



European Research Council
Executive Agency

Established by the European Commission

ERC Data Management Plan

Template



ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

European Research Council

Established by the European Commission

Project Acronym

Project Number

FACEDIFF	864694
----------	--------

Template for the ERC Open Research Data Management Plan (DMP)¹. The following sections should describe how you plan to make the project data Findable, Accessible, Interoperable and Reusable (FAIR). Each of the following five issues should be addressed with a level of detail appropriate to the project.

SUMMARY (dataset² reference and name; origin and expected size of the data generated/collected; data types and formats)

The purpose of the FACEDIFF data collection is to test the hypothesis that individual differences in facial expressivity equip individuals' differentially to engage with social environment, combining psychological (WP1), anatomical (WP2) and cross-species (WP3) methods.

The data types to be generated are survey and experimental results (e.g. questionnaire data and experimental responses: file types .csv, .sav, .xlsx), video recordings (.mkv) and related derived data and models from humans and rhesus macaques (*Macaca mulatta*). The project will also use two existing datasets. Firstly, for WP2b, the IXI dataset comprising MRI head/brain scans taken from 600 healthy individuals (London, UK: <https://brain-development.org/ixi-dataset/>). This dataset is freely available for scientific use. This data will be used to model facial muscle presence/absence, gross structure and gross symmetry with the associated demographic information. Secondly, for WP3b, existing long term macaque datasets from 2010 (e.g. Brent, Ruiz-Lambides, & Platt, 2017) will be used to calculate social network data. The total expected size is in the order of 100 TB (uncompressed macaque focal video observations estimated at 70 TB, and uncompressed videos of human experiments estimated at 30 TB). The data might be useful to other researchers interested in human and macaque behaviour.

1. MAKING DATA FINDABLE (dataset description: metadata, persistent and unique identifiers e.g., DOI)

¹ Based on '[Guidelines on FAIR Data Management in H2020](#)', version 3.0. 26.07.2016, Annex I

² Several datasets may be included into a single DMP.

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Data appropriate for sharing (estimated 250MB) will be shared via UoP Pure repository and will be allocated a DOI.

All data files will be named as work package_study_YYYYMMDD_Subject ID (e.g. WP1_study1a1_20200101_014)

Work packages = WP1, WP2, WP3, WP4

Studies = 1a1, 1a2, 1b, 1c, 2a, 2b, 3a, 3b

Subject ID = For human groups, number. For macaques, use fieldsite abbreviations.

For word documents: workpackage_study_title_YYYYMMDD

As the project spans different disciplines, clear contents pages and terminology information (glossaries) will be provided for users to navigate the repository effectively.

2. MAKING DATA OPENLY ACCESSIBLE *(which data will be made openly available and if some datasets remain closed, the reasons for not giving access; where the data and associated metadata, documentation and code are deposited (repository?); how the data can be accessed (are relevant software tools/methods provided?))*

Data to be made openly available will be deposited in the NTU repository with an open access licence. No specialist software will be required to access the data. Open licencing will be applied (e.g. CC BY 4.0). All anonymised datasets will be made openly accessible after publication, with the exception of the following:

Raw human data from WP1-3 (i.e. videos, photos, recorded facial dissections) in order to protect individual identity.

Video data from the captive macaque work (WP3a) due to contractual obligations with MRC Centre for Macaques.

3. MAKING DATA INTEROPERABLE *(which standard or field-specific data and metadata vocabularies and methods will be used)*

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

The project will span different disciplines (anatomy and psychology), so clear contents pages and terminology information (glossaries) will be provided for users to navigate the repository effectively. In addition, FACS (Facial Action Coding System: Ekman et al., 2002) terminology will be used to describe the facial expression and facial muscle data across the datasets.

4. INCREASE DATA RE-USE (*what data will remain re-usable and for how long, is embargo foreseen; how the data is licensed; data quality assurance procedures*)

It is expected that a CC BY 4.0 licence (attribution only) will be applied to the data that are being shared to permit the widest possible re-use.

It is expected that the complete dataset (with the restrictions outlined above) will be made available no later than the first online publication of the article or 12 months after the project end date whichever is earlier.

It is intended the dataset remains re-usable for a minimum of 10 years (in accordance with NTU Research Data Management Policy).

All staff working on the project will be fully trained in standardised data collection protocols and data back-up schedules.

5. ALLOCATION OF RESOURCES and DATA SECURITY (*estimated costs for making the project data open access and potential value of long-term data preservation; procedures for data backup and recovery; transfer of sensitive data and secure storage in repositories for long term preservation and curation*)

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

There are no additional costs associated with making the data freely available. The data will be deposited in NTU repository and data processing will be wholly management by the project research team. Open access publication fees have been included in the grant budget.

The PI (Prof. Waller) will have overall responsibility for data management in the project. For data collection, day-to-day data management responsibility will lie with the two Postdoctoral Research Fellows.

In accordance with NTU Research Data Management Policy, the data will be retained for ten years from whichever is the latest of:

- the completion of the research;
- the publication date of any research findings based upon the research data;
- the completion of any litigation proceedings, investigation on research misconduct, review of financial management or other formal enquiries.

After this ten year retention period, further retention will be reviewed.

During the project lifetime the collected data will be backed up on completion of each data collection session to the NTU's institutional One Drive service (remotely accessible, secure, automatically backed up). A dedicated shared drive will be set up for the project, enabling secure transfer of data. Access will be restricted to the project team. For data collected off-site, data will be saved to external USB hard drives encrypted using Microsoft Bitlocker software. Team members will be prompted to enter a password before the contents of an encrypted drive can be viewed.

DISCLAIMER. Please note that the ERC Data Management Plan is not a part of the Ethics Review. It is the responsibility of the Principal Investigator to inform the ERCEA Ethics Team of any ethics issues/concerns regarding the collection, processing, sharing and storage of data in relation to the project. The Principal investigator can also be asked to submit an Ethics Data Management Plan (Ethics DMP).