DATA FAIR NEWS AGGREGATOR

K. Rambabu[1], Barigela Kavya[2]

[1] **Assistant Professor(HOD) MCA, DEPT,** Dantuluri Narayana Raju College **, Bhimavaram, Andharapradesh**

Email id:- kattarambabudnr@gmail.com

[2]**PG Student of MCA,** Dantuluri Narayana Raju College **, Bhimavaram, Andharapradesh**

Email id:- barigelakavya75078@gmail.com

ABSTRACT

In this modern world we know that there are a lot of news websites and they cover news on various topics, out of which very few are of our interest. A new aggregator can be a tool to save a lot of time and with some modifications and filtrations we can fine tune it to show only news of our interest. A news aggregator is a very useful application to view information within minimum time. We can build a news aggregator application by scrapping some famous websites and serving the scrapped articles via Django on web or in any app. It is a web application which aggregates data from multiple websites. Then shows the articles on desired pages.

News is a very important start of the day. There are well known news sites online. Now, imagine opening several news websites every single day. The time wasted from reading from those websites and information gaining is also essential. It can give leverage over those who don't have it. Now, we have made this task really easy for everyone!!

In a news aggregator, you can read news from various websites. Then the news aggregator collects the article for you. And just after a click or two, you can view news of your choice.

1. INTRODUCTION

In the modern society that we are living in, we only want things easily for us. We do not want to go out and pick up a newspaper every time or even search google for the latest news, that is where out news aggregator web application comes in hand.

It will allow a user to view or read news conveniently without any hassle, all he/she has to do is just open our news website and he will be able to see all the latest news from two very famous and respected news websites and those news will al
so be fully trusted.

In today's world trusting a news is a difficult thing for us because of the spreading of fake news, so user will not have to worry about the authenticity of the news as news are from trusted platform.

At the same time, RSS provides condensed item information, a favourite for a news aggregator. Even through there are small discrepancies between RSS formats, they are still an effective solution for indexing articles.

A solution that uses web-crawling and Hyper Text Markup Language (HTML) parsing techniques is not adequate due to the fact that each website has its own structure. Also, even small changes to the Document Object Model (DOM) of the websites can have dire consequences to the parsed results. Another well- known method of fetching articles is based on specific APIs that the news platforms provide. An API only based solution would be limited given the fact that not all news outlets provide each other method. Nonetheless, these solutions can supplement the one based on RSS feeds.

The application delivers a couple of unique features as described below:

This system downloads articles from 2 predefined Reduces the time required to regularly visit websites by bringing information

into one place. Users can quickly access the feeds and they can then click on items they find interesting.
Simplicity in use, and fluidity of the graphical interface1. First, we'll setup our Django server
Although some existing news aggregators can offeer. Then, we'll integrate everything altogether So, let's much more complex interface and more advanced settings, they can be quite difficult to use, especially for a non-technical person.

• Spam removal – although we can say that most aggregator offer spam and ad filtering for articles, more and more have begun to add their own ads for profit generation. The developed aggregator does not alter the user experience by adding advertisements.
• Ability to filter and sort content according to user defined criteria.
• Provides an API that can be used by users to develop their own custom application. The fact is that most aggregators present in today's world does not provide a good interface to the user.
• In the second section we take a look on the related work and next section describes the system of the application. We present the application interface in Section 4. First, we discuss news aggregators and news websites and it cannot be configured to add more.
• News articles are in HTML format from news websites using a Python framework namely scrapy. The main goal is to extract different components of the webpage, such as title, the text of news, lead paragraph, publication data, the news author and the associated image. The news aggregator offers the option to organize the feeds into categories. Besides this, it also creates news summaries which decrease the required reading time. Flipboard is another well know application that manages RSS feeds. Flipboard offers the possibility of grouping feeds into categories as well as editing and adjusting them. The European Media Monitor (EMM) is a sizable project implemented by the European Commission's Joint Research Centre. EMM monitors social media from Europe in real-time gathering daily between 80 and 100.000 news articles in as much as 50 languages. In case websites have RSS support, EMM employs them, otherwise is reverts to HTML parsing.

