Submitted to NeuroInformatics 07/31/20

https://osf.io/5fg73/

Building FAIR functionality: Annotating event-related imaging data using Hierarchical Event Descriptors (HED)

Kay Robbins[1], Dung Truong[2], Alexander Jones[1], Ian Callanan[1], Scott Makeig[2]

[1]Department of Computer Science, University of Texas at San Antonio

[2]Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego

Abstract. In fields such as human electrophysiology, high-precision time series data is often acquired in complex, event-rich environments for interpretation of complex dynamic data features in the context of session events. However, a substantial gap exists between the level of event description information required by current digital research archive standards and the level of annotation required for successful meta-analysis or mega-analysis of event-related data across studies, systems, and laboratories. Manifold challenges, most prominently ontological clarity and extensibility, tool availability, and ease of use must be addressed to allow and promote sharing of data with an effective level of descriptive detail for labeled events. Motivating data authors to perform the work needed to adequately annotate their data is a key challenge. This paper describes the near decade-long development of the Hierarchical Event Descriptor (HED) system for addressing these issues. We discuss the evolution of HED, the lessons we have learned, the current status of HED vocabulary and tools, some generally-applicable design principles for annotation framework development, and a roadmap for future development. We believe that without consistent, sufficiently detailed and field-relevant annotations as to the nature of each recorded event, the potential value of data sharing and large-scale analysis in behavioral and brain imaging sciences will not be realized.

1. Introduction

The FAIR (Findable, Accessible, Interoperable, and Reusable) guiding principles formally articulated by Wilkinson et al. (Wilkinson et al. 2016) promote data-sharing and data stewardship with the goal of enabling scientific discovery, evaluation, and reproducibility. These general guidelines apply not only to datasets, but also to algorithms, tools, and workflows. FAIR is expressed in terms of scholarly digital research objects that can be identified with globally unique identifiers and characterized using metadata selected from formal vocabularies. Importantly, these digital objects should be machine-actionable, meaning that the objects themselves can also provide information with varying levels of detail to autonomous data explorers. Widespread development and adoption of FAIR standards across disciplines is needed to create a robust research ecosystem for supporting reproducible science.

1.1 Why HED?

Digital research objects referenced in the FAIR principles are generally larger units ─ specified at the level of a workflow or a study. Practical implementation of annotation standards and related tool development are necessarily left open to data providers and standards groups. Most current domain-relevant community standards supporting FAIR focus primarily on identification, location, top-level data organization, licensing, and data format specification. While standardization at these levels of detail is crucial, in many disciplines it is not sufficient to support meaningful meta-analysis (combining results or result statistics across studies) and mega-analysis (combining raw data or data features across studies) of high-precision time series data collected for cognitive neuroscience, psychology, and biomechanics, often in complex, event-rich environments (Boedhoe et al. 2019). Crucially missing from high-level annotation standards focusing on data organization and format are:

1. A system for specifying the exact nature of events occurring during the experiment (sensory, behavioral, environmental, and other) that may inform data analysis.
2. A standardized, actionable system for describing the relationship of experiment structure to the participant tasks and experimental goals.
3. A mapping relating events both to the experiment design and to participant tasks and experience.

We believe the Hierarchical Event Descriptor (HED) system has the potential to capture this information in both human- and machine-usable forms.

HED is an evolving framework that facilitates the description, annotation, validation, and extraction of events in experimental time series data. First developed at UCSD in the Ph.D. dissertation of Nima Bigdely-Shamlo (Bigdely-Shamlo 2014), the HED system has now been in development for nearly a decade (Bigdely-Shamlo et al.

2013) (Rognon et al. 2013) and has undergone several evolutionary steps (Bigdely-Shamlo et al. 2016) as developers and users gained practical experience with data sharing, annotation, and mega-analysis. This paper focuses on annotation of events in human electrophysiological brain imaging experiments, the discipline for which HED has been developed. However, as HED developers, we are increasingly focused on separating the standardization of vocabularies from the basic structure and tools for processing HED annotations. This independence from vocabulary makes HED tools applicable to other areas for which a discipline-specific annotation vocabulary can be built to extend the basic HED schema – areas as diverse as clinical neurophysiology, sports medicine, consumer behavior, and stock market economics.

1.2 How electrophysiological experiments are structured

To understand why HED is needed, one must understand the structure of experiments involving observation of behavior and physiology with a view to subsequent data analysis. Most human event-related experiments fall into one of three categories: controlled laboratory experiments, clinical assessments, or long-term monitoring. Laboratory experiments are organized in terms of participant task-design variables that are varied during the course of an experiment, typically in a pre-specified way, as the measured physiological signals and/or behavioral records are acquired. Increasingly, the behavioral records may also include continuous recordings of the task and participant environment, which need not be confined to a dedicated laboratory space (Makeig 2009).

1.2.1 Traditional stimulus/response experiments

Many laboratory electrophysiological experiments use a stimulus-response paradigm: perceptually distinct sensory stimuli are presented at precisely recorded times (often with abrupt onsets) and ensuing (and/or preceding) changes in the behavioral and physiological data streams are measured and modeled. Analysis extracts data epochs time-locked to some class of equivalent events so as to assess statistically the relationship of the event or event class to some measure of the physiological and behavioral data.

Traditionally, such experiments also record discrete participant action events such as button presses performed in response to stimulus presentations as motivated by the assigned participant task. Such button press events have typically been analyzed as "instantaneous events" (in both time and space). However, recorded behavioral responses may be more elaborated motor actions, with measurable temporal and spatial extent, whose detailed time courses can be captured by body motion capture and eye tracking systems and/or by response collection devices such as touch screens. Such "mobile brain/body imaging" experiments (Makeig et al. 2009) may allow examination of brain dynamics supporting a fuller range of natural human embodied cognition. In such paradigms, participant action events of interest may be identified during data collection or in *post hoc* analysis. Often, these actions are modeled and recorded in terms of an array of measures, for example the locations and timing of heel-strike and toe-off gait events, or the locations and timing of arm/hand movement onsets, acceleration maxima, and endpoints in reach-to-touch experiments.

Meta/mega-analysis across event-related imaging studies requires detailed specification of both sensory and behavioral events, as well as contextual information about the environment and recording parameters, task design, control variables, and the task behavioral imperatives prompted by different classes of experiment events. Additional challenges are posed by the recent trend towards more natural recording conditions in which participants may listen to fluent speech, watch movies, perform ambulatory tasks in a virtual, augmented, or actual reality laboratory environment, may interact in some ways with other participants, or even participate in real-life activities. The temporal complexity and detail of the electrophysiological, biomechanical, and other records may make their full interpretation dependent on knowing both the nature and timing of participant sensory experience, behavior, and intent.

Traditionally, event-related electrophysiology data has been analyzed by collecting experimental events in a few categories (e.g., *Targets* vs. *Nontargets*) based on some sensory features and/or task-related significance. However, the primary role of the brain can be viewed as informing and instigating behavioral action plans appropriate to the specific then-existing context – both the preceding event context and the subject's assessment of it. Such trial-to-trial variations have been less frequently studied than on-average brain/body responses to significant events, and yet flexibility in context-dependent responses to opportunities and challenges is intrinsic to human behavior. Unfortunately, equivalent event-within-context occurrences in most studies are rare. Hence, most studies do not have enough statistical power to support detailed hypotheses at more fine-grain levels of analysis. A straightforward response to this dilemma is to collect and analyze much larger quantities of data in each paradigm, a solution that is quite often impractical to propose or to carry out. An alternative is to draw existing data from previous studies. Applied to the diversity of events produced, observed, and annotated in these studies, both classical and new statistical methods may be used to reveal new, more detailed information about human cognition, behavior, and its supporting brain dynamics. To do this, however, having a detailed description of each recorded event is essential – hence the need for HED.

1.2.2 Clinical assessments

Electrophysiological recordings performed for clinical assessments are typically stored with detailed information about patient medical state in addition to standard subject trait metadata such as age, gender, and handedness. In much current clinical practice, events of interest are patterns in the recorded data that are assessed visually by the experienced clinician for a range of clinical signs. Mega-analyses across studies, for developing biomarkers for diagnostic applications, need machine-actionable versions of this complex metadata. Ideally, this metadata has been acquired as part of the experiment control and data recording process, and/or when the practitioner reads a clinical data recording and annotates the presence or absence of indicative features.

