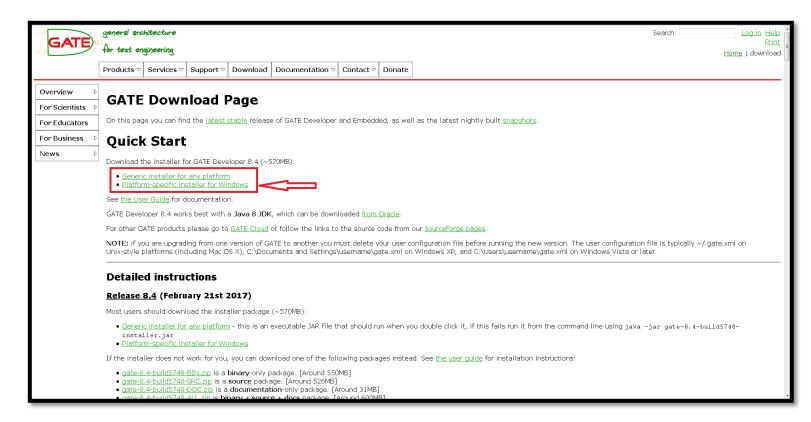
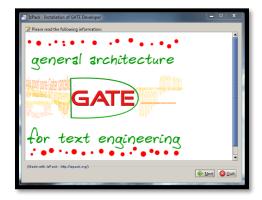
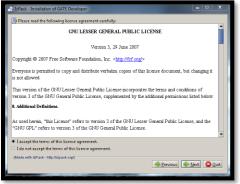
Maciej Medyk - CAT6930 - Gate in Natural Language Processing

Report on how to install and use Gate Developer

Gate Developer is an open source Information Extraction tool that allows to analyze text in order to extract named entities and relationships between words. Gate developer will allow you to analyze text in various document formats like pdf, txt, html, xml, etc. To install it you proceed to https://gate.ac.uk/download/





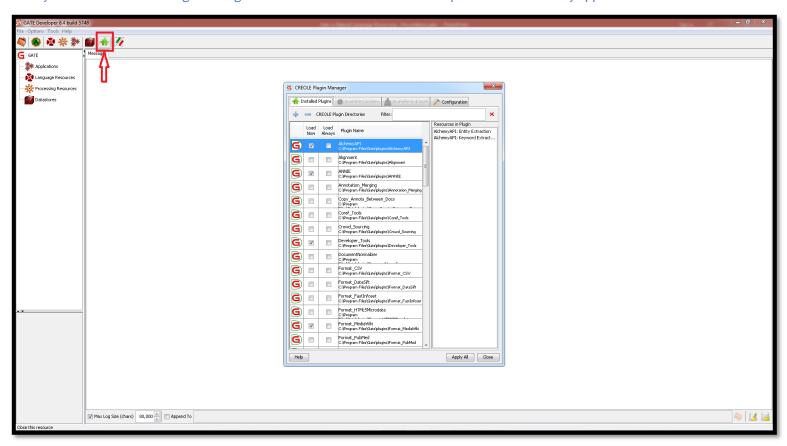




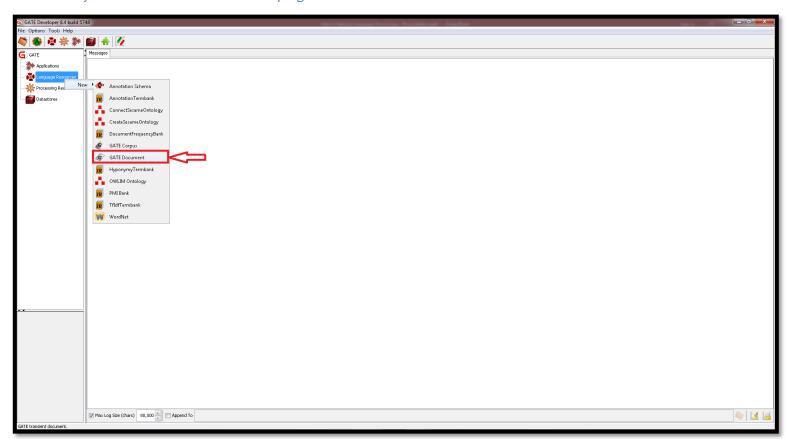
Once you install the software you will be able to access the software interface. Once loaded, the interface will be divided into left tool panel and main window. On the left tool panel, there are four categories: Applications, Language Resources, Processing Resources, and Datastores. In Applications, there is ability to create your new application or use ready made applications. There are three applications loaded by default (Annie, LingPipe, and OpenNLP) and many more available through Creole Plugin Manager. In Language Resources, there is ability to add documents from plain text, pdf documents, html, xml, csv, owl, and directly from urls. Processing Resourcess show what resources will be used to analyze text and Datastores offer a persistence service to the documents.

