

# eegBidsCreator or How to bidsify Embla

Nikita BELIY<sup>1</sup>

<sup>1</sup>Liege University, Belgium

February 5, 2019



## 1 Introduction

## 2 Usage

## 3 Conversion

## 4 Additional functionality

## 5 Additional tools

## 6 Plans

## 7 Reference

Introduction



Usage



Conversion



Additional functionality



Additional tools



Plans



Reference



# Introduction

# eegBidsCreator

gitlab:Methods/BIDS\_Tools/eegBidsCreator

- Reads the EEG files
  - Tested with RemLogic 3.4
- Creates BIDS folder structure
- Copy/Create and pre-fill JSON file
- Creates list of channels and events
- Convert to EDF+ or BrainVision format

Follows the BEP006

Introduction  
oo

Usage  
●oooooooooooooo

Conversion  
oooooooooooo

Additional functionality  
oooooooooooo

Additional tools  
oooo

Plans  
oo

Reference  
oooooooooooooooooooo

# Usage

# Usage

```
python3 eegBidsCreator.py [options] file_to_convert
```

- One-file executable will be available (at some point of time)

## Configuration

- Command line options
  - For often changed options
- Config. file
  - For more stable options

# Demonstration:

## COF098\_HN\_080518\_T233900

```
python3 eegBidsCreator.py -t HN -s 080518 -a T233900 -j Example/COF_HN.json -o output -c Example/COF_HN.ini  
/media/beliy/KINGSTON/Memodyn/ForTest/Emb/35d1ef92-1b65-42f3-aab4-b3d61076f448
```

```
output/  
|-- COFTest/  
|-- Test/  
|-- participants.tsv  
`-- sub-COF098_HN_080518_T233900/  
    '-- ses-080518/  
        '-- eeg/  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.ini  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.log  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_ObservationNotesData.eso  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_Video Clip 0001.wmv  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_biocMap.bcm  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_channels.tsv  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.edf  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.json  
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.edf  
            '-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.tsv  
`-- scans.tsv  
  
5 directories, 12 files
```

## Folder structure

## Task, session and acquisition

## Command line

-t HN -s 080518 -a T233900

Configuration file: [GENERAL]

TaskId = HN

AcquisitionId = T233900

SessionId = 080518

Subject Id	Session Id	Task & Acquisition Id
sub-COF098_HN_080518_T233900/	ses-080518/	
-- eeg/		
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.ini
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.log
-- sub-COF098_HN_080518_T233900	ses-080518	ObservationNotesData.eso
-- sub-COF098_HN_080518_T233900	ses-080518	Video Clip 0001.wmv
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.biocalMap.bcm
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.channels.tsv
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.eeg.edf
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.eeg.json
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.events.edf
-- sub-COF098_HN_080518_T233900	ses-080518	task-HN_acq-T233900.events.tsv
-- scans.tsv		

# BIDS specific files

- *participants.tsv*: List of subjects with age, sex and group
- *scans.tsv*: List of scans with start time
- *\_eeg.json*: Task description file
- *\_channels.tsv*: List of channels, with parameters
- *\_events.tsv*: List of events

```
output/
|-- COFTest/
|-- Test/
|--- participants.tsv
|--- sub-COF098_HN_080518_T233900/
    '-- ses-080518/
        '-- eeg/
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.ini
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.log
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_ObservationNotesData.eso
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_Video Clip 0001.wmv
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_biocalMap.bcm
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_channels.tsv
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.json
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.edf
            |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.edf
            '-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.tsv
        '-- scans.tsv
5 directories, 12 files
```

Created/copied files

# Participants and Scans files

## participants.tsv

Subject	Sex	Age
COF128	n/a	n/a
COF098	F	n/a
COF148	M	n/a
COF131	n/a	n/a
COF008	M	n/a
COF152	n/a	n/a
COF089	n/a	n/a
COF008	n/a	n/a
COF112	n/a	n/a
COF127	F	n/a
COF113	M	n/a

- Age from birth date and scan date
- Different age from different scans?