1.2.3 Long-term monitoring

Long-term monitoring applications often insert event markers into the data record *post hoc*, either manually as clinical notes and/or by integrating events noted in other, concurrently-recorded data streams including patient behavior and/or video recordings. Because of the long duration of these recordings, the context and environment as well as the patient state may change many times throughout the observation period. A central clinical objective is to note and advise clinicians of clinically relevant changes in patient condition. Scientific studies of such data should annotate and integrate as much of this information as possible to enable downstream analyses including development of more advanced diagnostic and state monitoring applications.

1.3 Community standards for neuroimaging data

The diverse annotation requirements of human electrophysiology and neuroimaging experiments make development of useful community standards quite challenging. The BIDS (Brain Imaging Data Structure) standards group (BIDS homepage n.d.) and its imaging modality subgroups are making a sustained and successful effort to implement FAIR standards (Gorgolewski et al. 2016) (Niso et al. 2018) at the level of data formatting and file organization, with emphasis on simplicity and efficiency. BIDS now has a large community of active user/developers and has quickly become the *de facto* standard for organizing human neuroimaging data. Many of the major neurophysiological software tools and archives now support or will soon support BIDS-formatted data. Work continues on BIDS standardization of formats for derived data as well as for auxiliary streams such as eye tracking data. Public data repositories such as *OpenNeuro* (OpenNeuro hompage n.d.) and computational portal sites such as *NEMAR* (NEMAR homepage n.d.) and *brainlife* (Brainlife homepage n.d.) organize their shared data in BIDS format. The BIDS group also supports development of some standardized processing pipelines (as containerized BIDS Apps) for fMRI and other types of data.

In 2019, BIDS (v.1.2.1) adopted HED as its event annotation standard to the extent that it allows (but does not require) users to incorporate HED event annotations as a column in their event spreadsheets. BIDS also allows researchers to include additional columns in the event files for “event codes” to accommodate the home-grown labels neurophysiology researchers typically use to categorize the different types of events in their experiments. Researchers can provide additional JSON files that include free-form text descriptions and/or HED annotations for each event code. Tools can use these JSON files to create full HED-based event annotations during analysis.

1.4 How to achieve analysis-enabling large-scale data sharing?

BIDS and the inclusion of HED event annotations are important steps in establishing open standards for *analysis-enabling* data event annotation. However, much practical work, on both study structure and event annotation, is still needed to achieve effective data-sharing for meta/mega-analysis in neurophysiological and other experiments involving event-related data analysis. The sharing gap is particularly evident for experiments using technologies that produce high-resolution time series records such as electroencephalography (EEG), magnetoencephalography (MEG), and electrocorticography (ECoG or iEEG). In such experiments, event-related brain or brain/body dynamics are the central focus of interest, yet experimental datasets released under current standards often do not contain essential information needed for large-scale analysis to reach its potential.

The barriers to achieving the goals of true analysis-enabling large-scale sharing and mega-analysis of neurophysiological and other event-related data are two-fold: *articulation* and *motivation*. Articulation barriers are technical in nature. The event annotation system must be sufficiently expressive to adequately retain the information needed for within- and between-study analyses, using standard tools and vocabulary rather than laboratory shorthand (e.g., Target) or meaningless (e.g., event-type 13) designations. Further, the annotation system must be capable of capturing both the structure and intent of the experiment, providing mappings of this structure information to the events and to the data in a form that is both human-comprehensible and machine-actionable, while also annotating task design-intended relationships of task events to one another, in particular intended relationships between sensory and behavioral events.

Motivational barriers are even more challenging, since this level of annotation rises well beyond current institutional data-sharing standards. The annotation process must provide clear value added, ideally enabling the researchers to thereby make use of tools that can extract more information from their data *in the context* of larger collections of existing data from a variety of related (not identical) experiments. In any case, the process must not prove so difficult that the perceived cost of performing annotation exceeds its perceived benefit.

Work on the HED system over the past decade has proceeded with an eye towards overcoming both the *articulation* and *motivation* barriers to achieving true data-sharing and enabling mega-analysis in the human electrophysiology and related fields. This paper describes the evolution of HED, explains how recent HED developments have enabled progress towards these goals, and presents a roadmap forward to make adequate annotation of event-related data a reality. We recount some lessons learned during the development and first application of HED to large-scale mega-analysis and reflect on what these lessons suggest as guidance for future development.

2. Hierarchical Event Descriptors (HED) – a history

HED is a hierarchically-structured vocabulary that allows experimental events, experiment design, and control variables to be annotated. Hierarchical event descriptors (HED *tags*) are path strings (much like directory paths on a computer). Event annotations (HED *strings*) consist of comma-separated lists of such strings. Building HED tag ontologies as shallow tree structures aims to provide meaningful organization of related terms into easily searched subcategories. Tag text and prefix matching are easily implemented, allowing effective search, extraction, and combination during annotation and analysis. Full HED string annotations can be validated for compliance against a specified vocabulary – a known HED schema, possibly extended by one or more schemas, defining term vocabularies relevant to a particular field of use. HED (HED homepage n.d.) is managed by the HED-standards organization (hed-standard homepage n.d.). In addition to online support for tag validation, open-source MATLAB and Python tools are available for validation and tag-based searching. A user-friendly GUI-Based MATLAB EEGLAB plugin is also available to support annotation.

The original HED-1 specification and supporting tools have undergone several structural revisions as first users gained experience with annotating data. The following sections describe those reorganizations, why restructuring was necessary, and what we have learned about how to effectively organize HED vocabularies.

2.1 HED-1: Path strings from a tree-structured controlled vocabulary

HED was introduced in 2013 to support annotation of experimental events in HeadIT (HeadIT homepage n.d.), an early public repository of EEG data hosted by the Swartz Center for Computational Neuroscience, UCSD (Bigdely-Shamlo et al. 2013). HED-1 was partially based on CogPO (Turner and Laird 2012). Event annotation in HED-1 was organized around a single hierarchy whose root was the *Time-Locked Event*. Users could extend the HED-1 hierarchy at its deepest (leaf) nodes to provide more details (HED-1 view n.d.). Several EEG studies were successfully annotated for distribution, and first analyses applying HED tools to the repository datasets were demonstrated.

Unfortunately, once we began annotating more complex datasets, we encountered a fundamental design flaw in HED-1 that we refer to as the "color Red" problem. This difficulty is illustrated in the next examples. For simplicity, we focus only on the annotation of stimulus shape and color, although HED supports descriptions of further details such as size and location.

Example 1. A HED-1 annotation for a visual stimulus presentation of a red triangle.

Time-Locked Event/Stimulus/Visual/Shape/Triangle,
Time-Locked Event/Stimulus/Visual/Uniform color/Red

HED-1 was able to successfully capture simple events, such as the stimulus of Example 1. However, efforts to tag more complex events – for example, the simultaneous presentation of *both* a red triangle and a green square yielded ambiguous HED strings, as shown in the next example.

Example 2. An ambiguous HED-1 annotation for a visual stimulus presentation consisting of both a red triangle and a green square:

Time-Locked Event/Stimulus/Visual/Shape/Triangle,
Time-Locked Event/Stimulus/Visual/Uniform color/Red,
Time-Locked Event/Stimulus/Visual/Shape/Rectangle/Square,
Time-Locked Event/Stimulus/Visual/Uniform color/Green

In HED-1, tags could appear in any order, but there was no mechanism for HED string parsers to recognize some tags as modifying other tags. Annotators solved this problem by extending the hierarchy:

Example 3. An unambiguous HED-1 annotation for the visual presentation of a stimulus consisting of *both* a red triangle and a green square:

Time-Locked Event/Stimulus/Visual/Shape/Triangle/Red,
Time-Locked Event/Stimulus/Visual/Shape/Rectangle/Square/Green

Other developers might address the same problem by emphasizing color before shape.

Example 4. An alternate annotation HED-1 annotation for the visual presentation of a stimulus consisting of a red triangle and green square presentation:

Time-Locked Event/Stimulus/Visual/Uniform color/Red/Triangle,
Time-Locked Event/Stimulus/Visual/Uniform color/Green/Square

Clearly, the complete tree hierarchy model of HED-1 was insufficient not only in terms of resolving ambiguity, but also in terms of expressiveness. If annotators wanted to include some level of detail about color in the annotation of an event representing a participant response action (e.g., to annotate, "The participant presses the red button."), they would be stuck including tags with leading *Time-Locked Event/Stimulus* strings unless they also added color terms to the *Participant response* branch of the schema. In this way, adjectives such as *Red* proliferated throughout the schema hierarchy, resulting in an explosion of replicate terms and an ambiguous parsing problem during downstream analysis.

