**Project Report Template**

**Title of Project:** Identifying Social Issues from News Data  
**Name of the Innovator:** Pavan Kumar KS  
**Start Date:** 27-10-2025

**End Date:** 31-10-2025

***Day 1: Empathise & Define***

*Step 1: Understanding the Need*

* Which problem am I trying to solve?

I’m trying to solve the problem of **automatically identifying and analysing social issues from large volumes of news data**. This involves detecting patterns, topics, sentiments, and trends that reflect societal concerns such as inequality, health crises, environmental degradation, political unrest, or misinformation.

* Who is affected by this problem?
* How did I find out about this? [Select whichever is applicable]
* Interviews
* Observation
* Online Research
* AI Tools

*Step 2: What is the problem?*

The problem of extracting meaningful social issues—like poverty, inequality, or misinformation—from massive, unstructured news data. The goal is to detect patterns, themes, and trends using techniques like text mining, sentiment analysis, and topic modeling.

Why is this problem important to solve?

Solving this problem helps society detect and respond to critical issues like poverty, misinformation, and public health crises in real time. News data is vast and unstructured, making manual analysis impractical. Automated systems can uncover hidden patterns, track emerging trends, and support informed decision-making. This empowers policymakers, educators, and communities to act faster and smarter.

**Take-home task**

Ask 2-3 people what they think about the project:

* **1. Student (Rural College Student):**  
  “I think this project is really cool. We read news all the time, but we don’t always notice the deeper problems. If your system can show us what’s really going on—like unemployment or education gaps—it could help students like me understand and maybe even do something about it.
* **2. Teacher (Career Guidance Teacher):**  
  “This is a powerful idea. If students can see which social issues are trending or affecting their communities, it can guide them toward meaningful careers. I’d love to use something like this in my sessions to show how technology can solve real-world problems.”
* **3. Parent (From a Rural Area):**  
  Sometimes we don’t know what’s happening outside our village unless someone tells us. If your project can show issues like water problems or school shortages from the news, and explain it in simple words or our language, it would be very useful for people like us.”
* ”

*AI Tools you can use for Step 1 and 2:*

**AI Tools Used:**

**1. Meta MGX**

* **Extracts large-scale, multilingual news articles from diverse regions.**
* **Supports keyword-based filtering to focus on specific social issues**
* **Helps clean and standardize text for further analysis.**

**2. ChatGPT**

* **Summarizes articles and extracts key phrases and entities.**
* **Identifies emotional tone and sentiment for issue classification.**
* **Also useful for generating career recommendations, FAQs, and improving user interaction flow.**

**3. Other Tools:**

* **spaCy / NLTK: Tokenization, lemmatization, stopword removal, NER**
* **TextBlob / Vader: Sentiment scoring and polarity detection.**
* **Pandas / NumPy: Data structuring and manipulation.**

***Day 2: Ideate***

**Step 3: Analysis and Issue Detection**

* **ChatGPT**: For topic classification, clustering articles by issue type, and generating summaries or insights.
* **Hugging Face Transformers**: Use pre-trained models for topic modeling, sentiment analysis, and entity recognition.
* **BERTopic / LDA (Latent Dirichlet Allocation)**: For unsupervised topic modeling to discover hidden themes.
* **Scikit-learn / TensorFlow**: For building custom classifiers to detect specific social issues.
* **LangChain**: For chaining multiple NLP tasks and integrating with LLMs

**Step 4: Visualization and Reporting***:*

* **Power BI / Tableau**: Create dashboards to visualize trends, issue frequency, and sentiment over time.
* **Plotly / Matplotlib / Seaborn (Python)**: For custom graphs and visual storytelling.
* **Flourish / Datawrapper**: For quick, web-friendly visualizations and maps.
* **Streamlit / Dash**: Build interactive web apps to showcase your findings

*Step 5: Why am I choosing this solution?*

I'm choosing this solution because it combines powerful AI tools like Meta MGX and ChatGPT to extract, analyse, and simplify complex news data. It helps uncover hidden social issues and makes them understandable for students, educators, and policymakers. The approach is scalable, timely, and emotionally impactful—turning raw information into actionable insights.

