



Gate In Natural Language Processing

Maciej Medyk

Presented In Class COT6930 – Natural Language Processing on April 06, 2017

Video Available On YouTube At The Link : <https://www.youtube.com/watch?v=SIvfZOmHjYM&t=8s>

Overview



- Introduction
- Installation
- Interface
- Build In Tools
- Demonstration
- Conclusion

Introduction

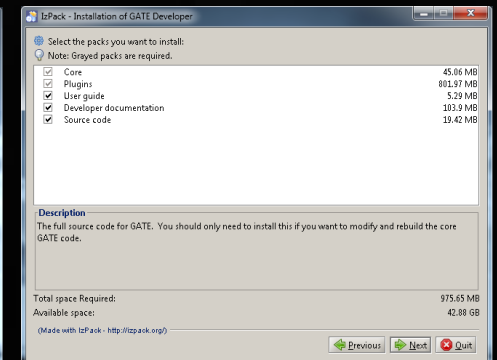
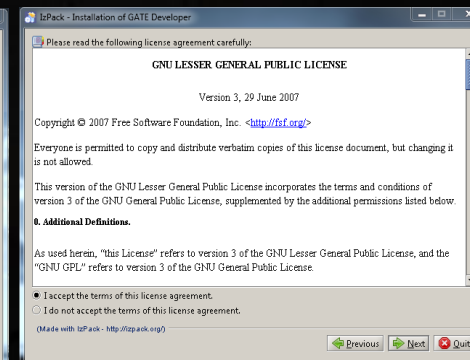
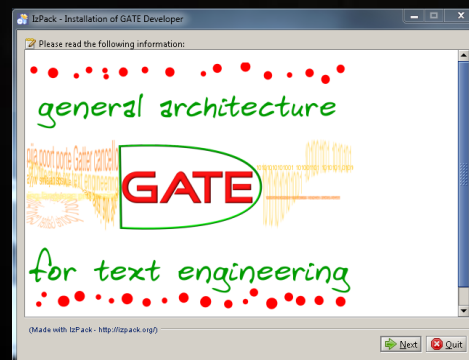
- Gate Developer is an open source Information Extraction tool that allows to analyze text in order to extract named entities and relationships between words.
- The concepts of the Gate Developer are:
 - **documents** to be annotated
 - **corpora** comprising sets of documents, grouping documents for the purpose of running uniform processes across them
 - **annotations** that are created on documents
 - **annotation types** such as 'Name' or 'Date'
 - **annotation sets** comprising groups of annotations
 - **processing resources** that manipulate and create annotations on documents
 - **applications**, comprising sequences of processing resources, that can be applied to a document or corpus

Installation

- To install visit link below : <https://gate.ac.uk/download/>
- There are Windows installer or universal installer using jar file that will require Java to be installed



