

## Introduction

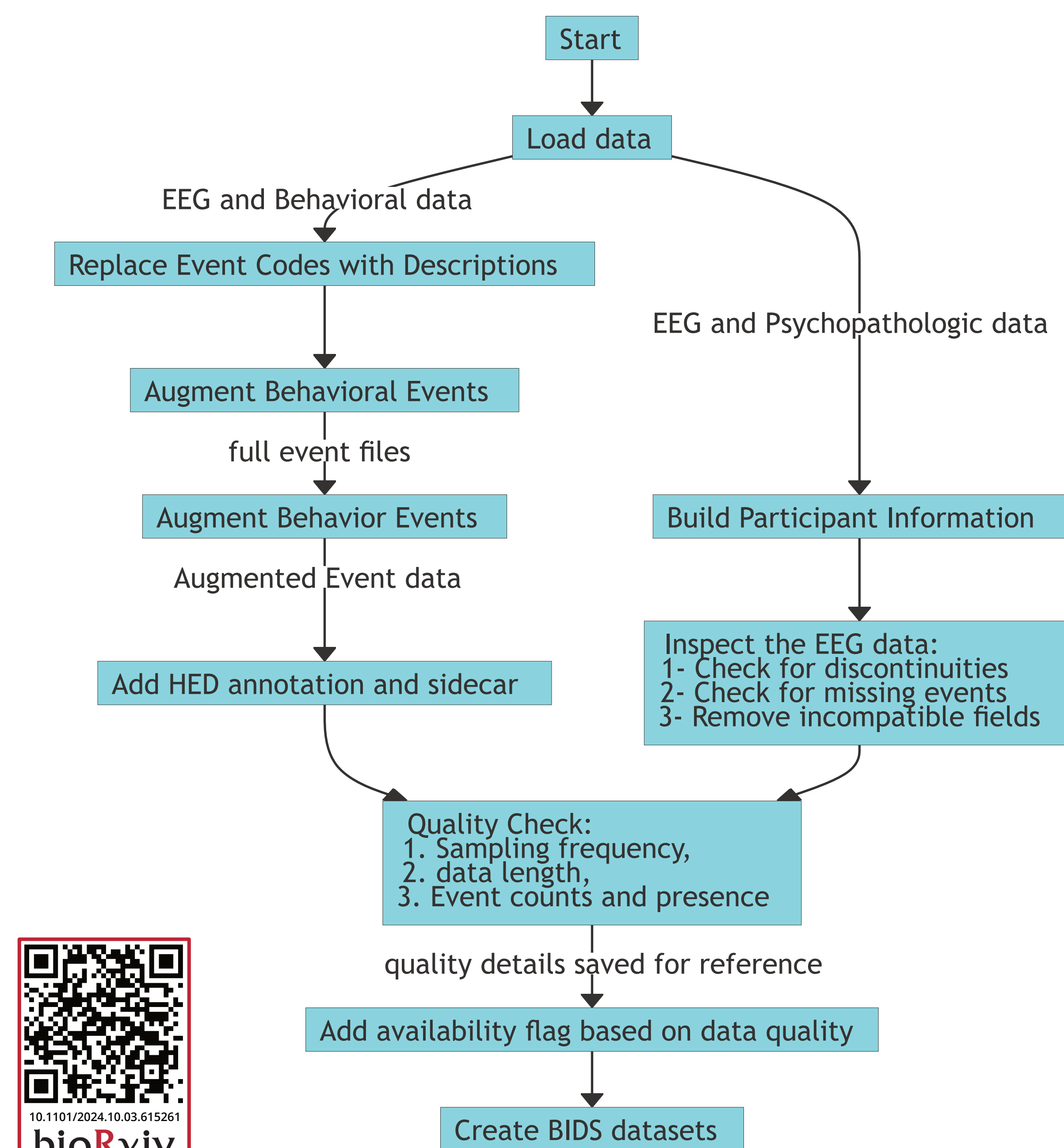
- The HBN project [1] spans ~5000 subs (5-21yo) with fMRI, EEG, Eyetracking, and psychopathological data.
- Data of ~2600 subjects have been released in the span of 9 Releases. EEG/Eyetracking data require significant curation to become "analysis-ready."

