Leiden University Data Management Plan[i] - FWN example[ii]

Please contact Leiden Research Data Office if you need help: datamanagement@library.leidenuniv.nl or http://www.library.leiden.edu/education-research/library-research/researchdata/support.html

Name and contact details	Name Surname
Name of project and group	Novel Targets to kill Drug Resistant Cells
Description of your research	This box provides a short description of what the research aims for.
Project duration	Start: *01-xx-2013* End: *01-xx-2017*
Names of people and their responsibilities for data management	PhD1: collection, storage, documentation, archiving PhD2: collection, storage, documentation, archiving Supervisor: storage, archiving External party: collection, storage, documentation, archiving
Funding body(ies)	European Research Council (ERC)
Grant number	Project reference: XX
Partner organisations	External party is a Dutch Medical Centre

About this Data Management Plan

Date written	02-04-2015
Date last update	02-04-2015
Version	1.0

Changes in this version of the Data Management Plan

Component	**Progress / Execution** *Please describe shortly what progress you have made, any questions or issues you have encountered and want to discuss, etc.*
1. Data collection	……
2. Data storage and back-up	……
3. Data documentation	……
4. Data access, sharing and reuse	……
5. Data preservation and archiving	……

1. Data collection

Describing the data you will be creating/collecting

1.1 **Will the project use existing or third party data ?**

- ☐ No
- ☐ Own / group previous research
- ☐ Academic collaborators
- ☐ Commercial collaborators
- ☐ Publicly available database / archive
- ☐ Specialist commercial data provider
- X Other: external party

Describe shortly provenance, type and format of this data. Are there any restrictions or requirements for use of third party data such as licensing conditions?

Genomic expression data (Microsoft Excel Spreadsheet) from External Party. Consent granted as part of joint project.

1.2 **What type()s of data will you collect or create, in what file format(s)?**[iii]

Note that not all formats are long-lived. For sustainable access you best use the formats recommended by data archives, see for examples: http://datacentrum.3tu.nl/fileadmin/editor_upload/File_formats/Preferred_formats.pdf or http://www.dans.knaw.nl/sites/default/files/file/EASY/DANS%20preferred%20formats%20UK%20DEF.pdf

Measurements: Western Blot Data (Jpeg), SRB Assay Data from FluoStarOptima (Excel), Cell Images (Tif/Jpeg). Necessary to generate new data in order to write thesis/publish articles.

1.3 **How will you collect and/or create your data?**

Please describe shortly. Name any relevant protocols and/or standard in your area of expertise.

Data is collected from the machines/equipment that generate the data (e.g. FluoStarOptima for SRB read outs or Las4000 for western blots) and subsequently stored on a portable USB drive before being backed up to an external hard-drive and common drive on university network storage.

1.4 **What tools, instruments, equipment, hardware or software will you use to capture, produce, collect or create the data?**

Please give the names of the tools and state if they are already available. If not, state how you intend to acquire them. If applicable, describe whether you use a paper or electronic labjournal.

Collect: FluoStar Optima (SRB), Las4000 (WB), Nikon Microscopes (Cell Images) (All Available)
Create: FluoStar Optima (SRB), Las4000 (WB), Nikon Microscopes (Cell Images) (All Available)
Process: Microsoft Excel, ImageJ (Fiji), Image Pro Microsoft Powerpoint (All Available)
View: Microsoft Excel, Image J (Fiji) (All available)
Analyse: Microsoft Excel, ImageJ (Fiji), Microsoft Powerpoint (All Available)

1.5 **What is the estimated size of the data?**

Please describe shortly. Stages to be adopted if relevant.

Data stage	Specification of type of research data	Software choice and file format	Data size now	Data size when project is finished
Raw data	Data directly obtained from experiments i.e. SRB data from FluoStar Optima, Gel Images from Las4000, Cell Images from Microscopes	*Software: Excel/ImageJ/Image Pro File Format: xls, xlsx, gel, tif*	*<1.49 GB*	
Processed data	Gel images processed in Image J, SRB data analysed in Excel, Nuclei counts generated by Image Pro attained from cell images.	*Software: Excel/ImageJ/Image Pro File Format: xls, xlsx, gel, tif, jpeg*	*<1.49 GB*	
Results	Summarised data in the form of powerpoint slides, excel sheets.	*Software: Powerpoint File Format:*	*<1.49 GB*	