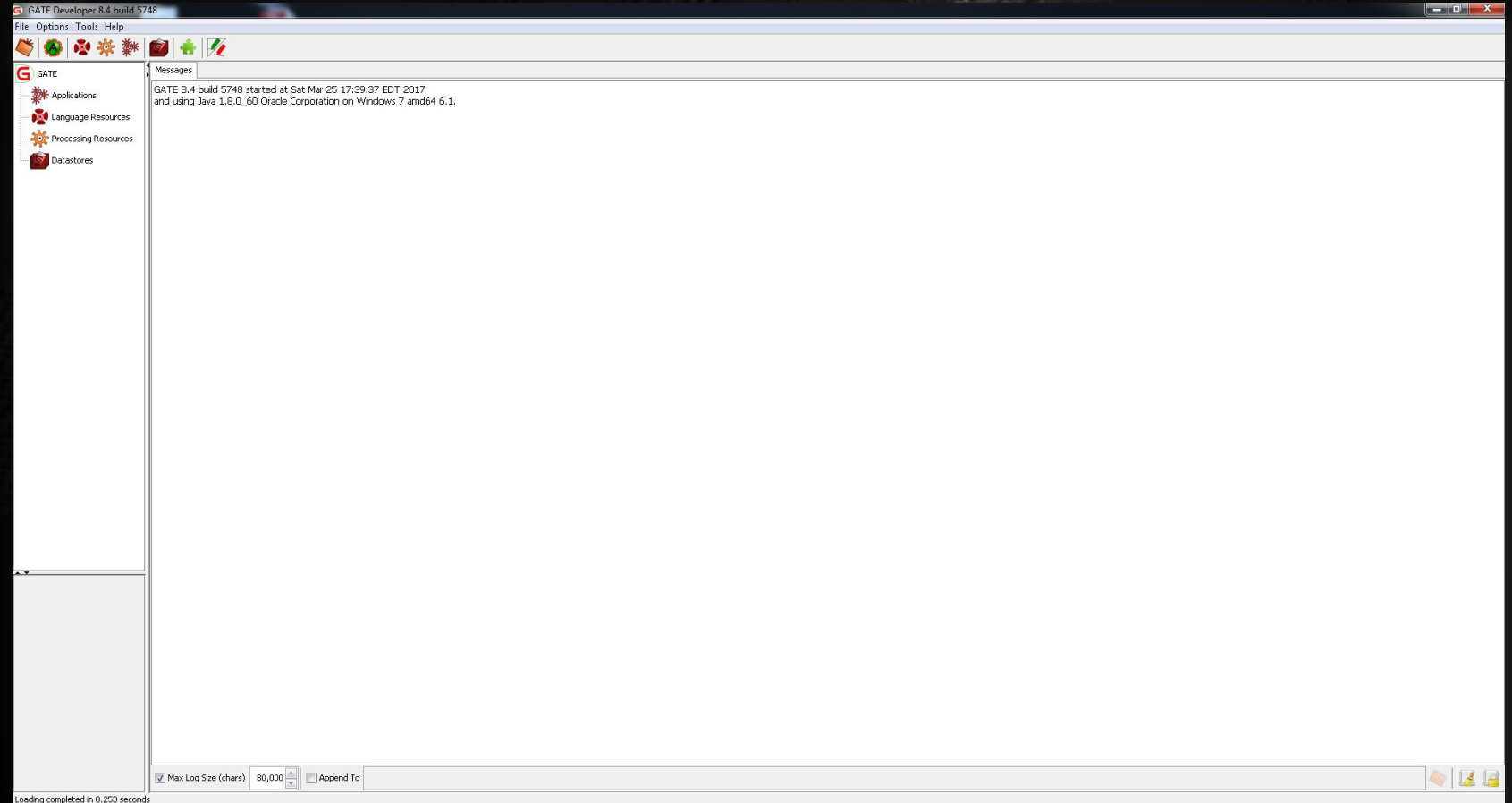
The screenshot shows the GATE Download Page. The header includes the GATE logo and navigation links: Products, Services, Support, Download, Documentation, Contact, and Donate. The main content area is titled "GATE Download Page" and "Quick Start". It provides instructions on how to download the installer for GATE Developer 8.4 (~570MB). Two links are highlighted in a red box: "Generic installer for any platform" and "Platform-specific installer for Windows". Below this, there is a "Detailed instructions" section for Release 8.4 (February 21st 2017), which lists the requirements for Java 8 JDK and provides links to the source code and documentation.