2. LITERATURE SURVEY AND RELATED WORK

Data Fairness in Aggregation:

Data fairness in newspaper aggregation refers to the equitable and unbiased collection and presentation of news articles from various sources.

Ethical Considerations:

Research often emphasizes the importance of ethical considerations in data aggregation, such as avoiding bias and ensuring diverse representation of sources.

Algorithmic Fairness:

Many studies explore algorithmic approaches to achieving fairness in selecting and ranking news articles. They often involve machine learning and natural language processing techniques.

Bias Mitigation:

Research may focus on techniques to mitigate bias in news aggregation, including algorithmic bias and editorial bias.

User-Centric Fairness:

Some research delves into user-centric fairness, considering how different users' preferences and backgrounds can affect their perception of fairness in news aggregation.

Trustworthiness and Reliability:

Ensuring that aggregated news articles come from reliable sources is another important aspect of data fair newspaper aggregation.

User Engagement:

Some studies may explore how user engagement and feedback can be leveraged to improve data fairness in newspaper aggregation platforms.

Privacy Concerns:

Privacy issues related to user data in news aggregation systems are also a subject of research.

To perform a comprehensive literature survey, you should search academic databases such as Google Scholar, IEEE Xplore, ACM Digital Library, and journals related to data science, information retrieval, and media studies. You can use keywords like "data fair newspaper aggregation," "algorithmic fairness in news aggregation," and "ethical considerations in news curation" to find relevant papers and articles. Be sure to check for the most recent publications to stay up to date with the latest research in this field.

3. EXISTING SYSTEM

The existing system of data fair newspaper aggregation typically involves collecting, curating, and presenting news articles and information from various sources in a transparent and unbiased manner. It aims to provide users with a comprehensive view of current events while maintaining data fairness and credibility. This can involve algorithms and manual curation to filter out biased or unreliable sources and present news from a diverse range of perspectives. Additionally, it may include features like fact-checking and source verification to ensure the accuracy of the information presented to users. However, specific implementations and methods can vary depending on the platform or organization providing the news aggregation service.

4. PROPOSED SYSTEM

A proposed system for fair newspaper aggregation would prioritize the following principles:
Diverse Source Selection: The system should gather news articles from a wide range of reputable sources, representing various perspectives and ideologies to avoid bias.
Algorithmic Transparency: Transparency in the selection and ranking of articles is crucial. Users should understand how articles are chosen and why they are presented in a particular order.
Editorial Guidelines: Establish clear editorial guidelines that ensure neutrality and avoid sensationalism or misinformation. Fact-checking and verification should be an integral part of the process.
User Customization: Allow users to customize their news feed based on their interests, but ensure that the system doesn't create echo chambers by exposing them to diverse viewpoints.