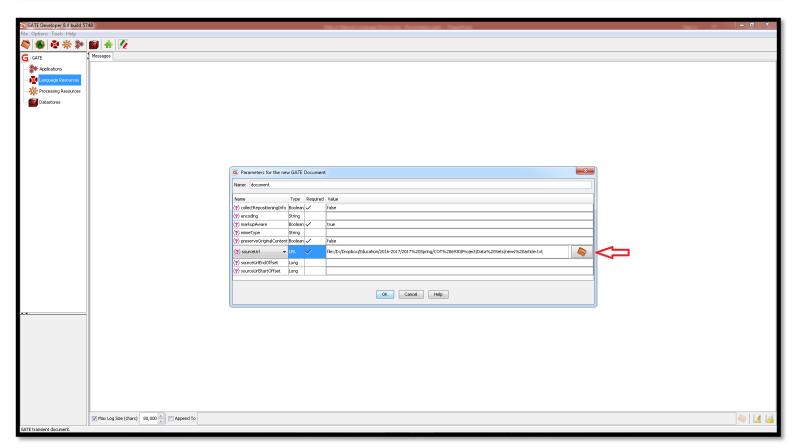


Here you can select Creole Plugin Manager that will allow to download more specialized made ready applications.

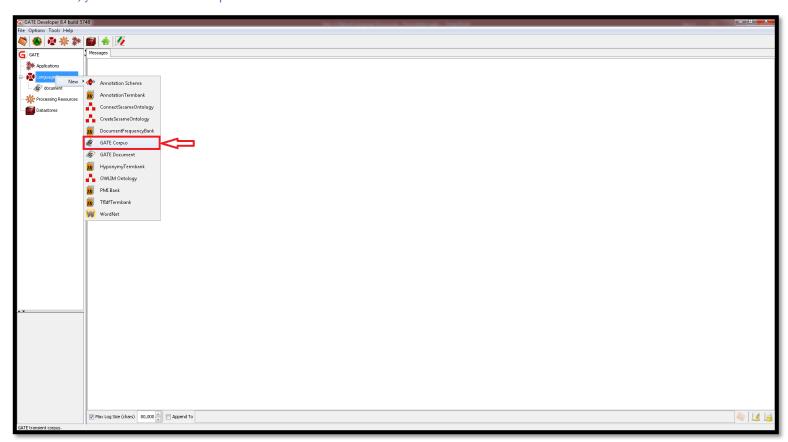


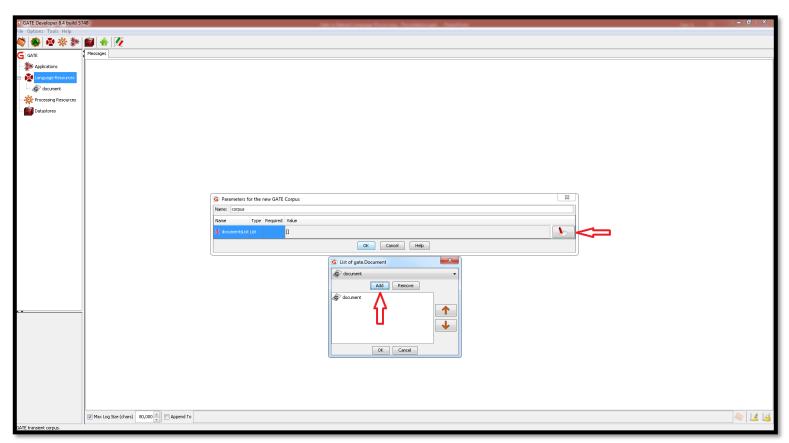
Here is how you load a text document into the program.