Interface

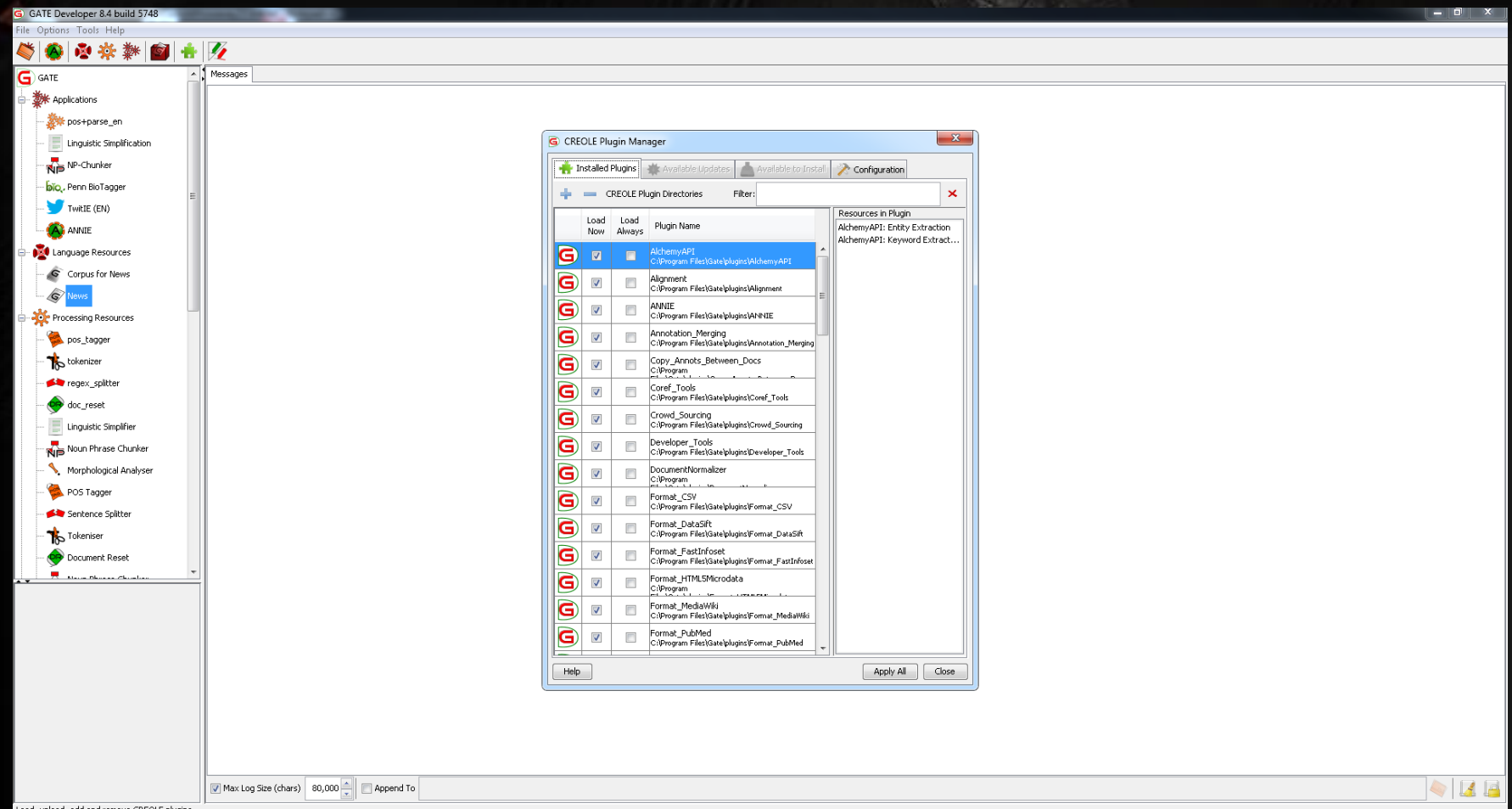
Gate has 4 categories

- Applications
- Language Resources
- Processing Resources
- Datastores



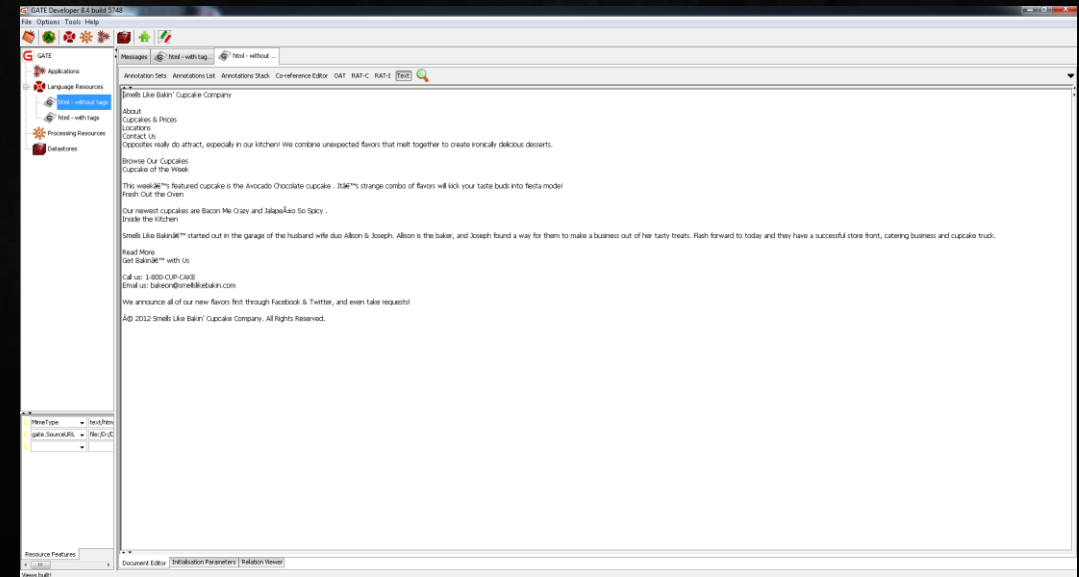
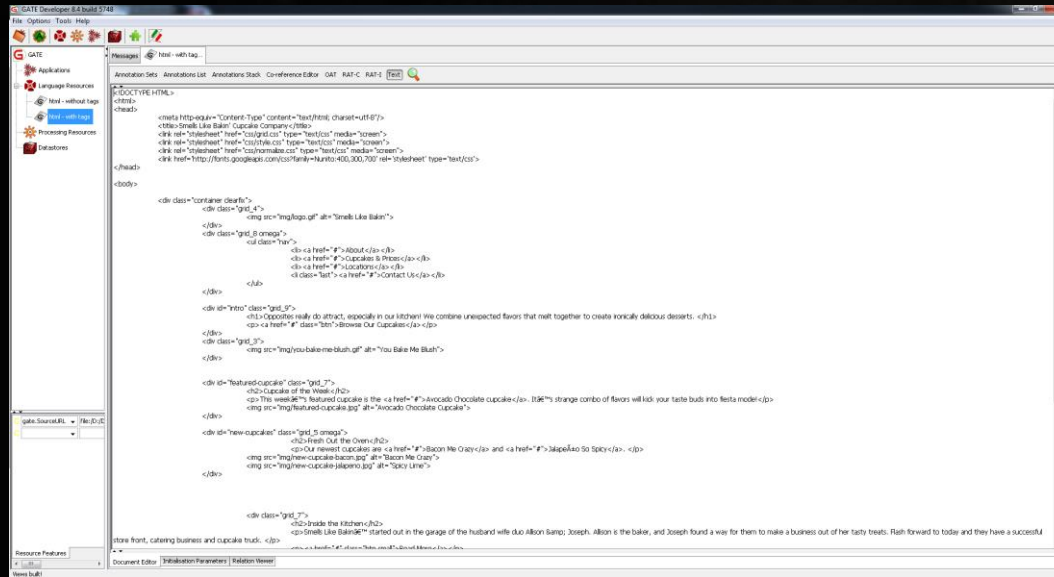
Build In Tools – Applications

- Annie
- LingPipe
- OpenNLP
- NP Chunking
- Sanford Parser
- TermRaider
- TwitIE
- TimeML



Build In Tools – Documents

- Gate allows to load various document types : txt , html , xml , pdf , mail , eml , sgml
- When loading tagged text like html or xml you have option of loading it with and without tags.



Demonstration

The background of the slide is a dark, moody image. It appears to be a large, circular, metallic structure, possibly a tunnel or a large wheel, with a bright blue horizontal line across the middle. The structure has a textured, metallic surface and some small, glowing orange lights. The overall atmosphere is mysterious and industrial.

Conclusion

Thank You