LinGUIne – Language Data Analytics Toolkit

Team: Pro-Nouns (Peter Dimou, Kyle Mullins, Peter Maresca, Matthew Talbot)
Coach: Robert Kuehl
Sponsor: Cecilia Ovesdotter Alm
Year: 2014

Basic Perspective

For users that are new to LinGUIne or to the Natural Language Processing domain.


Designed to be simple and straightforward.

Advanced Perspective

For users with domain experience or more technical knowledge.

Designed to show the user more complex layers of information.

Gives the user more control over their workflow.

Facilitates and encourages experimentation with project data.

Problem

Current Natural Language Processing tools require a high degree of technical proficiency

Visualizations of the results are very basic, if present at all

Important functions such as data annotation are often not part of the existing tools

Results are often unorganized or confusing

Features

- Open Source Software
- Extensibility through Plugin System
- Centralizes renowned NLP software such as Stanford NLP and NLTK
- Intuitive, user-friendly UI
- Accepts various forms of data, including social media data from Twitter
- Project Workspace keeps results organized

Process and Design

Iterative release cycle with thorough upfront requirements, risk assessment and architecture.

Extensibility

Plugins are used to add functionality to LinGUIne as future plugins

Existing NLP tools can be added to LinGUIne as future plugins

Future Work

- More import and export file formats
- Other social networks as data sources
- Scrape data from the web
- More analyses and visualizations
- Illinois Curator NLP software module
- Support for audio and video analyses
- Temporal alignment of multimodal data
- Trained model support
- Nested annotations