



UNBIASED

Brian Hao, Ricky Chan,
Erica Hendricks-Elliston, Phillip Foo
CSCI 499 Capstone
September 20, 2018



Problem

- News sources, through their word choices, are biased depending on the specific source and the specific story being covered.
- Individuals can be exposed to a less than accurate representation of a topic.

Example

The image shows a screenshot of a news article. At the top, there's a header for 'The New York Times' with navigation links like 'HOME', 'SEARCH', 'SUBSCRIBE NOW', and 'LOG IN'. Below this, the article title 'U.N. Withdraws U-2 Planes' is highlighted with a red box, with 'By REUTERS' underneath. To the right are social media icons. Below this is a 'FOX NEWS' banner with a search bar and a navigation menu including 'Home', 'Video', 'Politics', 'U.S.', 'Opinion', 'Business', 'Entertainment', 'Tech', 'Science', 'Health', 'Travel', 'Lifestyle', 'World', and 'On Air'. The main headline 'U.S. Suspends U-2 Flights Over Iraq' is prominently displayed, with 'U.S. Suspends' highlighted by a red box. Below the headline, it says 'Published March 11, 2003 • Fox News' followed by social media icons. At the bottom, there's a '13 WTHR' logo and a red bar with the text 'Weather Warnings (1)'.

ARCHIVES | 2003

U.N. Withdraws *U-2 Planes*

By REUTERS

FOX NEWS

Home Video Politics U.S. Opinion Business Entertainment Tech Science Health Travel Lifestyle World On Air

U.S. Suspends U-2 Flights Over Iraq

Published March 11, 2003 • Fox News

13 WTHR

Weather Warnings (1)

Iraq forces suspension of U.S. surveillance flights



What is the application?

- One stop news application for users looking to catch up on recent events.
- Intended for **general public**, specifically people that want an impartial view of the world,
- Unbiased is a **web application** that will **aggregate stories based on headlines**.



What can the application do?

- Aggregates news articles across many sources
- Compiles related articles into their own section/space.



Vision

- For the general public to come away with a more **accurate / impartial / nonpartisan / candid view** on **stories / events / the world**.



Implementation

- **Get Data**
 - Scrape news websites for articles
 - Utilize API(s) for news sources
- **Process (and Store) Data**
 - Group news articles that are covering the same event
 - Compare title and date to connect related articles
 - Store in graph DB to store relationship between articles
- **Display Data**
 - Format the data and display it in a user-friendly web-app

Technical Details P1

(Get Data)

- **News aggregation APIs:**
 - NewsAPI.org, Webhose.IO
 - Indexes articles from thousands of sources
 - Unified JSON return format
- **Individual news API:**
 - Use APIs provided by specific news websites
 - More messy



