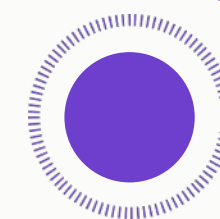




Text Analysis Program

Hannah Martinez



What the program will do:

Frequency

- Keep track of frequency of words in the database
- Hash Map

Store

- Tokenize and store corpora automatically
- List

Relationship

- Show the relationship between multiple words in the corpora.
- Graph

Concordance

- Show a list of words that appear before and after a specific word
- Trie

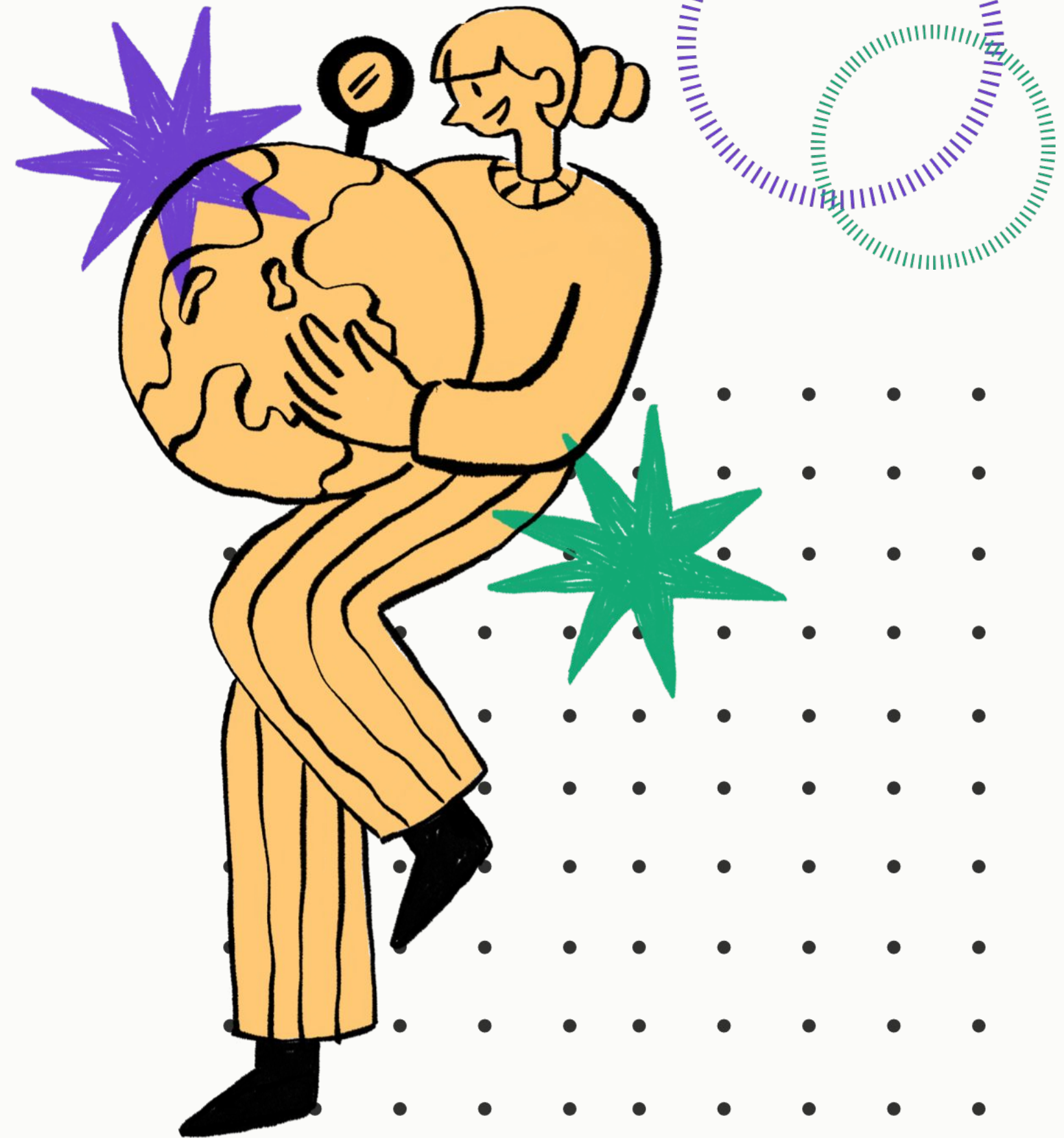
Breakdown of Corpora:

Sourcing

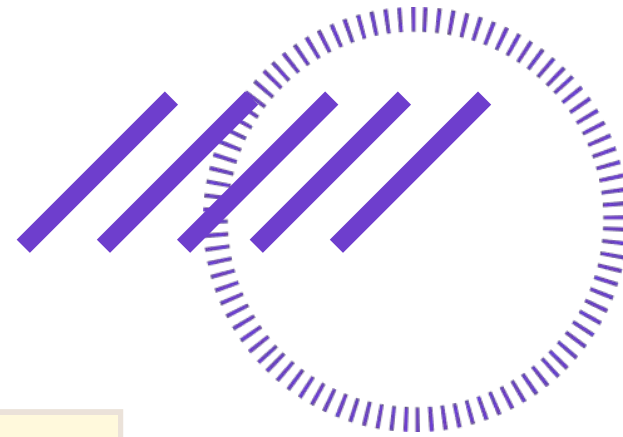
The corpora is built of old syllabus from the CS department at SDSU found under the Syllabus Catalog of the University.

User-Interface

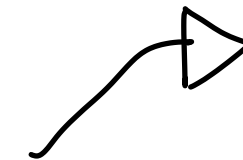
User can search for a string or a word, and will be given an analysis of how the phrase/word is used in the corpora.



GUI Mockup



select a department:



able to select
syllabi from
desired dept.

enter a word:



enter keyword

results for: grade

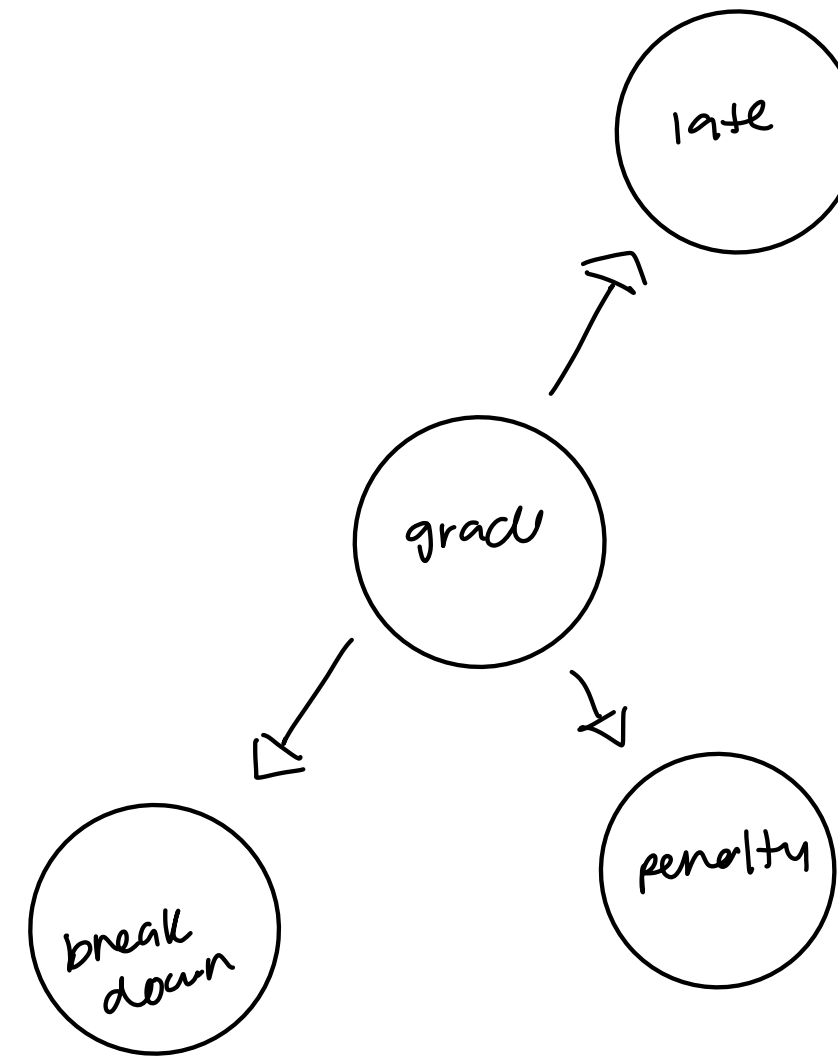
frequency: 24

most commonly used

before: after:

list of uses:

~~~~~	grade	~~~~~
~~~~~	grade	~~~~~
~~~~~	grade	~~~~~





# Actual Implementation

