Data extraction projet

List of students in the group:

- Alejandro JORBA
- Killian VARESCON
- Camille DUBOIS

Proposed topic for the project

We want to collect a dataset from at least one parliament to unlock further analysis of, example given, the votes of MPs (members of parliament), the topics discussed in the parliament, the composition of the assembly, etc...

Our high-end objective is to gather information from two to three parliaments (see below) which would also unlock cross-checking of e.g. alliances in EU vs. country-level parliaments, differences in topics tackled in each, statistic on the composition of the assembly, etc...

We found three parliaments that seem to provide access to open data – though each has its own format:

- The UK parliament: provides an API on their developer hub,
- The EU parliament: same,
- The French Assemblée Nationale: the open-data website doesn't seem to work properly ; we will consider this source depending on the evolution of the data availability.

Therefore, we predict that most of the data should be collected via APIs.

We will need to clean the heterogeneous data from each parliament and make it compatible. We have not figured out the data storage structure that we will use.