Tech Demo: AJAR Application

Academic Journal Article Recommender

General Assembly - DSI
September 8, 2020

Jason Lee

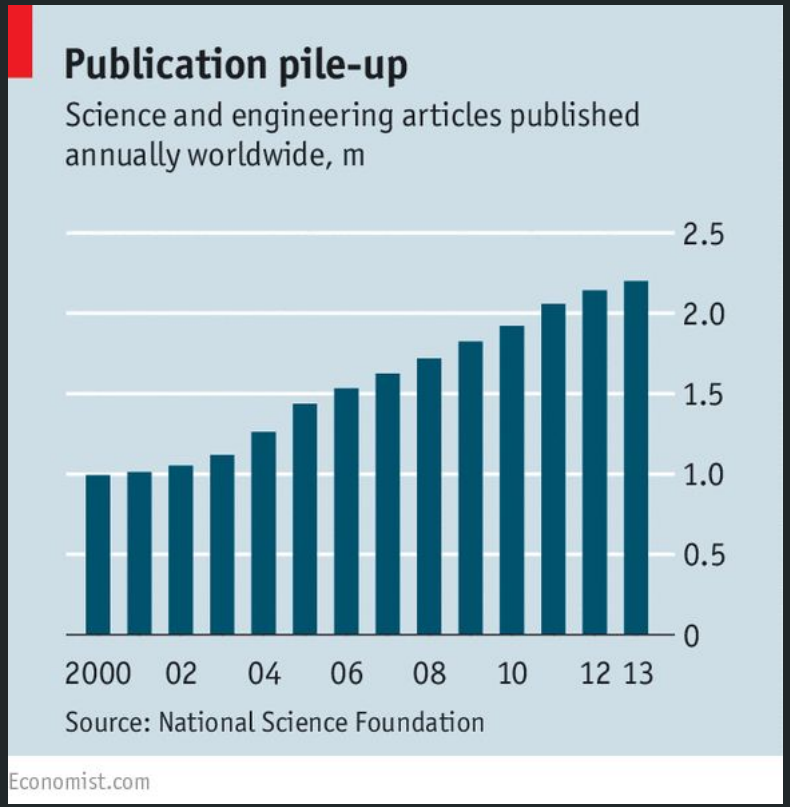
Introduction - Academic Literature Searching

Academic papers are the standard for sharing findings among the scientific community

However, taking time each week to browse through multiple journals can take many hours away from study or experiments

https://conductscience.com/

Introduction - Academic Literature Searching



Furthermore, the rate of publications is increasing each year

Estimates point to over 2 million articles published a year, across over 30,000 journals

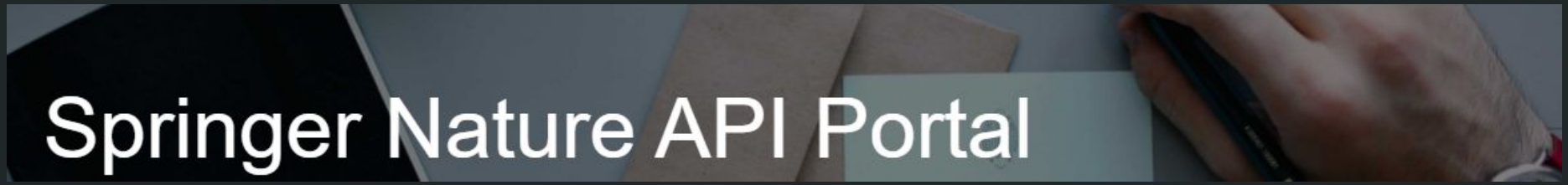
How can we help researchers effectively find relevant papers in a timely manner?

