



Sagol School
of Neuroscience
Tel Aviv University

A guide for sharing eye-tracking data

Data sharing contains two key components:

- (1) The format the data is saved in (folder and file structure, file naming and format)
- (2) The platform on which the data is shared, and its restrictions.

The following guide is for sharing data from **eye-tracking devices** in **BIDS format**, onto the **OSF platform** (<https://osf.io/>).

- For more information about OSF data and metadata management, see here: <https://help.osf.io/article/392-data-management>
- For more information about BIDS format, see: <https://bids-standard.github.io/awesome-bids/>

What is BIDS?

BIDS (Brain Imaging Data Structure) is a standardized system for organizing and storing data in a way that makes it easy to share and analyze. As it's widely used in neuroimaging, there has been progress in creating BIDS format for eye-tracking data in recent years. Akin to neural data, BIDS provides a consistent structure and naming convention for organizing the eye-tracking data and metadata, ensuring that anyone using the data knows exactly where to find each part of it and how to process it. This ensures that the data is consistent, facilitates collaboration and reproducibility, and simplifies the processing of the data.

Gaze Data

Eye-tracking data includes information about where and when a participant is looking during an experiment: eye position coordinates, gaze events, and possibly the experimental conditions under which these data were collected. **A key step is to anonymize the data** before sharing it with anyone. While conversion of the data to BIDS format focuses on the structure the data is shared in, **BIDS conversion does not handle the anonymization of sensitive data**. Thus, prior to processing the data for uploading, **it's your responsibility to ensure that the data does not contain any sensitive information about the participants**.

If you are collecting eye tracking data using EyeLink products, the gaze data is in EDF format (EyeLinkDataFile, not to be confused with European Data Format), which contain, on top of gaze information, participant identifiers and other sensitive data such as time and date, and computer details (e.g., the folder path).

If you are collecting data using Tobii / HTC Vive products, the gaze data is usually outputted in csv/tsv/JSON format, and might include sensitive data as well.

In sum, in both cases you need to make sure to:

- Remove participant identifiers
- Ensure no sensitive information (e.g., full names, medical identifiers) remains in the datasets post-conversion
- Make sure that any identifiable information is separated from the actual data (e.g., in a separate metadata file)

Key Components in BIDS Conversion

Key Components of BIDS for Eye-Tracking Data:

1. Folder Structure:

BIDS organizes data in a specific folder structure. For example, there will be a separate folder for each participant, and inside that folder, you might have data for different tasks or experimental conditions. These folders and files are named consistently, making it easier to navigate and use the data (see “Example Directory Structure for OSF” below).

2. File Naming:

Just like with other types of research data, BIDS has clear rules for naming files. For example, an eye-tracking data file might be named something like `sub-01_task-reading_eyes.tsv`, where:

- `sub-01` refers to the participant ID.
- `task-reading` refers to the specific task the participant was performing (e.g., reading a sentence).
- `eyes` refers to the specific data (in this case, the eye-tracking data).
- `.tsv` is the file format (Tab-Separated Values), a common format for storing structured data like eye-tracking measurements.

3. Metadata:

Along with the actual eye-tracking data, BIDS also standardizes how you store metadata (extra information). This includes things like the details of the experiment (e.g., the type of task, the conditions, the equipment used for eye-tracking), or information about the participants (e.g., age, gender). This metadata is typically stored in `.json` files that are

placed alongside the data files.

Convert Eye-Tracking Data to BIDS Format

We highly recommend you use the guides and tools in these links:

- https://www.fieldtriptoolbox.org/example/other/bids_eyetracker/
- <https://github.com/bids-standard/eye2bids>
- <https://www.nature.com/articles/s41597-024-03559-8>
- https://github.com/cpp-lin-lab/eye-tracking_cookbook

However, here's a high-level breakdown for each device:

Create the structure

HTC Vive Pro Eye (json/csv)

1. **Directory Structure:** Each participant's data should be placed in a folder like `sub-01`, `sub-02`, etc. under a folder named `eyetracking`.
2. **Metadata Files:**
 - `participants.tsv`: This file contains the list of participants (ID, age, gender, etc.).
 - `eyetracking.tsv`: This is where you store the raw eye-tracking data (e.g., gaze x/y coordinates, pupil size).
 - **Example file:** `sub-01/eyetracking/eyes.json` or `sub-01/eyetracking/eyes.csv`.
3. **Data Conversion:** If your data is in CSV or JSON, structure it into time columns (e.g., timestamp, gaze_x, gaze_y, pupil_size) with metadata (subject, session, task, etc.).

