CoBE LAB DATA MANAGEMENT GUIDELINES

Contents

V4.1 14/02/2021

Version	Summary of changes	Person responsible	Last amended
2.0	Second draft based on UK Data Archive's best practices, but tailored as much as possible to the needs of the CCAL research group at University of Exeter.	Myriam Mertens & Frederick Verbruggen	02/12/2012

Version	Summary of changes	Person responsible	Last amended
2.1	Incorporation of suggestions Open Access Team Exeter and Marine Renewable Energy Group. Omitted centralised CCAL back-up system until more resources have been identified.	Myriam Mertens & Frederick Verbruggen	21/01/2013
2.2	Correction of spelling errors and sources added. Comment about R added in section 2.1.3.	Myriam Mertens	17/07/2013
3.0	Scope of guidelines reduced to the Verbruggen lab. Revision and extension of existing guidelines (roles & responsibilities; data storage, documentation and formatting). Addition of guidelines on data sharing & archiving.	Myriam Mertens & Frederick Verbruggen	7/1/2015
4.0	New version of the guidelines tailored as much as possible to the needs of the new research group at Ghent University. Also major revision to ensure that data collection and processing will be carried according to Regulation (EU) 2016/679 and UGent's policies. Update of appendices and links to individual documents.	Frederick Verbruggen	05/11/2018
4.1	Update of the new lab name.	Frederick Verbruggen	14/02/2021

Sources: the text and content of this document are primarily based on and adapted from:

- Guidelines provided by Ghent University:
  - specific guidelines of the Faculty of Psychology and Educational Sciences (FPPW), Ghent University.
  - General university guidelines and information
- The UK Data Archive's data management guidelines:
  - Brochure 'Managing and Sharing Data. Best Practice for Researchers', May 2011

  - UKDA website, 'Create & Manage Data'
  - L. Corti, V. Van den Eynden, L. Bishop and M. Woollard, *Managing and Sharing Research Data. A Guide to Good Practice* (Los Angeles, 2014)
- The Marine Renewable Energy Group's data management policy (I. Ashton, H. Lloyd-Jones and A. Cowley, 'Developing Research Data Management Policy at Research Group Level', July 2013.
- Guidelines provided by University of Exeter:
  - http://www.exeter.ac.uk/research/researchdatamanagement/
  - https://ore.exeter.ac.uk/repository/bitstream/handle/10871/9255/ChecklistforOREdeposit13/_05/_09.pdf?sequence=2
  - http://as.exeter.ac.uk/it/
  - http://www.exeter.ac.uk/recordsmanagement/
  - http://www.exeter.ac.uk/research/toolkit/sharing/ip/ippolicy/
- The University of Cambridge's Managing Research Data pages
- Stanford University Data Management Services
- British Psychological Society, 'Code of Human Research Ethics', 2nd edition, 2014
- British Psychological Society, 'Code of Ethics and Conduct (2018)"
- Suggestions made by the University of Exeter's Open Access and Data Curation Team and Records Manager
- Suggestions made by Jan Lammertyn (FPPW Research data management support)

Authors: Dr Myriam Mertens and Prof. Frederick Verbruggen

Main overview

Proper research data management is integral to good research practice: it ensures that the data generated by lab members are stored securely, will be reusable in the future, and can be shared easily amongst collaborators. Moreover, it is an increasingly important part of funder and institutional requirements regarding open access to research.

In this document, you will find guidelines detailing how to manage data, and assigning roles and responsibilities. They are largely based on UGent's and the UK Data Archive's best practices, but tailored as much as possible to the needs of the Cognition, Behavior, & Ecology (CoBE) lab.

---

1. Roles and responsibilities

Research data management is the joint responsibility of the Principal Investigator (PI) and the team members (research assistants, PhD students, and postdocs).