## scans.tsv

```
eeg/sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.edf 2018-05-08T23:40:29
```

```
~
```

```
~
```

```
~
```

```
~
```

```
~
```

- lines are simply appended  
need manual sort/cleanup

Created/copied files

# Task file

## Command line

-j Example/COF\_HN.json

## Configuration file: [GENERAL]

JsonFile = Example/COF\_HN.json

### No input JSON:

```
"TaskName": "HN",
"DeviceSerialNumber": "f5e20dd8-7b52-442b-a227-932ef4f66566",
"Manufacturer": "RemLogic",
"SamplingFrequency": 200,
"RecordingDuration": 22811.810311,
"EEGChannelCount": 8,
"EOGChannelCount": 2,
"ECGChannelCount": 1,
"EMGChannelCount": 3,
"MiscChannelCount": 27
```

### With input JSON:

```
"TaskName": "HN",
"TaskDescription": "Horror Night",
"Instructions": "Play Justin Bieber on loop",
"DeviceSerialNumber": "f5e20dd8-7b52-442b-a227-932ef4f66566",
"Manufacturer": "RemLogic",
"SamplingFrequency": 200,
"RecordingDuration": 22811.810311,
"EEGChannelCount": 8,
"EOGChannelCount": 2,
"ECGChannelCount": 1,
"EMGChannelCount": 3,
"MiscChannelCount": 27
```

Created/copied files

## Task file (2)

- Checks if given task is same as in JSON file  
Insures correct task description
- Sampling frequency is common to all channels  
200 Hz and 500 Hz, lead to 1000 Hz
- If file has no “.json”: <file>\<taskId>.json will be loaded
- Missing fields reported in the log:

```
[INFO    ]:01/24/2019 16:18:55:root Creating eeg.json file
[WARNING]:01/24/2019 16:18:55:root JSON: Missing next required fields: ['EEGReference', 'PowerLineFrequency', 'SoftwareFilters']
[INFO    ]:01/24/2019 16:18:55:root JSON: Missing next recommended fields: ['CogAtlasID', 'CogPOID', 'InstitutionName', 'InstitutionAddress', 'InstitutionalDepartementName', 'HeadCircumference', 'TriggerChannelCount', 'EEGGround', 'EEGPlacementScheme', 'CapManufacturer', 'RecordingType', 'EpochLength', 'SoftwareVersions']
```

### **Created/copied files**

## channels.tsv

name	type	units	description	sampling_frequency	reference
Flattening	Resp.Flattening	x10^-3	Resp.Flattening-Cannula.Nasal	200	n/a
EKG	EKG	mV	EKG	200	n/a
Right Leg	EMG.Tibialis	uV	EMG.Tibialis-Leg.Right	200	n/a
Chin EMG	EMG.Submental	uV	EMG.Submental-Chin	200	n/a
Left Leg	EMG.Tibialis	uV	EMG.Tibialis-Leg.Left	200	n/a
REF	EEG	uV	EEG-REF	200	n/a
E1	EOG	uV	EOG-E1	200	n/a
E2	EOG	uV	EOG-E2	200	n/a
M1	EEG	uV	EEG-M1	200	n/a
C3	EEG	uV	EEG-C3	200	n/a
M2	EEG	uV	EEG-M2	200	n/a
Fz	EEG	uV	EEG-Fz	200	n/a
Cz	EEG	uV	EEG-Cz	200	n/a
Pz	EEG	uV	EEG-Pz	200	n/a
Oz	EEG	uV	EEG-Oz	200	n/a
Light BU	Luminance	n/a	Luminance-LDR	5	n/a

- Description formed from Embla's `<MainType>-<SubType>`
  - Units are SI, with scaling prefix, calculated automatically  
If there are no units, prefix is " $\times 10^{-k}$ ":  $5 \times 10^{-3} = 0.005$
  - Ref and filters should appear in .ebm, but do not

### Created/copied files

events.tsv

Time offset(s)	Duration(s)	Name	Response time	Score	Sample
210.916	0.00	LIGHTS-OFF	n/a	n/a	42183
210.916	30.00	SLEEP-S0	n/a	n/a	42183
211.376	11.34	RESP-RMI	n/a	n/a	42275
211.596	2.73	RESP-BREATH-I	n/a	n/a	42319
216.351	2.12	RESP-BREATH-I	n/a	n/a	43270
219.261	0.62	RESP-BREATH-I	n/a	n/a	43852
221.596	2.15	RESP-BREATH-I	n/a	n/a	44319
224.401	0.70	RESP-BREATH-I	n/a	n/a	44880
227.011	2.91	RESP-BREATH-I	n/a	n/a	45402
231.806	2.15	RESP-BREATH-I	n/a	n/a	46361
234.601	0.66	RESP-BREATH-I	n/a	n/a	46920
236.941	1.59	RESP-BREATH-I	n/a	n/a	47388
239.856	1.48	RESP-BREATH-I	n/a	n/a	47971
240.776	29.76	RESP-RMI	n/a	n/a	48155
240.916	30.00	SLEEP-S0	n/a	n/a	48183
242.007	0.30	SLEEP-RM	n/a	n/a	48401

- Response time is not defined in Embla
  - Score is defined but not written
  - Does someone have experience with events with score in Embla?