Introduction - AJAR Application

AJAR is a free online academic paper database that provides a number of services:

- An archive of nearly 20,000 recently published papers
- Personal accounts for researchers to save papers of interest
- Daily recommendations for new articles that match a user's saved papers

AJAR Methodology: Data Source

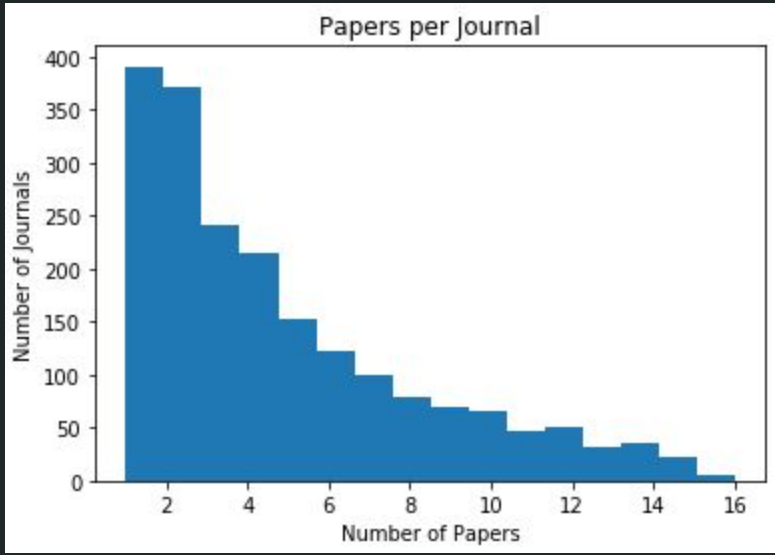


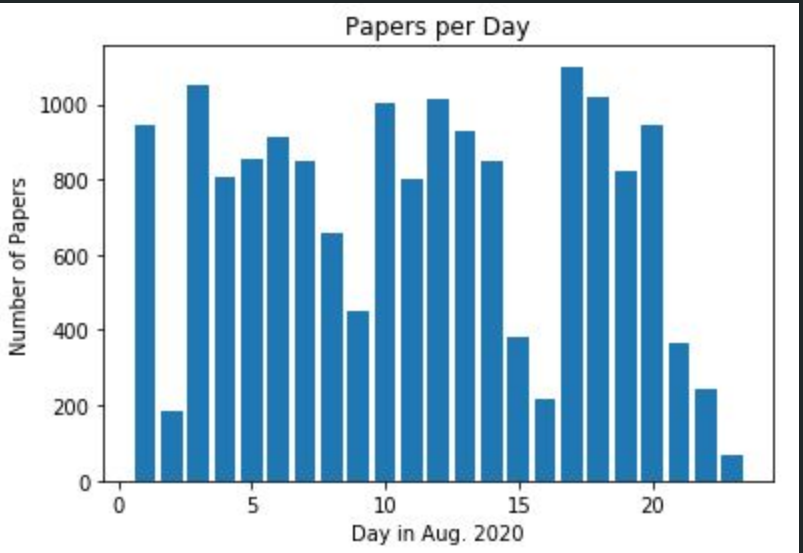
Articles are collated each day using the Springer API (https://dev.springernature.com/)

Data processing involves:

- Removing non-English articles
- Removing redactions/reprints/corrections
- Removing entries with incomplete information