1.1. The PI of the lab is responsible for:

- Setting up new research projects. This involves:
  - Creating a data management plan using DMPonline.be
  - Discussing the data management procedures with new team members at the beginning of projects
  - Assisting team members with data documentation where necessary
  - Establishing project acronyms for file and folder names (for joint projects)
- Overseeing projects. This involves:
  - Securely storing paper forms of completed studies in the faculty archive for research data.
  - Depositing data that support publications on which he is lead author

  - Approving team members' data files before they deposit
  - Depositing remaining data sets and project-level documentation at the end of research grants
  - Overseeing data disposal on lab computers at the end of PhD and grant-funded research projects
- Reviewing data management guidelines on a regular basis and in response to emerging issues and changing institutional or funder policies
- Maintaining a data management resources library with further guidance, templates, and key policy documents relevant to the lab

1.2. Team members are responsible for:

- Creating a data management plan for their own individual projects (if their research is not covered yet by a data management plan created by the PI)
- Creating documentation files for their experiments
- Storing and backing up data and other experiment files in the right format
- Securely storing signed consent and other paper forms of ongoing studies
- Handing over signed consent and other paper forms of completed studies to the PI
- Providing the PI with the definitive/final versions of their experiments' data, software and documentation files
- Depositing data that support publications on which they are lead author

1.3. PhD students are in addition responsible for:

- Writing a data management plan; see FPPW's doctoral regulations and DMP instructions
- Depositing remaining data sets and project-level documentation at the end of their doctoral research project
- Upload a data fact sheet to the UGent Biblio archive (for more information, see: https://www.ugent.be/pp/en/research/rdm/data-storage-fact-sheet.htm)

---

2. Data documentation, formatting and storage

2.1. File structure and data documentation

For each experiment, team members should store data and other files in one folder. Within each experiment folder, there can be several subfolders (depending on the nature of the experiment):

- Data
- Code
- Documentation

The PI and team members can group together the folders of experiments that are part of their larger research projects (i.e. grant-funded and PhD projects) in a 'project' folder.

2.1.1. Data

The data folder of an experiment will have two subfolders:

- *Raw*: raw files collected during the experiment. No other files should be present in this folder.
- *Processed*: any data that have been subjected to automated or manual processing routines. It is important to document the processing routines in an analysis document or in R (see section 2.1.3).

2.1.2. Code

The code folder of an experiment will have two subfolders:

- *Experiment*: the code and software that was used to run the experiment
- *Analysis*: the code that was required to analyse the data

2.1.3. Documentation

Detailed documentation of the data collection and data analysis process helps other researchers understand your data and can provide further evidence of data quality.

Team members should create sufficient *experiment/study- and data-level documentation* for each of their experiments, so that others can reuse or reanalyse the data and replicate the results. The PI can assist with experiment and analysis documentation, for example by providing team members with templates of research project summaries or by sharing R analysis scripts.

The documentation folder of an experiment should always include:

- *Experiment Documentation*: a text document with experiment and data information [as outlined in **Appendix B**]
- *Analysis Documentation*: a text document explaining how the data were processed and analysed, and including the information listed in **Appendix B**. If you use R, you can create the analysis document within R (using RMarkdown) and include a summary of data analysis outcomes.

In addition, the documentation folder of an experiment *could* also include:

- *Lab notes*: any notes taken during the experiment that are relevant for understanding and interpreting the data
- *An 'access and use conditions' document*: may be required to specify under what particular (i.e. restricted) conditions data can be accessed and used by other researchers (see Section 3, 'Data sharing and archiving' for more details)
- *Blank consent forms*: with information/debriefing sheets
- *Other documents*: experiment instructions, blank questionnaires, and any other documents used in a study/experiment. When depositing data, research outputs such as presentations and publications can also be included as a form of documentation.

Together with their project folders, the PI and PhD students should also create *project-level documentation* giving more detailed information about their research projects as a whole and thus the broader context of data collection: history, aims, hypotheses, etc. This can take the form of research reports (e.g. ERC End of Award Reports) and PhD theses, so it shouldn't require much additional work.

