasingly demand that research data are made accessible for other researchers in so-called trusted repositories (e.g., DANS).

2. A short introduction: Who, what, where, and when?

Who?

Publication packages

All members of the Department of Sociology (**staff**, **postdocs**, and **PhD students**) are responsible for ensuring that there is a publication package for every accepted article in a scientific journal or book (chapter) of which they are the first author. For more information about the publication packages, please see chapters 3 and 4. For **PhD students** and **bachelor** and **master** students special rules apply, please see sections 3.2 and 3.3.

Research packages

All principal investigators of the Department of Sociology are responsible for ensuring that there is a **research package** for their research projects. For more information about the research packages, please see chapters 5 and 6.

Access rights

By default, only the owner of the package has reading and writing rights. The coordinator and technical support director have reading rights.

⁸ <http://www.rug.nl/research/search/research-data-office/policy/documents/2015-research-data-policy.pdf>

⁹ http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code_of_Conduct_for_Scientific_Practice_2012.pdf

It is important that others involved in the research (co-authors/research colleagues/supervisors) can access the packages as well. Access can be requested via this form:

<http://myuniversity.rug.nl/infonet/medewerkers/ict/serviceesk/mutation-y-drive>

Access rights need to be requested per package. At present it is not possible for co-authors/ research colleagues/supervisors at other universities to access the packages.

What?

The packages contain all (quantitative and qualitative) data used in a certain study. Note that with 'data' we mean (primary and secondary) empirical data but also other types of data, such as meta-analyses. When for certain reasons it is not possible to include the data in a package (ethical guidelines, insurmountable logistical problems, privacy-related reasons or formal property rights), contact the coordinator.

Where?

Every member of the Department of Sociology has access to the safe digital environment of the university: the Y drive. On the Y drive a folder has been created for the publication and research packages (Y:\Staff\gmw\sociologie). The publication packages should be placed in the folder with your P-number which can be found in the folder of your research cluster (e.g., WALM). Contact the coordinator when you cannot find the folder with your P-number (for instance, when you recently started working at the department). Also contact the coordinator when you want to make a new research package.

All data on the Y drive are retained for at least ten years. The Y drive uses a backup system that makes sure files cannot be deleted (everything stored on that drive can be retrieved, even if someone edits or deletes data after it has been saved). When you leave the department you can request to keep your P-number so that you will still be able to access your package.

When?

It is wise to make the package as soon as possible after the completion of the research project. In any case, the package should be ready within three months after the publication or completion of the research project.



3. Publication packages

3.1 General information

Publication packages are made for every published scientific article or book(chapter) of which a department member is a **first author**.

On the Y drive a folder has been created for the publication packages (Y:\Staff\gmw\sociologie). The publication packages should be placed in the folder with your P-number which can be found in the folder of your research cluster (e.g., WALM). The name of the publication package should be 'PUB' followed by the surnames of the first, second, and last authors and year of publication (e.g. PUBJansenBakkerSmit2015).

Given the different types of research in the department it is difficult to provide an exhaustive list of information that should be included. A guideline is that a colleague who was not involved in the study should be able to understand your study and replicate your results based on the information in the package. For most studies the package would at least contain the following information (see also the checklist in Appendix 1):

- A *readme* file explaining which files are included in the package
- The published article or book (chapter)
- **Information that is needed for others to understand your study:**
 - Description of how and by whom the data were collected or created
 - Overview of questions that were asked (i.e., questionnaire)
 - Explanation of the variables (i.e., codebook)
 - Description of sample/participants
 - Information on whether your study was evaluated by an ethics committee (if so, include application and letter of approval)
 - A description of the contributions made by each author/co-author, for example 'devising and organizing the project', 'data collection', 'data analysis', 'critical reading of the text/feedback', and 'writing of the article/PhD thesis'
- **Information that is needed for others to replicate your results:**
 - Raw (or original) digital data¹⁰. Contact the coordinator when it is not possible to store the data in the publication package (ethical guidelines, insurmountable logistical problems, privacy-related reasons or formal property rights)
 - In the case of non-digital data state where the data are stored and how they can be accessed
 - Syntax 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
 - Syntax for created variables/scales and analyses or coding schemes for qualitative data
 - Other digital relevant research material: stimuli, instructional texts, experiment leader protocols, video material, software for simulation studies, ...