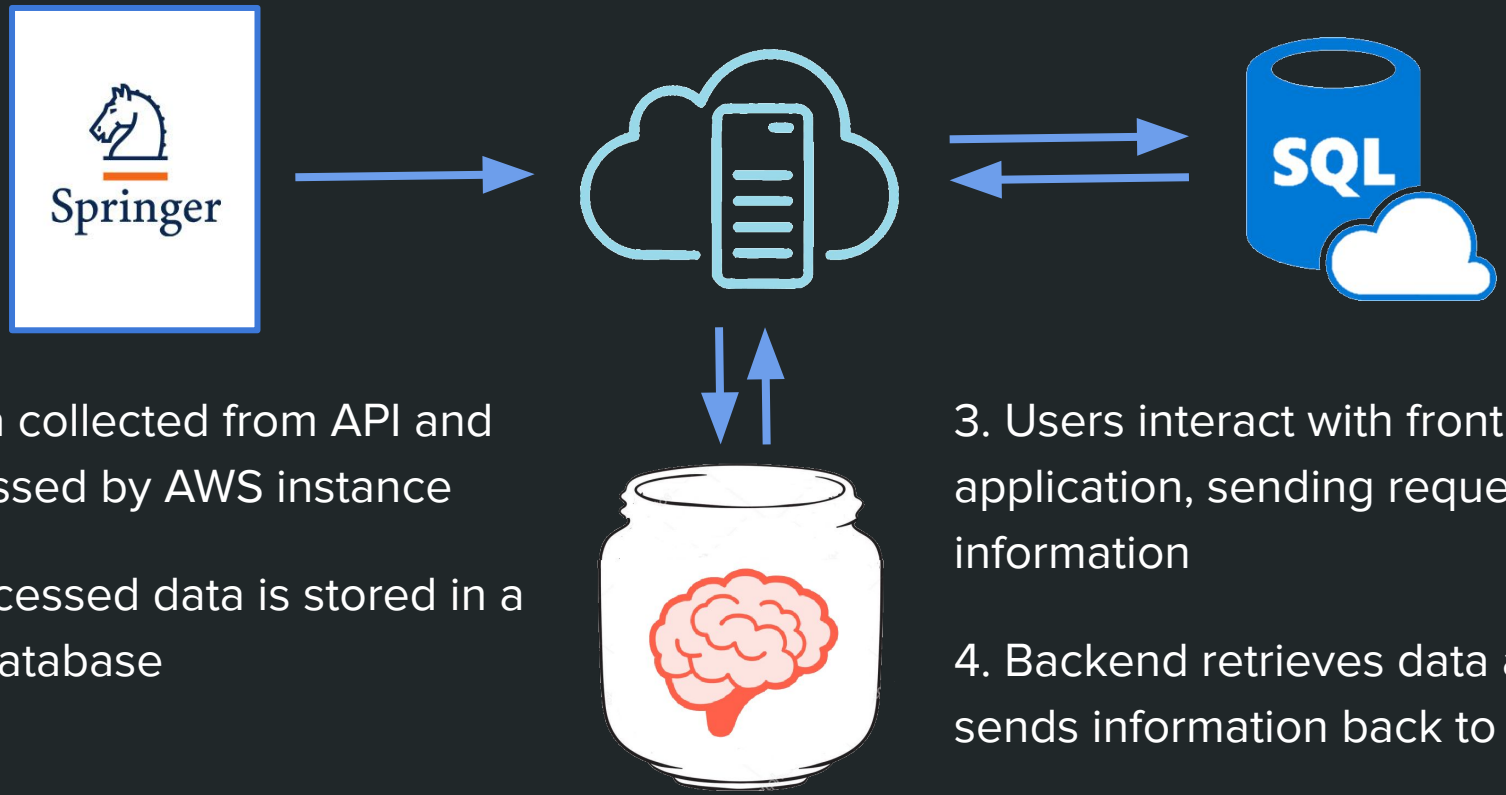
AJAR Methodology: Data Source





Database information for articles pulled for the month of August

AJAR Methodology: Process Diagram



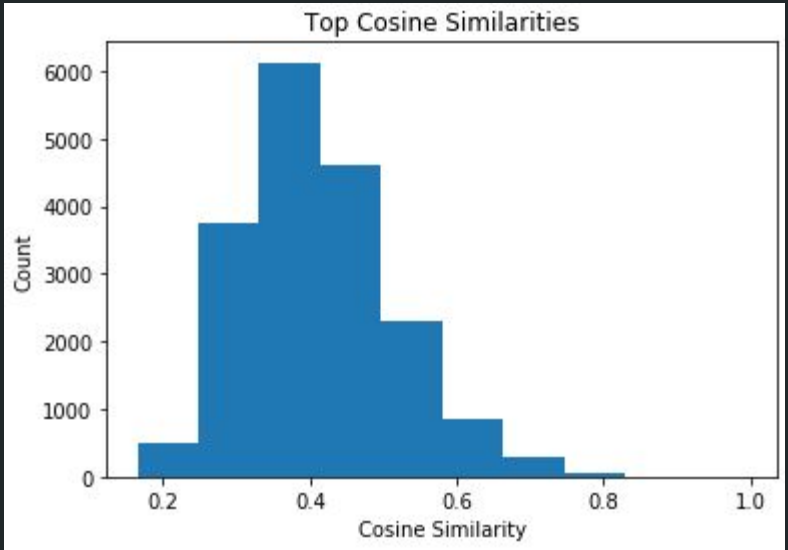
1. Data collected from API and processed by AWS instance