2.2. File formats, folder and file naming, and version control

2.2.1. File formats

Choosing appropriate file formats ensures longer-lasting digital data. Appropriate file formats for long-term usability of data are typically standard and open, non-proprietary formats. Team members can use the formats and software most suitable for their analyses, but–when practically possible–should convert their files to open or standard formats before offering data to repositories. Open or standard formats should also be considered for backups.

Recommended file types:

- Documentation files: (R)markdown, txt, and/or pdf
- Behavioural data: where possible, create a txt or csv copy
- Questionnaire data: csv file with the scores/ratings

- Digital image data: TIFF, PDF, or PNG
- Digital audio data: Free Lossless Audio Codec (FLAC)

See also recommended formats by the UK Data service.

2.2.2. Folder and file naming

Experiment folders and files should be named appropriately and consistently.

Experiment folder and file names should include the acronyms of both the larger research project and the specific experiment, and file names should in addition include any relevant specification. You should be able to tell what a file contains on the basis of its name rather than its location on the computer.

For example, the file named 'CIA_PrepTMS_ExpDoc.md' contains the experiment documentation of the 'Preparation TMS' experiment in the 'Control of Impulsive Action' project. For more folder and file name examples, see **Appendix A**.

Other recommendations: - Avoid special characters and spaces in filenames (underscores are fine!) - To ensure proper sequencing of files, use YYMMDD for dates, and 001, 002, 003, etc. for number sequences. - Keep the file names as short as possible

2.2.3. Version control

Version control should be used when there is, or likely will be, more than one version of a file, for example because it is edited or stored in multiple locations.

Version control helps you keep track of changes, and identify and locate the *master file* (the original from which working copies are created), *milestone versions* (significantly changed versions to keep), or the most recent version of a file. In this way, you know the correct version to work with and can revert to an older one if necessary. This is especially important in collaborative research.

Appendix C lists some *version control* strategies that the PI and team members can adopt.

Important points to keep in mind:

- Team members should keep a *single master file of raw data* that should be left untouched. For sharing with collaborators, editing, analysis etc. they should use working copies instead, so that the original files are not affected.
- Team members should *never delete raw data files*, even when experiment subjects are excluded or something went wrong with data collection. The only acceptable exceptions are raw data generated when they run themselves through a few trials to check whether the computer program works, or to give a demo.

2.3. Short-term data storage

Consistently storing and backing up data keeps them safe and recoverable.

Team members should regularly check the completeness, accuracy and integrity of all stored files for the experiments they conduct (e.g. checking that raw data files are complete, that there are no duplications, that all required documentation files are present and accurate, that there have been no inadvertent file deletions or modifications...).

2.3.1. Different storage options

Uncompressed experiment files can be stored locally (e.g. on a team member's PC), on shared network drives, or in the cloud (e.g. commercial platforms, such as OneDrive, Dropbox, Google Drive, or iCloud, but also cloud-based management services, such as Open Science Framework Open Science Framework (OSF)).

Network storage on shares refers to storing your data on UGent servers connected to your local computer over the UGent network. This is particularly useful for confidential data (for more info, see here). Each team member has personal disk space, but the team also has a shared network drive 'pp02_labverbruggen'. The PI can give each team member with an UGent account access to this drive.

The main advantage of cloud storage is that, as an online storage service, it gives access to your files from any computer connected to the internet. Commercial services also allow synchronisation across your various devices: files will be automatically stored on the cloud servers as well as on your computer's hard drive (if you have installed the Dropbox desktop application there). However cloud storage is should be avoided for *personal, confidential* or *sensitive data*, as discussed below.

2.3.2. Storage for the CoBE team

The collaborative nature of the lab's research also calls for central storage of data, software and documentation.

As described in the lab's Open Science Policy document, the CoBE team uses a 'co-pilot' system, whereby members of the research team will be asked to check each other's code and rerun (parts of) the analysis. This requires sharing of code and data throughout the project.

GitHub should be used for code sharing. The main advantage of the platform is that it allows version control; this makes it most suitable to jointly work on experiment and analysis software. See the lab's Co-pilot & GitHub guidelines for detailed information and practical guidelines.