It is important that others involved in the research (co-authors/supervisors) can access the package. Access can be requested via this form:

<http://myuniversity.rug.nl/infonet/medewerkers/ict/serviceesk/mutation-y-drive>

The package should be ready within three months after the article or book(chapter) is published. If staff members fail to make a publication package, they can be sanctioned (e.g., reduction of research time).

¹⁰ We strongly recommend storing the data in a sustainable format (see the preferred formats in Appendix 2). Data in sustainable formats can still be accessed when for instance certain software is no longer available.

3.2 PhD students

The policy of the university is that PhD students can only defend their thesis when there is a publication package for each chapter of their dissertation (excluding the introduction and conclusion). This implies that the graduate school will not send out the thesis to the reading committee before the packages are in place.

For PhD students the same guidelines as described in 3.1 apply. Please make a separate folder for each chapter of the dissertation (excluding the introduction and conclusion) and clearly indicate the chapter of the thesis in the name of the folder (e.g., Oldenburg_chapter1). Include a Word or PDF version of the chapter, also when the chapter is not yet published.

Send an e-mail to the coordinator¹¹ when you have completed the packages. The coordinator will check whether all necessary files are there¹² and inform the graduate school. Most packages are checked by the coordinator within a week, but if you are on a tight schedule it is wise to contact the coordinator beforehand.

3.3 Bachelor and master students

As of April 2017, in line with university guidelines, publication packages also need to be made for **bachelor** and **master** theses supervised by members of the Department of Sociology.

For bachelor and master students the same guidelines as described in 3.1 apply. Supervisors make a folder for their students within their own folder. Subsequently they give the student access to this folder via the following link:

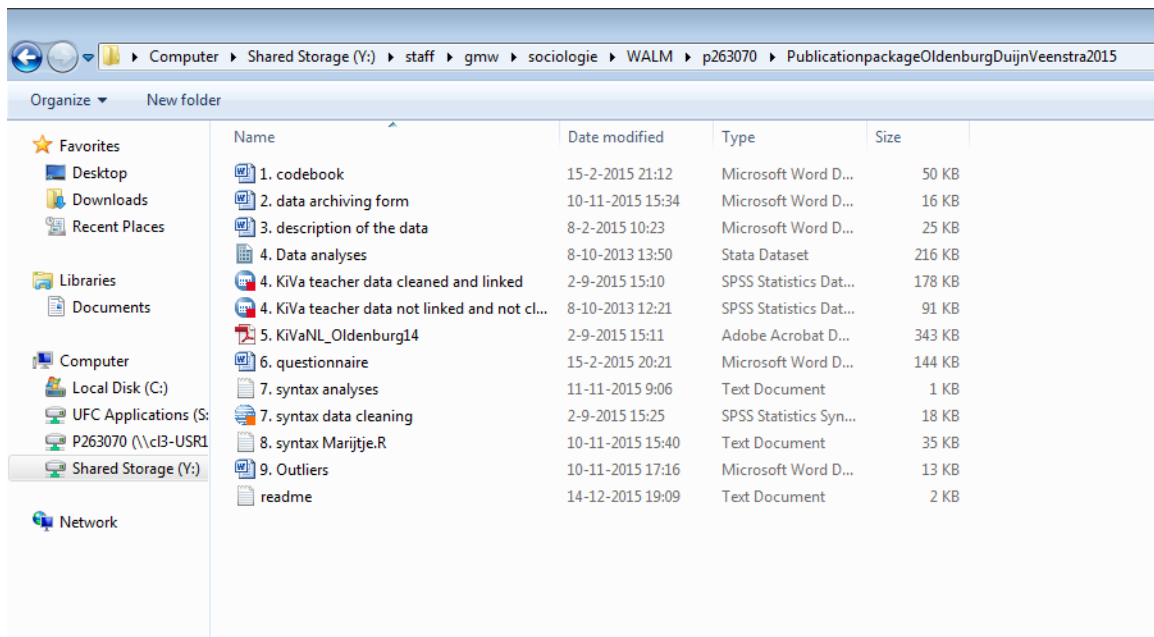
<http://myuniversity.rug.nl/infonet/medewerkers/ict/servicedesk/mutation-y-drive>

Students construct the package, but the supervisor is responsible for checking whether all files are present. When the thesis is finished, the supervisor prints and fills out the checklist in Appendix 1. This checklist needs to be signed and handed in at the 'onderwijsbalie' together with all the other documents (e.g. 'examenbriefje').

