HED Web

Release 0.0.1

HED Working Group

Jul 05, 2023

CONTENTS:

Downloads

- PDF docs
- Source code

Note: this is a work in progress. More information is coming.

OVERVIEW

This document contains the specification for third generation HED or HED-3G. It is meant for the implementers and users of HED tools. Other tutorials and tagging guides are available to researchers using HED to annotate their data. This document contains the specification for the first official release of HED-3G (HED versions 8.0.0-xxx and above.) **When the term HED is used in this document, it refers to third generation (HED-3G) unless explicitly stated otherwise.**

The aspects of HED that are described in this document are supported or will soon be supported by validators and other tools and are available for immediate use by annotators. The schema vocabulary can be viewed using an expandable schema viewer.

All HED-related source and documentation repositories are housed on the HED-standard organization GitHub site, https://github.com/hed-standard, which is maintained by the HED Working Group. HED development is open-source and community-based. Also see the official HED website https://www.hedtags.org for a list of additional resources.

The HED Working Group invites those interested in HED to contribute to the development process. Users are encouraged to use the *Issues* mechanism of the hed-specification repository on the GitHub hed-standard working group website: https://github.com/hed-standard/hed-specification/issues to ask for help or make suggestions. The HED discussion forum https://github.com/hed-standard/hed-specification/discussions is maintained for in depth discussions of HED issues and evolution.

Several other aspects of HED annotation are being planned, but their specification has not been fully determined. These aspects are not contained in this specification document, but rather are contained in ancillary working documents which are open for discussion. These ancillary specifications include the HED working document on spatial annotation and the HED working document on task annotation.

1.1 1.1. Scope of HED

HED (an acronym for Hierarchical Event Descriptors) is an evolving framework that facilitates the description and formal annotation of events identified in time series data, together with tools for validation and for using HED annotations in data search, extraction, and analysis. This specification describes the official release of third generation of HED or HED-3G, which is HED version 8.0.0.

Third generation HED represents a significant advance in documenting the content and intent of experiments in a format that enables large-scale cross-study analysis of time-series behavioral and neuroimaging data, including but not limited to EEG, MEG, iEEG, eye-tracking, motion-capture, EKG, and audiovisual recording. In principle, third generation HED might be extended or adapted to annotate events in any other type of ordered or time series data.

Specifically, the goal of HED is to allow researchers to annotate what happened during an experiment, including experimental stimuli and other sensory events, participant responses and actions, experimental design, the role of events in the task, and the temporal structure of the experiment. The resulting annotation is machine-actionable, meaning that it can be used as input to algorithms without manual intervention. HED facilitates detailed comparisons of data across studies.

HED annotations may be included in BIDS (Brain Imaging Data Structure) datasets https://bids.neuroimaging.io as described in Chapter 6: Infrastructure of the HED specification.

1.2 1.2. Brief history of HED

HED was originally proposed by Nima Bigdely-Shamlo in 2010 to support annotation in HeadIT and early public repository for EEG data hosted by the Swartz Center for Computational Neuroscience, UCSD (Bigdely-Shamlo et al. 2013). HED-1G was partially based on CogPO (Turner and Laird 2012).

Event annotation in HED-1G was organized around a single hierarchy whose root was the *Time-Locked Event*. Users could extend the HED-1G hierarchy at its deepest (leaf) nodes. First generation HED (HED-1G, versions < 5.0.0) attempted to describe events using a strictly hierarchical vocabulary.

HED-1G was oriented toward annotating stimuli and responses, but its lack of orthogonality in vocabulary design presented major difficulties. If *Red/Triangle* and *Green/Triangle* are terms in a hierarchy, one is also likely to need *Red/Square* and Green/Square* as well as other color and shape combinations.

HED-2G (versions 5.0.0 - 7.x.x) introduced a more orthogonal vocabulary, meaning that independent terms were in different subtrees of the vocabulary tree. Separating independent concepts such as shapes and colors into separate hierarchies, eliminates an exponential vocabulary growth due to term duplication in different branches of the hierarchy.