## Created/copied files

## EEG files

- Converted files from original Embla
- Contains data, events and channels description

```
sub-COF098_HN_080518_T233900/
`-- ses-080518/
    |-- eeg/
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.ini
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.log
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_ObservationNotesData.eso
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_Video Clip 0001.wmv
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_biocalMap.bcm
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_channels.tsv
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.edf
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.json
        |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.edf
        `-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.tsv
    `-- scans.tsv
2 directories, 11 files
```

Created/copied files

# Auxiliary files

- Original files not related to data or channels
- Detected by the extension:  
Must not have Embla's standard extensions
- Copied directly to output folder

```
sub-COF098_HN_080518_T233900/
`-- ses-080518/
    |-- eeg/
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.ini
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.log
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_ObservationNotesData.eso
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_Video Clip 0001.wmv
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_biocamMap.bcm
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_channels.tsv
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.edf
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.json
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.edf
    |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.tsv
    |-- scans.tsv

2 directories, 11 files
```

## Tool files

## eegBidsCreator files

- .ini: Conf file with all parameters used for conversion
- .log: A full log file
- Kept for transparency/repeatability

```
sub-COF098_HN_080518_T233900/
  -- ses-080518/
    -- eeg/
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.ini
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.log
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_ObservationNotesData.eso
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_Video Clip 0001.wmv
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_bicoalMap.bcm
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_channels.tsv
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.edf
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_eeg.json
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.edf
      |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_events.tsv
    -- scans.tsv

2 directories, 11 files
```

## Conversion

# Conversion

## Command line

```
-conversion {EDF,BV}
```

## Configuration file: [GENERAL]

```
Conversion = {EDF,BV}
```

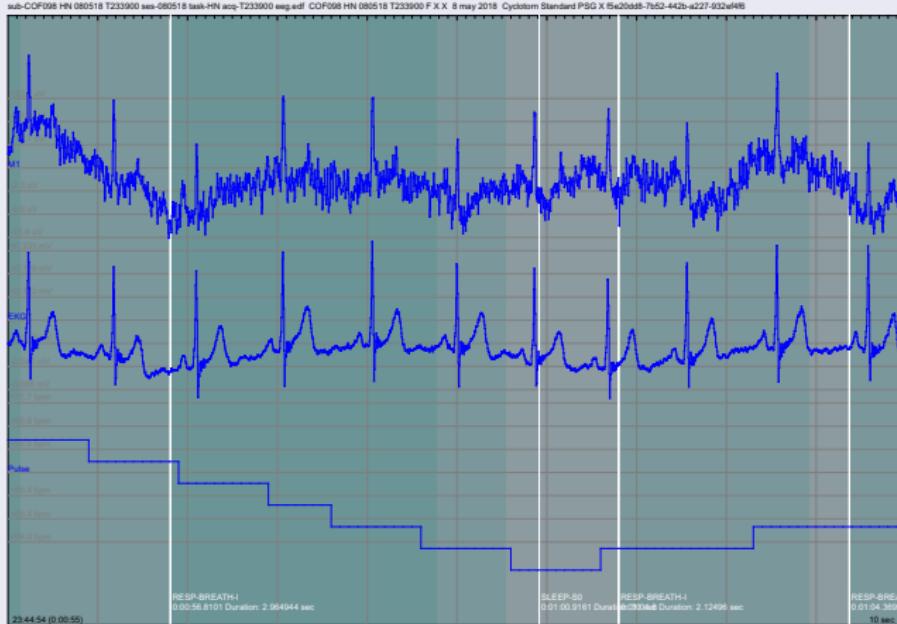
- Convert Embla to given format, EDF+ or BrainVision
- If not given, just copy Embla files

# EDF+

- Files:
  - *\_eeg.edf* – channels and data
  - *\_events.edf* – events
- Pros:
  - Compact – supports channels with different samplings
  - Portable – majority of tools read it
  - Open and well described
- Cons:
  - Not obvious to edit in place
  - Events support is not natural and not structured