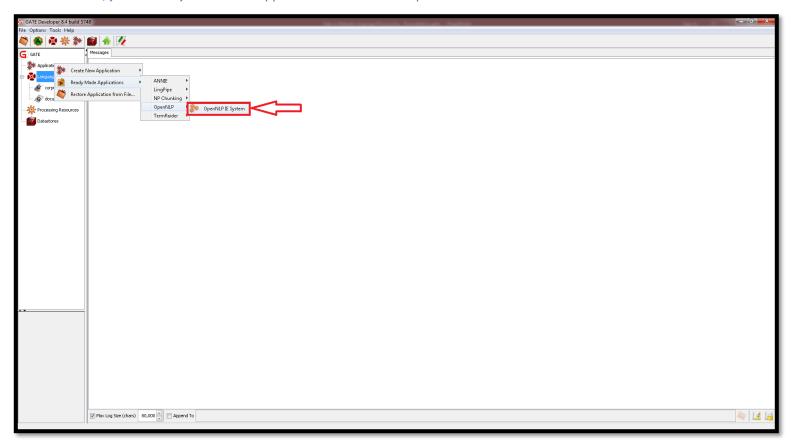


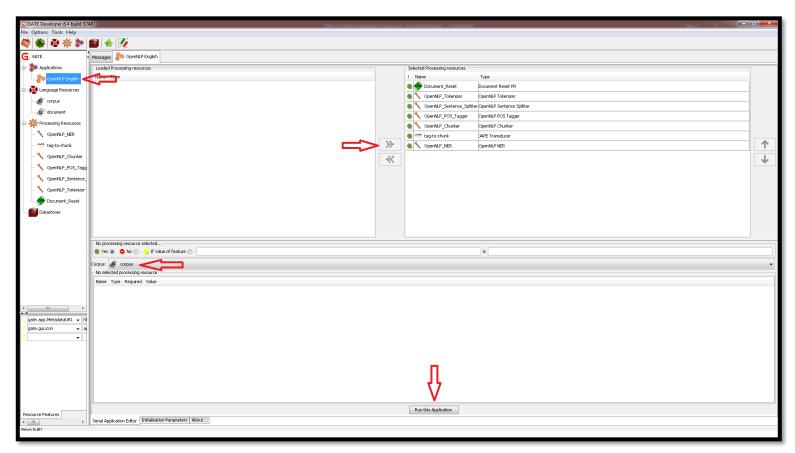
Afterwards, you have to create a corpus that will contain the document or set of documents.