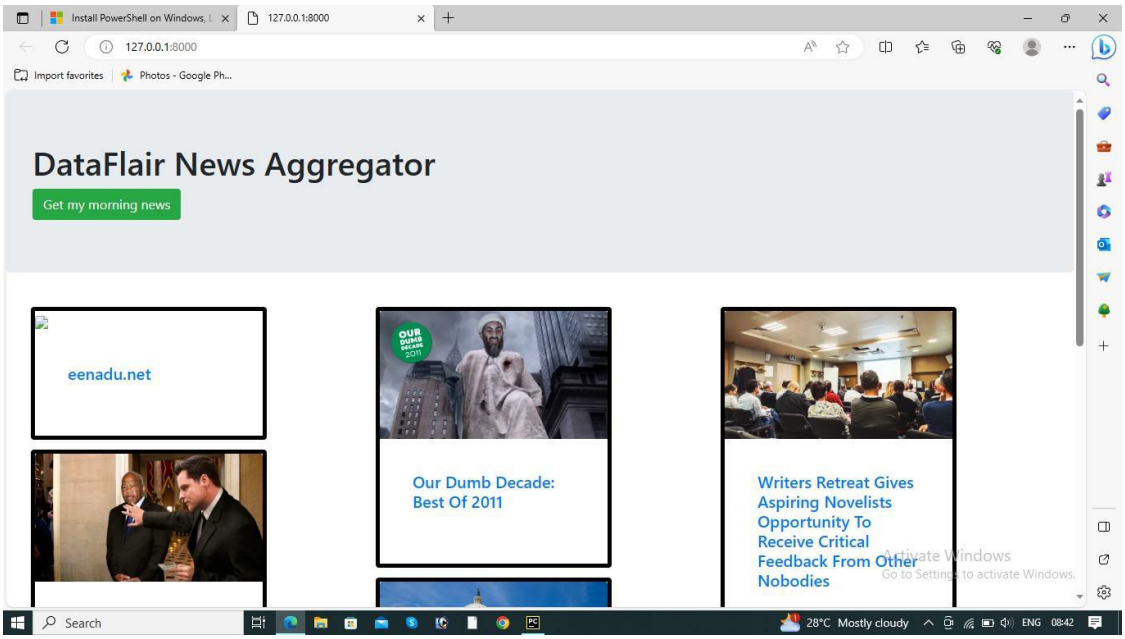
Data Privacy: Prioritize user privacy by collecting minimal data and giving users control over their data. Avoid profiling users for political or commercial purposes.

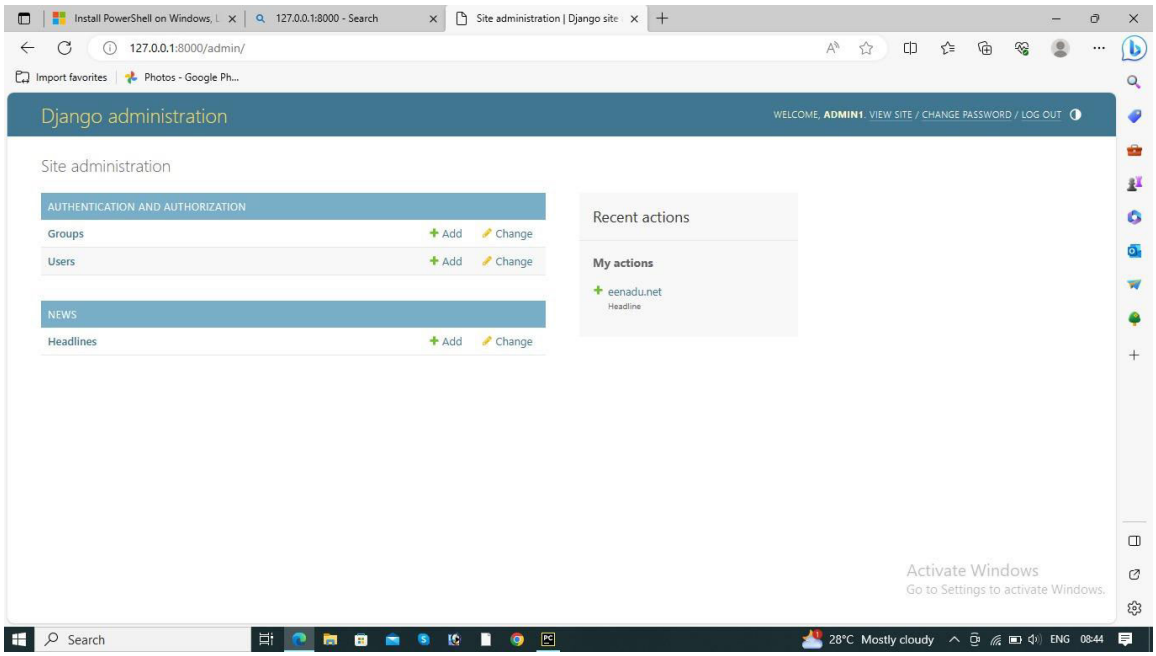
Community Feedback: Implement mechanisms for user feedback and moderation to identify and address biased or misleading content.

