

# Amazon Lightsail Containers Hackathon




Team – 3141  
Aggre-Gator

Anuj Koli  
Anurag Shenoy  
Ruchika Upadhyay  
Naveen Kumar  
Aman Pathak



# Introduction

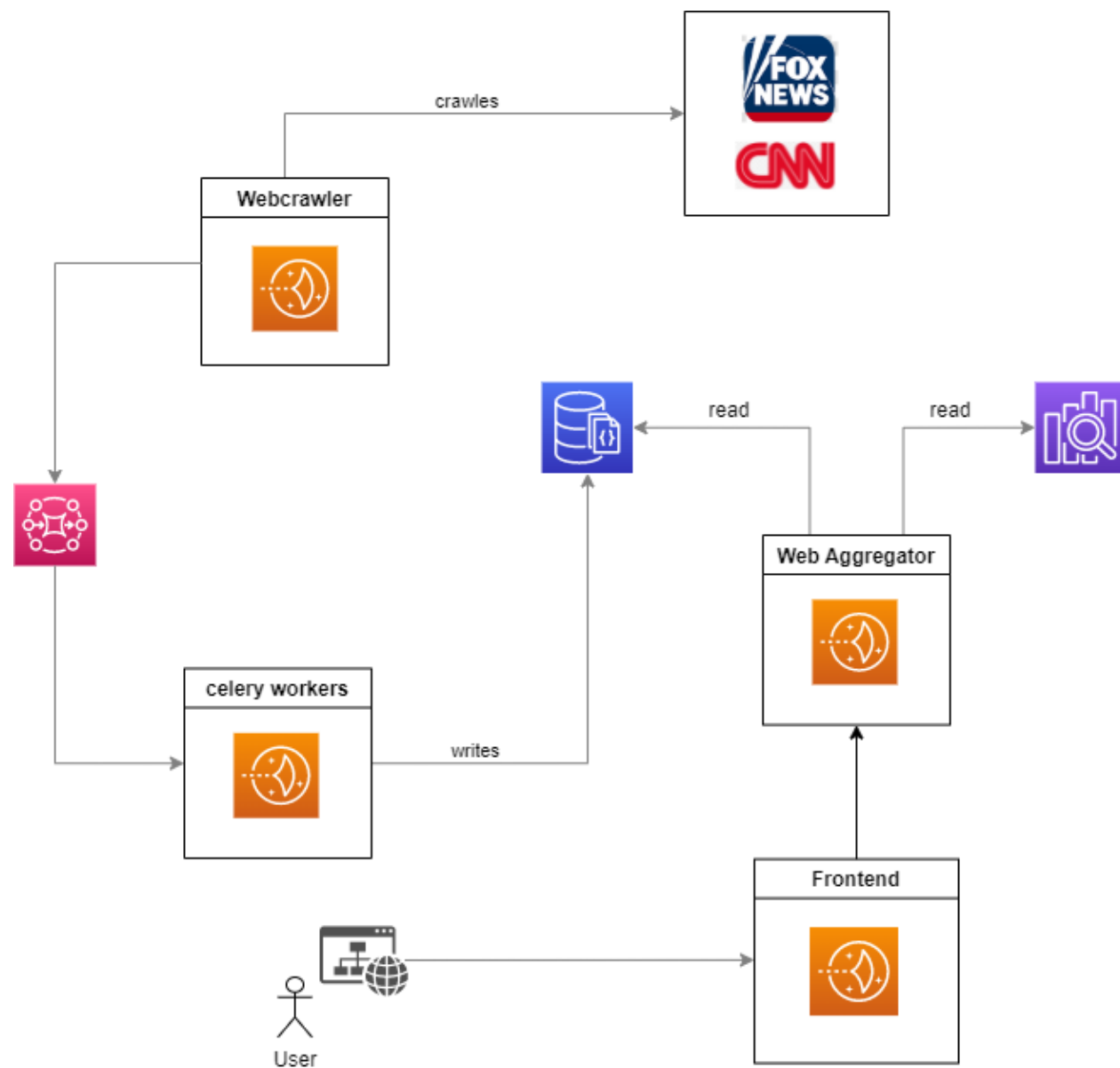
- This application uses AWS to perform long text searches on news articles on different newspaper websites to find the common news article and deliver them onto a single webpage.
  - In the First prototype, we are focusing on two popular news channels Fox News and CNN News.
  - The novelty of the project is that it focuses more on recent news and provides control to the user to choose which news source should he choose to read the article.
- 

+

•

○

# Architecture





# Frontend

---

Three Options – Home, Sources, Categories

---

Eight categories – Health, World, US, Business, Technology, Entertainment, Travel, Politics.