¹¹ Currently the coordinator is Beau Oldenburg.

¹² At this moment the coordinator will only check whether all necessary files are in place. See the checklist in Appendix 1. In the near future, the faculty's integrity committee will possibly check the packages more in-depth (e.g., rerun analyses).

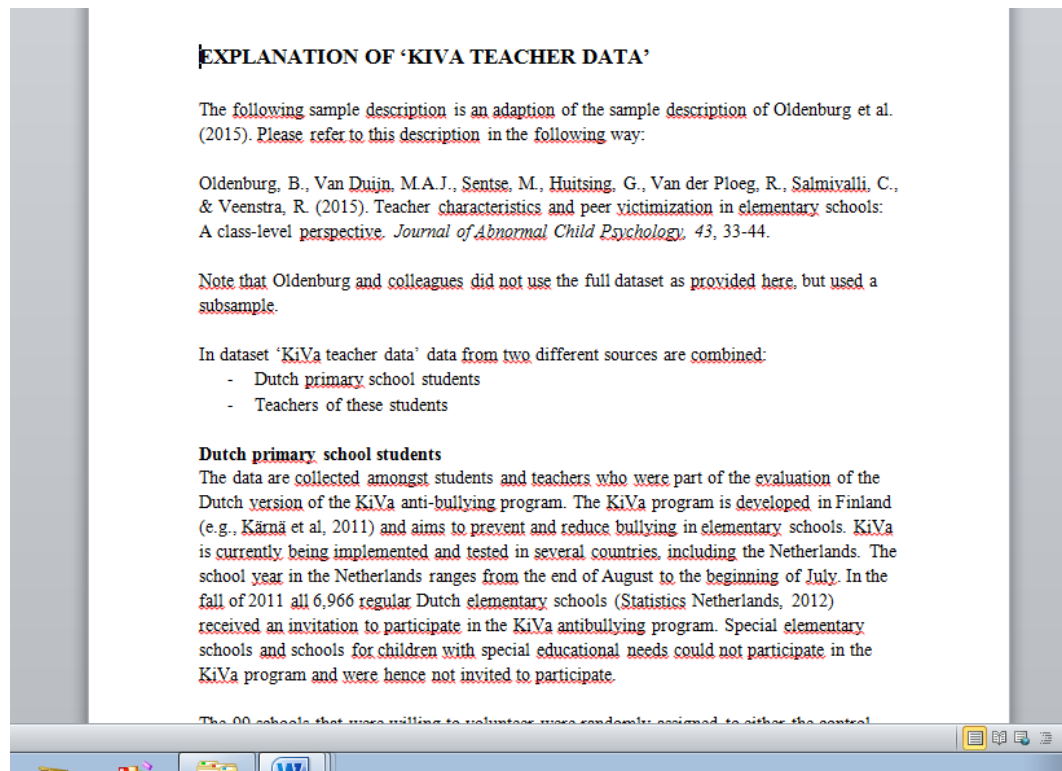
4. Example of a publication package



- 1. Codebook:** The codebook contains information on the variables in the data. In this codebook the name of the variables, a description of the variables, the source of the information (e.g., teachers or students), the value labels, original scales and references to the original scales are included.

Variabele	Beschrijving	Bron	Labels	Schaal	Referentie
<i>klasnr.</i>	Klas id nummer	KiVa student data			
<i>leerkrachtnr.</i>	Nummer leerkracht	Onderzoekers			
<i>schoolnr.</i>	Nummer school	KiVa student data			
<i>klas.</i>	Aan welke klas wordt lesgegeven	Leerkracht			
<i>datum</i>	Datum invullen vragenlijst	Leerkracht	(1) 11/12 juni 2012 (2) 18/19 juni 2012 (3) 25/26 juni 2012 (4) 2 juli 2012		
<i>laat</i>	Vragenlijst is na de leszing ingeleverd	Onderzoekers	(0) Niet te laat (1) Te laat		
<i>gebojaar.</i>	Geboortejaar	Leerkracht			
<i>uren</i>	Uren lesgeven per week	Leerkracht			
<i>jarenles</i>	Hoeveel jaren geeft u al les?	Leerkracht			

