

Summary: To profit from the possible benefits of the large-sized EPILEPSIAE database for the field of seizure prediction, a common annotation standard for eeg recordings is an essential prerequisite; multicenter contributions of data to such a project increase the needs for comparability and reliability of annotations.

Matthias Ihle¹, Carolin Gierschner¹, Christiane Lehmann¹, Anika Schinkel¹, Vincent Navarro², Michel LeVanQuyen³, Francisco Sales⁴, Nuno Silva⁵, Andreas Schulze-Bonhage¹

(1) Epilepsy center, University Hospital of Freiburg, (2) Hôpital de la Pitié-Salpêtrière, (3) Centre de Recherche de l'Institut du Cerveau de la Moelle, (4) Hospitals of the University of Coimbra, (5) Department of Informatics Engineering, University Coimbra

Funded by the European Union, the EPILEPSIAE project was established with the objective to develop a mobile seizure alarming device for epilepsy patients. For the development and validation of the seizure prediction algorithms, a multinational shared database with 300 EEG long-term datasets is currently being developed within the scope of the project.

As a joint project of several hospitals it is a crucial point to ensure the comparability of the eeg annotations at all project sites. Accordingly, a standardized EEG annotation protocol was developed which is presented here.

Methods:

In order to agree on a protocol, a common denominator with respect to the local situations with different prerequisites and approaches had to be found. The project members had to agree on the included events and their properties that are relevant for seizure prediction. In particular, a tradeoff between details of annotations and manpower available had to be made.

As a second step, the marker syntax had to be defined to ensure the automatic transfer of these annotations into the database.

Results:

It was decided that all EEG recordings had to be re-reviewed visually as there was a consensus that neither heterogeneous markers made according to the individual clinical requirements of the different centers nor presently available automated EEG analyzing software is sufficient to provide adequate reliability for the purpose of the project.

Markers defined were targeted at time points of interictal-ictal transitions, clinical vs. subclinical seizure patterns, sleep stages and spread. An annotation protocol was developed to ensure the reliability and comparability of the annotated eeg datasets from several clinics across Europe with differing premises. Consistency of annotations are to be controlled by reliability check.

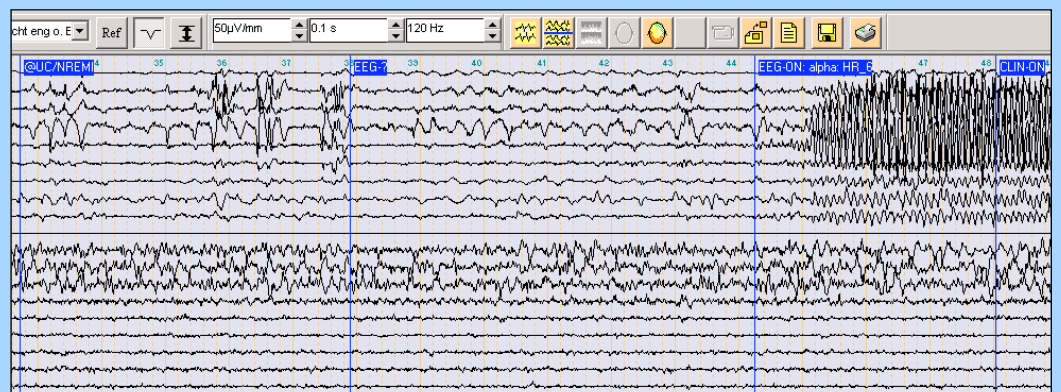
Annotations & Markers

Annotations are markers that are directly attached to an eeg file. They are usually made with client software of an EEG system.

Basically, a marker has two relevant properties: a timestamp and a textual content:

- Every marker belongs to a sample, respectively to the timestamp of that sample. In the screenshot on the left, this is depicted by the blue lines.
- The textual content of the markers is shown in the blue boxes besides the lines.

In the following, the textual content (syntax) of the markers of our annotation protocol is defined. We distinguish between markers for seizures, subclinical events and interictal spikes.