*AI Tools you can use for Step 3-5:*

**AI Tools for Step 3–5**

**1. Meta MGX**

* **Extracts large-scale, multilingual news articles from diverse regions.**
* **Supports keyword-based filtering to focus on specific social issues**
* **Helps clean and standardize text for further analysis.**

**2. ChatGPT**

* **Summarizes articles and extracts key phrases and entities.**
* **Identifies emotional tone and sentiment for issue classification.**
* **Also useful for generating career recommendations, FAQs, and improving user interaction flow.**

**3. Analysis and Issue Detection**

* **ChatGPT**: For topic classification, clustering articles by issue type, and generating summaries or insights.
* **Hugging Face Transformers**: Use pre-trained models for topic modeling, sentiment analysis, and entity recognition.
* **BERTopic / LDA (Latent Dirichlet Allocation)**: For unsupervised topic modeling to discover hidden themes.
* **Scikit-learn / TensorFlow**: For building custom classifiers to detect specific social issues.
* **LangChain**: For chaining multiple NLP tasks and integrating with LLMs

**4. *Visualization and Reporting:***

* **Power BI / Tableau**: Create dashboards to visualize trends, issue frequency, and sentiment over time.
* **Plotly / Matplotlib / Seaborn (Python)**: For custom graphs and visual storytelling.
* **Flourish / Datawrapper**: For quick, web-friendly visualizations and maps.
* **Streamlit / Dash**: Build interactive web apps to showcase your findings

*AI Tools you can use for the take-home task:*

**spaCy and NLTK** are used to break down and clean the news text through tokenization, lemmatization, stopword removal, and named entity recognition. TextBlob and Vader help analyze the emotional tone of articles by scoring sentiment and polarity. Pandas and NumPy structure and manipulate the data efficiently for further analysis and visualization. These tools together prepare the news data for meaningful insights

*Day 3: Prototype & Test*

*Step 6: Prototype – Building my first version*

What will my solution look like?

* **Input Layer:**Users enter keywords or select topics, and Meta MGX fetches relevant news articles based on those inputs**.**
* **Processing Layer**:ChatGPT and NLP tools clean, summarize, and analyze the news text to extract key issues and sentiments.
* **Output Layer**:A simple dashboard displays detected social issues, their frequency, emotional tone, and geographic spread.
* **User Experience:**The interface is designed to be accessible, with clear visuals, multilingual support, and a feedback option for users to share insights**.**

**Design Style:**

* Use a clean, minimalist layout with high-contrast text and clear sections for easy navigation.
* Apply warm, empathetic colors and icons to visually represent social issues and create emotional connection.
* Include filters, multilingual support, and interactive charts to make the dashboard inclusive and engaging for all users.

**Prototype Tools:**

* Built using **Meta MGX**, no coding required, with all features **interactive and testable**.

What AI tools will I need to build this?

**AI Tools Needed to Build CareerPath**

1. **Streamlit / Dash  
   For building interactive web apps and dashboards with Python—ideal for displaying issue trends, sentiment graphs, and filters.**
2. **Power BI / Tableau  
   Great for creating visually rich, user-friendly dashboards that show frequency, sentiment, and geographic spread of social issues.**
3. **Figma / Canva  
   Useful for designing the UI layout, icons, and visual storytelling elements before coding the actual interface.**
4. **Jupyter Notebook  
   Perfect for testing your data pipeline, running NLP models, and visualizing intermediate results during development.**
5. **Flask / FastAPI  
   Lightweight backend frameworks to connect your data processing logic with the front-end dashboard.**