**Purpose:** Curate HBN EEG data with rich task event information and personal metadata for downstream research.

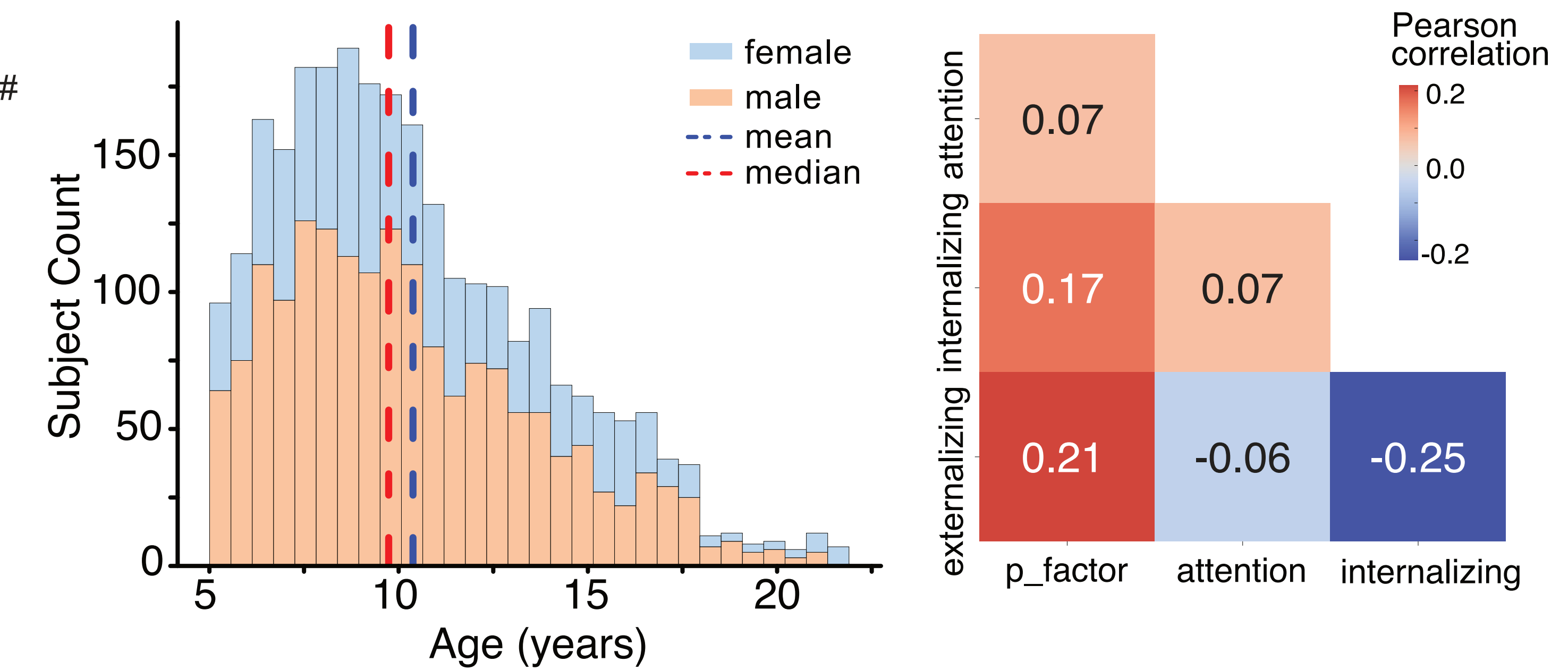