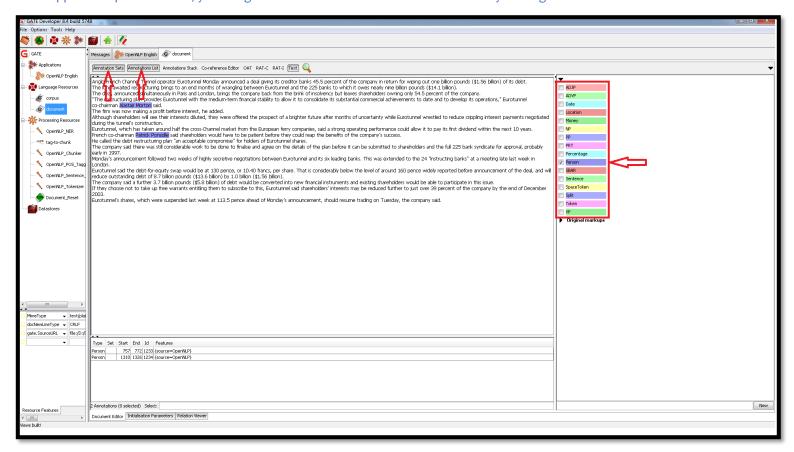


At this moment, you are ready to create an application. We will select OpenNLP

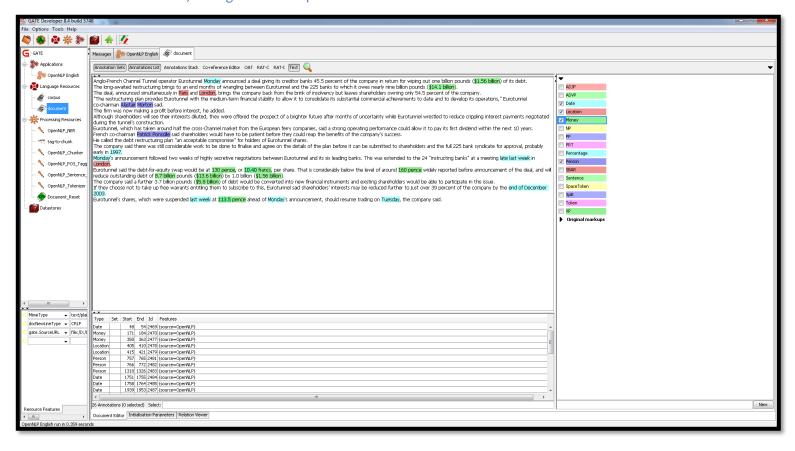




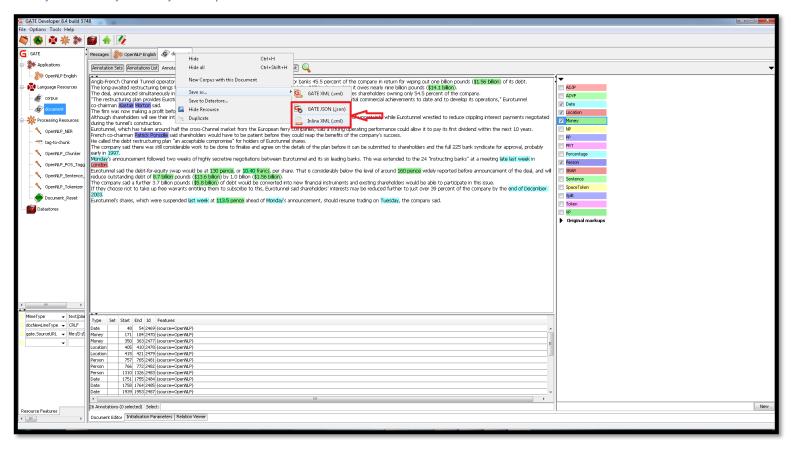
Once application process is run, you can go over the results for each text document by clicking on Annotation Sets and Lists.



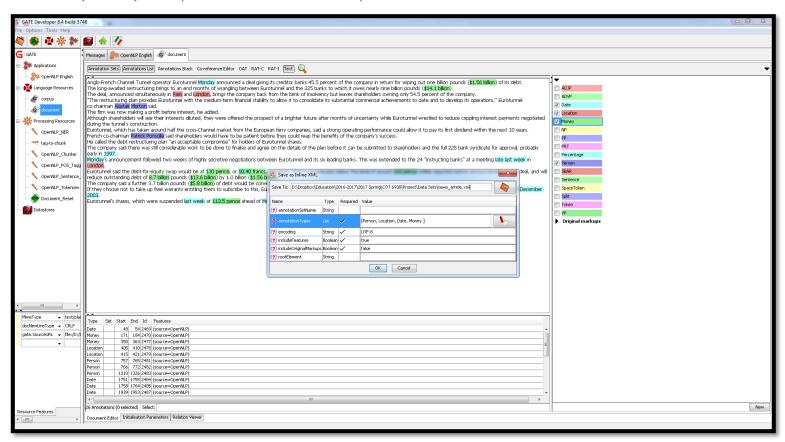
Once annotated and classified, the tag will become part of the text.