EDF+

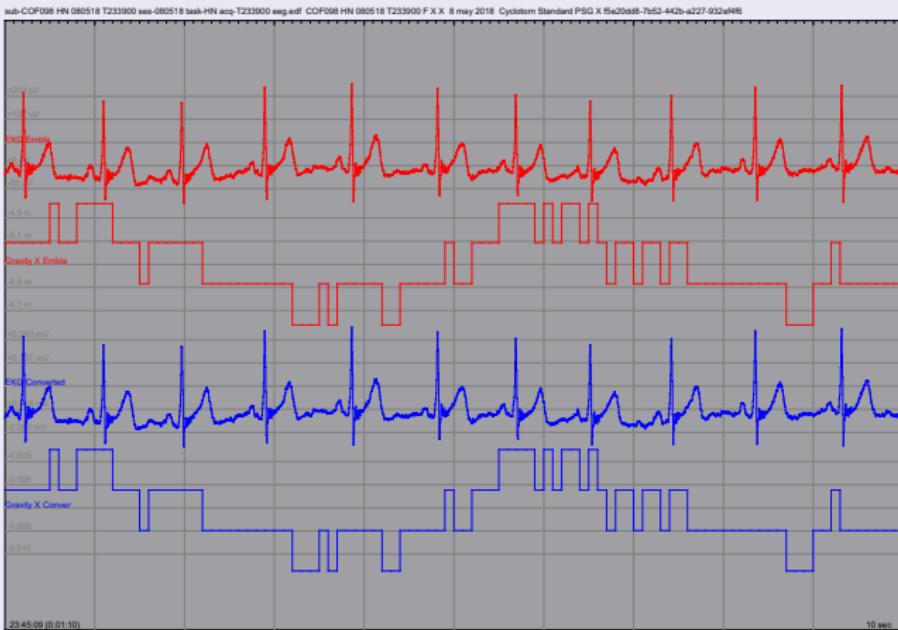
COF098\_HN\_080518\_T233900



EDF+

## Compared to Embla file

COF098\_HN\_080518\_T233900

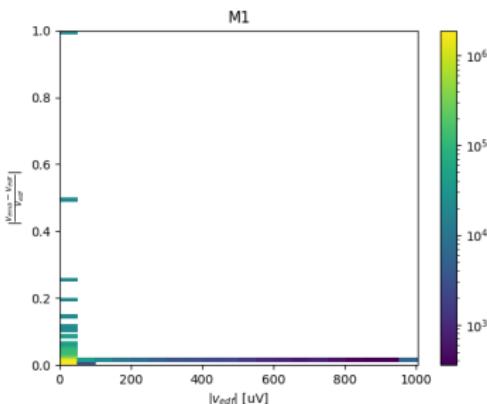


# EDF+

## Validation(1)

COF098\_HN\_080518\_T233900

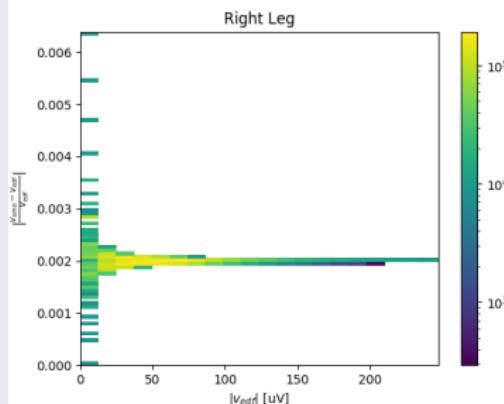
### M1



- $r = 30.675 \times 10^{-9} \mu V$
- $r' = 91.5555 \times 10^{-9} \mu V$

- Low ( $< 1\%$ ) constant relative deviation
- Large relative deviation for low values

### Right Leg



- $r = 91.8165 \times 10^{-9} \mu V$
- $r' = 7.62963 \times 10^{-9} \mu V$

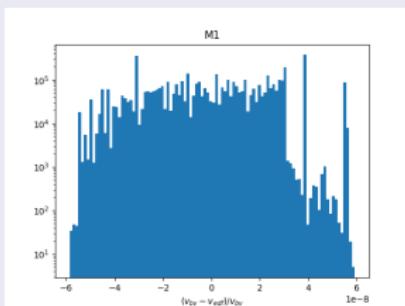
# Brain Vision

- Files:
  - `_eeg.vhdr` – header, with channels
  - `_eeg.vmrk` – events
  - `_eeg.eeg` – data
- Pros:
  - data identical to Matlab's `.dat`
  - easy readable/editable headers
  - open and described format
- Cons:
  - not supported by much tools (MNE supports it)
  - tools might implement only part fo standard
  - channels must have same frequency – oversampling
  - space-hungry: can inflate data ×10