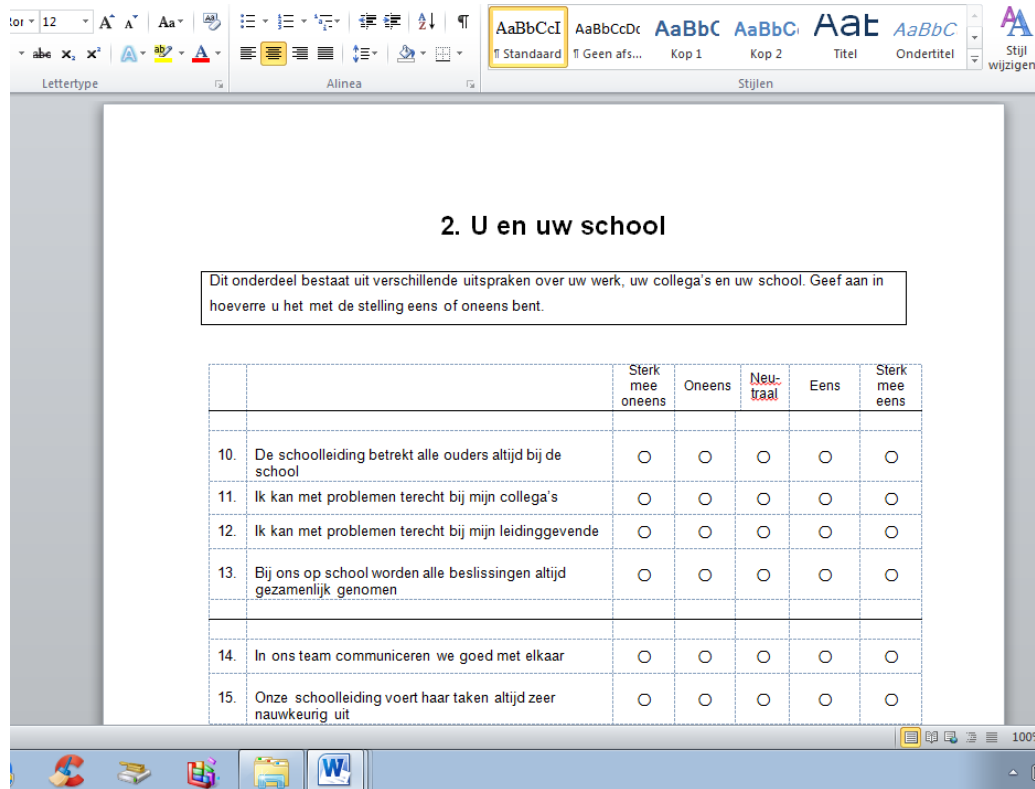
2. **Data archiving form:** This document contains information on whether the project was reviewed by an ethical committee. In addition, the contributions of each author are described. Furthermore, is described who collected the data. Finally, in this file is described that the data are stored open access at data archive DANS and that the paper questionnaires are archived at the GMW archive (and will be kept there for 10 years).
3. **Description of the data:** Because this was a rather complex data-collection (combining data from two sources) I made an extra file with an in-depth explanation of the data.



4. **Data analyses, KiVa teacher data cleaned and linked, and KiVa teacher data not linked and not cleaned:** I included the raw data and the data on which the publication was based in the publication package. It is important to always keep the original data file. This way the results can always be reconstructed.

5. **KiVaNL_Oldenburg14:** Published article.