Parentheses were introduced so that terms could be grouped. Tools for validation and epoching based on HED tags were built, and large-scale cross-study "mega-analyses" were performed. However, as more complicated and varied datasets were annotated using HED-2G, the vocabulary started to become less manageable as HED tried to adapt to more complex annotation demands.

In 2019, work began on a rethinking of the HED vocabulary design, resulting in the release of the third generation of HED (HED-3G) in August 2021. HED-3G represents a dramatic increase in annotation capacity, but also a significant simplification of the user experience.

New in HED (versions 8.0.0+).

1. Improved vocabulary structure

2. Short-form annotation

3. Library schema

4. Definitions

5. Temporal scope

6. Encoding of experimental design

Following basic design principles, the HED Working Group redesigned the HED vocabulary tree to be organized in a balanced hierarchy with a limited number of subcategories at each node. (See the expandable schema browser to browser the vocabulary and explore the overall organization. Chapter2:Terminology of the HED specification defines some important HED tags and terminology used in HED.)

A major improvement in vocabulary design was the adoption of the requirement that individual nodes or terms in the HED vocabulary must be unique. This allows users to use individual node names (short form) rather than the full paths to the schema root during annotation, resulting in substantially simpler, more readable annotations.

To enable and regulate the extension process, the root HED-3G head schema specified here includes, for the first time, *HED library schema* to extend the HED vocabulary to include terms and concepts of importance to individual user communities – for example researchers who design and perform experiments to study brain and language, brain and music, or brain dynamics in natural or virtual reality environments. The HED library schema concept may also be used to extend HED annotation to encompass specialized vocabularies used in clinical research and practice.

HED-3G also introduced a number of advanced tagging concepts that allow users to represent events with temporal duration, as well as annotations that represent experimental design.

1.3 1.2. Goals of HED

Event annotation documents the things happening during data recording regardless of relevance to data analysis and interpretation. Commonly recorded events in electrophysiological data collection include the initiation, termination, or other features of **sensory presentations** and **participant actions**. Other events may be **unplanned environmental events** (for example, sudden onset of noise and vibration from construction work unrelated to the experiment, or a laboratory device malfunction), events recording **changes in experiment control** parameters as well as **data feature events** and control **mishap events** that cause operation to fall outside of normal experiment parameters. The goals of HED are to provide a standardized annotation and supporting infrastructure.

Goals of HED.

1. **Document the exact nature of events** (sensory, behavioral, environmental, and other) that occur during recorded time series data in order to inform data analysis and interpretation.

2. **Describe the design of the experiment** including participant task(s).

3. **Relate event occurrences** both to the experiment design and to participant tasks and experience.

4. **Provide basic infrastructure** for building and using machine-actionable tools to systematically analyze data associated with recorded events in and across data sets, studies, paradigms, and modalities.

Current systems in neuroimaging experiments do not record events beyond simple numerical (3) or text (Event type Target) labels whose more complete and precise meanings are known only to the experimenter(s).

A central goal of HED is to enable building of archives of brain imaging data in a form amenable to new forms of larger scale analysis, both within and across studies. Such event-related analysis requires that the nature(s) of the recorded events be specified in a common language. The HED project seeks to formalize the development of this language, to develop and distribute tools that maximize its ease of use, and to inform new and existing researchers of its purpose and value.

1.4 1.3. HED design principles

The near decade-long effort to develop effective event annotation for neurophysiological and behavioral data, culminating to date in HED-3G, has revealed the importance of four principles (aka the PASS principles), all of which have roots in other fields:

The PASS principles for HED design.

1. **Preserve orthogonality** of concepts in specifying vocabularies.

2. **Abstract functionality** into layers (e.g., more general vs. more specific).

3. **Separate content** from presentation.

4. **Separate implementation** from the interface (for flexibility).

Orthogonality, the notion of keeping independently applicable concepts in separate hierarchies (1 above), has long been recognized as a fundamental principle in reusable software design, distilled in the design rule: *Favor composition over inheritance* (Gamma et al. 1994).