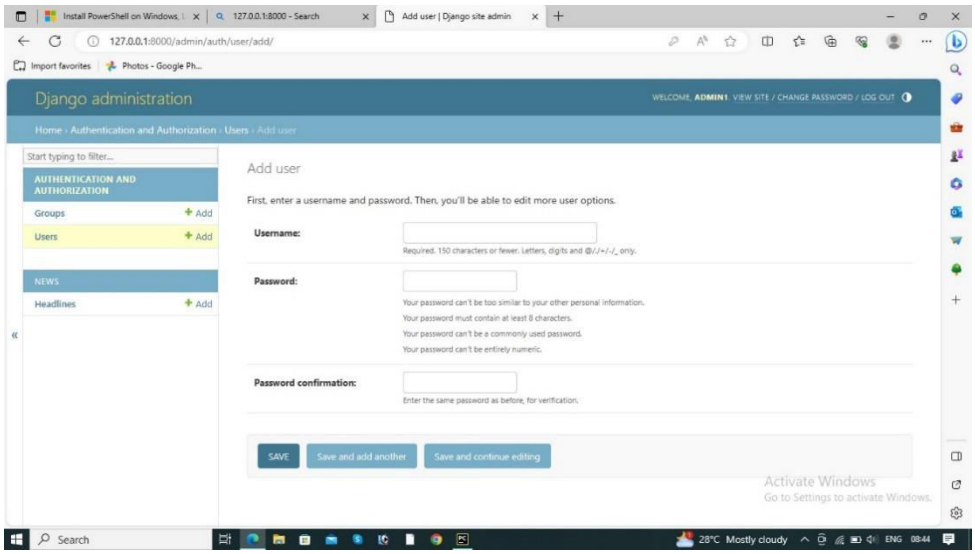
Transparency Reports: Regularly publish reports detailing the sources used, the selection process, and the algorithms used for article recommendations.

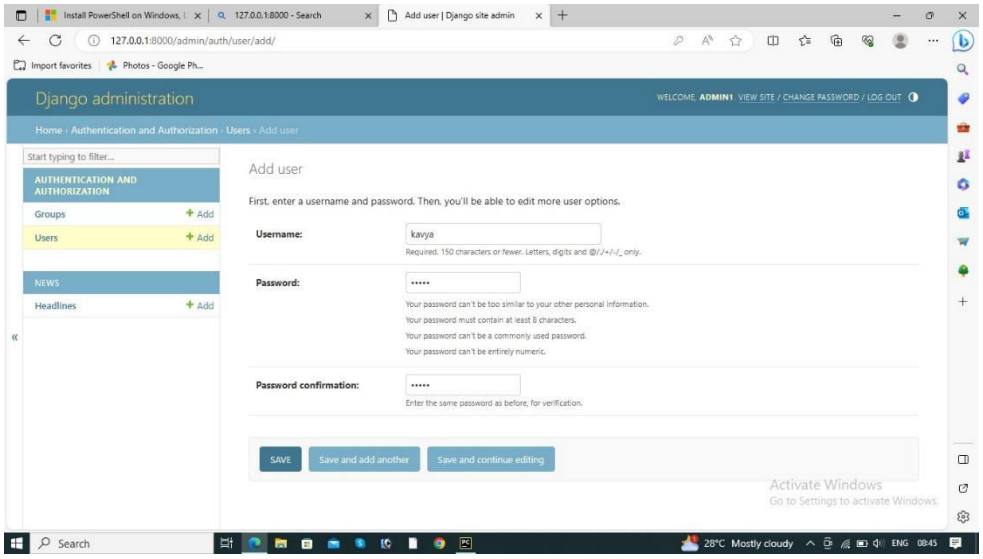
Human Oversight: Employ a team of human editors to oversee the system and make manual adjustments as needed to maintain fairness and accuracy