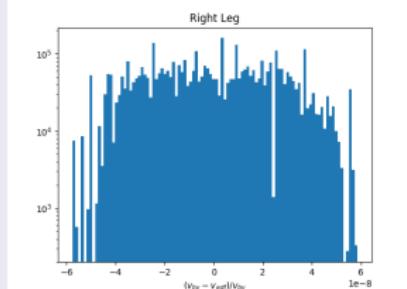
# Brain Vision Validation

COF098\_HN\_080518\_T233900

M1



Right Leg



- Negligible difference between EDF and BV  
Due float. point approximation

# Performance comparison

COF098\_HN\_080518\_T233900

- Duration: 6h20m12s
- Sampling: 200 Hz
- Channels: 41

	Files size	User time		Max. memory
		min	mean	
Embla	247 MB			
EDF	251 MB	48.26 s	49.06 s	1.97 GB
BV (int)	415 MB	67.46 s	68.29 s	2.34 GB
BV (float)	771 MB	75.74 s	77.21 s	2.67 GB

Introduction  
oo

Usage  
oooooooooooo

Conversion  
oooooooo

Additional functionality  
●oooooooooooo

Additional tools  
oooo

Plans  
oo

Reference  
oooooooooooooooooooo

## Additional functionality

# Data reduction

## Configuration file: [DATATREATMENT]

- *DropChannels*: list of channels to ignore
  - *StartTime/EndTime*: consider data between given time
  - *StartEvent/EndEvent*: consider data between given events
  - *IgnoreOutOfTimeEvents*: remove evts outside data period
  - *IncludeSegmentStart*: additional events at start of each segment
  - *MergeCommonEvents*: replace duplicated events by one
- 
- Gives additional control over data and events
  - Can produce loss of data

# Data reduction

Example: sub-COF098\_HN\_080518\_T233900



# Splitting in runs

- BEP006: Run = a non-intermittent period in which data for the subject(s) is continuously being acquired
- Corresponds to sequence in Embla: one emb file has several continuous data sequences

## Configuration file: [RUNS]

- *SplitRuns*: To do the splitting
- *MainChannel*: From which channel define sequences
- *MinSpan*: Ignore sequences shorter than given values

# Splitting in runs

Example: sub-COF098\_HN\_080518\_T233900



# Splitting in runs

Example: sub-COF098\_HN\_080518\_T233900

```
sub-COF098_HN_080518_T233900/
`-- ses-080518/
    |-- eeg/
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.ini
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900.log
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_ObservationNotesData.eso
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_Video Clip 0001.wmv
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_biocalMap.bcm
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_een.json
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-1_channels.tsv
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-1_eeg.edf
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-1_events.edf
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-1_events.tsv
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-2_channels.tsv
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-2_eeg.edf
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-2_events.edf
        |   |-- sub-COF098_HN_080518_T233900_ses-080518_task-HN_acq-T233900_run-2_events.tsv
    `-- scans.tsv
```

- *run-x* appended to EEG data and events files
- each run appears in *scans.tsv*
- *run-x* appears even if only one run exists
- how treat events in-between runs?

# Anonymization

- BEP006: Implicitly requires anonymization of subject name and recording date

## Configuration file: [ANONYMIZATION]

- *Anonymize*: To do anonymization
- *StartDate*: Fake-date of recording
- *SubjName*: Fake subject name
- *BirthDate*: Fake subject birth date
- If set to “None”, act as these fields are non existing in data recording

# anonymization

example: sub-cof098\_hn\_080518\_t233900



# Logging

## Parameters

### Command line

- *--log Level*: Verbosity level
- *--logfile Logfile*: Secondary log file
- *-q, --quiet*: Suppress output to terminal

### Configuration file: [LOGGING]

- *LogLevel*: Verbosity level
- *LogFile*: Secondary log file
- *Quiet*: Suppress output to terminal

- log file placed in the output directory

# Logging

## Verbosity levels

- CRITICAL: There no critical messages  
Supress all messages
- ERROR: Error was produced and stopped execution  
All files, except logfile will be removed
- WARNING: Something unusual found in file, but no error  
Duplicated channels, missing events or events without channels etc.
- INFO: General information on execution  
This is default level
- DEBUG: Detailed information about each step  
A lot of useless information printed, to use only if there some errors