2. Processed data is stored in a SQL database

3. Users interact with front end application, sending requests for information

4. Backend retrieves data and sends information back to user

AJAR Methodology: Recommender



Language features created using a Tfidf Vectorizer applied to the abstracts for each paper

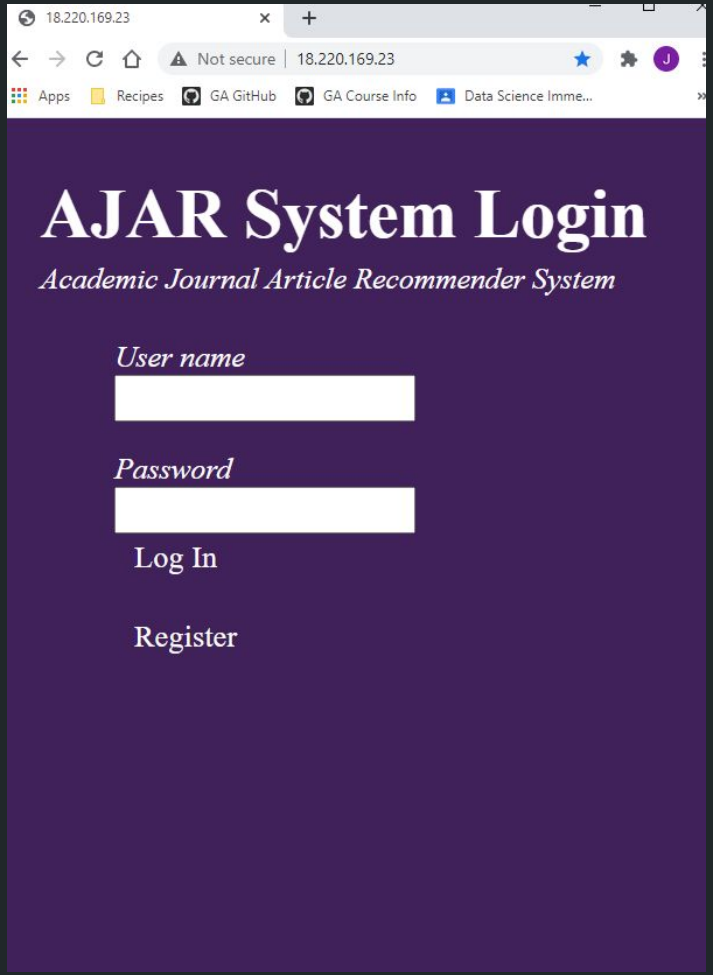
Cosine similarity is calculated to determine best matches between favorited papers and papers published each day

AJAR Demonstration

AJAR can be found online at the ip address

18.220.169.23

Please feel free to explore the app on your own machine during the demonstration



Conclusions and Next Steps

AJAR is a tool to help researchers across all disciplines spend more time doing research, while still effectively keeping apace of advances in their fields. While a fully effective application in its current form, there are still several possible avenues for improvement:

- More sophisticated recommendations (feature engineering, BERT)
- Incorporating additional sources of data beyond Springer
- Metadata analysis (author tracking, for example)

Through your support, we can continue to build this platform to assist in academic pursuits in labs across the world

Questions?

Thanks for listening!