2.2 HED-2: Orthogonality and abstraction

The "color Red" problem demonstrated that adjectival concepts such as color and location are most typically descriptive properties or attributes rather than categorical object subtypes. In other words, *Attribute* tags should be in a separate category from nominative tags representing objects such as *Square* and *Triangle*. Attributes should also be separated from high-level event concepts such as *Stimulus presentation* and

Experiment control. This realization led to HED-2, a major HED redesign focused on removing ambiguity and improving expressiveness (Bigdely-Shamlo et al. 2016).

The idea of the redesign was *to* ***group independently varying "orthogonal" concepts into separate hierarchies***. In HED-2, top level hierarchical groupings roughly correspond to nouns (*Event*, *Participant*, *Paradigm*, *Experiment context*), adjectives (*Attribute, Sensory presentation*), and verbs (*Action*). HED-2 also introduced arbitrary levels of nested parentheses to group attributes with the items they modify (HED-2 view n.d.).

Example 5. A HED-2 annotation for a visual stimulus presentation consisting of both a red triangle and a green square. (The text to right of arrows is commentary, not part of HED. The extra spacing is for presentation and ignored by HED.)

Event/Category/Experimental stimulus,	⇐ It is a stimulus event
Sensory presentation/Visual,	⇐ The stimulus modality is visual
(*Item/2D shape/Triangle*, *Attribute/Visual/Color/Red*),	⇐ A red triangle is displayed
(*Item/2D shape/Rectangle/Square*, *Attribute/Visual/Color/Green*)	⇐ A green square is displayed

HED-2 has many other refinements including the addition of *unit classes* and validation of all the *units* associated with numerical values that are specified in the BIDS standard.

2.3 HED and BIDS: Potential for widespread adoption

As HED-2 enables annotation of events at levels of detail required for effective event-related data mining, in 2019 the BIDS standards (v1.2.1) adopted HED-2 as its (to date only) method for detailed annotation of events beyond freeform text description (Niso et al. 2018) (Pernet et al. 2019) (Holdgraf et al. 2019). The rapid adoption of the BIDS standards in the worldwide neuroimaging community means that a mature and stable HED platform, including adequate and simple-to-use tools, could soon find widely increasing use and applications.

2.4 HED Tools

During the evolution of HED-1 and HED-2, several software tools were developed to build HED usability (hed-standard homepage n.d.). Work on the software infrastructure assured the independence of HED validators from the particular version of the HED schema used for the annotation. This ***separation of implementation from interface*** allows any appropriately formatted controlled vocabulary to be validated without changing the validation infrastructure. Tools were also developed to assist users in annotating events in existing datasets, principally CTAGGER (Rognon et al. 2013), a user-friendly GUI-based tool for annotating events in EEG or other experiment data. A *pop_hedepoch* function for EEGLAB (Delorme and Makeig 2004) and other open-source tools allowed users to select EEG or other time series data epochs time-locked to events annotated with specified combinations of HED tags.

2.5 First application to EEG mega-analysis

A large-scale multi-study mega-analysis of HED-2 tagged data (Bigdely-Shamlo et al. 2019a) demonstrated that time-locked features of trial-averaged event-related potentials (ERPs) associated with HED strings containing the same HED tags were significantly more similar than ERPs not associated with such tags. Without the work performed in this project to add detailed and consistent HED tag annotations to the event records, these cross-study comparisons would have been highly laborious if at all feasible. This experience with large-scale, automated analysis showed us both the strengths and limitations of HED-2 annotation, leading to our current effort to further restructure and evolve the HED system to address usability limitations, a development process to build and release HED-3.

2.6 HED development supporting organization and formats

HED development currently uses the HED organization repository (hed-standard homepage n.d.) with the Github fork-pull mechanism for proposing schema changes and updating schema vocabulary. The underlying HED schema is stored in *XML* format for all machine processing purposes. The schema is also stored in a human-readable WYSIWYG *.mediawiki* format, making it easier for users to edit. We have built functions to convert back and forth between *.mediawiki, XML*, and *JavaScript/HTML* formats. We also support an online schema validator that checks various aspects of schema structure and an online converter that converts between *.mediawiki* and *XML*. A J*avascript*/*HTML* tool has also been built to display the schema in an interactive, expandable form in web browsers, facilitating the review of ongoing changes. In the next section, we describe the new HED-3 design and development.

HED-3: Performing analysis-enabling annotation

HED-3 clarifies and simplifies the structure of the top-level HED tag hierarchy to better-support annotation and readability. It also begins to expand the scope of the high-level HED schema and the precision of the HED syntax. In particular, HED-3 better supports specification of *experiment design, structure,* and *intent* as well as the *participant task and expectation* as part of the annotation process. Why should this information be an essential part of event annotation? Because the aspects and attributes of experimental events that are most important to document and apply in subsequent analysis are their relationships to the *participant task* (and thereby, to participant cognition and to experimenter purpose) as well as to the event temporal context. These relationships are intrinsically connected to the *experiment design and structure*.

A key concept underlying HED-3 is its **unique mapping rule**. The HED-3 requires that the individual nodes (terms) appear no more than once in the schema. While this requirement may somewhat complicate the schema-design process, it permits great improvements in usability for HED end-users as explained below. Further, HED-3 introduces the concept of user community-specific **Schema Libraries** and also further develops the HED tools to maximize ease of use. Though extensive, this reorganization does not impact the use of HED in BIDS, since the HED tools can validate against any specified HED schema.

3.1 Enhancing user usability.

A structural review of HED-2 revealed many opportunities for making HED more user-friendly and enhancing its usability. These changes (accomplished or proposed) affect four distinct stages of HED-related activity: 1) annotating data, 2) reviewing existing data annotations, 3) performing analysis of HED-annotated data, and 4) designing and extending HED schema vocabularies.

3.1.1 Annotating data

During the process of annotating data, users must be able to easily find suitable tags in the schema and to immediately understand the relationship of these tags to other tags in the hierarchy. Thus, top-level HED tags should to be conveniently organized into coherent, compact, and balanced hierarchies for easy selection by users. When tagging data, users must be able to browse the schema vocabulary conveniently and select tags (for example, in an expanding accordion view) to add to an event annotation. The user must also be able to easily study the tags that are currently assigned to an event and take advantage of "smart" suggestions as to which additional tags to include in the annotation.

Sparse hierarchies. Currently the prolixity of the HED-2 schema itself was a barrier to ready usability. A large set of legacy tags had been added, more through a process of accretion than through strategic planning, as new and more complex experiments were tagged. As a result, some branches of the HED schema hierarchy became quite deep and detailed while other branches remained relatively bare.

Usability guidelines for bullet points and pull-down lists suggest limiting sub-categories to seven, and certainly to no more than ten items (Carliner 1987). In developing HED-3, we have therefore pared the main HED schema hierarchy and reduced the number of top-level category tags to around seven: *Event* (5), *Item* (4), *Attribute* (7), *Action* (3), *Participant* (4), *Custom* (0), where numbers in parentheses here give the number of subcategories in the current HED-3. Schema details are evolving as the release date for HED-3 approaches. For an expandable view of the HED-3 hierarchy, see (HED-3 view n.d.).

Supporting tool. The current CTAGGER ("community tagger") tool for annotating data provides a graphical user interface (GUI) to assist HED users in the annotation process. The main interface consists of two parts: a list of event items (e.g., event types) to be annotated, and the HED schema. Users can browse the HED schema in an expanding accordion view and select tags to add to the appropriate event annotations. The GUI is designed to ease visual inspection of tagged events and to guide users during tagging by showing recommended tags. Various features are added to make the tagging and reviewing process more efficient including, but not limited, to tag search capacity, error highlighting and quick editing, copying tags, and a report tool that displays shared and unique tags among selected tagged events. A planned improvement for CTAGGER is a "smart" suggestions feature that recommends additional tags to include in the annotation based on previously-selected tags.

3.1.2 Reviewing tagged data

In HED-2, strings created during the annotation phase are built and reviewed only in the fully elaborated **Annotation view** format (see Example 5 above). However, when researchers want to review how events have been annotated, the complete HED strings can be long and difficult to read or grasp quickly. Potential further inclusions – links to global identifiers and/or other items that could make HED annotations more readily "machine-actionable" – would further complicate HED string review and annotation. Thus, we realized that HED-3 should include a streamlined HED string representation for visual inspection, while maintaining a one-to-one correspondence between the readable "short-form" HED strings and the underlying full "long-form" HED string contents. The HED-3 **Reading view** refers to a shorthand HED string display syntax optimized for quickly viewing the contents of completed HED tag strings. Its concise representations are designed to be much easier to read and review.