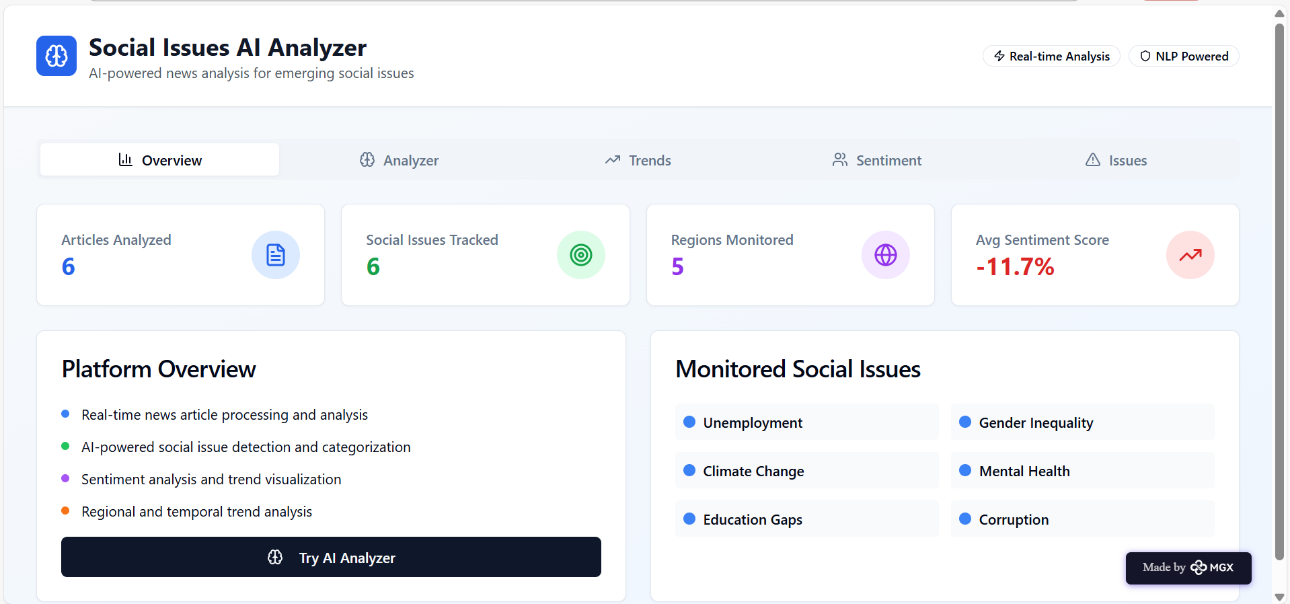
What AI tools I finally selected to build this solution?