Anonymised and non-confidential data can also be shared via GitHub or OSF; for confidential data the PP02/UGent share (see above) should be used.

2.3.3. Backups

To prevent the loss of data and other files, everybody should back up their own computer systems (including all office and personal computers where experiment files are held) at regular intervals. The backup process can be automated by using backup software (for example Apple's Time Machine or Windows Backup).

Often *incremental backups* are used because they require less storage space and are quicker to perform; older backups are then deleted in favour of newer backup files once the disk is full. As it is best not to overwrite old backups with new ones, have your backup software notify you when the backup disk is full and old backups are deleted, and then select a new backup disk.

Full system backup files should be stored offline, on external hard drives (avoid USB sticks for backups). Removable storage media should be properly organised and labelled (indicating content and date) to facilitate restoration when needed. It is recommended that the PI and team members regularly verify their backup files by trying to restore them.

Note that all data and files stored on the UGent network are automatically included in various backup systems.

2.3.4. Data security

- **Computers**: To improve security, all office, lab and personal computers where experiment files are stored should be locked with a password and protected against viruses and malware. For more

information, see UGent's pages on Information Security: https://helpdesk.ugent.be/security/en/ and https://www.ugent.be/en/facilities/ict/information-security

- **Cloud storage**:
  - All team members using cloud storage are strongly encouraged to use two-step verification (when possible) for additional protection to their accounts.
  - On lab computers, team members should not install desktop applications nor set the login screen on the cloud-storage websites to 'Remember me'. This is to prevent other users of the lab computers from accessing team members' accounts, and also to avoid syncing while running an experiment.
  - The PI and team members should *NEVER* store *personal, confidential* or (commercially or otherwise) *sensitive data* in the cloud (unless proper encryption is used). Crucially, only *anonymised research data* that cannot reveal the identity of living individuals, or be linked to individual participants by data recipients outside the lab, can be stored unencrypted in the cloud.
  - See the **Glossary** for more information about *personal, confidential* and *sensitive data*, and *anonymised research data*.
- **External hard drives**:
  - External hard drives should be *encrypted* and stored securely (e.g. in a room that can be locked). When not using their computers for a longer period (e.g during holidays), the PI and team members should not store their hard drive in the same location as their computer system.
  - As physical storage media have a limited life, files should be transferred to a new hard drive every 2 to 5 years (with the files on the old drive properly erased, i.e. by overwriting them).
- **Paper-based files**:
  - Paper-based files containing participants' details, such as *signed consent forms or brain stimulation screening questionnaires*, have to be stored securely (i.e. in a locked filing cabinet) to avoid disclosure of personal data. Moreover, it is important to keep them separate from raw research data files to prevent data recipients outside the lab from linking research data to individual participants.
  - Team members should store *completed questionnaires* of ongoing studies separately and securely (i.e. in a locked filing cabinet). Upon completing their study, they should accurately convert the questionnaire responses into digital tabular data (which should not contain any personal identifiers).
  - Team members should *NEVER* write the participant number or code on the consent forms or questionnaires.
  - Under some circumstances, an extra paper-based file can be used to link individual names to participant numbers or codes; this file should also be stored securely, and be destroyed once data collection is finished. Note that when individual names can be linked to participant numbers or codes (e.g. via a paper-based file), the data files are not fully anonymised (which has implications for storage, sharing, and archiving). Thus, these links should be avoided when not strictly necessary.
  - Team members should hand signed consent forms, payment sheets, brain stimulation screening questionnaires, and other paper questionnaires to the PI upon completing a study.
  - The PI should in turn safely store the paper forms received in the faculty archive for research data. After a retention period of at least 5 years, the archive will contact the PI with instructions to securely dispose of the material.

---

3. Data sharing and archiving

3.1. Legal and ethical issues regarding data sharing

Data sharing touches upon a number of important legal and ethical issues involving *intellectual property rights (IPR)*, *personal, confidential* and *sensitive data*, and informed consent that should be taken into account.