Example 6. HED-3 Reading view for a visual stimulus presentation event consisting of both a red triangle and a green square.

Sensory event, Visual, (Triangle, Red), (Square, Green)

Example 6 contains the same information as the HED-2 annotation in Example 5, provided it has an accompanying schema that satisfies the HED-3 unique mapping rule. In HED-3, each term or node can appear at most once in the HED schema. This allows HED-3 tools to read, present, and translate HED strings interchangeably between Annotation and Reading views. The Reading view may also be useful for displaying string information *during* annotation. Example 6 uses a single end node to represent each tag, but for visual clarity the user can use as much of the relevant tag suffix as desired (e.g., *Sensory/Visual* rather than just *Visual*). Example 7 shows the full HED-3 Annotation view of the tag string presented in Example 6.

Example 7. HED-3 Annotation view for a visual stimulus presentation event consisting of both a red triangle and a green square of Example 6.

Event/Sensory event,	It is a sensory environment event
Attribute/Sensory/Visual,	The sensory modality is visual
(*Item/Object/Shape/2D/Triangle*, *Attribute/Sensory/Visual/Color/Red*),	A red triangle appears
(*Item/Object/Shape/2D/Rectangle/Square*, *Attribute/Sensory/Visual/Color/Green*)	A green square also appears

Notice that although the HED-2 and HED-3 annotations in Examples 5 and 7 are similar, the underlying HED-3 schema of Example 7 has been slightly reorganized to satisfy uniqueness. The event type of Example 7 indicates that it is an environmental sensory event occurring in the participant's field of view, rather than specifying it is an experimental stimulus as in Example 5. In Example 7, the relationship of this sensory event to the intent of the experiment (e.g., that the event is an integral part of the experiment structure) should be further specified by additional tags (as discussed below). This separation between sensory events and experiment design is in recognition that brain and behavioral dynamics may be affected by sensory input in complex ways.

3.1.3 Analyzing annotated data

The HED-3 string **Analysis view** refers to the internal representation of HED strings used by tools during data search and analysis. The HED-2 tools generally store HED strings in a canonical form, with the string tags organized in a specific order, all in lower case with separating white space removed. Some tools retain pointers within the original string to facilitate building user messages.

The most common data search and analysis task is to identify a set of events, across multiple studies, satisfying some (possibly complex) set of search criteria and to extract data epochs time-locked to those events for further analysis. Subsequently applied analysis tools might further distinguish epoch subgroups or consider relationships among the retrieved epochs in light of more detailed HED string differences.

In addition to supporting common types of data search and collection operations, HED-3 is designed for future applications using more extensive knowledge-integration techniques including natural language processing. We anticipate the HED schema being used to store additional metadata, including globally unique identifiers and links to external resources and knowledge bases. Such links, once created by domain experts, need not be visible during standard HED annotation and review. In fact, globally unique identifiers assigned to HED schema nodes could be stored in separate external databases. (Bigdely-Shamlo et al. 2016) described the natural correspondence between HED schema elements and the Resource Framework Description (RDF standard 2015) for interchange of web-linked data. We have shown that converting HED strings to RDF mappings is a practical possibility.

3.1.4 Extending the HED schema vocabulary

One of the difficulties we encountered in HED-2 was the tendency, when faced with a new concept, to add overly-specific terms and jargon to the schema hierarchy – for example, adding musical terms to tag events in music-based experiments, video markup terms for experiments involving movie viewing, traffic terms for experiments involving virtual driving, and so forth. Clinical fields using neuroimaging also have their own specific vocabularies of terms for noting data features of clinical interest (e.g., *seizure*, *sleep stage IV*). Including these discipline-specific terms quickly makes the top-level HED schema unwieldy and less usable by the broader user community.

In building the HED-3 schema, we have tried to remove terms with an overly-specific scope of use. To accommodate the needs of individual research and clinical subfields, HED-3 instead introduces the concept of **HED Library schema**. To use a programming analogy, when programmers write a Python module, its code does not become part of the Python language. Instead the module becomes part of a library that is used in conjunction with core modules of the programming language. Similar to the design principles imposed on function names and subclass organization in software development, HED Library schema must conform to some basic rules:

1. Every node name must be unique *within* a Library schema.
2. Node names should be meaningful and readily understood by most users.
3. If possible, no schema node should have more than 7 subordinate nodes.
4. Terms that are used independently of one another should be in different sub-trees (orthogonality).

As in Python programming, we anticipate that many different HED schema libraries may be defined and used, in addition to the top-level HED schema, to annotate details of events in experiments designed to answer questions of interest to particular research or clinical communities. Since it would be impossible to avoid naming conflicts across schema libraries that may be built in parallel by different user communities, HED-3 will support schema library namespaces. Using a syntax similar to Python (yet to be fully specified), users will be able to add library tags qualified with namespace designators. For example, in a music/brain imaging experiment, annotation of the presentation of an unusually lengthened note might use the musical term "*fermata*" via a music schema library.

Development of discipline-specific schema libraries, maintenance and extension of the standard HED schema itself, and building effective tutorials on use of HED all necessitate development of graphic tools that can present HED **schema library overviews.** Various mechanisms are implemented or planned to enhance presentation of top-down **Schema view** information that includes suitable graphics to facilitate efficient use of HED schema libraries by users.

3.2 Improving HED expressive capabilities

In addition to usability improvements, HED-3 is introducing a number of structural enhancements that will allow annotators to capture much richer information about the experiments in a form that is both understandable to humans and machine-actionable. The information includes the nature and structure of the experimental control variables, the temporal organization of the experiments, and detailed contextual information describing the conditions present when each event occurs. This section explains these enhancements and illustrates their effects with several examples.

3.2.1 HED-3 *Definition* tags

Typical experimental events share common characteristics. In HED-2 these common characteristics must be tagged separately for each event or event type, making the tagging process repetitive. HED-3 introduces the *Definition* tag to facilitate tag reuse. Users first specify a tag group containing one *Definition* tag and one *Label/#* tag, with other tags that form the definition. Here # represents a placeholder for the name that the user gives to identify this definition. *Definition* tag groups can appear anywhere in a HED event file, but usually definitions are consolidated into an event marking the beginning of the recording. The definition names must be unique for a given dataset.

Within the dataset, users can use the *Label/#* (with # replaced by the actual definition name) as a placeholder for all the tags that appear in the definition. The use of "defined names" can greatly improve the readability of the tag strings. HED parsers will automatically expand these names, replacing the *Label/#* with the contents of the entire corresponding definition tag group (without the *Definition* tag) during validation and analysis.

Example 8. Definition of the setup for viewing to reduce repetition during tagging.

(*Definition, Label/ViewSetup,*	⟸ Define *ViewSetup* as
Visual,	⟸ a visual presentation
(*Screen,*	⟸ on a screen
Distance/100 cm,	⟸ that is 100 cm away
Width/84 cm, Height/68 cm))	⟸ with dimensions 84 cm × 68 cm

Once defined, the tag *Label/ViewSetup* can be used as a shorthand within the dataset to represent the tags of Example 8.

Definition tag groups are useful not only for improving readability and efficiency, but also to provide an essential foundation for specifying events with temporal scope, as explained in the next section.

3.2.2 Representing events with temporal scope

HED-2 introduced *Onset* and *Offset* tags to represent the start and end of events of finite duration, but did not specify how tools should handle such events. In practice, matching *Onset*-*Offset* pairs during analysis proved impossible or ambiguous when the *Onset*-*Offset* pairs were temporally overlapping or when different *Onset*-*Offset* pairs occurred within a dataset. Analysis based on HED-2 annotations thus required filtering out the *Offset* events, thereby removing any information about event duration.

To address these problems, HED-3 introduces the notion of the temporal scope of an event and specifies how HED-based tools should handle event scope annotation to support analysis of events that occur over time. HED-3 events are assumed to occur at a single point in time (i.e., they are "point" events) unless they are given an explicit temporal scope (i.e., they are "scoped" events).

The most direct HED-3 method of specifying scoped events uses *Onset* and *Offset* tags with defined names. Using this method, an event with temporal scope actually corresponds to two point events. The start of the temporal scope is marked by the event in which a (*Label/#*, *Onset*) appears. The end of that temporal scope is marked by the event in which a (*Label/#*, *Offset*) appears. Here # represents the name that appears in the corresponding *Definition*. The following examples illustrate the definition and use of a scoped event representing the playing of a movie.

Example 9. Definition of a label to represent playing a movie on a screen.