			ppt/pptx		
	Other…				

2. Data storage and security Ensuring that all research data are stored securely and backed up or copied regularly during your research	
2.1	**Where will you store your data?** *Please describe how safe storage is guaranteed. Specify your method if your data is collected and / or transported in different locations / countries.*
	X On university departmental network storage (J:) ☐ On university personal network storage (P:) ☐ In a Virtual Research Environment (Sharepoint) X Physical storage (e.g. USB, external hard drive) ☐ Cloud service (e.g. SURFdrive) ☐ Other, namely: …
	Data is stored on a portable USB drive, university network storage and an external hard drive. The portable USB drive is used to collect all data and transfer this to the network storage and external hard drive. All of these are in different locations, thus minimising the chance for loss of all data at any given time, since backup is possible from a second/third location.
2.2	**Will your data be backed up?** *Please specify shortly for each storage device frequency, location of backups and who is responsible.* *Describe how you can restore your data in the event of data loss and who is responsible.*
	Yes, there will be regular back-upping. External Hard-disk backup > Weekly J-drive (University Network Storage) > Weekly Portable USB > Weekly
2.3	**Are there any commercialisation, ethical or confidentiality restrictions about handling your data?** *Please specify shortly.*
	Contractual obligations ☐Requirements by law : protection of personal data (e.g. privacy law) : specify in 4.1 ☐Requirements by law : copyright, intellectual property : specify in 4.1 ☐Ethical restrictions (e.g. ethical review) : specify in 4.1 ☐ Commercial considerations (e.g. patentability) ☐ Formal security standards ☐ No requirements X Other, namely: funder requirements
	"The European Research Council expects its funded researchers to retain files of all research data used during the course of their work, and that they be prepared to share this data with other researchers whenever it is not bound by copyright restrictions, by confidentiality agreements, or by contractual clauses." http://www.bath.ac.uk/research/data/policy/funder-data-policies.html
2.4	**How will access to the data be managed during the project?** *Please specify for each storage device, from different locations / countries.*
	The European Research Council (ERC) has authority to grant access, additional or otherwise, to my data. Also, see question 4.2.
2.5	**What are the main risks to data security?** *Please list risks, e.g. accidental deletion, falling into the wrong hands.* *Please describe what would happen if the data get lost or become unusable.*
	Accidental deletion, loss of USB/external hard disks, theft. If data became unusable or got lost, then this would prove highly detrimental to future analysis, experimental planning and the publication of

	generated data.
2.6	**What measures do you take to comply with the security requirements and to mitigate the risks?** *Describe how you can restore your data in the event of data loss and who is responsible.* *If applicable, please describe procedures to ensure personal data are handled confidentially and who is responsible.*
	☐ Access restrictions ☐ Encryptions ☐ Data processing ☐ De-identification / Anonymization X Regular back-ups ☐Master copy stored on university network storage ☐Master copy stored elsewhere ☐ Other, namely: …
	External Hard-disk backup > Weekly J-drive (University Network Storage) > Weekly Portable USB > Weekly
2.7	**How do you differentiate between raw and processed data?** *Please explain shortly why you (do not) differentiate.*
	☐ I will not differentiate ☐ I will create a new file for processed data ☐ I will create a new file for processed data and I will lock raw data ☐ Other, namely: …
	Differentiation is dependent upon the type of data. For example, for western blot data, processed data is named differently. "IMAGE" is placed at the end of the file name to indicate processing and thus differentiate from raw image data for blots. However, for SRB assay results/BCA assay results, the raw data is clearly labelled within the same excel file; i.e. processing is done in a separate excel sheet within the same file.
2.8	**Is there any non-digital data or outputs that the project will generate? Where will these outputs be stored?** *Please specify shortly **and describe who is responsible for storage of these outputs**.*
	Non-digital date comprises paper lab books which are stored on campus. Books are placed in a locker to which I have a key. All data discussed in the lab book is given more or less the same name as the file in question when written in the lab book, thus all data in the lab book/electronic storage can be traced and easily referred to.
2.9	**Do you expect to have any supplementary costs for storage not covered by the project budget?** *Please specify*
	.....