ABC News
abc-news



Aftenposten
aftenposten



ANSA.it
ansa



Ars Technica
ars-technica



Associated Press
associated-press



Axios
axios



ABC News (AU)
abc-news-au



Al Jazeera English
al-jazeera-english



Argaam
argaam



Ary News
ary-news



Australian Financial Review

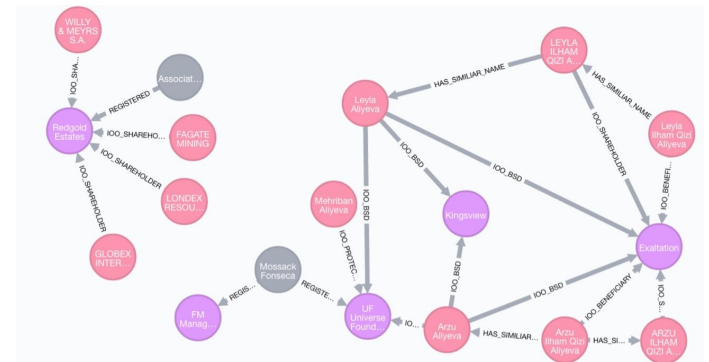


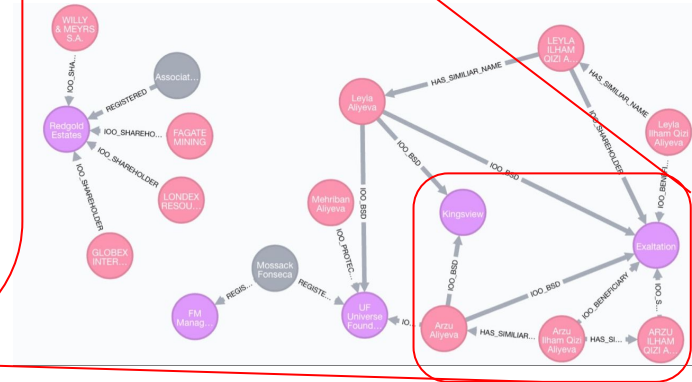
BBC News
bbc-news

[View all \(there's over 30,000 more!\)](#)

- 
- JavaScript

- 





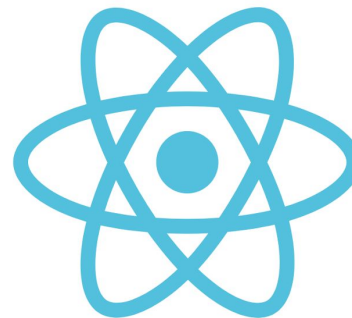
Technical Details P3 (Display Data)

- Front-end: HTML, CSS, React.js

HTML



CSS



React JS

GUI Mockup



Technical Problems

- How do we plan to use the APIs and the languages in website?
 - Use APIs to capture trending articles via JavaScript

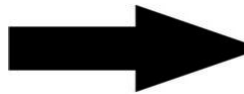
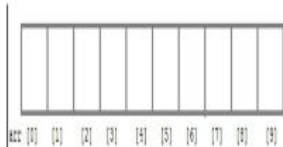


Technical Problems

- How can we accurately group articles into their own section/space?

- Parser

{JSON}



```
author: null,  
title: "Evangelical Leaders Are Frustrated at GOP Caution on Kavanaugh  
Allegation",  
description: "They say religious conservatives may feel little motivation to  
vote in the November elections unless Senate Republicans defend and confirm  
Judge Kavanaugh.",  
url: https://www.nytimes.com/2018/09/20/us/politics/brett-kavanaugh-  
religious-voters.html,  
urlToImage: https://static01.nyt.com/images/2018/09/19/business/00scotus-  
evangelicals-reed/merlin_138966603_07d91755-e0b0-4a0d-88c9-ca64e49e4052-  
facebookImage.jpg,  
publishedAt: "2018-09-20T15:30:04Z",  
content: "\"The worst thing that can ever happen to any woman or man who has  
been a victim is to shut them down and not listen to them,\" said Marjorie  
Dannenfelser, president of the Susan B. Anthony List, a group that opposes  
abortion rights. \"A tragic piece of this is ... [+2673 chars]\""
```



Technical Problems

- How to deal with articles that may fall under two groups?
 - Show them twice
 - Check if already grouped then don't group again



Technical Problems

- How much customizability to give user?
 - Rank their top news sources and store in database
 - Rank their favorite news category

```
source: {  
  id: "the-new-york-times",  
  name: "The New York Times"  
},
```




Technical Problems

- How do we store the data?
 - Article database
 - Archive
 - User database
 - General info
 - Preferences



Project Goals

- Basic working news aggregation platform that is user friendly
- Be able to pull from the top 5 ranking news sites
- Be able to sort all the articles from the news sites into groups with relative accuracy
- Be able to sort the articles displayed (most recent, trending, etc)



Future Features

- Allow users to be able to mark articles as favorites (?)
 - create an account, authentication, protect accounts, secure password+username, security (via database)
- Allow users to report an article as unrelated / incorrectly placed
- Archive of news articles of the same events based on what was trending for a specific date or range of dates
- Comment Section



Metrics

- How do we measure success?
 - General website functionality (Opening links, sorting articles, etc)
 - Basic user convenience
 - How accurately are we able to group the news articles?



Conclusion

- News site for users to go to that will combine all news outlets into a single source which aggregates all headlines of the same topic.
- Compare differing headlines of the same topic in order to determine the most accurate information.
- Work will be distributed equally, such that everyone will learn the workings of both front and backend. Initially focusing on creating a proper backend.

Thank you.

We will now be happy to answer any questions.