

DMP title

Project Name Towards natural brain-computer interfaces: representation learning for identifying the temporal encoding of realistic video footage in EEG - DMP title

Project Identifier G081722N

Grant Title G081722N

Principal Investigator / Researcher Alexander Bertrand

Project Data Contact alexander.bertrand@kuleuven.be

Description Brain-computer interfaces (BCIs) facilitate interaction between the brain and a computer or machine. Electroencephalography (EEG) is by far the most popular non-invasive BCI modality, because it is cheap, mobile, and it has an excellent temporal resolution to track neural responses that are time-locked to a sensory stimulus. However, traditional BCI paradigms heavily rely on synthetic and controlled sensory stimuli and an active participation of the user, making it very hard to integrate such paradigms in practical “everyday-life” use cases. If BCI technology would be able to cope with uncontrolled and natural sensory stimuli, it would be able to naturally blend in with normal behavior and activities of the user. Inspired by recent breakthroughs towards decoding EEG responses to speech, we have the challenging goal to design a new data-driven methodology to identify and quantify the temporal coupling between natural video footage and its EEG responses. To this end, we will leverage representation learning techniques from the field of computer vision, and rethink them in combination with recent insights in EEG decoding and visual BCI. We will also investigate whether and how this framework can be used to track visual and spatial attention to natural stimuli. Not only would this be a game-changing tool for various experiment-driven research fields in neuroscience and medical sciences, it would also pave the way towards many new BCI applications in various domains.

Institution KU Leuven

1. General Information

Name applicant

Alexander Bertrand & Tinne Tuytelaars

FWO Project Number & Title

G081722N Towards natural brain-computer interfaces: representation learning for identifying the temporal encoding of realistic video footage in EEG

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data
- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

1) Type of data: 64-channel EEG recordings + metadata

Format: EDF, BIDS, mat, binary (depending on processing stage)

Volume: 185GB (200hour of 64-channel EEG data downsampled to 250Hz sampled at 16bit)

How created: by performing EEG recordings on different participants, while they watch specific video stimuli

2) Type of data: video files with stimuli for the EEG experiments

Format: any video file format (e.g. MP4)

Volume: 15GB

How created: public videos (e.g. youtube, etc.)

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- Yes

Privacy Registry Reference: [G-2022-4765-R2\(MAR\)](#)

-name of participant Short description of the kind of personal data that will be used:

-age

-gender

-contact details (e-mail and phone)

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- Yes

PRET reference: [G-2022-4765-R2\(MAR\)](#)

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

The protocol for the EEG recordings will be explained in a document.

There will be a readme file explaining the content and structure of the recorded EEG files, and how to run the scripts to prepare the files in Python and/or Matlab

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

5. Data storage and backup during the FWO project

Where will the data be stored?

The master copy will be kept on STADIUS' and/or PSI's central storage servers (encrypted and backed up) during the research but also after the project is finished. For heavy computations, the data will be hosted on the Vlaamse Supercomputer Centrum (VSC).

Some parts of the EEG data may be made publicly available, after consent from the participants. In this case, we will post it on a research data platform such as, e.g., Zenedo.org

How is backup of the data provided?

The storage used by the research group STADIUS is backed up daily. The backup process is managed by the IT division of the department.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

The data generated in the project will be stored on storage facilities of the research unit, with a total size of 8TB, of which 4 TB is still available.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

The (maintenance) costs are covered by the division.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data stored in the ESAT servers has access regulated by an access control list (ACL) that grants: read-write access to the project owner read-only access to specific users The ACL is managed by the project owner. Client computers can access the data using: SMB2 (or higher) from specific IP ranges NFSv4 from specific (IT managed) systems.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All mentioned data will be kept for at least 5 years after the end of the project.

Where will the data be archived (= stored for the longer term)?

After the period of 5 years if the data generated during the project is not longer in use locally, it will be removed from the STADIUS servers. Note that some parts of the dataset may remain publicly available via public online repositories (see above)

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The (maintenance) costs are covered by the division.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

The (anonymized) EEG data might be made available to the public for re-use, after obtaining consent from the participants in the study.

Where/how will the data be made available for reuse?

- In an Open Access repository

When will the data be made available?

- Upon publication of the research results

Who will be able to access the data and under what conditions?

the data will be publicly available to anyone

What are the expected costs for data sharing? How will the costs be covered?

we will use a free online research data sharing platform such as zenodo.org

8. Responsibilities

Who will be responsible for data documentation & metadata?

The PI, Alexander Bertrand and the researchers working on the project

Who will be responsible for data storage & back up during the project?

The data storage and back up is the responsibility of the IT support of ESAT

Who will be responsible for ensuring data preservation and reuse ?

The PI, Alexander Bertrand

Who bears the end responsibility for updating & implementing this DMP?

The PI, Alexander Bertrand, bears the end responsibility of updating & implementing this DMP.