6. Questionnaire: I included the questionnaire as presented to the respondents in the publication package.

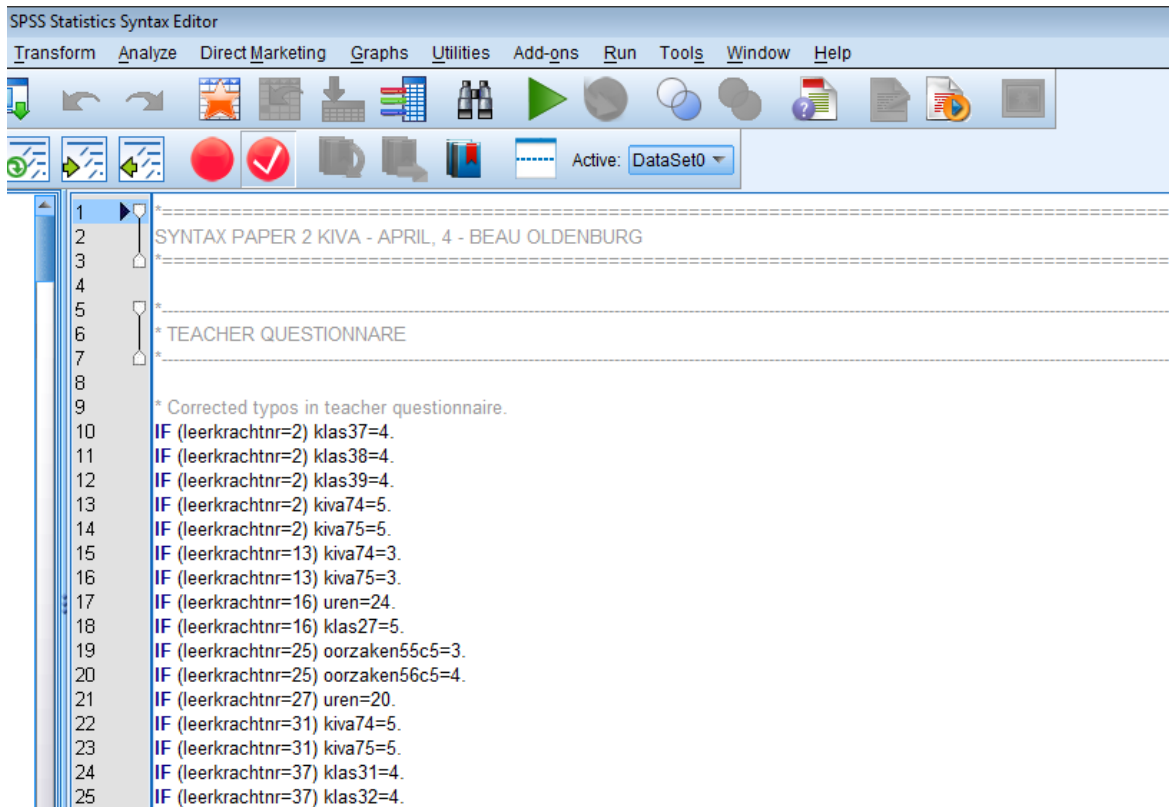


2. U en uw school

Dit onderdeel bestaat uit verschillende uitspraken over uw werk, uw collega's en uw school. Geef aan in hoeverre u het met de stelling eens of oneens bent.

		Sterk mee oneens	Oneens	Neu- traal	Eens	Sterk mee eens
10.	De schoolleiding betreft alle ouders altijd bij de school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11.	Ik kan met problemen terecht bij mijn collega's	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12.	Ik kan met problemen terecht bij mijn leidinggevende	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13.	Bij ons op school worden alle beslissingen altijd gezamenlijk genomen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14.	In ons team communiceren we goed met elkaar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15.	Onze schoolleiding voert haar taken altijd zeer nauwkeurig uit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Syntax analyses & syntax data cleaning: I cleaned the data using SPSS and analyzed the data using Stata.



```

SPSS Statistics Syntax Editor
Transform Analyze Direct Marketing Graphs Utilities Add-ons Run Tools Window Help
Active: DataSet0

1
2 SYNTAX PAPER 2 KIVA - APRIL, 4 - BEAU OLDENBURG
3
4
5
6 * TEACHER QUESTIONNAIRE
7
8
9 * Corrected typos in teacher questionnaire.
10 IF (leerkrachtnr=2) klas37=4.
11 IF (leerkrachtnr=2) klas38=4.
12 IF (leerkrachtnr=2) klas39=4.
13 IF (leerkrachtnr=2) kiva74=5.
14 IF (leerkrachtnr=2) kiva75=5.
15 IF (leerkrachtnr=13) kiva74=3.
16 IF (leerkrachtnr=13) kiva75=3.
17 IF (leerkrachtnr=16) uren=24.
18 IF (leerkrachtnr=16) klas27=5.
19 IF (leerkrachtnr=25) oorzaken55c5=3.
20 IF (leerkrachtnr=25) oorzaken56c5=4.
21 IF (leerkrachtnr=27) uren=20.
22 IF (leerkrachtnr=31) kiva74=5.
23 IF (leerkrachtnr=31) kiva75=5.
24 IF (leerkrachtnr=37) klas31=4.
25 IF (leerkrachtnr=37) klas32=4.
  