The PI and other team members should *NEVER* formally or informally share:

- Data and other files for which they don't own *copyright* or don't have permission to share from the copyright holders
- Data and other files that contain *personal*, *confidential* or *sensitive information* without research participants' written consent

Although this is strictly not necessary, we will also ask inform research participants' in the information sheet about sharing *anonymised research data*.

In other words, data sharing requires researchers to know who owns the *IPR* of their data and how the General Data Protection Regulation (GDPR), which covers the use of personal data, applies to them. It also has important implications for the way in which informed consent forms should be drawn up.

See the **Glossary** for more details about *IPR*, *personal, confidential* and *sensitive data*, and *anonymised research data*. For more information about data sharing and informed consent, and a template consent form, see **Appendix D**.

3.2. Sharing active data

Team members and the PI can use the network share, private OSF projects, or private GitHub repository's to informally share appropriate *active* (i.e. not yet archived) data and other files with each other, and with trusted researchers beyond the lab, such as reviewers and grant collaborators.

Team members should consult the PI first before sharing *active data* with external researchers.

When informally sharing *active data* and other files, it is important to prevent accidental changes to or deletion of those files. If you use cloud services (such as Dropbox) for sharing, its website lists two main ways to *avoid unauthorised file changes or deletions*:

- Don't invite people to a shared folder, but send them a link to the folder or file you want to share, which they can then click to view and download a copy of the file(s). Remove the link if you no longer want the folder or file to be accessible from the link.
- If you do invite people to shared folders, give them view-only permissions. In this way, members of your shared folder will be able to see the latest version of the files in the folder, but will not be able to add, delete or edit files.

3.3. Archiving completed data sets

It is lab policy to formally share completed research data sets by depositing them in a data repository. This raises the data's profile and enables their proper citation by other researchers. Moreover, depositing data ensures their long-term preservation and fulfils growing funder and institutional requirements regarding open access to research.

3.3.1. Who?

Lead authors, whether they be PI, PhD student or post-doc, should deposit the completed data sets (with accompanying software and documentation) supporting their scientific publications (journal articles, book

chapters. . .).

The PI and PhD students should in addition deposit project-level documentation and any remaining data identified for archiving from their grant-funded research and PhD projects respectively.

3.3.2. Where & when to deposit?

Deposits supporting publications When publishing papers, lead authors from the lab should simultaneously make the underlying data available by making their OSF projects (see above) public.

Deposits at the end of PhD and grand-funded projects Upon completion of their grant-funded or PhD research projects, the PI and PhD students should make sure that they comply with funder and/or institutional requirements regarding data deposits:

- PhD students should upload a data storage fact sheet as well as project-level documentation (i.e. their PhD thesis) to Biblio
- The PI should also offer any remaining project data selected for archiving to OSF.
- In addition, if their funder requires them to make their data available via a particular repository, PhD students and the PI should register their projects, and provide (links to) project-level documentation as well as links to the data on OSF in this funder-designated repository within the required time frame.

3.3.3. Before depositing data

Before making data and accompanying files public in OSF, team members should seek the PI's approval. The PI grants permission after confirmation that all the appropriate steps outlined in this document have been followed. If necessary, he can ask team members to edit files before making our projects public.

The PI and other team members should consider the following questions before offering data for archiving:

- **Do all collected data sets need to be preserved?**
  - The PI and other team members should in the first instance focus on archiving the data that support their publications.
  - For deposits at the end of grant-funded or PhD research projects, the PI and PhD students should select remaining data for preservation in compliance with funder and/or institutional requirements.
- **Is it legally, ethically and commercially appropriate to share the data?**
  - The PI and other team members should ensure that there are no issues regarding intellectual property rights, disclosure of (commercially or otherwise) sensitive, confidential and personal information, or informed consent that would preclude them from making their data open access.
- **Do the data require a restricted level of access?**
  - If there are legal/ethical/commercial issues, repositories can restrict and regulate access and use of data, so that these data can still be deposited and shared under certain conditions.
  - The PI and other team members should also consider whether they want the repository to delay making their data publicly available for a certain period, for example until they have published the research outputs associated with these data.
  - If applicable, the PI and team members can indicate that restricted access is required in an 'access and use conditions' file (see Section 2.1.3., 'Data documentation').
- **Are the data and documentation files properly structured and formatted, and are data appropriately documented?**
  - The PI and other team members should check whether the data sets selected for archiving comply with repository requirements regarding file structure, formatting and documentation (if they have followed the lab's data management guidelines, this should normally be the case), and whether file contents are accurate.
  - Depositors should ensure, in sum, that other researchers are able to understand and reuse their data.