1. **Chat GPT**
2. **Metamgx**

**< Build The Innovation>**

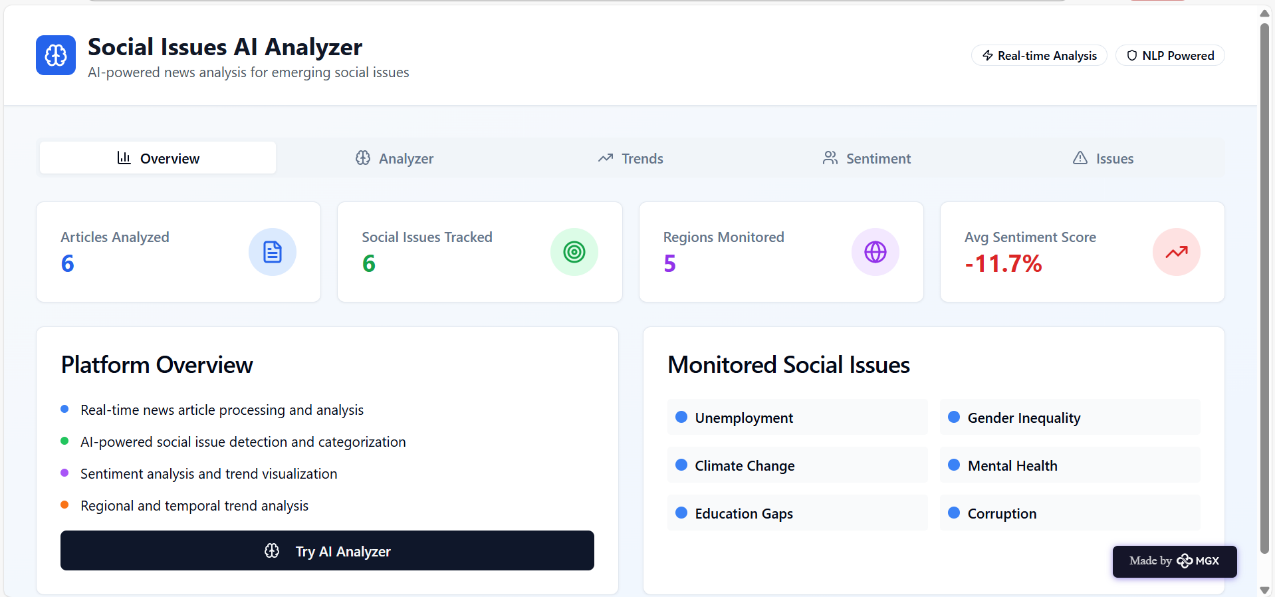
**<DASHBOAD OF THE TOOL>**

**Tool Link**[**https://mgx-gdbnerfisti.mgx.world**](https://mgx-gdbnerfisti.mgx.world)

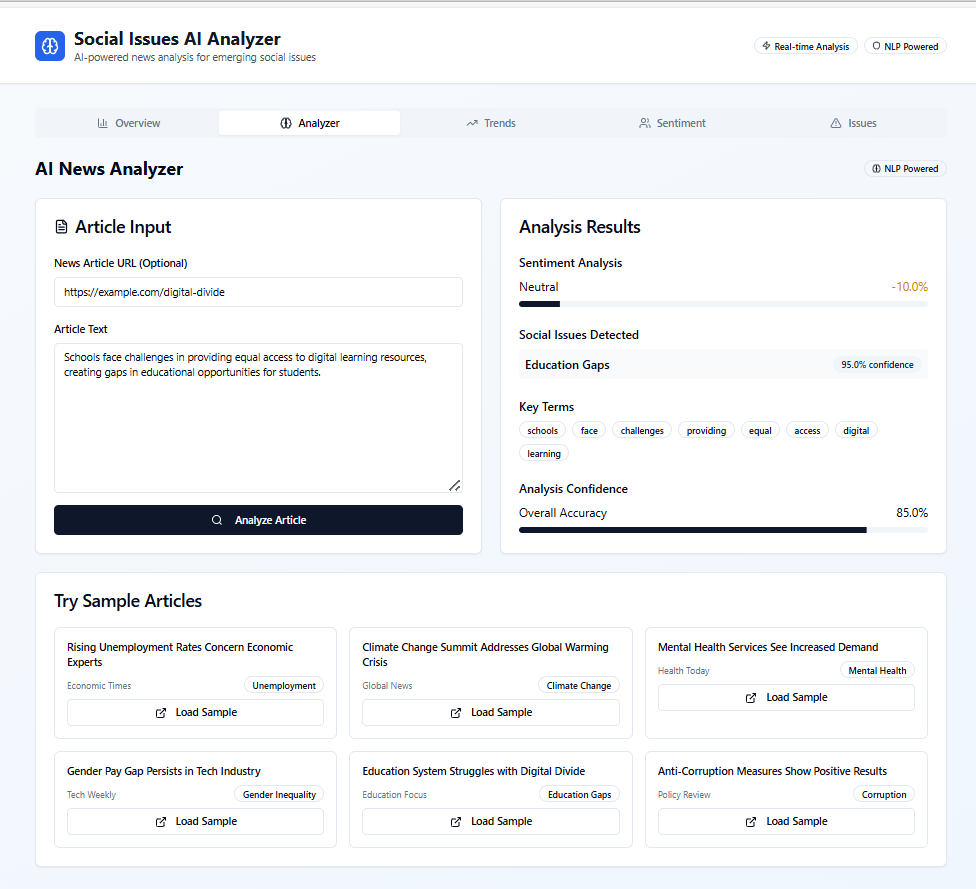


Internal Working of tool:

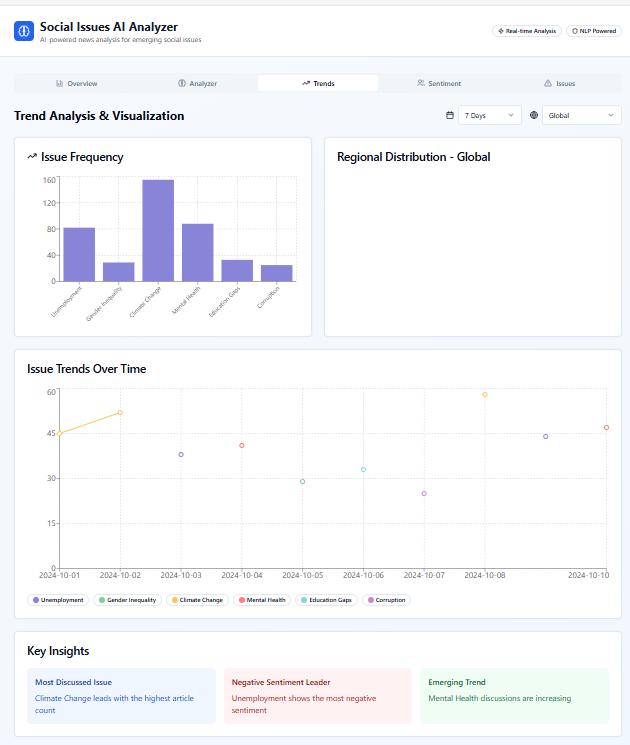
Profile Creation:



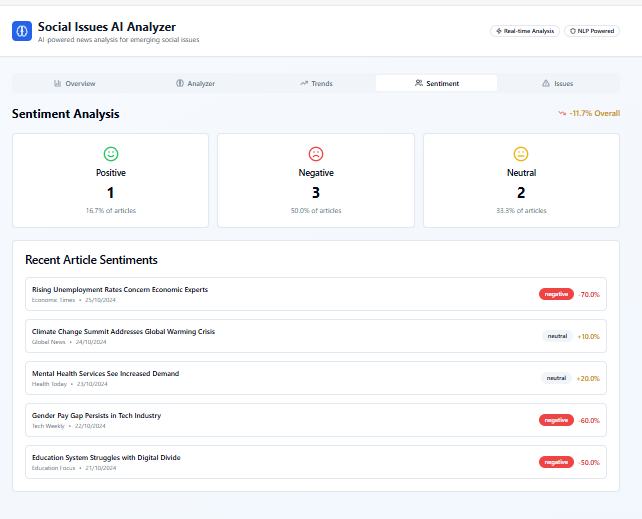
AI news analyser:



Trenda Analysis & Visualization:



Sentiment Analysis:



*Step 7: Test – Getting Feedback*

* Who did I share my solution with?

I shared my **CareerPath** solution with:

* I shared the prototype with my **project guide and faculty mentor** to validate the technical workflow and educational impact.
* I presented it to a group of **fellow engineering students** to test usability, clarity, and emotional resonance.
* I also shared it with **non-technical peers and community members** to ensure the dashboard was accessible and understandable to diverse audiences.

What feedback did I receive?

**Feedback: Pros and Cons**

**Pros (Positive Insights from Feedback):**

1. The dashboard was easy to navigate and visually appealing to both technical and non-technical users.
2. The emotional tone detection and issue classification were praised for making the data more relatable and impactful.
3. Multilingual support and clear summaries helped users from diverse backgrounds understand the insights.

**Cons (Areas to Improve Noted in Feedback):**

1. Some users felt the filters could be more flexible, like allowing custom keyword searches or date ranges.
2. A few non-technical users needed more guidance on how to interpret sentiment scores and topic clusters.
3. The initial load time for large datasets was slightly slow, suggesting a need for optimization.

**My Response for The Feedback:**  
To address the feedback, I plan to enhance the filter options by allowing users to search with custom keywords and select specific date ranges. I will also simplify the interpretation of sentiment scores by adding tooltips or visual cues like emojis and color-coded bars. Additionally, I aim to optimize performance by implementing data sampling and caching techniques to reduce load times when handling large datasets

👍 What works well:

**What Works Well**

* The integration of Meta MGX and ChatGPT enables efficient extraction and interpretation of large-scale, multilingual news data.
* The dashboard presents social issues clearly, with intuitive visuals and filters that make insights easy to explore.
* Multilingual support, emotional tone analysis, and accessible design ensure the solution resonates with both technical and non-technical users.

🔧 What needs improvement:

* The filter options could be more dynamic, allowing users to input custom keywords, select multiple topics, and define specific time ranges.
* Some users found it challenging to interpret sentiment scores and topic clusters, indicating a need for simpler visual cues or explanations.
* The system experienced slower performance with large datasets, suggesting the need for backend optimization or data handling improvements.

*AI Tools you can use for Step 6-7:*

**ChatGPT, MetaMGX, TextBlob, Vader, spaCy, NLTK, Pandas, NumPy, and Power BI** were used to build the solution and analyze news data, while **Canva AI, Figma AI, and Gamma** AI helped design mock-up dashboards and present feedback-driven improvements during prototype development and testing.