```


8. Syntax Marijtje: Syntax used for checking outliers, making plots, etc.

9. Outliers: Overview which outliers were eventually deleted.

Readme file: A file explaining which files are in the package.

5. Research packages

Research packages contain all data, syntaxes, and other important information belonging to a larger research (i.e., projects where several publications are based on one dataset such as the KiVa project). All principal investigators of the Department of Sociology are responsible for ensuring that there is a research package for their research projects. A benefit of making a research package is that individual researchers can to a large extent refer to the research package when making their publication packages.

On the Y drive a folder has been created for the research packages (Y:\Staff\gmw\sociologie). Contact the coordinator when you want to make a new research package.

Given the different types of research in this department it is difficult to provide an exhaustive list of information that should be included. A guideline is that a colleague who was not involved in the project should be able to understand your research. For most research projects the package would at least contain the following information:

- A *readme* file explaining which files are included in the package
- The published articles or books (chapters) within the project
- A description of the contribution made by each researcher. Types of contribution include 'devising the project', 'organizing the project' and 'data collection'
- Information on who owns the data
- Overview of questions that were asked (i.e., questionnaire)
- Explanation of the variables (i.e., codebook)
- Description of sample/participants
- A description of whether the project was evaluated by an ethics committee (if so, include application and letter of approval)
- Raw (or original) digital data¹³. Contact the coordinator when it is not possible to store the data in the research package (e.g., due to ethical guidelines, insurmountable logistical problems, privacy-related reasons or formal property rights)
- In the case of non-digital data state where the data are stored and how they can be accessed
- Syntax 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
- Syntax for created variables/scales or coding schemes for qualitative data
- Other digital relevant research material: stimuli, instructional texts, experiment leader protocols, video material, software for simulation studies, ...

It is important that others involved in the research (co-authors/supervisors) can access the package. Access can be requested via this form:

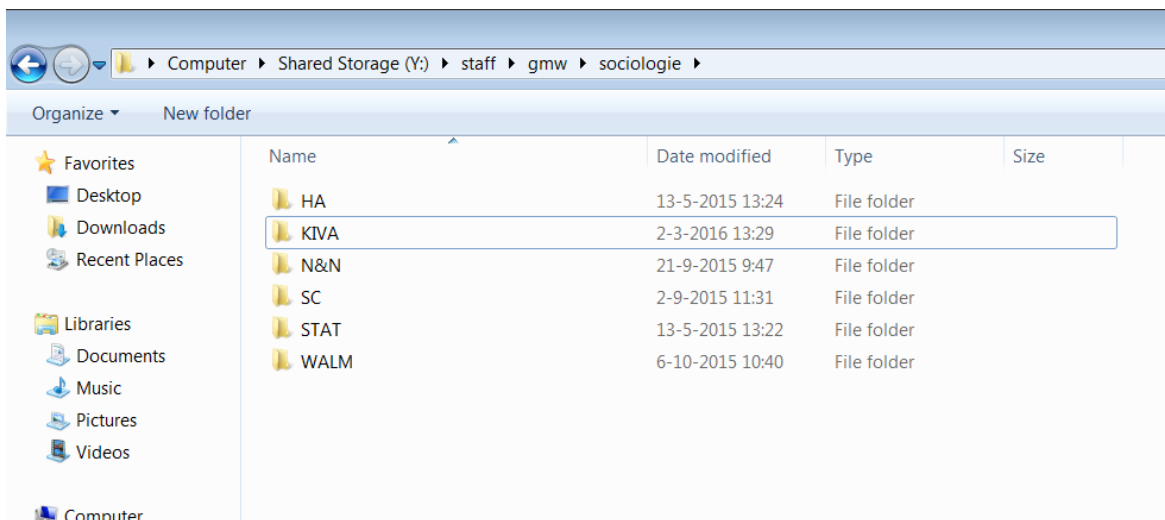
<http://myuniversity.rug.nl/infonet/medewerkers/ict/servicedesk/mutation-y-drive>

¹³ We strongly recommend storing the data in a sustainable way (see the preferred formats in Appendix 2). Data in these formats can still be accessed when for instance certain software is no longer available.

