*General description:*

There are **two** kinds of input:

* A list of URLs that contains the documents to be analyzed (Academic client).
* A list of projects’ descriptions brought by the team (Business client).

For both cases, the system is expected to receive the information in Excel/CSV files for it to be adapted to the predictive model’s dataframe.

*Structure*

For the URL input, the expected structure is:

* In each row, a URL from Scopus. From the URL, Scopus database can be queried with Bibliometrix and the required information (Title, Abstract, Keywords) is retrieved in a Bibtex file that can be used for bibliographic statistics.

For the project list input, the expected structure is:

* In each row, there is the title of the project, the summary and keywords determined by the working team that manages the project. This option does not allow bibliographic statistics thanks to the lack of Scopus references.