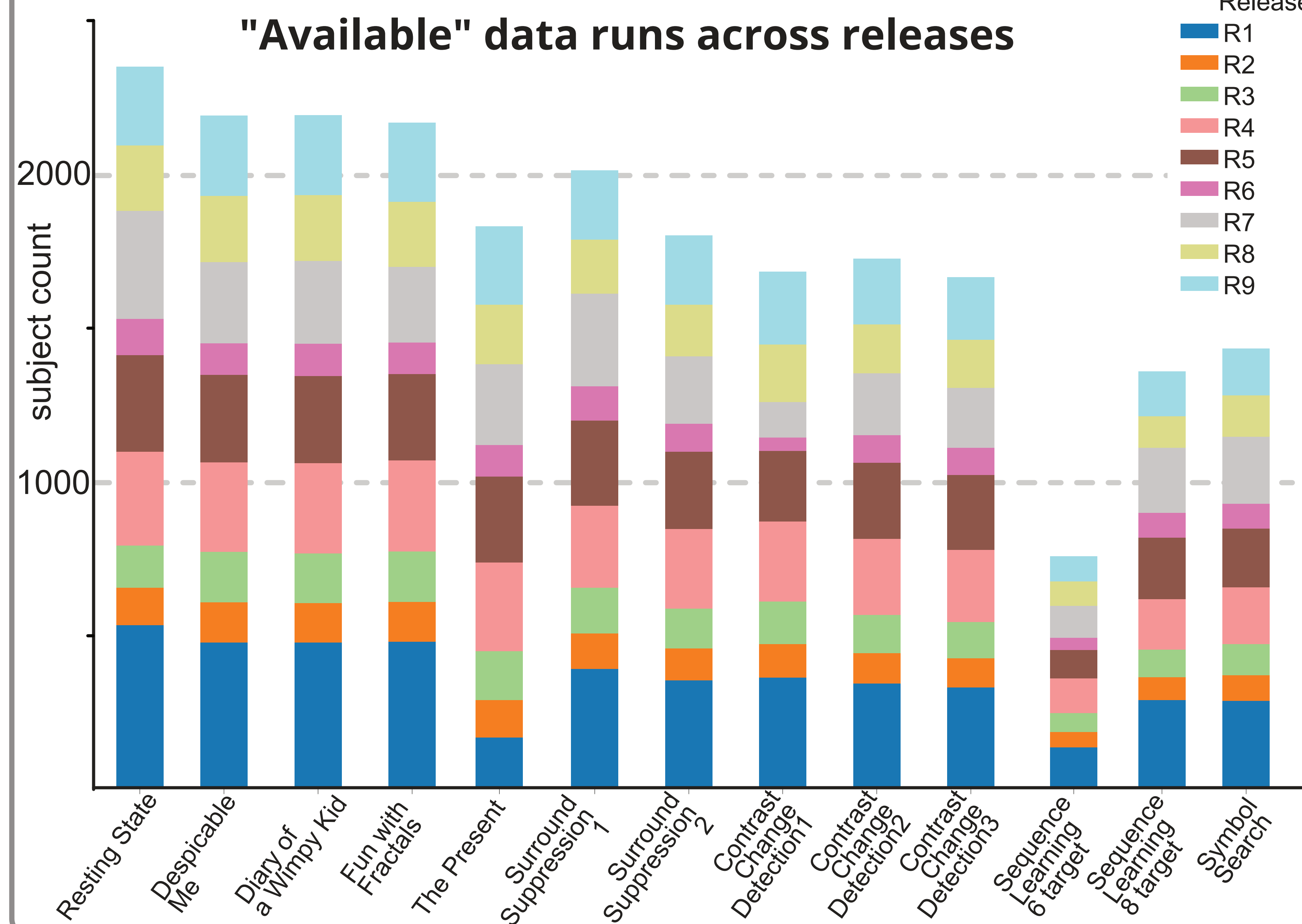
BIDS & Hierarchical Event Descriptors (HED) provide a FAIR structure for data sharing.