# Logging

example: sub-cof098\_hn\_080518\_t233900

```
[INFO ]:01/25/2019 15:29:00:root >>>>>>>>>>>>>>>
[INFO ]:01/25/2019 15:29:00:root Starting new bidsifier
[INFO ]:01/25/2019 15:29:00:root <<<<<<<<<<<<<<<<
[INFO ]:01/25/2019 15:29:00:root File: /media/beliy/KINGSTON/Memodyn/ForTest/Emb/35d1ef92-1b65-42f3-aa
b4-b3dd1076f448/
[INFO ]:01/25/2019 15:29:00:root Output: /home/beliy/Memodyn/eegBidsCreator/output/
[INFO ]:01/25/2019 15:29:00:root Detected embla format
[INFO ]:01/25/2019 15:29:00:root Patient Id: COF098_HN_080518_T233900
[INFO ]:01/25/2019 15:29:00:root Session Id: 080518
[INFO ]:01/25/2019 15:29:00:root Task Id: HN
[INFO ]:01/25/2019 15:29:00:root Acq Id: T233900
[INFO ]:01/25/2019 15:29:00:root JSON File: Example/COF_HN.json
[WARNING]:01/25/2019 15:29:00:root Found 2 files with same identification. They will be removed.
[INFO ]:01/25/2019 15:29:00:root EEG will be saved in /home/beliy/Memodyn/eegBidsCreator/output/sub-CO
F098_HN_080518_T233900/ses-080518/eeg/
[INFO ]:01/25/2019 15:29:00:root Reading channels
[INFO ]:01/25/2019 15:29:00:root Updated common sampling frequency to 200 Hz
[WARNING]:01/25/2019 15:29:00:root Channel 'Heart Rate_DR': Starts 63.71 sec later than recording
[WARNING]:01/25/2019 15:29:00:root Channel 'Heart Rate_DR': Starts 63.71 sec later than other channels
[INFO ]:01/25/2019 15:29:00:root Duration: 6:20:12.410000
[INFO ]:01/25/2019 15:29:00:root Cropping start time: from 2018-05-08T23:40:29.230000 to 2018-05-09T00
:00:00
[INFO ]:01/25/2019 15:29:00:root Cropping end time: from 2018-05-09T06:00:41.640000 to 2018-05-09T01:0
0:00
[INFO ]:01/25/2019 15:29:00:root Reading events info
[INFO ]:01/25/2019 15:29:19:root Creating eeg.json file
[WARNING]:01/25/2019 15:29:19:root JSON: Missing next required fields: ['EEGReference', 'PowerLineFreque
ncy', 'SoftwareFilters']
[INFO ]:01/25/2019 15:29:19:root JSON: Missing next recommended fields: ['CogAtlasID', 'CogPOID', 'In
stitutionName', 'InstitutionAddress', 'InstitutionalDepartmentName', 'HeadCircumference', 'TriggerChanne
lCount', 'EEGGround', 'EEGPlacementScheme', 'CapManufacturer', 'RecordingType', 'EpochLength', 'Software
Versions']
[INFO ]:01/25/2019 15:29:19:root Creating channels.tsv file
[INFO ]:01/25/2019 15:29:19:root Creating events.tsv file
[INFO ]:01/25/2019 15:29:19:root Converting to EDF+ format
[INFO ]:01/25/2019 15:29:19:root Creating events.edf file
[INFO ]:01/25/2019 15:29:19:root Creating eeg.edf file
[INFO ]:01/25/2019 15:29:19:root Timepoint 1: Duration 1:00:00
[INFO ]:01/25/2019 15:29:25:root Copying auxiliary files
[INFO ]:01/25/2019 15:29:25:root >>>>>>>>>>>>>>>
[INFO ]:01/25/2019 15:29:25:root Took 25.025172966 seconds
[INFO ]:01/25/2019 15:29:25:root <<<<<<<<<<<<<
```

Introduction  
oo

Usage  
oooooooooooo

Conversion  
oooooooo

Additional functionality  
oooooooooooo

Additional tools  
●ooo

Plans  
oo

Reference  
oooooooooooo

## Additional tools

# Explore Emb files

## With Python3

```
>from DataStructure.Embla.Channel import EbmChannel  
>ch = EbmChannel("Chin EMG.ebm")
```

List all parameters:

```
>print(ch)
```

```
Version 4.0
Time [datetime.datetime(2018, 5, 8, 23, 40, 30, 590000),
0000)]
Channel 4
Sampling 200000
Gain 92
SCount 2
DBLsampling 200.0
RateCorr 1.8837764942380986e-05
RawRange [-0.00300855, 0.00300855, 0.0]
TransRange [-0.0001, 0.0001, 0.0]
Channel_32 [4, 4]
ChannName Chin EMG
CalFunc
CalUnit V
CalPoint [8 entries]
SubjectName COF098_HN_080518_T233900
SubjectId COF098_HN_080518_T233900
Data [0 entries]
DataGuid b240c504-b6a6-47cd-aff1-d1645ec6a570
SigType EMG.Submental
LowHigh [-0.0001, 0.0001]
SigRef
SigMainType EMG.Submental
SigSubType Chin
Endian <
Wide True
_stream <_io.BufferedReader name='/media/belyi/KINGSTON/Memo
12344440/Chin_EMG.ebm'...
```