5. RESULTS AND DISCUSSION SCREEN SHOTS




Install PowerShell on Windows, L
127.0.0.1:8000 - Search
Site administration | Django site
127.0.0.1:8000/admin/
Import favorites
Photos - Google Ph...
Django administration
WELCOME, ADMIN1. VIEW SITE / CHANGE PASSWORD / LOG OUT
Site administration
AUTHENTICATION AND AUTHORIZATION
Groups + Add Change
Users + Add Change
NEWS
Headlines + Add Change
Recent actions
My actions
+ eenadu.net
Headline
Activate Windows
Go to Settings to activate Windows.
Search
28°C Mostly cloudy
ENG 08:44


Install PowerShell on Windows, L
127.0.0.1:8000 - Search
Add user | Django site admin
127.0.0.1:8000/admin/auth/user/add/
Import favorites
Photos - Google Ph...
Django administration
WELCOME, ADMIN1. VIEW SITE / CHANGE PASSWORD / LOG OUT
Home › Authentication and Authorization › Users › Add user
Start typing to filter...
AUTHENTICATION AND AUTHORIZATION
Groups + Add
Users + Add
NEWS
Headlines + Add
Add user
First, enter a username and password. Then, you'll be able to edit more user options.
Username:
Required. 150 characters or fewer. Letters, digits and @/./+/-/_ only.
Password:
Your password can't be too similar to your other personal information.
Your password must contain at least 8 characters.
Your password can't be a commonly used password.
Your password can't be entirely numeric.
Password confirmation:
Enter the same password as before, for verification.
SAVE
Save and add another
Save and continue editing
Activate Windows
Go to Settings to activate Windows.
Search
28°C Mostly cloudy
ENG 08:44



6. CONCLUSION AND FUTURE SCOPE

CONCLUSION

In conclusion, data fair newspaper aggregation offers a promising solution for accessing diverse news sources and staying well-informed. By utilizing data fairness principles, this approach ensures a balanced representation of perspectives and minimizes bias in news curation. It promotes transparency, credibility, and inclusivity in the dissemination of information, ultimately contributing to a more informed and engaged society. However, ongoing efforts are needed to refine algorithms, address potential challenges, and ensure continuous improvements in the quest for fair and unbiased news aggregation.

FUTURE SCOPE

A data-driven newspaper aggregator has a promising future, with several potential areas for growth and development:

Personalization: Enhance user experience by offering personalized news recommendations based on individual preferences and behavior, using AI and machine learning algorithms.

Multimedia Integration: Incorporate videos, podcasts, and interactive content to cater to evolving multimedia news consumption trends.

Real-time Updates: Provide up-to-the-minute news updates through continuous data scraping and processing.

Fact-Checking and Verification: Integrate tools for real-time fact-checking to combat misinformation and fake news.

Niche Aggregation: Specialize in specific niches like science, technology, finance, or entertainment to cater to niche audiences.

Localized Content: Offer hyper-localized news by partnering with local news outlets and using geolocation data.

Subscription Models: Explore premium subscription models for an ad-free, in-depth news experience.

User-generated Content: Allow users to contribute content, reviews, and opinions to foster community engagement.

Data Visualization: Implement data visualization tools to help users understand complex news stories and trends better.

AI Journalism: Experiment with AI-generated news articles, although ethical considerations should be a priority.

Partnerships: Collaborate with news outlets, publishers, and content creators to expand your content offerings.

Data Analytics Services: Offer data analytics and insights to news organizations and businesses to generate revenue.

7. REFERENCES

1) R. E. Radu, O. Grigorescu and R. V. Rughiniş, "Security News Aggregator," *2019 18th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, 2019, pp. 1-8, doi:10.1109/ROEDUNET.2019.8909609.

2) S. Khan, T. Khan, C. Prasad, A. Khatri and I. Khan, "Intelligent News Aggregator and Validator," *2019 International Conference on Nascent Technologies in Engineering (ICNTE)*, 2019, pp. 1-5, doi: 10.1109/ICNTE44896.2019.8945945.