3.3.4. After depositing data

Once it is confirmed that data and accompanying files have been archived in OSF and associated research outputs have been published, team members may choose to dispose of the corresponding files on their computer hard drives, network shares, or cloud accounts, and thus only hang on to active and unpublished data if they wish.

If they need or want to dispose of data, it is recommended that the PI and PhD students refrain from deleting experiment files *at least* until they have fulfilled all institutional and/or funder data deposit requirements at the end of their grant or PhD project, and have published the associated research outputs.

The PI will decide on and oversee the disposal of data on lab computers at the end of grant-funded and PhD projects.

---

4. Glossary

Data documentation: Data documentation enables users to understand research data, as it documents their creation, meaning, content, structure and manipulations. It also forms the basis of catalogue metadata compiled by data repositories in order for users to find and cite research data. Data documentation includes information about research data at different levels: the project level, the experiment or study level, and the data level. (by 'study', we mean a meaningful (sub)group of experiments, e.g. an experiment and follow-up experiments).

Experiment/study-level and data-level documentation: This is the responsibility of team members conducting experiments. Elements of data-level documentation can often be embedded within a data file itself (e.g. variable labels within a data file). Experimenters should nevertheless also prepare supporting documents to provide sufficient experiment-/study- and data-level documentation. These files should enable other researchers to replicate the study or reanalyse the data.

Project-level documentation: Applicable to research that is part of a larger project. Project-level documentation includes information about the broader context of data collection: the project history, aim, objectives, hypotheses, etc. It can take the form of research reports (e.g. ERC End of Award Reports), so it shouldn't require much, if any, additional work. This is the responsibility of PIs and PhD students in charge of grant-funded and PhD research projects respectively. If project-level documentation is not applicable (e.g. for pilot experiments), data documentation is restricted to Experiment/study-level and data-level documentation.

Version control: Version control entails a variety of strategies that researchers can adopt to keep track of different versions of files. Also see **Appendix C**.

Master file and milestone versions: A master file is an original file that you always want to be able to go back to, and from which working copies are created for editing. For example, if you are editing a digital image, you would not want to lose the original picture. Milestone versions are versions of a file that are changed to such an extent–compared to previous versions–that you want to keep them (e.g. the first draft, revised draft, final draft of a file). Also see **Appendix C**.

Incremental backup: Some backup programs, such as macOS' Time Machine, make incremental backups to a backup disk. This means that it makes an initial full backup of your system, and subsequently backs up the files that have changed since the last backup. It saves hourly backups for the last 24 hours, daily backups for the past month, and weekly backups beyond that. As noted above, avoid deletion of old back-ups.

Active data: Experiment data that have not yet been archived, i.e. that have not yet been deposited in a data repository for long-term preservation.

Intellectual Property Rights, including copyright: Before sharing data, researchers should know who owns the intellectual property rights (IPR) of these data. Copyright is an important IPR that prevents

unauthorised copies and publishing of original work that exists in written or recorded form. Researchers should therefore never share data or other materials via the cloud, data repositories, or other means unless they own copyright or have permission for sharing from the copyright holders.

More details on UGent's Intellectual Property policy can be found in the Valorisatiereglement and in the (Onderzoeksreglement AUGent)[https://www.ugent.be/intranet/nl/reglementen/onderzoek/reglementen/onderzoeksreglementaugent.pdf/at_download/file] (both in Dutch).