Tobii (json/csv)

1. **Directory Structure:** Same as for Vive Pro Eye: `sub-01`, `sub-02`, etc., under an `eyetracking` folder.
2. **Metadata Files:**
 - `participants.tsv`: Similar structure to the Vive Pro Eye data.

- `eyetracking.tsv`: Gaze positions, pupil sizes, and timestamps.
 - **Example file:** `sub-01/eyetracking/tobii_data.json` or `sub-01/eyetracking/tobii_data.csv`.
3. **Data Conversion:** Convert CSV/JSON to TSV format, ensuring each row corresponds to a specific timepoint with gaze position and pupil data.

EyeLink (.edf)

1. **Directory Structure:** Same format, but the data is stored in `.edf` files under the subject folder.
 - **Example:** `sub-01/eyetracking/sub-01_task_name.edf`.
2. **Metadata Files:**
 - `participants.tsv`: Standard participant metadata.
 - `eyetracking.tsv`: This can either contain specific timestamps for the events you're interested in (e.g., fixation onset) or you can keep the raw `.edf` files.
3. **Data Conversion:**
 - Convert `.edf` files into BIDS-compatible formats, extracting relevant eye-tracking data (e.g., gaze positions, pupil size, etc.).
 - Tools like `pyedf` can help you process and extract the data from EDF files into a more manageable format.

Then, BIDS format requires standardized naming conventions. For example:

- **Subject:** `sub-01`, `sub-02`, etc.
- **Task:** `task-name`.
- **Eye-Tracking File:** Should be named according to the device and task, such as:
 - `sub-01_task-experiment_eyetracking.csv`
 - `sub-01_task-experiment_eyetracking.json`

Prepare the Data for Upload

1. Create Folders and Files:

- Follow the directory structure from BIDS. At the root, you should have:
 - `dataset_description.json`: Describes your dataset (e.g., name, license, etc.).
 - `participants.tsv`: Participant metadata.
 - `sub-01/`, `sub-02/`, etc., with the eye-tracking files and metadata.

2. Metadata:

- You must ensure that each data file has associated metadata (e.g., timepoint, participant ID, task name) to ensure data is interpretable.

3. Review BIDS Guidelines:

- Ensure compliance with BIDS for Eye-Tracking to make sure your data adheres to the best practices for sharing.

Upload Data to OSF

We recommend you use this guide: <https://help.osf.io/article/392-data-management>

Once your data is properly organized and formatted:

1. Create an OSF Project:

- Log into OSF and create a new project.
- Choose "Open" or "Restricted" sharing depending on your preference.

2. Upload Files:

- Use the **OSF File Manager** to upload your BIDS-formatted eye-tracking data.
- Ensure the structure is maintained, with the correct folder hierarchy (`sub-01`, `sub-02`, etc.) and metadata files (`participants.tsv`, etc.).

3. Versioning and Tracking:

- OSF supports versioning, so if you make changes or updates to the data, you can upload new versions while preserving the original versions.

4. Linking to Repositories:

- If you have any analysis code or related neuroimaging data, link those repositories to the OSF project.

Document Your Dataset

- Ensure that you include a **README** file in the root directory of your project that describes:
 - The eye-tracking data collection methods (e.g., device, sampling rate).
 - The structure of the dataset (how to navigate the folders, meaning of each file).
 - Any additional information or disclaimers about the data (e.g., data quality, exclusions).

Example Directory Structure for OSF

```
/my_project
/sub-01
  /eyetracking
    sub-01_task-experiment_eyetracking.tsv
    sub-01_task-experiment_eyetracking.json
/sub-02
  /eyetracking
    sub-02_task-experiment_eyetracking.tsv
    sub-02_task-experiment_eyetracking.json
dataset_description.json
participants.tsv
README.md
```

Follow this doc to ensure compliance:

<https://bids-specification.readthedocs.io/en/stable/common-principles.html>

Does somewhere else in the doc **:Commented [AM1]** refer to this? It's an important piece of the guide and it's good to state it exists beforehand

Commented [RH2R1]: Now it does (above)