Archiving Data: A Cognitive Neuroscience use-case

Gorka Fraga González
Andrew Clark
Alexis Hervais-Adelman
Susanne Weber?
Melanie Röthlisberger?
Leonard Held
https://doi.org/10.5281/zenodo.XXXXXX

What is data archiving?

Brief definition of what we understand as data archiving. See UZH ZB resources

What is an Open Research Data repository?

Here shortly specify what an Open Research Data repository is. Ref to web for searching in repositories database

Why use an ORD repository?

Explain the obvious. Highlight that the information is not necessarily public access but researchers may enquire /send an application to access the data (restricted access)

How do you archive data?

First see how much your data follows the FAIR principles – Short recap and link to ZB resources on "How fair is your data' . Find a suitable repository and see if there are particular requirements
Proceed
This is better illustrated with an example of a project with a relatively complex data.
——–

Use case: a neurolinguistics project

This primer describes a use case for the upload of a neurolinguistics project into an Open Research Data (ORD) repository. The reference project involves and Speech-in-noise comprehension task in which participants have to (...). The project's data involves electroencephalography and behavioral data, as well as auditory stimuli presented in the experiment. The repository is xxxx

Think of a data structure system

First search for any exising standards in the field. Here we follow BIDS (insert short descript), why ? (xxx) Best to struct the project like this from start

Organize your data

This is best done at the start of a project.
[Insert Folder tree of the project here]

Decide what needs to be archive

Not everything must or should be archived.
Schematic of minimum project elements

- Data

    – Raw

    – Preview/thumbnails

    – Derivatives

- Metadata

    – .JSON

    – Tables

- Code

- Documentation

    – Filenaming convention

    – Metadata/data specifications

    – Procedures

    – READMEs

Selecting a repository

What makes a good repository for my data ? In this case (...)
Workflow chart: File naming, organisation, metadata.

Archiving

Here Andre fills in some details about how in this particular repo we need to proceed or additional files/struct required by it

Common Issues and Troubleshooting

Here I would describe issues with this project. e.g. in this case we don't have a problem of too large data set, the main problems can come from manipulations into the raw data (from source to 'raw', which is not really raw). (see knownIssues.txt file)

Take home message

Technical Box 1 - Archiving scripts for stimuli presentation

Here issues with version control and scripts to run experiments (e.g., what files where presented, etc)

Technical Box 2 - Challenges of EEG experiments

Here Andrew inserts common operations done to raw EEG data

More information