Personal, confidential and sensitive data: Research with human participants can generate data containing personal, confidential or sensitive information, the sharing of which may breach legal and ethical obligations. UGent's guidelines for the classification of information and data can be found here

- *Personal data*: In a GDPR context, personal data "means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (Article 4).
- *Sensitive personal data*: Sensitive personal data are (Article 9):
  - personal data revealing racial or ethnic origin, political opinions, religious, philosophical beliefs, or trade-union membership
  - genetic data, biometric data processed solely to identify a human being, or health-related data
  - data concerning a person's sex life or sexual orientation.
- *Confidential data*: data containing identifying information that an informant gives in confidence, that two parties agree to keep confidential, i.e. secret.

Researchers should not share personal, sensitive or confidential data without research participants' written consent. Disclosure of (sensitive) personal and confidential data without consent constitutes a breach of the GDPR and psychologists' duty of confidentiality towards research participants (see e.g. the British Psychological Society's Code of Ethics and Conduct.

It is best that team members do not collect personal and sensitive data that are not necessary for their research. They should not disclose or share participants' personal details on the signed consent forms and payment sheets, nor experimental data/questionnaire responses in which participants can be individually identified. Some more information about the GDPR can be found on UGent's website: https://www.ugent.be/en/news-events/news/gdpr-en.htm

As outlined in the policy framework for research data management at Ghent University, data can also be sensitive from a commercial or institutional point of view. More specifically, "when research data are research results that are subject to technology transfer as defined by the Ghent University Regulations on Research Transfer, they must be reported to the TechTransfer Office of Ghent University prior to any kind of publication, in order to examine whether these research data can are subject to technology transfer or are protected by intellectual property rights, such as patents."

Anonymised research data: Researchers can share anonymised research data, where (in)direct personal identifiers (such as names) revealing the identity of living individuals have been removed, and which data recipients are unable to link to individual participants. Anonymised research data are exempt from the GDPR, and anonymising data maintains confidentiality. It is still good practice to explicitly inform subjects about the sharing of anonymised research data when seeking their consent to participate in experiments.

Encryption Encryption is used to safely store or move files, as it prevents unauthorised access. You can encrypt individual files, folders or entire storage devices.

Team members should encrypt the information on their external hard drives, as these can easily get lost or stolen. When using macOS' Time Machine, this can be done by checking the 'Encrypt backups' box for the backup disk. There are also external hard drives available with built-in encryption. **Caution: if you forget your password, you will not be able to recover the files on your encrypted backup drive!**

5. Appendices

Appendix A. Folder structure and file naming

Examples from the 'Ctrl-ImpAct' ERC research project:

Folder names:

- Project folder: *Ctrl-ImpAct*
- Experiment Folder: *CIA_PrepTMS*

This would create the following structure:

- *CtrlImpAct*
  - *CIA_PrepTMS*
    - *Data*
    - *Code*
    - *Documentation*

File names: File names are also based on the experiment folder. For example:

- Data:
  - Behavioural data: *CIA_PrepTMS_Behavioural_s1.txt*
  - TMS data: *CIA_PrepTMS_Tms_s1.txt*
- Code:
  - Final, bug-free version of program file: *CIA_PrepTMS.py* (during debugging stage, one can use a suffix; e.g. *CIA_PrepTMS_d1.py*. Usually, you can delete the d-files).
  - Analysis program: e.g. *CIA_PrepTMS_ChoiceAnalysis.R*
- Dcoumentation folder:
  - Experiment documentation: *CIA_PrepTMS_ExpDoc.md*
  - Lab notes: *CIA_PrepTMS_LabNotes.md*
  - Analysis documentation: *CIA_PrepTMS_AnalysisDoc.pdf*
  - Informed consent: *CIA_PrepTMS_InformedC.md*
  - Debriefing: *CIA_PrepTMS_Debrief.rtf*
  - ...

