Alternative Visualizations of News Searches

Carlson Eddie (edc4), Tom Lehmann (tlehmann), Kimberly Nguyen (kimmyyhn),
Derek Reinelt (dreinelt), Raymond Zhang (rayz)

Background: The KNOW Project

The KNOW project (https://depts.washington.edu/knowjsis/home/index.php) aims to create an easily-browsable and searchable collection of news sources from around the world. The collection is curated by members of the KNOW project; this imparts value not found with standard search engines or other online news collections. Currently, users of the KNOW project are directed to a "news list", which groups available news sources by country.

The Problem: Fragmented, Difficult Searching of Sources

Users currently attempting to use the KNOW project encounter obstacles to easy and efficient use. These problems can be largely divided into two sets.

The first set of problems arises because the KNOW project does not currently have a unified search engine that combines all of the news sources. Users must access a news source's web site and attempt searches there. This means that users cannot:

- easily search all news sources at once – users must access each news source's site individually.
- find related articles on different news sources – news sources are not linked to each other.
- use a unified interface to search all new sources – different news sources have different interfaces, and users must learn these interfaces before being able to do their searches.

In addition, even when users are able to search a news source and receive results, these results are often presented in the form of a list, with little to no control over how the elements of the list are sorted or presented. This often means that users cannot:

- group articles by location of publication, date of publication, or by other criteria.
- determine at-a-glance which other topics are related to their search terms.
- find articles similar to those that are returned on the list.

All of these problems combined means that users currently do not enjoy a quick, simple, and efficient way of performing a search and interpreting the results with the push of a single button.

Our Solution: A Unified Search Engine with Alternative Visualizations

We propose designing and implementing a unified search engine for many of KNOW's news sources. Major components of our design include:

- a unified search engine, empowering users to search all of KNOW's news sources with a single query.

- alternative visualizations, allowing users to sort and understand their search results in different ways based upon their needs.

Presenting a unified search engine to users not only reduces the number of tasks they must perform to find articles on a given topic, easing their use of the KNOW database, but allows them to combine information from multiple sources to gain new insights and further their research.

We plan on making many alternative visualizations of search results available to users. Examples include a timeline-based view, where news articles are arranged chronologically so that users can see how news coverage has changed over time, and a map-based view, so users can identify where certain topics are being highlighted in the media. We hope that these alternative visualizations allow users to more quickly and easily find what they are looking for, or discover trends that they hadn't noticed before.

By implementing a search engine with these features, we hope to make the searching experience much easier and more useful to users, increasing the use of the KNOW project, and benefiting the community as a whole.

Implementation:

We plan to implement this unified search engine with three modules. The first module takes a user query, sends that query to individual news sources in KNOW's database, and collects results. The second module analyzes the collected results, and computes connections between the results based on the visualization mode that the user requests. The third module takes the results and connections, and displays it in an interactive fashion to the user.

What interests our group about this project can be divided into two major parts. First, we are interested in the implementation of search engine backend that connects to each individual news source to submit a search query; designing a program to connect with multiple different interfaces may be challenging for us. Second, we are interested in the user interface portion of this project; in the past, members of our group have found user interfaces to be more difficult to design and implement in an orderly fashion.

We currently plan to use Java to implement our search engine, as all of us are familiar with Java, and it poses less challenges for quick and efficient program implementation than other languages; issues such as memory management are already handled, and standard libraries provide pre-existing functionality.

Risks:

Major risks in designing and implementing our project lie in connecting our program to different news sources with different search interfaces. We plan on mitigating this by intentionally restricting the scope of our search engine to a set of news sources with similar search interfaces. Further risks lie in the difficulty of implementing user interfaces in an orderly fashion; user interface code often becomes a messy tangle that is hard to debug and nigh-impossible to fix or update. We plan on mitigating this through careful design, separating our user interface into distinct modules that interact with each other in well-defined ways, and by working closely with each other to ensure that one person does not write code that interferes with somebody else's work.