(*Definition, Label/PlayMovie*, ⟸ Define the label *PlayMovie*
Sensory event, Visual, Movie, Screen) ⟸ to be playing a movie on a screen

Example 10. Use of a defined label to annotate a scoped event. The event stream then includes presentation of a *Star Wars* movie clip. These events are tagged using the *PlayMovie* label defined above (Example 9):

Sensory event, ⟸ A sensory event (event category)
(*Label/PlayMovie*, *Onset*, ⟸ The *Star Wars* clip starts playing
(*Label/Star Wars*, *ID*/3284)) ⟸ The clip is nicknamed and identified

. . . [The *Star Wars* clip is playing] . . .

Sensory event, ⟸ A sensory event (event category)
(*Label/PlayMovie*, *Offset*) ⟸ The Star Wars clip ends.

Only one label can appear with the *Onset* and the *Offset* tags. Tools match only these labels to determine the scope of the temporal event. That is, (*Label/XXX*, *Onset*) marks the start of a temporal event *XXX*. That event ends at the next point event containing either (*Label/XXX*, *Offset*) or a (*Label/XXX*, *Onset*). In the latter case, this second event marks the end of the previous *XXX* event and the start of a new *XXX* temporal event. Many other events can occur between these two events. The effect of this temporal event on these intervening events is discussed in the next section.

Example 10 may appear to violate the "one label only with *Onset* and *Offset* tags" rule. Notice that the *Label/Star Wars* does not appear at the same level as the *Onset* tag in Example 10. Instead it is in a tag group enclosed in parentheses. Generally, any number of additional tags can be included in the *Onset* event, but for clarity they should be grouped in parentheses. This tag group should not be repeated in the *Offset* event. All of the tags in this tag group are assumed to be applicable for the temporal duration of the event.

The *Duration* attribute tag is an alternative method for specifying an event with temporal scope. The start of the temporal scope is the event in which the *Duration* tag appears. The end of the temporal scope is implicit and may not coincide with an actual event appearing in the dataset record. Instead, tools calculate when the scope ends in the dataset record. Tags grouped with an event *Duration* tag are applied as *Context* during the event scope as explained below. *Duration* tags do not need a defined label.

3.2.3 *Context* tags and scoped events

To make effective use of the information provided by scoped events (i.e., events having annotated duration), the ***Analysis View*** requires tag remapping for events occurring *during* the scoped event. That is, events that occur between the *Onset* and *Offset* pairs for *PlayMovie* should inherit the information that a particular movie is playing without requiring the user to explicitly enter those tags for every intervening event. However, events occurring during an ongoing movie presentation may elicit somewhat different data effects than the actual onset or offsets of the event. HED-3 introduces the *Context* tag to capture this distinction. Analysis tools should evoke a preprocessor that remaps all scoped events into *Onset*, *Offset*, and *Context* events. Example 11 shows a situation in which the subject presses a key, thereby generating an event during playing of the *Star Wars* movie clip of Example 10:

Example 11. An event occurring within the temporal scope of a *PlayMovie* event. The *Context* tag group for this key press event is automatically added during preprocessing.

Subject action,	⇐ This is a subject action event.
Press, Key,	⇐ The subject presses a key.
(*Context*, (*Label*/*PlayMovie*, (*Label*/*Star Wars*, *ID*/3284)))	⇐ The *Star Wars* movie clip is playing.

Events within the temporal scope of nested *Onset-Offset* pairs (and/or *Duration* tags) should inherit *Context* tags from any number of still-occurring scoped events. During analysis tools insert all of these tags in the same *Context* tag group.

The *Context* tag group of an event may also hold many other different types of information including metadata, experiment environmental information, subject state changes, and participant task condition. Instead of tagging each individual event with all the applicable contextual information, the researcher defines the contextual information in *Definition* tag groups, and then uses labelled *Onset-Offset* event pairs (or events with a *Duration* tag) to mark the boundaries of the respective contextual information (e.g., onsets and offsets of long stimulus sequences, changing task conditions). The burden is left to the HED tools to apply the contextual tags appropriately to events upon request during analysis. A library of mapping tools is being developed in both MATLAB and python to facilitate analysis.

3.2.4 Including metadata as *Context* information

In using HED-2 annotations for large-scale data analysis, we found several pieces of information were missing in the HED records. Static metadata such as subject IDs were usually stored somewhere other than in the data recording itself or in the event list. Analysts had to write code for each study used in the analysis to incorporate such data features. In BIDS-archived data, subject IDs may be placed in a recommended (though not required) *participant.tsv* file at the top level of the study hierarchy. A cleaner approach from the viewpoint of analysis is to store relevant metadata directly in the event files. The following examples illustrates how HED-3 allows study or subject metadata to be specified as part of a scoped event, allowing analysis tools to make use of it during data selection and processing.

Example 12. Subject metadata and experimental setup are defined as *SetupInfo*.

(*Definition, Label/SetupInfo,*	⟸ Dataset metadata is defined here
Subject/S2058, Gender/Female,	⟸ This is female subject S2058
Indoors, Sitting,	⟸ She is seated indoors, her EEG recorded
Scanner/Biosemi-64),	⟸ using a 64-channel Biosemi system

Example 13. *SetupInfo* is used as a scoped event.

Experiment control,	⟸ It is an experimental control event
(*Label/SetupInfo, Onset)*	⟸ that is the start of a scoped event.

If no (*Label/SetupInfo*, *Offse*t) event occurs in this data recording, the end of the scoped event defined in Example 13 is the end of the recording. The HED remapping tools copy the information defined in Example 12 into the *Context* tag group of every subsequent dataset event.

If the experimental setup information changes during the experiment, for example by having the subject at some point continue performing their task while standing, then the *Sitting* tag could be moved out of the *SetupInfo* definition in Example 12 and used in the *Onset* tag group of Example 13: (*Label/SetupInfo, Onset, Standing*). A later event tagged with (*Label/SetupInfo, Onset, Standing*) would then mark the moment when the subject stood and invoke an implicit *Offset* of the previous *SetupInfo Onset* event. The new *Standing* context information would be applied to all subsequent events until a corresponding *SetupInfo Offset*, a new *SetupInfo Onset*, or the end of the file occurs.

By making local (changing as the recording progresses) and global (whole session) metadata available in a consistent format across different studies, HED-3 tagging can greatly facilitate meta/mega-analysis. Automated tools may now automatically address a myriad of statistical questions across studies. For example, in studies for which the *Age* trait is provided, is age a significant factor in explaining subject differences in some physiological or behavioral measure? We are developing tools to copy standard BIDS metadata into an initial *SetupInfo* event; for other data formats parallel tools may be developed as needed.

3.2.5 *Condition* tags, control variables, and experimental design

Laboratory experiments generally are performed specifically to understand how measures of interest change as specified independent or control variables are varied. Encapsulating information about prevailing control variable settings using *Condition* tags allows HED-3 to capture this intentional aspect of an experiment in a machine-analyzable form. The *Condition* tag can be used in concert with the HED-3 event scope mechanism described above to record experiment control variables as they are varied during the experiment. The following examples use a Rapid Serial Visual Presentation (RSVP) experiment to illustrate how HED-3 can document this aspect of experimental design.

In RSVP experiments, a sequence of images are presented in succession on a screen at some constant rate and the brain or behavioral response is measured. Our example experiment was designed to measure whether the rate of image presentation affected EEG image response characteristics. The researchers defined a single independent or control variable (presentation rate), which alternated between two level categories (slow and fast).

The experiment used a counter-balanced experimental design. That is, in some experiment blocks the subject saw the slow presentation sequence first and in other blocks the fast presentation. The slow and fast presentation levels were not fixed but were instead selected at random from three different rates for each level. Example 14 shows how to use HED-3 to define the two-level *Rate* control variable.

Example 14. The definition of experiment control variables using a *Condition* tag.

(*Definition, Label/SlowPresentation*,	Define the label *SlowPresentation*
(*Condition*,	as a *Condition* control variable
Rate/#, Visual, Screen))	representing a visual presentation rate
(*Definition, Label/FastPresentation*,	Define the label *FastPresentation*
(*Condition*,	as a *Condition* control variable
Rate/#, *Visual*, *Screen*))	representing a visual presentation rate

Notice that the definitions of *SlowPresentation* and *FastPresentation* in this example do not reference a specific rate. Instead, the # place holder is used. When the specific temporal event is tagged using these definitions, the actual values should be provided.

Example 15. The start of a temporal event representing a slow presentation.

