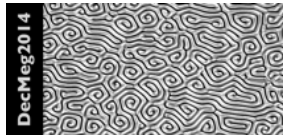


[Customer Solutions](#)[Competitions](#)[Community ▼](#)[Christophe](#)[Logout](#)

Thank you for accepting the rules.



DecMeg2014 - Decoding the Human Brain

2 months to go

Monday, April 21, 2014

\$5,000 • 104 teams

Sunday, July 27, 2014

Dashboard

[Home](#)[Data](#)[Make a submission](#)[Information](#)[Description](#)[Evaluation](#)[Rules](#)[Prizes](#)[Timeline](#)[Forum](#)[Leaderboard](#)[My Team](#)[My Submissions](#)

Forum (27 topics)

Clarification to rule on No External Datasets
13 hours ago

Publication restrictions
yesterday

Team
yesterday

Using shared code
yesterday

Segmenting the signal and gaining features
yesterday

Clarification please: Public and Private split on the leaderboard
2 days ago

[Competition Details](#) » [Get the Data](#) » [Make a submission](#)

Data Files

| File Name | Available Formats |
|-------------------|----------------------------------|
| train_01_06 | .zip (1.42 gb) |
| train_07_12 | .zip (1.43 gb) |
| train_13_16 | .zip (970.13 mb) |
| test_17_23 | .zip (1.64 gb) |
| random_submission | .csv (31.72 kb) |

The training data consist of 9414 trials, i.e. MEG recordings and the class labels (Face/Scramble), from 16 subjects (subject01 to subject16). The test set comprises 4058 MEG recordings from 7 subjects (subject17 to subject23), without class labels. For each subject **approximately** 580-590 trials are available. Each trial consists of 1.5 seconds of MEG recording (starting 0.5sec before the stimulus starts) and the related class label, Face (class 1) or Scramble (class 0). The data were down-sampled to 250Hz and high-pass filtered at 1Hz. 306 timeseries were recorded, one for each of the 306 channels, for each trial. All the pre-processing steps were carried out with [mne-python](#). The trials of each subject are arranged into a 3D data matrix (trial x channel x time) of size 580 x 306 x 375. For each subject in the train set, the 3D data matrix and the associated vector of 580 class labels are saved into a file named *train_subjectXX.mat*. The data of the subjects in the test set are saved in files named *test_subjectXX.mat*. These files are in Matlab v5.0 format and can be easily read from Octave/Matlab, Python (see [scipy.io.loadmat\(\)](#)), R (see the [R.matlab](#) package), Java (see JMatIO) and from other programming languages.

Notice that the all train and test files are conveniently grouped into 4 large zip files: train_01_06.zip, train_07_12.zip, train_13_16.zip and test_17_23.zip.

The MEG sensors are spatially arranged as described in the file [Vectorview-mag.lout](#), which is a 2D approximation of the actual layout (see the picture in the [description page](#)). The 306 sensors are grouped, three at a time, in 102 locations. At each location two orthogonal gradiometers and one magnetometer record the magnetic field induced by the brain currents. The magnetometer measures the z (radial) component of the magnetic field, while the gradiometers measure the x and y spatial *derivative* of the magnetic field.

Notice that if the number in the sensor name ends with "1" then it is a magnetometer.

If it ends with "2" or "3", then it is a gradiometer. For example, "MEG 0113" is a gradiometer and "MEG 0111" is a magnetometer.

Simple Python and Matlab code examples to load the data and to create a submission file, are available [on GitHub at this link](#). Examples of code for other programming languages may be available in future. Contributions are welcome.

File descriptions

- **train_XX_XX.zip** - a collection of the files of certain subjects of the trainset (4-6 subjects each zip file).
- **test_17_23.zip** - a collection of all the test data from subject17 to subject23.
- **train_subjectXX.mat** - the training data of subject XX.
- **test_subjectXX.mat** - the test data of subject XX.
- **random_submission.csv** - a sample submission file in the correct format.

Data fields

Each file train_subjectXX.mat has the following fields:

- **X** : the 3D data matrix, *trial* x *channel* x *time* (580 x 306 x 375).
- **y** : the vector of the class labels (580), in the same order of the trials in X. The classes are **0** (Scramble) and **1** (Face).
- **sfreq** : the sampling frequency in Hz, i.e. 250.
- **tmin**: the time, in seconds, of the beginning of the trial from the stimulus onset, i.e. -0.5
- **tmax**: the time, in seconds, of the end of the trial from the stimulus onset, i.e. 1.0.

Each file test_subjectXX.mat has the same fields as above with the exception of the absence of field **y** and the addition of this field:

- **Id**: a numerical identifier of each trial of the test set. The identifier is of 5 digits (SSTTT). The first two digits (SS) indicate the subject number (17 to 23) and the last three digits the trial number (000 to 580). See the [Submission File](#) instructions for details on how to use this field for preparing a submission.