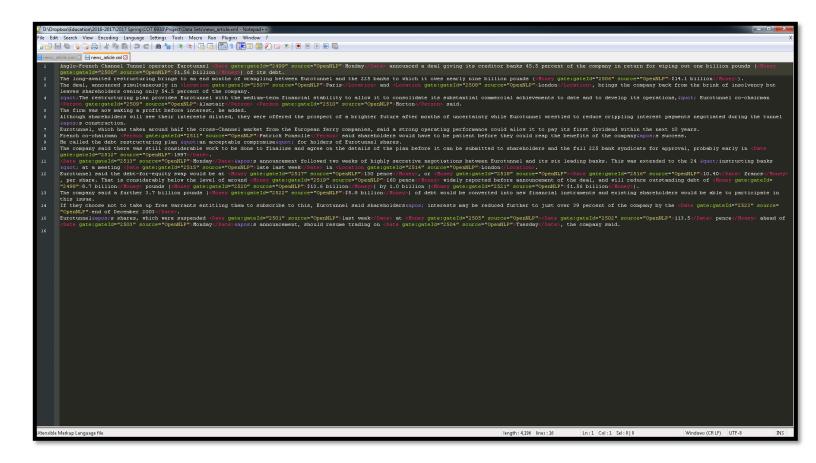


Now you are ready to save your output to the XML or JSON file.

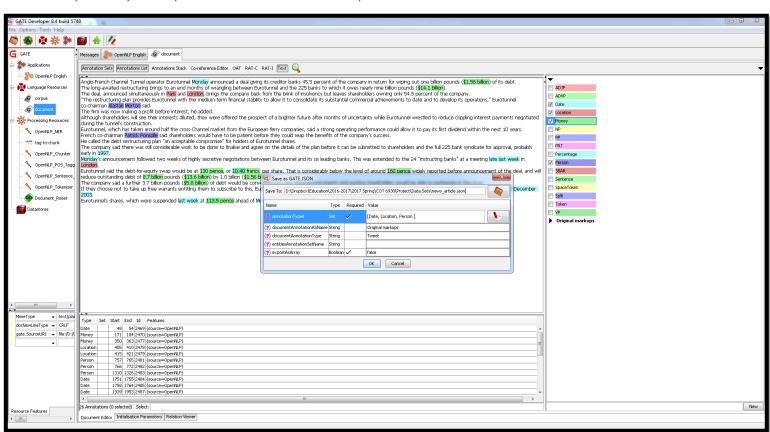


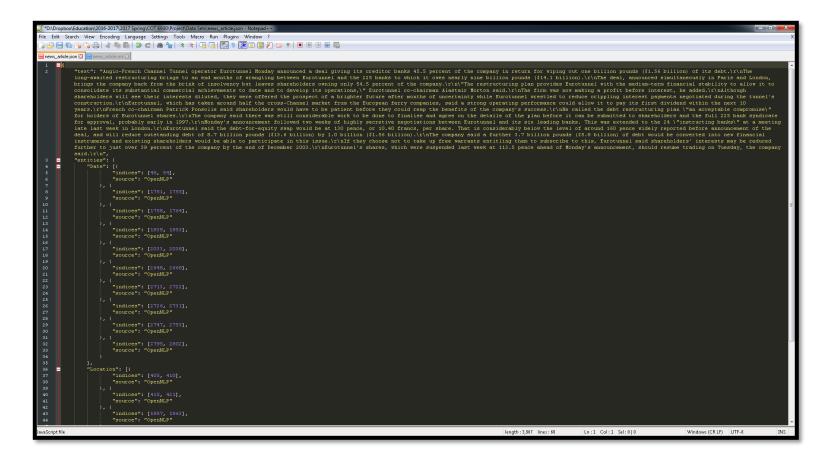
Here is how you save your output as XML file and how the output looks like in file.