Seizure markers:

```
<seizure> ::= "@" <seizure_type> "/" <vigilance_state>
<eeg-on> ::= "EEG-ON" <seizure_pattern> ":" <electrode_list>
| "EEG-ON: NA" /* eeg onset (not available) */
<eeg-?> ::= "EEG-?" /* earliest eeg change */
<eeg-off> ::= "EEG-OFF" /* eeg offset */
| "EEG-OFF: NA"
<clin-on> ::= "CLIN-ON" /* clinical seizure onset */
| "CLIN-ON: NA"
<clin-?> ::= "CLIN-?" /* first behavioural alteration */
<clin-off> ::= "CLIN-OFF" /* clinical seizure offset */
| "CLIN-OFF: NA"
<prop1> ::= "PROP1:" (<electrode_list> | "NA")
<prop2> ::= "PROP2:" (<electrode_list> | "NA")
```

Subclinical event markers:

```
<sub-on> ::= "SUB-ON" /* subclinical event onset */
| "SUB-ON: NA" /* onset not available */
<sub-off> ::= "SUB-OFF" /* event offset */
| "SUB-OFF: NA" /* offset not available */
```

Interictal event marker:

```
<interictEvt> ::= "type" <number> ":" <interictal_pattern>
<amp_max> <field_extension>
• <amp_max> ::= <electrode_name>
• <field_extension> ::= "(" <electrode_list> ")" | NULL
```

Other production rules:

```
<electrode_list> ::= <electrode_name>
| <electrode_name> ", " <electrode_list>
| "all"
| "left hemisphere" | "right hemisphere"
<electrode_name> ::= STRING
<number> ::= INTEGER
```

Marker definition & syntax

Seizures:

- <seizure> is the mandatory first marker for every seizure. All other markers found between this and the next <seizure>-marker contain information about the current seizure.
- The timestamps for the different onset/offset markers are defined by the timestamp of the according markers.
- If an onset/offset cannot be determined (e.g. video failure), the 'NA' variant distinguishes such markers that are *not available* from markers that were forgotten to be set.

Propagation:

- The <prop1>-marker defines the electrodes being involved in an early propagation within the first 10 seconds after the eeg onset.
- The marker should be set about 10 seconds after the eeg onset if the duration of the seizure is long enough.
- "NA" indicates that there is no propagation for the seizure.
- The <prop2>-marker contains a list of electrodes that are involved in a late propagation (after the first 10 seconds)
- The marker should be set after the <prop1>-marker directly before the eeg offset marker.
- If the duration of the seizure is less than 10 seconds a late propagation is not defined (-> PROP2: NA).

Subclinical events:

- Subclinical events are important for seizure prediction, particularly when analyzing intracranial recordings.
- For reasons of manpower, just 10 subclinical events are annotated per recording day.

Interictal spikes:

- For every dataset there are different types of spikes/interictal events, characterized by:
 - their pattern/morphology: <interictal_pattern>
 - the electrode with the max. amplitude: <amp_max>
 - the field electrodes of the event: <field_extension>
- For every type of interictal activity just one event is annotated exemplarily (for reasons of manpower). The different types are numbered (<number>).
- The timestamp of the marker determines the peak timestamp of the event.

```
<seizure_type> ::=
UC: unclassified
SP: simple partial
CP: complex partial
SG: secondarily generalized
```

```
<vigilance_state> ::=
'?' unclear
'A' awake
'NREM1' sleep stage 1
'NREM2' sleep stage 2
'NREM3' sleep stage 3
'NREM4' sleep stage 4
'REM' REM
```

- invasive (only) recordings: either awake (according to video) or unclear
- The vigilance state is determined 10 seconds before either the eeg or the clinical seizure onset.

```
<seizure_pattern> ::=
'ad' amplitude depression
'lafa' low amplitude fast activity
'rs' repetitive spiking
'alpha' rhythmic alpha waves
'beta' rhythmic beta waves
'delta' rhythmic delta waves
'subdelta' rhythmic subdelta waves
'theta' rhythmic theta waves
'sw' sharp waves
'ce' cessation of interictal activity (only)
```

```
<interictal_pattern> ::=
'sp' spike
'lafa' low amplitude fast activity
'ps' polyspikes
'sw' sharp waves
```

```
Basic data types:
INTEGER, STRING, NULL
```