## Data curation

EEG data was augmented with the behavioral responses (seperately shared on HBN), complemented with psychopathological factors, and series of quality chekcs.



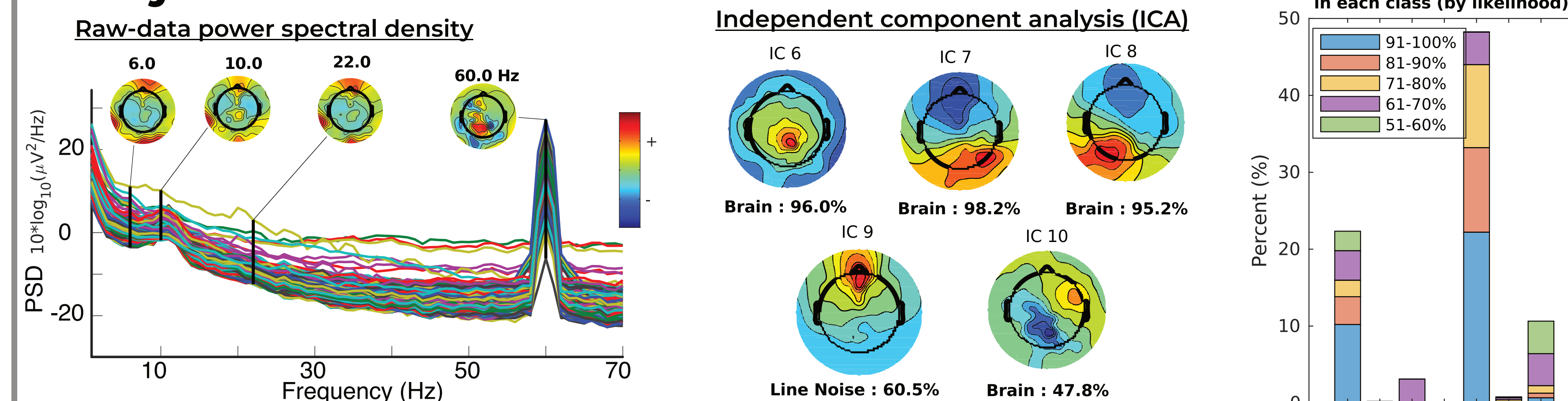
## Data summary



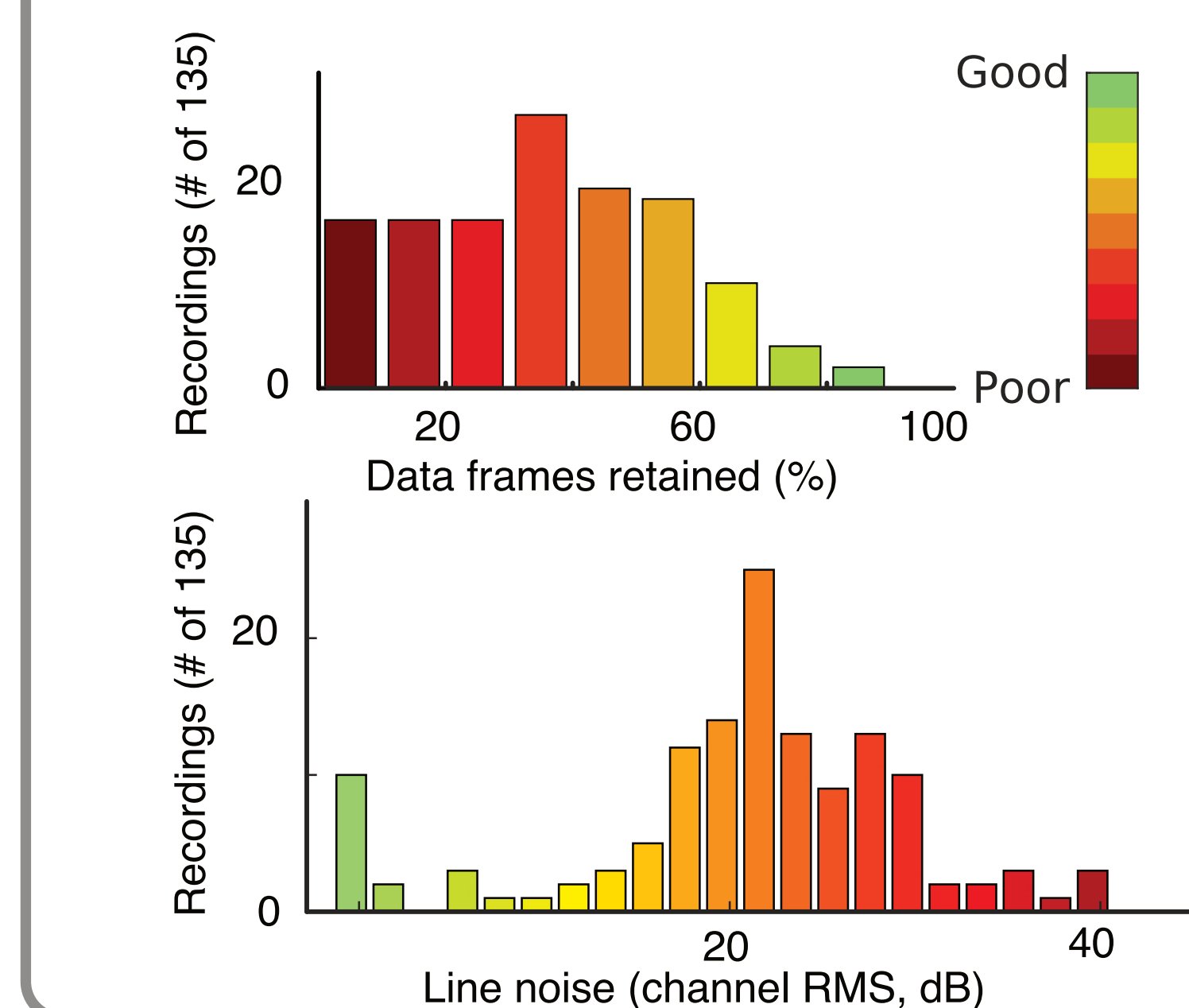
- Three active tasks** (5 runs) with rich HED tags.
- Three passive tasks** (7 runs), resting state, watching videos, SSVEP
- ~2600 subjects** have "available" EEG data.
- Psychopathology metrics** give insight into the developmental state of participants.