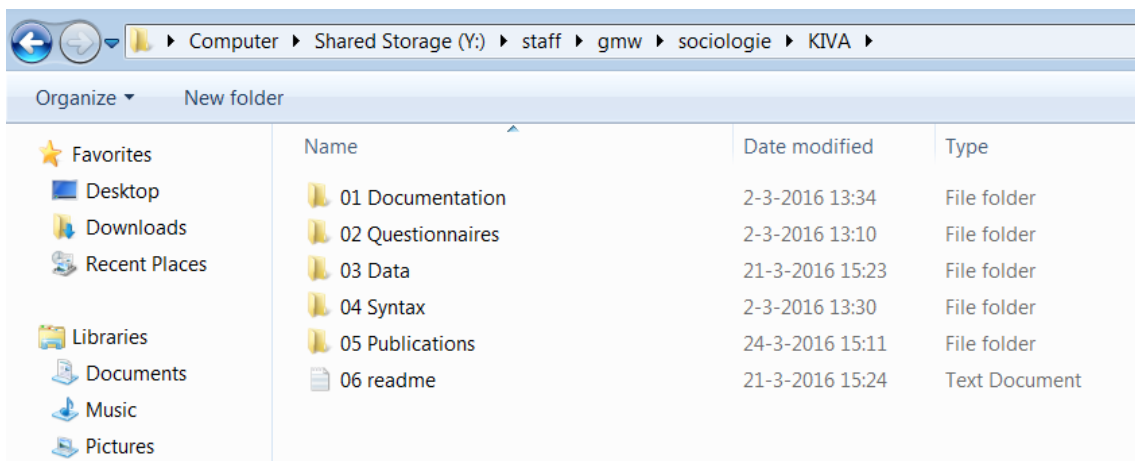
The package should be ready within three months after the research project is completed. If owing to unforeseen circumstances data collection has not yet ended by that date contact the coordinator, the deadline may be extended. If principal investigators fail to make a research package, they can be sanctioned (e.g., reduction of research time).

6. Example of a research package

The research package of the KiVa project is used as an example. The KiVa package can be found in the following location:

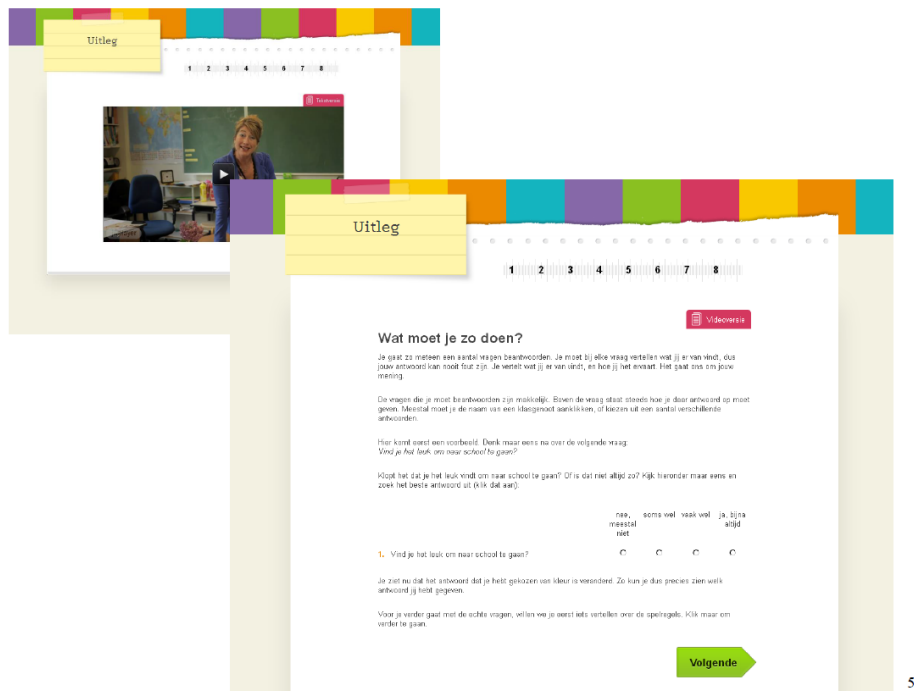


The package contains the following folders:



- 1. Documentation:** We have combined the KiVa research data management plan and the KiVa codebook in one document so that anyone who was not directly involved in the KiVa project can find all necessary information in one document. The documentation contains among other things information on how the data are handled, who owns the data, a description of the contribution made by each researcher, a description of the sample, and information on the variables. Please contact b.oldenburg@rug.nl or d.r.veenstra@rug.nl if you would like to use this documentation as an example.