Sensory event,	It is a sensory event
(*Label/SlowPresentation*,	under the Slow Presentation condition
Rate/1 Hz, Onset))	with an actual presentation rate of 1 Hz

If each level of the control variable had represented a fixed value, the actual values would have been included in the definitions. In this experiment, however, presentation rate (in the *SlowPresentation* or *FastPresentation* conditions) is selected at random from three rate value choices. Hence, we defer specifying actual rate values until the actual *Condition* event *Onset* tags when the specific rate for the task block is specified.

3.2.6 *Block* tags and experiment temporal organization

A typical experiment usually consists of a sequence of subject task-related activities interspersed with rest periods and/or off-line activities such as filling in a survey. *Block* tags can be used to represent this temporal organization in a manner similar to the way *Condition* tags are used to represent the control variables in the experimental design.

The temporal organization of the RSVP experiment whose condition variables were defined in Example 14 was as follows. One of the two experimental conditions was selected (e.g., *SlowPresentation* or *FastPresentation*). The subject performed 50 RSVP trials under the selected condition and then was given a brief rest period. After performing another block of 50 trials, the subject was given an offline survey to fill in. The survey was followed by two more blocks of 50 trials each under the other (fast vs. slow presentation) experimental condition. This experiment structure has two types of time blocks – viewing images, and taking the survey. Example 16 shows the definitions of these block types using *Block* tags.

Example 16. Define the temporal blocks for the RSVP experiment.

(*Definition, Label/ViewImage, Block*) ⟸ Define a block called *ViewImage*

(*Definition, Label/TakeSurvey, Block*) ⟸ Define a block called *TakeSurvey*

Block definitions could include other relevant informational tags; these are omitted here for simplicity. After defining the blocks and conditions, the researcher can document both the actual experimental structure of a dataset using scoped events as illustrated by the excerpt of a sample event file, as shown in Example 17.

Example 17. Excerpt of an event file with HED-3 documentation of experimental structure. Experiment times are here shown as whole seconds for visual simplicity. Also, each event should have an event category tag (e.g., *Experimental control*) which is omitted here for simplicity.

Time(s)	HED tags	Explanation
105	(*Label*/*SlowPresentation*, *Rate*/1 *Hz*, *Onset*)	Enter the *SlowPresentation* condition (Ex. 14) using the 1-Hz image presentation rate
120	(*Label*/*ViewImage, Onset*)	Begin a *ViewImage* block
480	(*Label*/*ViewImage, Offset*)	End that *ViewImage* block
585	(*Label*/*ViewImage, Onset*)	Begin another *ViewImage* block
880	(*Label*/*ViewImage, Offset*)	End that *ViewImage* block
900	(*Label*/*SlowPresentation, Offset*)	End the *SlowPresentation* condition
1100	(*Label*/*Take*Survey, *Onset*)	Begin the offline survey block
1350	(*Label*/Survey, *Offset*)	End the offline survey block
1540	(*Label*/*FastPresentation*, *Rate*/5 *Hz*), *Onset*)	Enter the *FastPresentation* condition using a 5-Hz image presentation rate
1650	(*Label*/*ViewImage, Onset*)	Begin a *ViewImage* block
1920	(*Label*/*ViewImage, Offset*)	End that *ViewImage* block
2050	(*Label*/*ViewImage, Onset*)	Begin another *ViewImage* block
2325	(*Label*/*ViewImage, Offset*)	End that *ViewImage* block
2400	(*Label*/*FastPresentation, Offset*)	End the *FastPresentation* condition

At analysis time, HED tools map the block and condition information into *Context* tags for the events occurring between the above paired *Onset-Offset* tags. Additional block-relevant information could be included in the *Onset* tag group, such as the kind of target image the subject is to focus on in that block.

3.2.7 Example dataset structure viewer

In HED-3, structural *Block* and *Condition* tags can be used to provide a wealth of information to automated data search and analysis tools. For example, a data repository could use this information to automatically produce a visualization of the dataset structure via a repository data browsing application. Figure 1 below shows a mock-up visualization for the Example 17 event file.

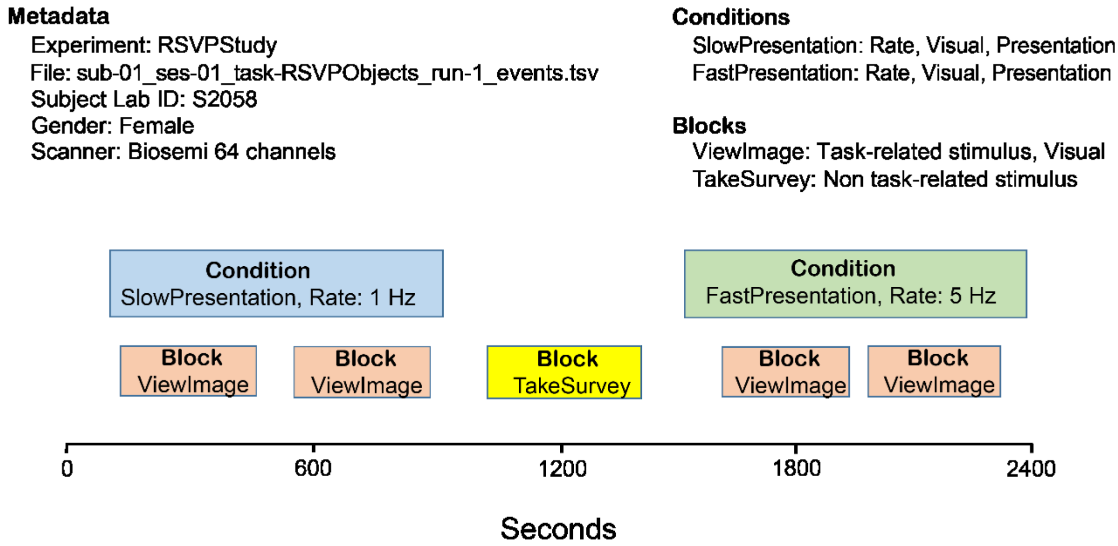

Figure 1. Mock-up of an experiment timeline overview automatically extracted from the event file of Example 17 using the definitions of the *Condition* and *Block* definitions from Examples 14 and 16.

Such a timeline viewer application might be used by researchers to verify that the experiment was actually conducted according to the documented specification. The availability of such annotations might also encourage researchers to more completely document items they might otherwise ignore or forget to tag (such as their administration of a survey between the two groups of blocks as Figure. 1). More details (e.g., selected types of trial events) might be optionally included in the timeline display when/as space permits.

Importantly, the scoped *Condition* and *Block* events make the vital information about experimental structure available at the event level during analysis. Without this information, analysts need to hand-craft code to manually map the needed information in each study. For example, an automated tool could easily be written that would test whether there was a significant difference in given measure (e.g., some power ratio between two brain areas) across all repository studies for conditions that varied a specified control variable (e.g., image presentation rate). One might also test which subject traits were factors accounting for feature variance or whether certain available subject metadata variables such as participant age influenced response times under certain conditions.

The structural information provided by *Condition* and *Block* tags could also be invaluable in analysis of clinical and long-term recording studies that monitor changes in the continuous signal record. Ideally, some or all of the structural tags (as well as many trial event tags) for future experiments using HED tagging could be recorded by the experiment control program as the experiment is conducted and the data are saved.

3.2.8 Documenting task-event-intent relationships – open questions

Often in human neuroimaging, and particularly in human EEG/MEG experiments, the participant ***task*** rather than the stimulus sequence varies across conditions. For example, in one *Condition* the participant may be asked to respond with a button press only to one type of presented stimuli, while in another *Condition* they are to respond only to some other type. Here the stimulus presentation parameters themselves do not change between conditions. HED-3 can express what a subject actually did (e.g., pressed the red button) but does not yet have good semantics for expressing complex relationships and causal linkages between events mediated by the structure of the user task. A complete record of events in an experiment should capture both the detailed intentions and expectations of the subject (as specified in subject task instructions) as well as the stimulation details (as produced by the experiment design).

Our original plan in developing HED-1 and HED-2 was to incorporate the CogPO (Turner and Laird 2012) list of task paradigms, with hopes of linking HED event descriptions to task databases such as the Cognitive Atlas (Cognitive Atlas homepage n.d.). We have removed the resulting *Paradigm* tags from HED-3, however, because the available paradigm nomenclature is not standardized. Text descriptions of tasks in the *Cognitive Atlas* vary in specificity and use only broadly-defined and sub-field specific terminology. These descriptions do not, at present, represent machine-actionable information.

