

IDEAL

Spreadsheet

Tony Ardura, Austin Burnett, Rex Lacy, Shawn Neumann

Project Description

IDEAL is utilized to provide visualization aids for researchers making use of large web archive files.

This information will be processed through Hadoop with a spreadsheet like interface.

Why would we do things this way?

Interface, iteration #1

- Using,HTML / CSS, jQuery and javafx

Interface consists of selectable archives

Filter options to restrict output

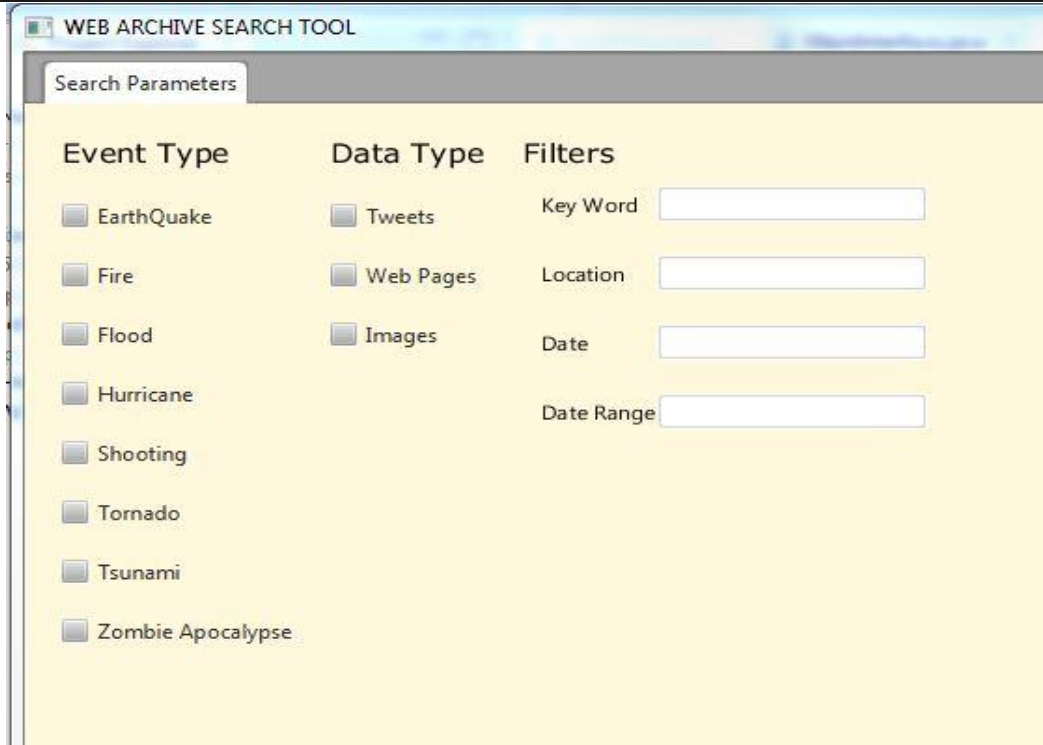
Selectable list items take user directly to result

Iteration # 1

An early idea for a filter tool

Addresses filter needs

But not client specs



The screenshot shows a web application window titled "WEB ARCHIVE SEARCH TOOL". Below the title bar is a tab labeled "Search Parameters". The main content area is divided into three columns: "Event Type", "Data Type", and "Filters".

Event Type	Data Type	Filters
<input type="checkbox"/> EarthQuake	<input type="checkbox"/> Tweets	Key Word <input type="text"/>
<input type="checkbox"/> Fire	<input type="checkbox"/> Web Pages	Location <input type="text"/>
<input type="checkbox"/> Flood	<input type="checkbox"/> Images	Date <input type="text"/>
<input type="checkbox"/> Hurricane		Date Range <input type="text"/>
<input type="checkbox"/> Shooting		
<input type="checkbox"/> Tornado		
<input type="checkbox"/> Tsunami		
<input type="checkbox"/> Zombie Apocalypse		

Iteration #2


More current version:

Web page of results

Ability to continue search

Ability to further refine

IDEAL Spreadsheet

Subject	People Talking	Images Related	Tweets Related
Floods, China	Obama		https://twitter.com/GlobalFloodNews/status/440318944302231552

Add New Search

Iteration #3

Client Feedback

On the right track.
Eliminate check
boxes.

Add feature to apply
filters by dragging
through cells
(similar to excel)

The screenshot shows a web browser window with the title "IDEAL Interface". The address bar displays the file path: file:///C:/Users/spn2460.viz1/Dropbox/Capstone/checkBox.html. The browser's taskbar shows several open applications, including "Scholar Gateway We...", "My VT", and "Inbox - spn2460@vt...".

