
PROJECT PROPOSAL

WHICH TASKS ARE COMPLETED? WHICH ARE PENDING?

Below is the original list of tasks, with a status next to each.

Task	Assigned To	Est. Hours	Status
Create dot-matrix, cosine similarity, bag-of-words representation, mean average precision, and stop-word functionality in JavaScript	Devin Burke	15	Completed
Create pre-defined Cranfield methodology data sets, measures, and relevance judgments	Joshua Ray	5	In Progress
Implement Cranfield data sets in JavaScript using mean average precision as a measure	Joshua Ray	5	In Progress
Create algorithm to score and rank sentences based on implemented techniques, and adjust according to success based on Cranfield evaluation	Devin Burke	10	In Progress
Determine methodology for filtering page document to only applicable content (i.e., filter out ads, navigation, etc.)	Joshua Ray	5	Completed
Create underlying Chrome browser extension and configurability of what percent of the document should be returned	Joshua Ray	15	In Progress
Research and implement algorithm for calculating reading time-savings based on summary versus entire document	Devin Burke	5	Completed
		60	

ARE YOU FACING ANY CHALLENGES?

We have resolved and/or are in the process of resolving the following challenges:

- We had to create a method to identify and omit non-sentence-based content from consideration, such as captions of images, tables of data, navigation, advertisements, basic headings, etc.
- We had to create a method to identify when a period is and isn't the end of a sentence, such as: 1.23, J. Edgar Hoover, Figure 1., St., is...bad, Ph.D., etc.
- We are in the process of throttling the number of sentences returned for very long pages, so we aren't returning, e.g., 70 sentences for a 700-sentence page.