Associating a dataset with a well-known paradigm detailed in the *Cognitive Atlas* or elsewhere remains possible in HED-3 using various informational tags. Currently, HED-3 users can define a *Task* with a *Label* and add tags that specify concepts relating to the task (essentially, listing keywords pertaining to it). This approach is not a true answer to task specification, though it may allow searching across studies for task keywords of interest. *ReproNim* (Keator et al. 2013), a project to harmonize clinical and behavioral data terminology and to provide semantic linkage of BIDS datasets to NIDM terminology, is a promising effort that might provide useful tools for specifying tasks and rectifying naming of metadata variables across studies.

There remain many open questions about how to relate task design and user intent and expectation so as to annotate linkage between experiment events when their conditional linkage involves a complex set of rules (e.g., the subject is expected to press the left button when the most recent stimulus is …). This process is difficult even for well-known tasks such as the *N*-back continuous performance task (Kirchner 1958) in which a subject is to indicate whether each current stimulus matches the one presented *N* stimulus presentations earlier. The issue of event temporal relationships exposes deeper neurological questions. Both brain and behavioral dynamics are shaped not only by intentions but also by prevailing participant expectations, including expectations generated by immediately preceding events. The notion of defining "neighborhoods of influence" for events and using this information in automated tools is likely possible to implement but has been deferred to future HED development. Work to build the necessary HED-3 (or perhaps HED-4) infrastructure is just beginning.

3.2.9 Other prospects for future automation of the tagging process

Development of supporting tools for assisting in annotation is ongoing and critical for good annotation. We have experimented with various “Wizards,” guide systems for setting up and annotating experiments. However, the HED system itself has evolved more rapidly than the applicable tools. Thus, much work needs to be done in this area.

Another important goal is to carry out more of the tagging process in earlier stages in the experiment execution, particularly by making generation of HED tags for stimulation events an active responsibility of experimental control programs. We hope to work with major control program maintainers to add this option. It might also be possible to build some automated tagging facility to capture the logic of the experiment control program, including intended functional relationships between experiment-delivered stimulus events and intended participant motor action events (e.g., “Push the button whenever you see a red square.”). Development of a way to map control programs into “meta-scripts” would facilitate the incorporation of HED annotations of task structure without requiring careful work and understanding by experimenters, but may itself prove difficult. Unfortunately, information gathered from the event log itself would yield only stochastic information (experiment control rules with some degree of uncertainty), and the tools required would be control-program specific. Thus, HED will need a system for specifying task design that is simple enough for any investigator to learn and use easily.

4. Design lessons and a roadmap forward

Our near decade-long effort to develop effective event annotation for neurophysiological and behavioral data, culminating to date in HED-3, has led us to several useful insights (aka four PASS principles, below), all of which have roots in other fields:

1. **P**reserve orthogonality of concepts in specifying vocabularies.
2. **A**bstract functionality into layers (e.g., more general vs. more specific).
3. **S**eparate implementation from the interface (for economy of effort).
4. **S**eparate content from presentation (so as to maintain a unique, sufficient internal representation, while presenting data to users in forms they can more readily review and understand).

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). Similarly, making validation code independent of the schema as per (3) allows redesign of the schema without having to re-implement the validators.

Abstraction of functionality into layers (2) and separation of content from presentation (4) 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.

4.1 HED now and in future.

Much of our design of HED-3 has benefitted from our experience in performing a large, cross-study, HED-tag based mega-analysis (Bigdely-Shamlo et al. 2019a) (Bigdely-Shamlo et al. 2019b) (Robbins et al. 2020) in which we learned not only what worked in HED-2 and what didn't, but what questions we wanted to ask and couldn't answer, as well as what approaches to more complete event descriptions might be most feasible.

We hope to release HED-3 publicly on September 1, 2020; its development is open-source and available for review at the hed-standard repository (hed-standard homepage n.d.). An expandable viewer is available (HED-3 view n.d.). The release will include tools for converting between short-form and long-form views of HED-3 tag strings and for HED string validation. We plan to complete implementation of event temporal scope mapping for studies archived in BIDS format soon following the initial release, with updating of current data search and analysis tools to follow. Support for Library schema is also planned. More sophisticated task definition and event linkage annotations will hopefully be integrated in HED-4.

4.2 HED community development.

HED development and HED coding of now archived and soon-to-be archived data must have substantial and sustained research community contributions to be successful. HED will not achieve its major aim of enabling meta/mega-analysis of now rapidly accumulating neurophysiological data archives without adoption and active use and exploitation, as well as creative further contributions by diverse communities of researchers and clinicians. Communities of researchers in areas such as clinical neurophysiology and music psychobiology have already expressed interest in developing discipline-specific HED vocabularies for EEG event annotation. To become involved in using and further developing HED, researchers must be convinced that it is an important part of assuring the legacy and increasing the total value of their data – both to themselves and their students and as well to others.

Currently, there remains much work to do to integrate schema libraries into the HED system. Tools for building, documenting, versioning, and making available HED schema libraries all must function smoothly to be attractive for use in practice. Having available open-source tools for performing useful analysis making use of HED information, as well as well-annotated HED-informed data archives linked to computing resources supporting relevant tool libraries should increase interest in using HED. A careful, ambitious and enthusiastic tutorial campaign will also be needed to allow HED annotation to become sufficiently widespread to reach "critical mass" momentum. However, our experience has taught us that performing truly useful data annotation is not trivial, even given good tools and tutorials.

We have begun the process of enlarging the HED user community and annotating studies in HED-3 for archiving, retrieval and computation via NEMAR (NEMAR homepage n.d.), a DATCOR (integrated data-tools-compute) resource for human electrophysiological data that we and collaborators are now building. NEMAR will also act as a portal to OpenNeuro (OpenNeuro hompage n.d.), the NIMH-supported archive of human neuroimaging data of all modalities (Gorgolewski et al. 2017).

The EEGLAB (Delorme and Makeig 2004) computational portal to the XSEDE high-performance computing network (Martínez-Cancino et al. 2020) via the Neuroscience Gateway (Sivagnanam et al. 2020), soon to be integrated with NEMAR, will allow intensive, high-performance processing of HED-tagged, BIDS-formatted data without requiring voluminous data transfer and data copy management. EEGLAB tools are now incorporating HED as a basic foundation to support analysis and re-analysis of individual studies as well as meta/mega-analysis of archived data across studies.

Other MATLAB tool environments such as Fieldtrip (Oostenveld et al. 2011) may be able to easily incorporate handling of HED tag information by using or adapting HED MATLAB library functions. The Neuroscience Gateway also supports several other neuroscience tool environments applicable to NEMAR-archived data.

While our own research has focused on analysis of scalp EEG data, the HED system is equally applicable to any human neuroimaging experiment, and HED library schema extending the top-level HED schema vocabulary to include modality-specific terms should be straightforward to build and integrate. Also, future use of HED need not be restricted to neuroimaging – any time series or time sequence in which timing of events is recorded should be able to be usefully represented using HED.

Though human electrophysiological data in the form of scalp EEG was the first noninvasive human brain imaging modality, progress in its analysis and interpretation has long lagged behind technical developments in data acquisition. In the clinical neurophysiology field, visual inspection of the raw channel records is still the most prevalent mode of information extraction, while in cognitive neuroscience, study of details in event-related response averages across similar classes of events has long dominated practice and teaching.

While substantial progress has been made in the past twenty years toward extracting more of the rich information about human brain dynamics contained in electrophysiological recordings (EEG, MEG, iEEG), much collected data has not been mined using the resultant analysis approaches. Further, applications of machine learning to electrophysiological data are still in their infancy and require availability of well-annotated data to deliver accurate markers and new understanding of how the brain supports human behavior and experience, normal and pathological. We believe that, given sufficient care, interest, and continued investment, the HED system can, should, and will play an important role in this evolution. Further expansion of use of HED annotation to many types of time series and time-ordered data also appears a useful possibility.

Acknowledgments:

The authors would like to acknowledge the seminal contributions of Nima Bigdely-Shamlo to the development of HED-1 and HED-2. The authors would also like to thank Arno Delorme for his comments and work in integrating HED into EEGLAB. The authors would also like to express their appreciation for the support and work of Jonathan Touryan of the Army Research Laboratory as well as Tony Johnson and Michael Dunkel of DCS Corporation in creating the large HED-2 data corpus that enabled much of the HED evolution.

Funding:

This work was sponsored by the United States Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by the U.S. National Institutes of Health (NS047293; EB023297; MH120037), and by a gift to UCSD from The Swartz Foundation (Old Field NY).