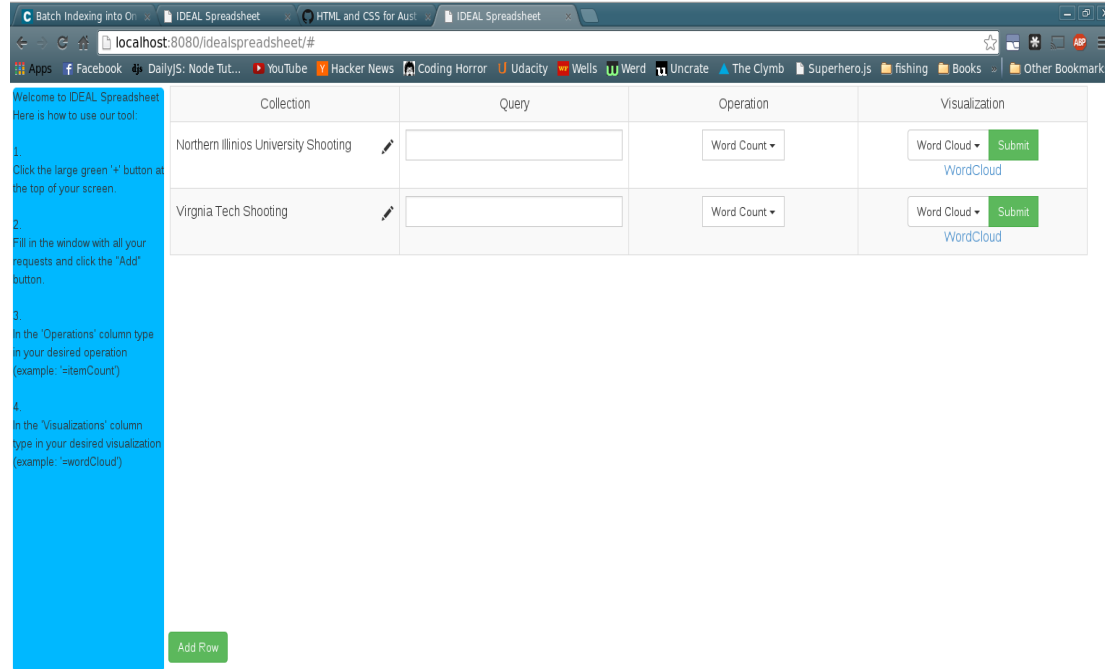
The main content area is titled "IDEAL File Search Interface" and contains the following elements:

- Select Warc File:** A dropdown menu with "Zombie Apocalypse" selected.
- Media Source:** Four radio buttons for "Facebook", "Twitter Post", "Web Page", and "Images", all of which are currently unselected.
- Advanced Filtering:** Four input fields for "Author", "Date", "Location", and "Key Word".
- Search Now:** A button located below the "Author" field.
- Table:** A table with four columns: "File Name", "Facebook Posts", "Tweets", and "Images". Each column contains an empty input field.

Iteration #4, Final Version

Combines spread sheet type feel with ease of use functionality.

Can be modified easily to accommodate future users.



The screenshot shows a web browser window displaying the IDEAL Spreadsheet application. The browser's address bar shows the URL `localhost:8080/idealspreadsheet/#`. The application interface is divided into a left sidebar and a main content area. The sidebar, which is blue, contains a welcome message and four numbered instructions: 1. Click the large green '+' button at the top of your screen. 2. Fill in the window with all your requests and click the "Add" button. 3. In the "Operations" column type in your desired operation (example: `=itemCount`). 4. In the "Visualizations" column type in your desired visualization (example: `=wordCloud`). At the bottom of the sidebar is a green "Add Row" button. The main content area features a table with four columns: "Collection", "Query", "Operation", and "Visualization". The table contains two rows of data. The first row is for "Northern Illinois University Shooting" and the second is for "Virginia Tech Shooting". Each row has a query input field, a dropdown menu for the operation (currently set to "Word Count"), and a visualization section with a "Word Cloud" dropdown, a "Submit" button, and a "WordCloud" label below it.

Collection	Query	Operation	Visualization
Northern Illinois University Shooting	<input type="text"/>	Word Count	Word Cloud <input type="button" value="Submit"/> WordCloud
Virginia Tech Shooting	<input type="text"/>	Word Count	Word Cloud <input type="button" value="Submit"/> WordCloud

Problems Faced

- Requirement Ambiguity
- Project Scope
- UI / Backend Balance

Learned

- Get rid of ambiguity as fast as possible
- Better group organization
- More research

Looking Ahead

- Help section
- SOLR integration
- More operations

Demo Video

<https://www.youtube.com/watch?v=f0gsRTUxQNs&feature=youtu.be>

Questions?