Retrieve data between *t\_start* and *t\_end*:

```
>ch.GetValueVector(  
timeStart=t_start,  
timeEnd=t_end)
```

With *raw=True*, returns raw  
(unscaled) values

# Run over all Embla in folder

Example/run\_COF\_example.sh

- Simple bash script to run over all files in given directory
- Task, session and acquisition from Subject Id  
Cofitage samples only
- Correct JSON file loaded automatically
- Prints Embla process id and exit code
  - 0 – all good
  - 1 – errors has been found

# Scripts for testing/performance

- *Example/Validation\_mne.py*
  - Compare Embla EDF, converted EDF+ and converted BV files
- *Example/Performance\_Test.sh*
  - Runs tools 1000 times and store time and memory statistics
  - Similar to python's *timeit* tool

Introduction  
oo

Usage  
oooooooooooo

Conversion  
oooooooo

Additional functionality  
oooooooooooo

Additional tools  
oooo

Plans  
●○

Reference  
oooooooooooooooo

# Plans

# Plans

- Provide converted dataset to BIDS community  
Need an JSON files for BH&BL of Cofitage
- Provide conversion to CRC MatLab format
- Implement plugin architecture  
Or at least allow user modules at key point of script
- Implement writer for Embla  
Allow data control

Introduction  
oo

Usage  
oooooooooooo

Conversion  
oooooooo

Additional functionality  
oooooooooooo

Additional tools  
oooo

Plans  
oo

Reference  
●oooooooooooo

# Reference

# Command line arguments

## Id arguments

**-s sesId:** The Id of session, **Optional**

BEP006: a non-intermittent period in which the subject is wearing the EEG cap

**-t taskId:** The Id of the task **Mandatory**

BEP006: a short mnemonic, not necessarily an exact description of the task

**-a acqId:** The Id of acquisition **Optional**

BEP006: a custom label one may use to distinguish different set of parameters used for acquiring the same task

# Command line arguments

## Input/output arguments

**-o OutDir:** Path to the output directory **Optional**

Place where all files will be stored, default id “.”

**-c ConfFile:** Path to conf file **Optional**

INI file with all parameters

Cmd line options override config file options

# Command line arguments

## Logging arguments

`--log Level`: How much of info will be printed **Optional**

One of CRITICAL, ERROR, WARNING, INFO, DEBUG

`--logfile Logfile`: Path to a clone of log file **Optional**

If you prefer to read it in a fixed location instead of temporary one

`-q`: Suppress the output to terminal **Optional**

Do not affect the log files

# Command line arguments

## Configuration files and conversion

**-j jsonfile:** Path to task JSON file required by BIDS **Optional**

File describing the task.

**--conv {EDF,BV}:** Converts EEG files in given format **Optional**

Only EDF+(*EDF*), and Brain Vision (*BV*) are implemented. Brain Vision .eeg file should be identical to .dat file from Matlab conversion

# Command line arguments

## Configuration file

*-c conffile: Path to conf file*

Optional

- Standard .ini file
- Section name in brackets:  
*[GENERAL]*
- settings key=value:  
*LogLevel=INFO*
- All parameters for script,  
even given in cmd line
- Cmd line overrides ini file
- Copied into dest. folder

*eegBidsCreator.ini*

```
[GENERAL]

;;Custom label to distinguish different session inside same acquisition
;;SessionId=My_Session
;

;;Custom label identifying the task, should be short mnemonic.
;TaskId=My_Task

;;Custom label to distinguish different set of parameters used
;AcquisitionId=My_Acquisition

;

;;A path to json file containing metadata about the EEG recording
;JsonFile=My_description.json

;;Path to destination folder, which will contain BIDS formatted
;;ady files of same session, these files will be erased
;OutputFolder=.

;

;;Verbosity level of standard output, one of [DEBUG, INFO, WARNING]
;;less verbose, default is INFO
;LogLevel=INFO

;

;;Path to the log file what will contain all the messages
;LogFile=.

;

;;Select format to convert, one of [BrainVision,EDF]
;Conversion=
```

# Configuration file

Section: GENERAL

## Sample ID

- *SessionId* =  
Custom label to distinguish different session for same subject
- *TaskId* =  
Custom label identifying the task
- *AcquisitionId* =  
Custom label to distinguish different set of parameters used for acquiring the same task

# Configuration file

## Section: GENERAL

### Input and output control

- *OutputFolder = .*  
Path to destination folder, which will contain BIDS formatted data.
- *Conversion =*  
Select format to convert, one of [BrainVision,EDF]
- *CopySource = yes*  
To copy original files into source directory
- *MemoryUsage = 2*  
Estimated memory allowance for the process in GB, increasing could increase the speed of execution

