

Davis Shepherd	dshep
Jason Uanon	juanon
Matthew Kraus	krausmt

KNOW VT (Visualization Toolkit)

Project Vision

The KNOW Visualization Toolkit (VT) comprises a suite of visualization tools that will be used by students of the Jackson School to analyze current events published by internet news sources throughout the world. The KNOW project allows students to provide annotations about news articles from approximately 182 news sources. Students have annotated 3500-4000 articles, however the difficulty is in analyzing this data set. The problem is that students cannot interact with the annotations that they have painstakingly created. The Visualization Toolkit proposed here addresses this problem by allowing students to transform annotation data into meaningful correlations to help students learn about the world and generate research questions.

Although this proposal team has not exhaustively researched whether similar tools exist, the well known General Social Survey (GSS) has a few similarities. The GSS allows researchers to correlate data and perform scientific research about American society; similarly, the KNOW project seeks to allow researchers to ask interesting questions, but instead the data are annotated news articles published throughout the world. This novel approach will allow researchers to correlate news articles about the same subject but which might have differing perspectives because they are published by different agencies in different countries.

High-Level Software Architecture

The existing KNOW project appears to be web centric, therefore, primary development will be concerned with web site development. As the web site is dynamic, PHP is a good candidate because it allows powerful server-side interaction with the back end database (in this case, containing news article annotations). JQuery, which is a robust JavaScript library, will allow content to be served to the web page interactively. To allow users to visualize data, Highcharts (www.highcharts.com) will be used as a foundation to allow interactive JavaScript based charts and graphs to be rendered on the KNOW web site. Using the Google Maps API, students will be able to place social data on maps. Of course custom visualization tools may need to be developed to cater to the specific needs of the KNOW project, but when possible, existing technologies will be employed. Either way, the main engineering challenge of the Visualization Toolkit project is integrating existing and new technologies to serve the needs of researchers.

Risks

Perhaps the most serious challenge will be in limiting the scope of the project. The existing KNOW project is grand in scope yet still in its infancy in terms of features and functionality. Coordinating ambitious desires of the International Studies department within the time allotted will be difficult. One major concern is the data set itself. Is the database normalized to 3NF or BCNF, or is the data set semi-structured? Is the existing data in a usable form? How should extremely non-structured data be interpreted? The project team will need to investigate these questions before even deciding what can feasibly be done in ten weeks. For example, an ill formed database might require reworking in order to perform meaningful queries and to use data for visualizations.

To minimize these risks, the project will build a visualization toolkit that is relatively small but modular, reusable, and well documented. It would be undesirable to have the project scrapped from

the KNOW web site due to lackluster development efforts. This will require constant coordination with the KNOW project team.

Conclusion

The KNOW Visualization Toolkit will provide students of the Jackson School of International Studies with an arsenal of dynamic web-driven content to aid them in their research. Using a modular approach, the Visualization Toolkit will give students the resources to visualize correlations in news stories generated globally, as well as contrast their differences, helping them generate important research questions.