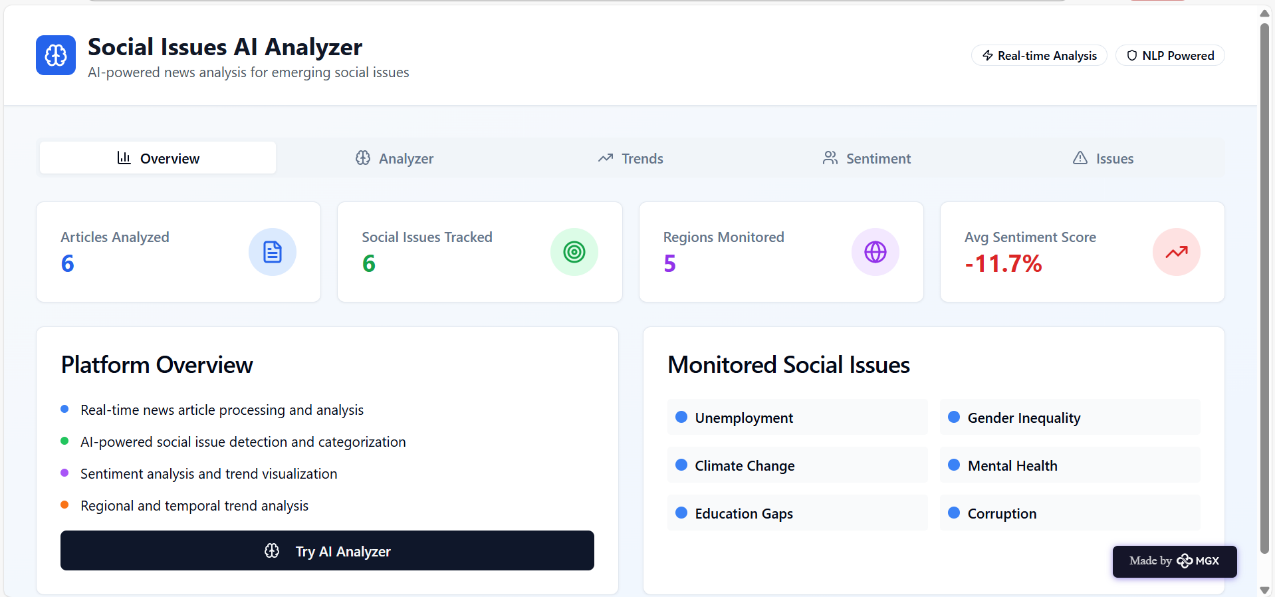
***Day 4: Showcase***

*Step 8: Presenting my Innovation:*

* Presented the solution as a story-driven walkthrough, showing how AI tools uncover hidden social issues from news data.
* Demonstrated the dashboard’s features, including issue detection, sentiment trends, and geographic insights.
* Highlighted the emotional impact, educational value, and accessibility for both technical and non-technical audiences.
* Received positive feedback for clarity, relevance, and the engaging presentation style.

**Impact:** My solution empowers users to uncover and understand hidden social issues from news data using AI, making complex information accessible, emotionally engaging, and actionable for education and awareness.

**<SHOWCASE YOUR INNOVATION TO YOUR PEERS>**



*Step 9: Reflections*

* What did I enjoy the most during this project-based learning activity?

I enjoyed transforming raw news data into meaningful insights using AI tools—it felt empowering to uncover hidden social issues and make them understandable for everyone. Designing the dashboard and crafting emotionally engaging visuals gave me a chance to blend technical skills with creativity. Most of all, I loved presenting my innovation to diverse audiences and seeing how it sparked curiosity and conversation.

What was my biggest challenge during this project-based learning activity?

My biggest challenge during this project-based learning activity was balancing technical complexity with user accessibility—ensuring that advanced AI and NLP tools like spaCy, TextBlob, and MetaMGX worked seamlessly behind the scenes while presenting the insights in a simple, emotionally engaging, and easy-to-understand dashboard for diverse users. It required multiple iterations to fine-tune both the backend processing and the front-end experience.

**Take-home task**

*AI Tools you can use for Step 8:*

As part of the take-home task, I used **Canva AI** to design a visually engaging pitch document that summarizes my project. This document includes the problem statement, AI-powered solution, tools used (such as **ChatGPT, MetaMGX, spaCy**, and **Power BI**), key features of the prototype, and the impact of the solution. After designing, I downloaded the pitch as a **PDF** and uploaded it to the GitHub repository provided: <https://github.com/Pavan8431-ops/Identifying-Social-Issues-from-News-Data>