# Configuration file

## Section: LOGGING

### Logging control

- *LogLevel = INFO*  
Verbosity level of standard output, one of [DEBUG, INFO, WARNING, ERROR, CRITICAL], from more verbose to less verbose
- *LogFile =*  
Path to the clone log file, if you prefer to read it in a fixed location instead of temporary one
- *Quiet = no*  
Set to “yes” to suppress output to terminal

# Configuration file

## Section: DATATREATMENT

### Data control

- *DropChannels =*  
A comma separated list of channels to ignore
- *StartTime =/EndTime =*  
Crop data to the interval specified by StartTime and EndTime. Time must be specified in format:YYYY-MM-DD HH:MM:SS.ffffff
- *StartEvent =/EndEvent =*  
Crop data to the interval defined by StartEvent and EndEvent
- *IgnoreOutOfTimeEvents = yes*  
Ignore events occurring outside the data taking period
- *IncludeSegmentStart = no*  
Add additional event for start of each segments
- *MergeCommonEvents = yes*

# Configuration file

Section: RUNS

## Runs splitting

- *SplitRuns = no*  
Set to “yes” to split the data in runs instead of one continuous record
- *MainChannel =*  
Name of channel which will define the splitting. Must be defined if SplitRuns is set to “yes”
- *MinSpan = 0*  
Minimal time span (in mins) to keep a run

# Configuration file

## Section: ANONYMIZATION

### Anonymization control

- *Anonymize = yes*  
To anonymize data, by changing subject name and start date of recording to given values
- *StartDate = 1973-03-01*  
Force recording to start at this date, in format YYYY-MM-DD. If set to None, start date will not be anonymized
- *SubjName = John Doe*  
Force Subject name to this value. If set to None, the name will not be anonymized
- *BirthDate =*  
Force Subject Birth date to this value, in format YYYY-MM-DD. If set to None, the birth date will not be anonymized

# Configuration file

Section: BRAINVISION

## Output control

- *Encoding = UTF-8*

Encoding: either UTF-8 or ANSI, specifies encoding used for .vhdr and .vmrk files

- *DataFormat = IEEE\_FLOAT\_32*

DataFormat: one of IEEE\_FLOAT\_32, INT\_16, specifies the number format for .eeg file

- *Endian = Big*

Endian: either Little or Big, specifies Endianess of written data

# Configuration file

Section: EDF

## Output control

- *DataRecordDuration = 10*

Duration of data record segment, in seconds. Shorter duration slightly reduces file size, but increases time to read file

# Backup slides

# Measured values decoding

- Values stored as raw integer  $h \in [-32768, 32767]$
- Original values retrieval:
  - Embla:  $v = h \times g \times 10^{-9}$ , where gain  $g$  an integer
  - EDF:  $v = h \times r$ , with resolution  $r$   
$$r = \frac{\text{ph\_max}}{32767}$$
  - $g = \text{int}(r \times 10^9)$  – Values are rounded up

## Example: M1

	raw value $h$	Gain $g$	$\text{ph\_max}$	Final value $v$
Embla	32767	31	$1005.13 \times 10^{-6}$	$1015.777 \times 10^{-6}$
EDF	32767	$30.675 \times 10^{-9}$	$1005.13 \times 10^{-6}$	$1005.13 \times 10^{-6}$

# Embla conversion to EDF

- Embla uses original range  $r$  to calculate gain  
Transform range  $r'$  to convert to EDF
- $h' = \text{int}(v_{\text{Embla}}/r')$
- $v' = \text{int}(h \frac{r}{r'})r' - \text{Second rounding}$

## Example: M1

	raw value $h$	resolution $r$	ph_max	Final value $v$
Embla	32767	$31 \times 10^{-9}$	$1005.13 \times 10^{-6}$	$1015.777 \times 10^{-6}$
Embla'	32767	$91.5555 \times 10^{-9}$	$3000 \times 10^{-6}$	$1015.808 \times 10^{-6}$
EDF	32767	$30.675 \times 10^{-9}$	$1005.13 \times 10^{-6}$	$1005.13 \times 10^{-6}$
Embla	2	$31 \times 10^{-9}$	$1005.13 \times 10^{-6}$	$0.620 \times 10^{-9}$
Embla'	1	$91.5555 \times 10^{-9}$	$3000 \times 10^{-6}$	$0.916 \times 10^{-9}$
EDF	2	$30.675 \times 10^{-9}$	$1005.13 \times 10^{-6}$	$0.614 \times 10^{-6}$