Appendix B. Experiment and Analysis documentation

The documentation folder of an experiment should always include:

- *Experiment Documentation*: a text document with the following information:
  - Project information: a short summary of the larger research project of which the experiment is a part, so as to explain the wider context of data collection
  - Study/Experiment information, including:
    - Experiment code(s) as mentioned on informed consent forms
    - Research question
    - Experiment context (full name of experimenter, location, date and time, credit or paid participants)
    - Data collection methods (description of experiment, data collection protocols; instruments, software, and hardware used...)
  - Data file information, including:
    - Structure of data files (overview of all data files–e.g. behavioural, TMS, SCR...–and the relationship between them)

        * Content of data files: names, labels and descriptions for all variables and their values
    - Quality control measures: e.g. piloting of experiments, timing checks, calibration procedures (e.g. Eyelink), checking of manual data entries (e.g. in case of questionnaires), etc.
- *Analysis Documentation*: a text document that explains how the data were processed and analysed. It should include the following information:
    - How many subjects were included? Did you exclude subjects in the final analysis. If so, why?
    - Which variables were analysed? Did you exclude variables? If so,why?
    - Did you exclude trials when analysing the data (e.g. incorrect trials, trials following an error, mean +2.5 SD, etc.)?
    - Did you modify data or create derived/constructed variables from the original data? If so, explain the logic.
    - When R is used, the analysis document can be created within R and include a summary of data analysis outcomes. Examples can be provided by the PI.

Appendix C. Version control strategies

Git and GitHub

Version control can be done with Git or GitHub. Note that a local repository suffices for version control. See the lab's Git & GitHub guidelines for more information and links to relevant documentation: https://github.com/CoBE-lab/cobe-copilot-guidelines

Manual version control:

- To uniquely identify and distinguish between different versions of a file, incorporate version information in the file name:
    - Include a version number (number each successive version of a document sequentially, indicating major and minor changes by whole and decimal numbers respectively: e.g. v1...), or a date (in a format that allows easy sorting: 'yymmdd'). For example:
        * CIA_PrepTMS_ExpDoc_v1
        * CIA_PrepTMS_ExpDoc_180930
    - For files edited by multiple authors, also include author initials as an identifier. For example:
        * CIA_PrepTMS_ExpDoc_v1_FV
        * CIA_PrepTMS_ExpDoc_180930_FV
- Save files under a new file name (i.e. with a new version number or date, and if applicable, initials) *before* editing them.
- If you need to keep track of a file's history, include a version control table within the file (such as the one at the beginning of this document) or in a separate document. Record for every version of the file what changes were made, when and by whom, so that you know how different versions differ from each other and who contributed what.
- For the files you are responsible/lead author for, you might have to manually merge edits from multiple co-authors into a new version (with a new file name).
- Decide which versions of a file to keep, and how to organise them. Things to keep in mind:
    - Keep a single master file, especially for files where loss of the original would present a problem, such as raw data. Store them separately and leave them untouched (so do not overwrite or delete!).
    - For sharing with collaborators, editing, analysis etc. use working copies of these master files instead, so that the original files are not affected and you can always go back to them, if necessary, to start from scratch.
    - Team members are responsible for the master files of experiments they conduct: only they should have write access to these master files.
    - You may not have/want to keep all old versions of files, but identify milestone versions to hold on to.

  - You can store non-current versions of files separately from current versions.
- Synchronise files held on different computers: e.g. by using Dropbox, which will automatically sync files on devices connected to the cloud storage service.

Appendix D. Data sharing and informed consent

It is important that informed consent forms, while discussing personal data protection and confidentiality, don't preclude data sharing.

Taking the nature of the lab's research into account, consent forms should inform participants that:

- The information and data provided will be used for statistical research and research purposes only.
- No data or responses will be published in which they can be individually identified.
- No personal details of individual participants will be revealed to persons outside the research project; signed consent and payment forms will be securely stored and disposed of after a 5-year retention period.
- Anonymised research data will be registered and archived in an institutional/specialist data repository so as to make them available to other researchers.