

Data management plan of the Division of Toxicology, Wageningen University

1. Purpose

The aim of this data management plan is to set guidelines on the storage and traceability of data with the ultimate goal to allow reviewers and authorized colleagues to trace back the data from raw data to the published forms. The guidelines within this data management plan apply to all data produced within projects performed at the division of Toxicology. This includes data derived within BSc and MSc thesis projects, PhD and postdoc projects as well as other research projects.

2. Data management roles

All students and personnel at the division of toxicology are responsible for adequate storage of the raw and processed data that they produce according to the protocols provided in Appendix I to III. These appendices indicate the guidelines the storage of data obtained with tissue culture (appendix I), analytical methods (appendix 2), and computer models (appendix 3). To facilitate the traceability, metadata files that contain information on where the data underlying a publication or thesis can be found, need to be created and stored as highlighted in section 5. These metadata files should be send together with the proofs of accepted papers or with the final version of a report or thesis to the supervisors. The role of the supervisors is to evaluate the quality of the raw and processed data during regular meetings as well as the adequacy of metadata files. Technicians are responsible for the backup of storage of all data files, including the raw, processed and metadata files, beyond the timespan of the projects.

3. Sharing and ownership

The data created at the division of Toxicology within ongoing research projects will be owned by the division of Toxicology. Sharing of data will be arranged on a case-by-case basis, requiring consent of the head of the department, prof. I.M.C.M Ivonne Rietjens. Sharing and storage of experimental data obtained at other facilities (e.g. other research groups at the University, RIKILT, or companies) will be performed according to the specified agreements in the project proposal. When these data obtained at other facilities are linked to scientific papers or reports of PhD students or Postdocs from the division of Toxicology, long term storage should preferably be at the division of Toxicology. If this is not possible, the metafile provided with each publication should indicate where these data are stored at the other facility.

4. Short and long-term storage of data

Appendix I-III provide the guidelines on how and where to store the raw and processed data related to different experiments at the division of Toxicology. Depending on the type of data, short-term storage will be on the desktop computers of individual researchers. Data that should be provided for long-term storage include all raw data files and processed data files underlying specific publications (i.e. scientific publications, (project) reports, or MSc and Bsc theses). Long term storage of all data will be at least 5 years beyond the time of publication.

5. Documentation and metadata

During the research documentation of the data will be recorded in paper lab journals. Apart from the experimental procedures and conditions applied, a references should be made in the lab journals to the electronic files of the raw and processed data files. Lab journals will be kept at the documentary room for 5 years beyond the end of each project.

In addition, metadata files will be created with each publication and thesis to facilitate the traceability of the raw and processed data files that are linked to the publications. For scientific publications and intermediate reports for PhD and postdoc projects, these metadata files are prepared in Microsoft Word and should include: your name, date of paper acceptance (if applicable), full reference of accepted publication (if applicable), dates of data collection for each study, names of the raw and

processed data files and locations where they can be found within the long term storage, name of person who collected data for each study.

For BSc and MSc thesis projects similar Word files need to be created that include: your name, title of thesis, supervisor of thesis, dates of data collection for each study, names of the raw and processed data files and locations where they can be found within the long term storage. The metadata files are stored on [N:\Data Management](#) according to appendix IV and should be send together with the proofs of accepted papers or with the final of a report to the supervisors to be checked. BSc and MSc students should send this metadata file with the final version of the thesis to the daily supervisor. All metafiles are kept for short term on N, but are converted every week to a hard disk and backup hard disk.

6. Data management plan for PhD/postdoc projects

In the initial research plan of a PhD candidate or postdocs, it should be outlined how data are collected and stored, and if there are plans to share the data with others by filling out the table in Appendix V. An initial version of the data management should be discussed with the project supervisors and the plan needs to be submitted with the research plan to the graduate school. However, the plan should also be maintained in time to reflect changes in procedures while the research is ongoing. Within the project-specific data management plan references can be made to group procedures leaving only project-specific issues to be addressed.