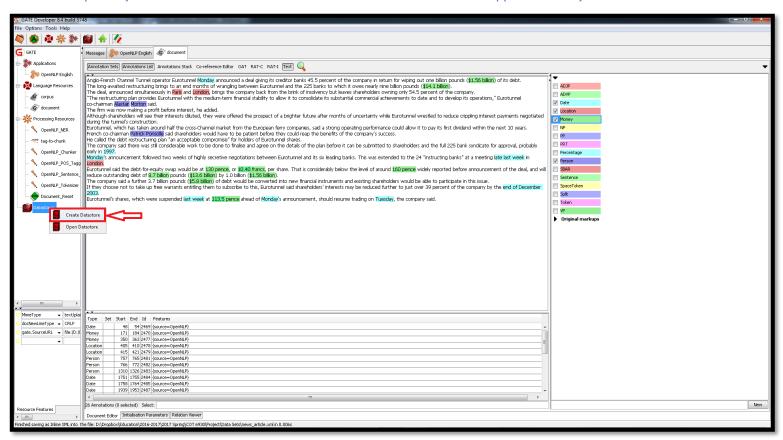


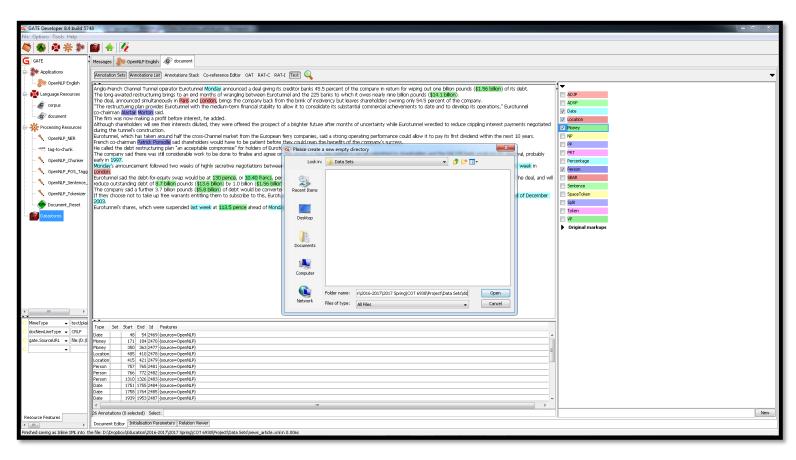
Here is how you save your output as JSON file and how the output looks like in file.

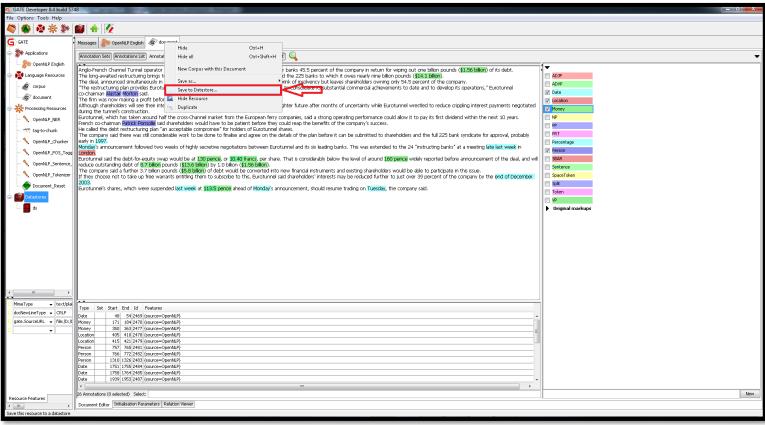




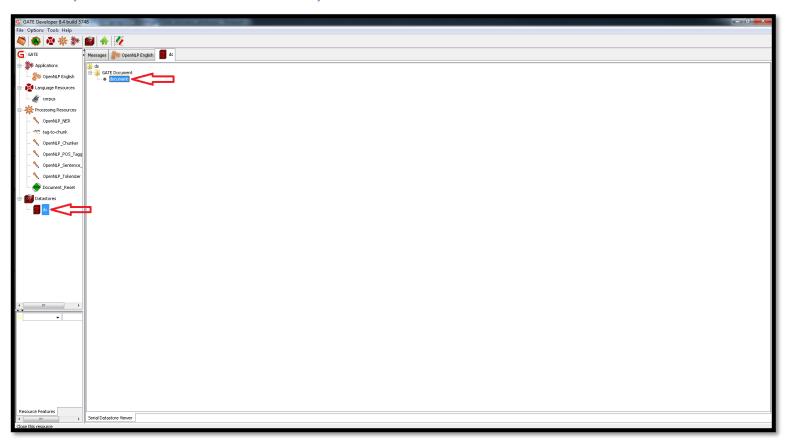
You can also persist your work in Datastore for later use with all the annotations that application and you made.







To restore your source document with all annotations you do click on Datastore and choose which file to restore



Now you can see restored document under Language Resources