3. Data documentation Documenting your data to help future users to understand and reuse it	
3.1	**How will files be named?** *Please describe shortly.*
	Files will be named as follows: "Experiment Name/Number--Date(Day/Month/Year)--Author (if appropriate)"
3.2	**How will folders be named and structured?** *You are invited to draw a folder structure and describe it shortly.*
	Data (Folder) Type of Experiment (e.g. IC50 Assays; Folders) Experiment Name & Date (Folders or Files) Experiment Name/Date/Attempt (Files)

	Data Summary (Final results for experiment summarised in ppt; Files)
3.3	**How do you handle version control to maintain all changes that are made to the data?** *Please explain your choice shortly. Remember to also document any deletion of data, if applicable.*
	☐ No version control (e.g. original files are overwritten) ☐ Version control software, namely: … X Data/version number in filename/folder X 'Track changes' feature in software ☐ By saving the script with which I process my data ☐ Other, namely: …
	For Microsoft word documents, the track changes feature is used to maintain changes made to documents. For data analysed in Excel, a new spreadsheet is created within the same file to preserve the original data and allow different methods of analysis/versions to be tracked. Image data analysed using Image J is saved with the suffix "image" after editing, whilst the original image file is left untouched.
3.4	**What metadata standard will be used, if any?**[iv] *Please explain why you use this standard (most used in my discipline, required by the data archive where I will deposit my data). Please outline how the metadata will be created (read me file, spreadsheet, in the data). If no standard exist, please specify which metadata is needed to understand the data.*
	☐No metadata standard is used ☐ Generic metadata standard (e.g. Dublin Core) X Standard automatic Windows metadata (e.g. from Word, Excel) ☐ Specialised metadata standard, namely: … X Other metadata standard, namely: …
	Standard automatic Windows metadata from Word/Excel/Powerpoint/Image J will be used. Information on data purpose/means of creation/time + date/authorship will additionally be listed in the file name (e.g. Knockdown experiment II-Western Blots-23/1/15-XX)
3.5	**What supporting information / documentation will you create to enhance understanding of the data ?** *Please describe shortly how peers should be able to understand the data. Examples are a readme.txt, lab journals, a codebook, survey questions etc. Is there a standard for documentation in your field? Describe at what moment in your research process you will add the documentation necessary to make sure the data is understandable for peers.*
	All data discussed in the lab book is given more or less the same name as the digital file in question when written in the lab book, thus all data in the lab book/electronic storage can be traced and easily referred to. All data stored electronically is eventually dated, discussed and described in a paper lab journal.

4. Data access, sharing and reuse Managing access and security, sharing your data	
4.1	**Are there any restrictions placed on sharing / reuse of some / all of your data?** *Please account for not sharing your data. Reasons may be ethical, commercial, security-related, protection of personal data rules, intellectual property, copyright,*
	"The European Research Council expects its funded researchers to retain files of all research data used during the course of their work, and that they be prepared to share this data with other researchers whenever it is not bound by copyright restrictions, by confidentiality agreements, or by contractual clauses." http://www.bath.ac.uk/research/data/policy/funder-data-policies.html
4.2	**With whom will you share your data at which stage in your research? You can use the table below.** *Please state any sharing requirements, e.g. funder data sharing policy. Please describe shortly how you will share your data: on request, pro-actively, etc.. Please specify how your data can be accessed.*

	Would not share with anyone	Would share with my immediate collaborators	Would share with others in my research centre or at my institution	Would share with scientists in my field	Would share with scientists outside of my field	Would share with anyone
Immediately after the data has been generated	N/A	Yes-for advice or guidance	Yes-for advice or guidance	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP
After the data has been normalized and/or corrected for errors	N/A	Yes-for advice or guidance		Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP
After the data has been processed for analysis	N/A	Yes-for advice or guidance	Yes-for advice or guidance	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP
After the data has been analysed	N/A	Yes-for advice or guidance	Yes-for advice or guidance	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP	Data is not in appropriate form for scientific publication. Not appropriate in terms of privacy/IP
Immediately before publication	N/A	Yes-for advice or guidance	Yes-for advice or guidance	Yes; for article/publication review purposes prior to publication.	Data is not yet publically available. Not appropriate in terms of privacy/IP	Data is not yet publically available. Not appropriate in terms of privacy/IP
Immediately after the findings derived from this data have been published	N/A	Yes-data are already publically available/published	Yes-data are already publically available/published	Yes-if data are already publically available/published	Yes-if data are already publically available/published	Yes-if data are already publically available/published

Based on: Interview worksheet, Jake Carlson, Purdue University Libraries / Distributed Data Curation Center

……

4.3 **If intending to share any part of the data, do your participant consent forms include information about intentions for sharing, retention of data and steps taken to protect participants privacy and confidentiality?**

X Not applicable.