Abstraction of functionality into layers (2) and separation of content from presentation (3) are well-known principles in user-interface and graphics design that allow tools to maintain a single internal representation of needed information while emphasizing different aspects of the information when presenting it to users.

Similarly, making validation and analysis code independent of the HEDschema (4) allows redesign of the schema without having to re-implement the annotation tools. A well-specified and stable API (application program interface) empowers tool developers.

DOCUMENTATION

2.1 1. HED publications

Explanation of the history, development, and motivation for third generation HED:

> Robbins, K., Truong, D., Jones, A., Callanan, I., & Makeig, S. (2020, August 1). Building FAIR functionality: Annotating events in time series data using Hierarchical Event Descriptors (HED). https://doi.org/10.31219/osf.io/5fg73

Detailed case study in using HED-3G for tagging:

> Robbins, K., Truong, D., Appelhoff, S., Delorme, A., & Makeig, S. (2021, May 7). Capturing the nature of events and event context using Hierarchical Event Descriptors (HED). BioRxiv, 2021.05.06.442841. https://doi.org/10.1101/2021.05.06.442841

2.2 2. Working documents

Mapping of HED terms and their descriptions to known ontologies is:

> HED-3G Working Document on Ontology mapping https://drive.google.com/file/d/13y17OwwNBlHdhB7hguSmOBdxn0Uk4hsI/view?usp=sharing

Two other working documents hold portions of the HED-3G specification that are under development and will not be finalized for Release 1:

> HED-3G Working Document on Spatial Annotation https://docs.google.com/document/d/1jpSASpWQwOKtan15iQeiYHVewvEeefcBUn1xipNH5-8/view?usp=sharing

> HED-3G Working Document on Task Annotation https://docs.google.com/document/d/1eGRI_gkYutmwmAl524ezwkX7VwikrLTQa9t8PocQMlU/view?usp=sharing

2.3 3. Schema viewers

The HED schema is usually developed in `.mediawiki` format and converted to XML for use by tools. However, researchers wishing to tag datasets will find both of these views hard to read. For this reason, we provide links to three versions of the schema. The expandable HTML viewer is easier to navigate. Annotators can also use CTAGGER, which includes a schema viewer and tagging hints.

Table 1: HED web-based schema vocabulary viewers.

Viewer	Link
Expandable HTML	https://www.hedtags.org/display_hed.html?version=8.0.0
Mediawiki	https://github.com/hed-standard/hed-specification/blob/master/hedwiki/HED8.0.0.mediawiki
XML	https://github.com/hed-standard/hed-specification/blob/master/hedxml/HED8.0.0.xml

2.4 4. HED Websites

The following is a summary of the HED-related websites

Table 2: HED websites.

Description	Site
Information and documentation	
HED organization website	https://www.hedtags.org
HED organization github	https://github.com/hed-standard
HED specification repository	https://github.com/hed-standard/hed-specification
Examples of HED annotation	https://github.com/hed-standard/hed-examples
HED documentation website	https://github.com/hed-standard/hed-standard.github.io
HED Python resources	
Python code repository	https://github.com/hed-standard/hed-python
Python validator and tools	https://github.com/hed-standard/hed-python/tree/master/hedtools
HED JavaScript resources	
HED JavaScript code	https://github.com/hed-standard/hed-javascript
BIDS validator	https://github.com/bids-standard/bids-validator
HED Matlab resources	
Matlab source code	https://github.com/hed-standard/hed-matlab
Annotator resources	
CTAGGER executable jar	https://github.com/hed-standard/hed-java/raw/master/ctagger.jar
CTAGGER repository	https://github.com/hed-standard/CTagger
Java repository	https://github.com/hed-standard/hed-java
Online HED tools	
Online website	https://hedtools.ucsd.edu/hed
Docker deployment	https://github.com/hed-standard/hed-python/tree/master/webtools/deploy_hed

INDICES AND TABLES

- genindex
- modindex
- search