- 2. Questionnaires:** This folder contains screenshots of the online questionnaires as presented to the respondents.



- 3. Data:** This folder contains the cleaned data. The raw data could due to privacy reasons not be included in the package, but we describe where these data can be found.
- 4. Syntax:** This folder contains all syntax that are used to clean the different waves.
- 5. Publications:** We included the KiVa publications in this package.

The screenshot shows a Windows File Explorer window with the following table of files:

Name	Date modified	Type	Size
Overview	21-3-2016 16:39	Microsoft Word D...	14 KB
OldenburgBosman2016	12-2-2016 10:42	Adobe Acrobat D...	124 KB
vdPloegSteglichSalmivalliVeenstra	3-11-2015 10:14	Adobe Acrobat D...	258 KB
OldenburgVanDuijn2015	15-10-2015 11:37	Adobe Acrobat D...	343 KB

- 6. Readme file:** A file explaining which files are in the package.

Appendix 1. Checklist publication packages

CHECKLIST PUBLICATION PACKAGES

Name:	P-number:	Package:	Date:
-------	-----------	----------	-------

<p>General</p> <ul style="list-style-type: none"> <input type="checkbox"/> Co-authors/supervisors access <input type="checkbox"/> Folder in right place <input type="checkbox"/> Name clear <input type="checkbox"/> Clear structure <input type="checkbox"/> Read me file <input type="checkbox"/> Publication (+ publication in Pure)
<p>Understanding study</p> <ul style="list-style-type: none"> <input type="checkbox"/> Description of how and by whom the data were collected or created <input type="checkbox"/> Questionnaire <input type="checkbox"/> Codebook <input type="checkbox"/> Sample/participants description <input type="checkbox"/> Ethics approval <input type="checkbox"/> Contributions per author
<p>Replicating results</p> <ul style="list-style-type: none"> <input type="checkbox"/> Digital data (raw) <input type="checkbox"/> Digital data (analyzed) <input type="checkbox"/> Syntax cleaning <input type="checkbox"/> Syntax analyses
<p>Future data</p> <ul style="list-style-type: none"> <input type="checkbox"/> Questionnaires in GMW archive <input type="checkbox"/> Digital data archived

Appendix 2. Sustainable data formats

Text processing

Preferred formats

- pdf/A
- Open Document Format: .ODT

Accepted formats

- RTF
- Postscript
- Word

Spreadsheets

Preferred formats

- Open Document Format: .ODS
- pdf/A

Accepted formats

- MS Excel
- Lotus 1-2-3
- Quattro Pro

Data files

Preferred formats

- SPSS portable files
- SAS transport files
- STATA transport files

Pictures

Preferred formats

- Tiff
- JPG (100% quality)

Databases

Preferred format

- ANSI SQL

Accepted formats

- MS Access 95 and higher
- dBase III+ en hoger
- DataPerfect
- Filemaker Pro 3.0 and higher
- Paradox

Plain text

Preferred format

- UNICODE with Byte Order Mark (UTF-8, UTF-16 of UTF-32)

Accepted formats

- ASCII (7 bit)
- ISO 8859 character sets
- MS-DOS codepages
- MS-Windows codepages
- Apple codepages

Mark-up

Accepted formats

- XML (including XHTML) with DTD or scheme
- SGML (including HTML) with DTD

Audio / video

Preferred formats

- Audio: MP3 (256 kbps), WAV, AIFF
- Video: MPEG-2, MPEG-4 AVC/H264, lossless AVI, QuickTime DV

University of Groningen
Department of Sociology

Director of Research:
prof. dr. René Veenstra
Phone: +31 5036240
Email: d.r.veenstra@rug.nl

Visiting Address:
Grote Rozenstraat 31
9712 TG GRONINGEN

Mail address:
Grote Kruisstraat 2/1
9712 TS GRONINGEN