☐ Yes. *Please specify the relevant formula in the consent form.*

……

4.4 **Who has authority to grant (additional) access to your data?**

Please describe shortly.

☐ Only you
☐ A colleague from the project, namely: …
☐ Supervisor
X Funder
☐ Collaborator / research partner organisation
☐ Other, namely: …

The European Research Council (ERC) has authority to grant access, additional or otherwise, to my data.

4.5	**How will you manage copyright and Intellectual Property Rights issues?** *Who owns the data? How will the data be licensed for reuse? Please describe shortly your choices and their consequences.*
	The University of Leiden owns IP rights to data produced in this project.
4.6	**What is the audience for reuse?** *Please list possible audiences and purposes. Consider who might use it now and who might use it later.*
	Future Postdocs/PhD Students/Master students

5. Data preservation and archiving Preserving your data	
5.1	**Which criteria will you use to decide which data has to be archived?** *Please shortly describe your choices.*
	☐ Type of data (raw, processed) and how easy it is to reproduce it ☐ Relevance of content for others ☐ Usability of format for others ☐Data underlying publications ☐Verification of research ☐Available time ☐Available money ☐Other, namely: …
	According to ERC regulations, all data needs to be made available during the project, and I assume this is also the case once the project is finished. Therefore, all data (raw and processed) should be archived. The purpose of this is for possible research verification and verification of data underlying publications.
5.2	**How long should your data be preserved? Are there any requirements regarding the disposal of data?** *State obligations you have by law, funder, university, etc. if any.* *Describe how you will dispose of the data, e.g. how you will get approval, what people and/or tools you need, etc.*
	"The European Research Council expects its funded researchers to retain files of all research data used during the course of their work, and that they be prepared to share this data with other researchers whenever it is not bound by copyright restrictions, by confidentiality agreements, or by contractual clauses." http://www.bath.ac.uk/research/data/policy/funder-data-policies.html No specific time limit stated, but data should be made available at least during course of project/foreseeable future.
5.3	**Which data repository is appropriate for archiving your data?** *Please describe shortly. Does this archive have a 'data seal of approval' or another form of certification?*
	☐ Discipline specific (international) repository, namely ... ☐ 3TU.Datacentrum ☐ SurfSara ☐ DANS ☐ Other (international) repository, namely : X Other, namely: Europe PubMed Central
	"Europe PubMed Central" http://erc.europa.eu/sites/default/files/document/file/ERC_Open_Access_Guidelines-revised_2013.pdf
5.4	**Does the archive have specific requirements concerning file formats, metadata etc.** *Provide relevant urls to the documentation on these requirements. Describe how you intend to meet those requirements, e.g. converting the file formats, providing supplementary documentation.* *Will there be extra costs to prepare your data for archiving? Please specify. See http://www.data-archive.ac.uk/media/247429/costingtool.pdf*
	……
5.5	**What costs (if any) will your selected repository charge? Who pays?** *Please state the costs in euro's and the institution that pays for it.*
	Currently unknown. Universiteit Leiden will be responsible for payment should this be necessary.
5.6	**Who is responsible for the data after the project ends?** *Please state a position and the current person in that position.*

	Supervisor.

[i] This template is based on the 3TU data management plan, the University of Bath data management plan and the Data Management Checklist of the University of Western Sydney.

[ii] This example is based on a plan written R.P. McLaughlin MSc, as part of the data management training for PhD's from the LACDR, Autumn 2014. We thank Ronan for his kind permission to copy the information from his plan into this example.

[iii] Data types can be : documents (text, MS Word), spreadsheets, field notebooks, diaries, questionnaires, transcripts, surveys, codebooks, audiotapes, videotapes, photographs, (transcribed) test responses, models, algorithms, measurements, simulations, observations, software source code, computational model output, etc.
Think of the different stages (for instance : video recording, transcript, annotation, lists of typological features ....).

[iv] See http://www.dcc.ac.uk/resources/metadata-standards or http://en.wikipedia.org/wiki/Metadata_standards or the relevant repository.