## Dataset metrics on NEMAR.org

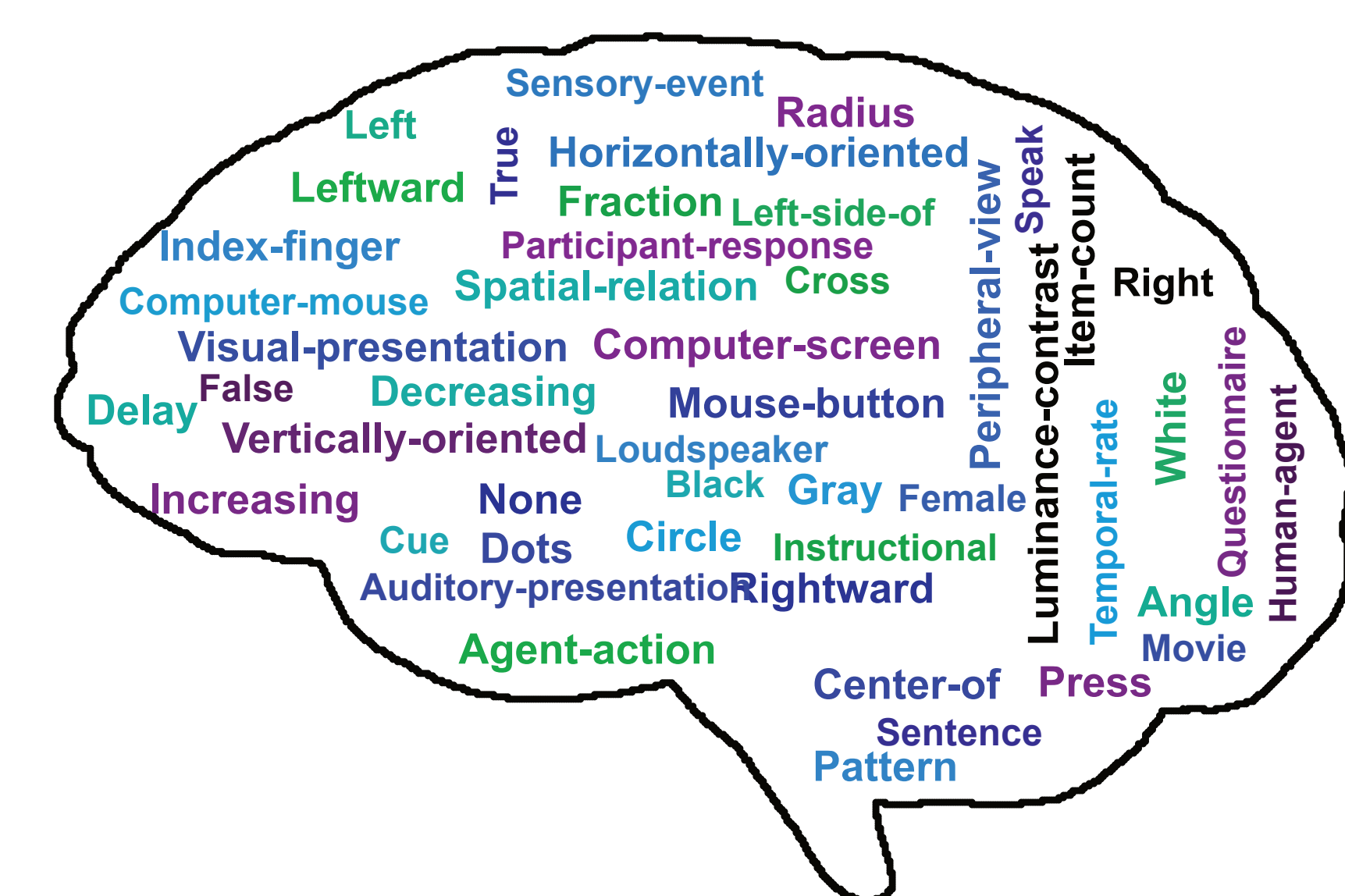
### Subject-level metrics



### Dataset-level metrics



### Hierarchical Event Descriptors (HED) word cloud



- Subject-level metrics** provide granular data-quality insights.
- Dataset-level metrics** give an overall view of the dataset quality, helping researchers decide which datasets to use.
- Hierarchical Event Descriptors (HED)** annotation provide transparent annotation for experiments and event markers.

## Conclusions

- HBN-EEG is the largest and most comprehensive open EEG dataset.
- Free not-for-profit compute/storage resources on **NEMAR** and **NSGportal**[2] helps analyze these datasets.