The views and the conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S Government. The U.S Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

BIDS homepage (n.d.) Brain Imaging Data Structure. https://bids.neuroimaging.io/

Bigdely-Shamlo N (2014) Combining EEG Source Dynamics Results across Subjects, Studies and Cognitive Events. PhD, University of California, San Diego

Bigdely-Shamlo N, Cockfield J, Makeig S, Rognon T, La Valle C, Miyakoshi M, Robbins KA (2016) Hierarchical Event Descriptors (HED): Semi-structured tagging for real-world events in large-scale EEG. Front Neuroinform 10: . https://doi.org/10.3389/fninf.2016.00042

Bigdely-Shamlo N, Kreutz-Delgado K, Robbins K, Miyakoshi M, Westerfield M, Bel-Bahar T, Kothe C, Hsi J, Makeig S (2013) Hierarchical Event Descriptor (HED) tags for analysis of event-related EEG studies. In: 2013 IEEE Global Conference on Signal and Information Processing. pp 1–4

Bigdely-Shamlo N, Touryan J, Ojeda A, Kothe C, Mullen T, Robbins K (2019a) Automated EEG mega-analysis II: Cognitive aspects of event related features. NeuroImage 116054 . https://doi.org/10.1016/j.neuroimage.2019.116054

Bigdely-Shamlo N, Touryan J, Ojeda A, Kothe C, Mullen T, Robbins K (2019b) Automated EEG mega-analysis I: Spectral and amplitude characteristics across studies. NeuroImage 116361 . https://doi.org/10.1016/j.neuroimage.2019.116361

Boedhoe PSW, Heymans MW, Schmaal L, Abe Y, Alonso P, Ameis SH, Anticevic A, Arnold PD, Batistuzzo MC, Benedetti F, Beucke JC, Bollettini I, Bose A, Brem S, Calvo A, Calvo R, Cheng Y, Cho KIK, Ciullo V, Dallaspezia S, Denys D, Feusner JD, Fitzgerald KD, Fouche J-P, Fridgeirsson EA, Gruner P, Hanna GL, Hibar DP, Hoexter MQ, Hu H, Huyser C, Jahanshad N, James A, Kathmann N, Kaufmann C, Koch K, Kwon JS, Lazaro L, Lochner C, Marsh R, Martínez-Zalacaín I, Mataix-Cols D, Menchón JM, Minuzzi L, Morer A, Nakamae T, Nakao T, Narayanaswamy JC, Nishida S, Nurmi EL, O'Neill J, Piacentini J, Piras F, Piras F, Reddy YCJ, Reess TJ, Sakai Y, Sato JR, Simpson HB, Soreni N, Soriano-Mas C, Spalletta G, Stevens MC, Szeszko PR, Tolin DF, van Wingen GA, Venkatasubramanian G, Walitza S, Wang Z, Yun J-Y, Thompson PM, Stein DJ, van den Heuvel OA, Twisk JWR (2019) An empirical comparison of meta- and mega-analysis with data from the ENIGMA Obsessive-Compulsive Disorder Working Group. Front Neuroinform 12: . https://doi.org/10.3389/fninf.2018.00102

Brainlife homepage (n.d.) Brainlife cloud platform. https://brainlife.io/

Carliner S (1987) Lists: The ultimate organizer for engineering writing. IEEE Transactions on Professional Communication PC-30:218–221 . https://doi.org/10.1109/TPC.1987.6449088

Cognitive Atlas homepage (n.d.) Cognitive Atlas Collaborative Knowledge Base. www.cognitiveatlas.org

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134:9–21 . https://doi.org/10.1016/j.jneumeth.2003.10.009

Gamma E, Helm R, Johnson R, Vlissides J, Booch G (1994) Design Patterns: Elements of Reusable Object-Oriented Software, 1 edition. Addison-Wesley Professional, Reading, Mass

Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, Poline J-B, Rokem A, Schaefer G, Sochat V, Triplett W, Turner JA, Varoquaux G, Poldrack RA (2016) The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data 3:160044 . https://doi.org/10.1038/sdata.2016.44

Gorgolewski KJ, Esteban O, Schaefer G, Wandell BA, Poldrack RA (2017) Openneuro: A gree online platform for sharing and analysis of neuroimaging data. In: Organization for Human Brain Mapping 2017. Vancouver, Canada

HeadIT homepage (n.d.) HeadIT EEG Dataset repository. headit.ucsd.edu

HED homepage (n.d.) Hierarchical Event Descriptors. http://www.hedtags.org/

HED-1 view (n.d.) HED-1: Hierarchical Event Descriptor Expandable Schema View (Deprecated). http://www.hedtags.org/display_hed.html?version=1.3

HED-2 view (n.d.) HED-2: Hierachical Event Descriptor Expandable Schema View. http://www.hedtags.org/display_hed.html

HED-3 view (n.d.) HED-3: Hierachical Event Descriptor Expandable View (Development Version). http://www.hedtags.org/display_hed_restruct.html?version=reduced

hed-standard homepage (n.d.) HED-Standards Organization. https://github.com/hed-standard

Keator DB, Helmer K, Steffener J, Turner JA, Van Erp TGM, Gadde S, Ashish N, Burns GA, Nichols BN (2013) Towards structured sharing of raw and derived neuroimaging data across existing resources. NeuroImage 82:647–661 . https://doi.org/10.1016/j.neuroimage.2013.05.094

Kirchner WK (1958) Age differences in short-term retention of rapidly changing information. J Exp Psychol 55:352–358 . https://doi.org/10.1037/h0043688

Makeig S (2009) Mind monitoring via mobile brain-body imaging. In: Schmorrow DD, Estabrooke IV, Grootjen M (eds) Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience. Springer, Berlin, Heidelberg, pp 749–758

Makeig S, Gramann K, Jung T-P, Sejnowski TJ, Poizner H (2009) Linking brain, mind and behavior. International Journal of Psychophysiology 73:95–100 . https://doi.org/10.1016/j.ijpsycho.2008.11.008

Martínez-Cancino R, Delorme A, Truong D, Artoni F, Kreutz-Delgado K, Sivagnanam S, Yoshimoto K, Majumdar A, Makeig S (2020) The open EEGLAB portal interface:High-performance computing with EEGLAB. NeuroImage 116778 . https://doi.org/10.1016/j.neuroimage.2020.116778

NEMAR homepage (n.d.) NEMAR: Neuroelectromagnetic Data Archive & Tools Resource. https://nemar.org

Niso G, Gorgolewski KJ, Bock E, Brooks TL, Flandin G, Gramfort A, Henson RN, Jas M, Litvak V, T. Moreau J, Oostenveld R, Schoffelen J-M, Tadel F, Wexler J, Baillet S (2018) MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. Scientific Data 5:180110 . https://doi.org/10.1038/sdata.2018.110

Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. In: Computational Intelligence and Neuroscience. https://www.hindawi.com/journals/cin/2011/156869/. Accessed 4 Dec 2017

OpenNeuro hompage (n.d.) Openneuro data-sharing platform. https://openneuro.org

Pernet CR, Appelhoff S, Gorgolewski KJ, Flandin G, Phillips C, Delorme A, Oostenveld R (2019) EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Scientific Data 6:103 . https://doi.org/10.1038/s41597-019-0104-8

RDF standard (2015) W3C Resource Framework Description Standard. w3.org/TR/2015/REC-rdfa-core-20150317

Robbins KA, Touryan J, Mullen T, Kothe C, Bigdely-Shamlo N (2020) How sensitive are EEG results to preprocessing methods: A benchmarking study. IEEE Trans Neural Syst Rehabil Eng 28:1081–1090 . https://doi.org/10.1109/TNSRE.2020.2980223

Rognon T, Strautman R, Jett L, Bigdely-Shamlo N, Makeig S, Johnson T, Robbins K (2013) CTAGGER: Semi-structured community tagging for annotation and data-mining in event-rich contexts. In: 2013 IEEE Global Conference on Signal and Information Processing. pp 5–8

Sivagnanam S, Yoshimoto K, Carnevale T, Nadeau D, Kandes M, Petersen T, Truong D, Martinez R, Delorme A, Makeig S, Majumdar A (2020) Neuroscience Gateway enabling large scale modeling and data processing in neuroscience research. In: Practice and Experience in Advanced Research Computing. Association for Computing Machinery, Portland, OR, USA, pp 510–513

Turner JA, Laird AR (2012) The cognitive paradigm ontology: design and application. Neuroinformatics 10:57–66 . https://doi.org/10.1007/s12021-011-9126-x

Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3:160018 . https://doi.org/10.1038/sdata.2016.18