Appendix I. Guidelines storage data obtained with tissue culture

Storage name	System	Data Output (program)	Data extension	Experiment Code	Storage			Size
					Time span	Who	Where	
RAW : YY-MM-DD_R_INI_X*	Luminometer	EF****	xls	L	short	researcher	personal desktop	10 Mb
	Spectrofotometer	EF	xls	S				20 Mb
	Fluorimeter	EF	xls	F	middle	researcher	N: Scratch/Backup**	20 Mb
	PamStation 12	Image Folder	TIF	PS12				30 Gb
	RotorgeneQ	RQ software	rex	RQ	long/back up daily	lab manager	ext. hard disk room 1022***	30 Mb
Processed: YY-MM-DD_P_INI_X	Luminometer	Excel	xls	L	short	researcher	personal desktop	2 Gb
	Spectrofotometer	Excel	xls	S				2 Gb
	Fluorimeter	Excel	xls	F	middle	researcher	N: Scratch/Backup**	2 Gb
File should contain a link to the raw data file(s)	PamStation 12	BioNav	Bn5	PS12				long back up daily
	RotorgeneQ	Prism/word	Pzf/doc	RQ	2 Gb			
Presented: YY-MM-DD_PR_INI_X	Luminometer	Prism	pzf	L	short	researcher	personal desktop	2 Gb
	Spectrofotometer	Prism	pzf	S				2 Gb
	Fluorimeter	Prism	xls	F	middle	researcher	N: Scratch/Backup**	2 Gb
File should contain a link to processed file and publication	PamStation 12	BioNav	Bn5	PS12				long/back up after acceptance
	RotorgeneQ	Prism/word	Pzf/doc	RQ	2 Gb			

*YYYY-MM-DD: Year-month-day; R: Raw data; P: Processed data; C: Computer Model; M, Meta Data; PR: presented data; INI: unique wur-name (haan003); X: Experiment code

** Data files that are used for a publication (i.e. scientific publication, project report, BSc and MSc theses), should be provided together with a metadata file on N:Scratch/Backup when proofs are obtained for a publication or with the final version of a report

*** Data on N: Scratch/Backup will be automatically converted to an external hard disk every week. This hard disk will be maintained and backed up by Laura de Haan

****: Excel formatted

Appendix II. Guidelines storage data obtained with analytical methods

Storage name	System	Data Output (program)	Data extension	Experiment Code	Storage			Size
					Time span	Who	Where	
RAW : YY-MM-DD_R_INI_X*	Waters prep HPLC (lab 0019)	Empower		H	short	automatic	HPLC server	10 Gb
	Waters 600/717 (lab 0019)	Empower		H				20 Gb
	Waters Alliance (lab 1018)	Empower		H	middle	automatic	HPLC server	20 Gb
	Waters UPLC (lab 1018)	Empower		U				30 Gb
	Waters UPLC H-class (lab 1018)	Empower		U	long/back up daily	lab manager	ext. hard disk room 1022***	30 Gb
	LC-MS			MS				
Processed: YY-MM-DD_P_INI_X File should contain a link to the raw data file(s)	Integrated peaks (all HPLC, UPLC and LC-MS systems)	Excel	xls	H, U, or MS	short	researcher	personal desktop	2 Gb
					middle	researcher	N: Scratch/Backup**	2 Gb
					long back up daily	lab manager	ext. hard disk room 1022***	2 Gb
Presented: YY-MM-DD_PR_INI_X File should contain a link to processed file and publication	e.g. Michaelis-Menten curves or other fitted data (all HPLC, UPLC and LC-MS systems)	Prism	pzf	H, U, or MS	short	researcher	personal desktop	2 Gb
					middle	researcher	N: Scratch/Backup**	2 Gb
					long/back up after acceptance	lab manager	ext. hard disk room 1022***	2 Gb

*YYYY-MM-DD: Year-month-day; R: Raw data; P: Processed data; C: Computer Model; M, Meta Data; PR: presented data; INI: unique wur-name (haan003); X: Experiment code

** Data files that are used for a publication (i.e. scientific publication, project report, BSc and MSc theses), should be provided together with a metadata file on N:Scratch/Backup when proofs are obtained for a publication or with the final version of a report

*** Data on N: Scratch/Backup will be automatically converted to an external hard disk every week. This hard disk will be maintained and backed up by Laura de Haan.

Appendix III. Guidelines storage computer model files

Storage name	System	Data Output (program)	Data extension	Experiment Code	Storage			Size
					Time span	Who	Where	
RAW : YY-MM-DD_C_INI_X *	PBK/D models	Barkley Madonna	mmd	PBKD	short	researcher	personal desktop	250 kb
					middle	researcher	N: Scratch/Backup**	250 kb
					long/back up daily	lab manager	ext. hard disk room*** 1022	250 kb
Presented: YY-MM-DD_PR_INI_X	PBK/D models	Prism	pzf	PBKD	short	researcher	personal desktop	2 Gb
					middle	researcher	N: Scratch/Backup**	2 Gb
					long/back up after acceptance	lab manager	ext. hard disk room 1022***	2 Gb
File should contain a link to model file and publication		Excel	xls	PBKD	middle	researcher	N: Scratch/Backup**	2 Gb
					long/back up after acceptance	lab manager	ext. hard disk room 1022***	2 Gb

*YYYY-MM-DD: Year-month-day; R: Raw data; P: Processed data; C: Computer Model; M, Meta Data; PR: presented data; INI: unique wur-name (haan003); X: Experiment code

** Data files that are used for a publication (i.e. scientific publication, project report, BSc and MSc theses), should be provided together with a metadata file on N:Scratch/Backup when proofs are obtained for a publication or with the final version of a report

*** Data on N: Scratch/Backup will be automatically converted to an external Hard disk every week. This hard disk will be maintained and backed up by Laura de Haan.