---

Two Sources – CNN News and Fox News.

---

Frontend is written in Python using Flask.

---



CNN

## La subvariante BA.2 de ómicron ahora es la cepa dominante de covid-19 en Estados Unidos, según los CDC

La subvariante de ómicron, BA.2, más contagiosa, es ahora la cepa dominante en Estados Unidos, según los CDC. La subvariante es responsable de más del 50% de los nuevos contagios de la semana pasada.

👤 Author: Alexandra Ferguson

🕒 Date Published: 2022-03-29T20:54:12Z



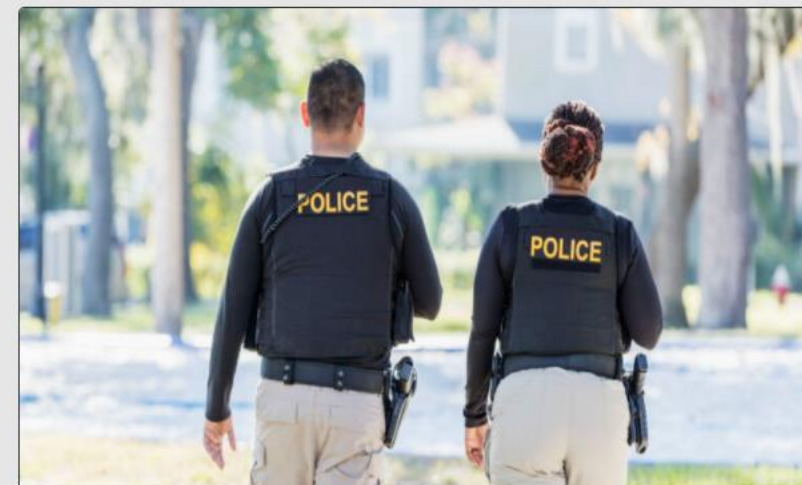
Reuters

## No current talks with Tesla about potential union vote -UAW president - Reuters.com

United Auto Workers (UAW) President Ray Curry said on Tuesday he has had no discussions with Tesla <a href="https://www.reuters.com/companies/TSLA.O" target="\_blank">(TSLA.O)</a> or its Chief Executive Elon Musk about a potential union vote by workers at the ...

👤 Author: None

🕒 Date Published: 2022-03-29T20:52:00Z



Ars Technica

## A controlled experiment focuses on improved policing method

A method called "procedural justice" policing appears to work in the real world.

👤 Author: John Timmer

🕒 Date Published: 2022-03-29T20:51:55Z





+

•


○

# Backend

- Flask is also used in Backend.
- Using beautiful soup, urllib and requests for web scraping.
- We are using Elastic Search to store the scraped fox news and CNN news articles for high performance text searching.
- Extract all the details of the news source like title, author, image, content etc.



# Working

- The web scraper extracts all the information like content, author, title, and image from the news source and the frontend will display them according to the defined format.
  - We are using DocumentDB as a permanent data store to store the articles as documents.
- 

+

•

○

# Future Proposal

- Log in and ML algorithm for user searched data to show relevant content as soon as the user opens the website.
- Integration with Alexa Search History which deals with the news.



# TEAM



**Koli Anuj**  
**25**

**The University of  
Florida(MS)**



**Anurag Shenoy**  
**26**

**The University of  
Florida (MS)**



**Aman Pathak**  
**24**

**The University of  
Florida (MS)**



**Ruchika Upadhyay**  
**27**

**New York University  
(MS)**



**Naveen Kumar**  
**25**

**New York University  
(MS)**



# Thank You!

# Questions?