Appendix IV. Guidelines for storage of meta data files

Storage name	Type of publication	Data Output (program)	Data extension	Identifier	Storage			Size
					Time span	Who	Where	
RAW : YY-MM-DD_M_INI_X*	Scientific publication	Word	docx	DOI number publication	short	researcher	personal desktop	250 kb
	PhD thesis	Word	docx	ISBN number thesis	middle	researcher	N: Scratch/Backup**	250 kb
	Project reports	Word	docx	REP	long/back up daily	lab manager	ext. hard disk room*** 1022	250 kb
	MSc thesis report	Word	docx	MSc				
	BSc thesis report	Word	docx	BSc				

*YYYY-MM-DD: Year-month-day; R: Raw data; P: Processed data; C: Computer Model; M, Meta Data; PR: presented data; INI: unique wur-name (haan003); X: Experiment code

** Data files that are used for a publication (i.e. scientific publication, project report, BSc and MSc theses), should be provided together with a metadata file on N:Scratch/Backup when proofs are obtained for a publication or with the final version of a report

*** Data on N: Scratch/Backup will be automatically converted to an external Hard disk every week. This hard disk will be maintained and backed up by Laura de Haan.

Appendix V. WGS format for a Data Management Plan for research projects and guidance for implementation at the Division of Toxicology, Wageningen University

Heading	Guidance	Guidance for implementation at the Division of Toxicology
Organizational context	<ul style="list-style-type: none"> • Name • Supervisor 	
Short description of your research	Give two or three lines to explain what is not obvious from the project title	
Data management rules	<p>[Project specific rules]</p> <ul style="list-style-type: none"> • Who has control over the data (PhD or post-doc him/herself as a researcher, or are there other colleagues involved), what is the role of your supervisor? • Is there a person in the research group with a specific responsibility for data management? • Who is responsible for your data after you have left? 	Describe according to the “Data management plan of the Division of Toxicology, Wageningen University”, provided above.
What type of research data will be produced?	<p>Choose from:</p> <ul style="list-style-type: none"> • Raw data; (e.g. data from experiments or observations) • Derived / processed data • Models (including data from simulations) • Documentation of the research process, such as lab notebooks 	Describe according to appendix I to III.
Software choices (*)	<ul style="list-style-type: none"> • What software will you use to create, analyse and visualize your data? • Are these choices common practice in your field? 	Describe according to appendix I-III.
What is the amount of the data, and how will the amount increase in time.	Give an estimate in (mega – giga – tera) Bytes	Describe according to appendix I-III
Sharing and ownership (*)	<ul style="list-style-type: none"> • Do you expect that others may be interested to re-use you data, and do you have plans to share it with them? • Are there specific funder’s requirements to share you data, or to impose an embargo? • If other parties (outside your group or outside Wageningen UR) are involved in this research, are there agreements how the data will be used and shared? • Are there privacy or security issues, and if there are, how are you dealing with them? 	Describe according to the “Data management plan of the Division of Toxicology, Wageningen University”, provided above.

Documentation and metadata (*)	<p>How will your data be documented</p> <ul style="list-style-type: none"> • During your research • For long term storage • For data sharing 	Describe according to appendix IV
Short term storage (*)	<ul style="list-style-type: none"> • In what format will you store your data? • Where will the data be stored physically and how will it be backed up? • Do you follow the common practice in your research group, and if not, why? • What system for directory- and file names, and for version control do you intend to use? • Where do you document what the different files are, and how they were created or processed (paper or electronic lab notebook)? 	Describe according to appendix I-III
Long term storage	<p>Do you plan to store your data for the long-term after the conclusion of your research? If not, argue why. Is there a common practice in your field or do you intend to use the services provided by Wageningen UR¹?</p> <p>Guidance: Long term storage is at least recommended if</p> <ul style="list-style-type: none"> • data underlies publications • long term storage and sharing is required by funders, or if there is a legal requirement • it is likely that other may want to re-use the data 	Describe according to appendix I-III