

# **Mitigating bias in media: A technological approach to informed decision making**

**Author: Tolu Ojo**

## **Introduction:**

In today's digital age, the high often overwhelming volume of information we encounter on digital news and social media platforms makes it challenging to discern the reliability and credibility of information. This report outlines an innovative solution to this problem, a platform designed to empower users to assess bias in articles or videos by providing insights into the background and affiliations of authors and companies associated with media content in addition to fact checking the content. To prevent the problem of media saturated with misinformation and disinformation.

## **Problem Statement:**

The rise of misinformation and biased content on social media has led to a growing need for tools that help users evaluate whether they can trust the information. In addition, users susceptible to misinformation enables manipulation of. Existing fact-checking mechanisms are often time-consuming and may not provide a holistic view of an author's affiliations, limiting users' ability to make informed judgments.

Disinformation, the deliberate spread of misleading often fabricated content, is a powerful tool that is becoming less and less noticeable to the naked eye with increased adoption of AI-generated content. AI generated content enables the generation of photorealistic images and videos, mimicking voice audio. When these tools are used for malicious intent has the ability to target communities and altering the worlds perception of real-life events. Evidence of this was seen in 2016 American elections where social media platforms like Facebook were used by far-right activists to target voters with misinformation about political candidates [1]. In 2020-2021 Covid misinformation was putting lives at risk [2] and now in 2024 with the rise of AI generated it is proposed to have a huge impact on the world's perception of international conflict and wars such as the Russian invasion of Ukraine and in America intensifying voter suppression [3]. This has the huge potential to impact to threat on democracy internationally.

Metrics taken from users who encountered misinformation about COVID-19, provide an insight to the trend in user behaviours when faced with misleading content. Results from this showed in both years 2020 and 2021 54.5 % of the respondents nothing, while the percentage of respondents acted. However out of those that acted only 15.5% of them used a fact checking site and only 10.5% blocked or reported it to social media [2]. This small counter to misinformation indicates there is a need for easy access to fact checking and bias checking of sources to actively fight the spread of misinformation and disinformation.

Existing technology called Squash built in 2021, proposed automated fact checking platform powered by an algorithm that converts audio from live speeches to text file given by politicians or bodies of government and compares against a databased of published fact checks providing the user with a conclusion as to whether the information it accurate. This

platform described as a 'truth meter' was able at times to detect a truth, however a limitation of this platform was the inaccuracy due to a small learning database, and the insufficiency of published fact checks and the inability to fact check at high speeds [4]. The ability to fact-check at high speeds is key to combatting the spread of disinformation and misinformation that This spread is accelerated with the high speed of information sharing across networks, Squash however did not extend the truth, automated fact checking tools to social media platforms which is what makes the proposed platform Infoguard stand out. Whilst legislation and policies are being implemented to restrict disinformation and misinformation with more ethical AI [5]. They are essentially playing catch up with fast-growing tools like Chat-GPT and DALL-E, this calls for a preventative measure as well as a protective measure to combat the spread of disinformation and misinformation.

### **Proposed Solution:**

Introducing "InfoGuard," a revolutionary platform designed to analyse articles and videos for biases, affiliations, and credibility. Encouraging users to be critical about the information they consume.

### **Aims:**

Users can make informed decisions about the content they consume, promoting media literacy and critical thinking skills among users.

Provides a transparent view of potential biases and affiliations creating a more accountable media landscape .

### **User Persona:**

Our target users are individuals who seek reliable information and want to be aware of potential biases in the media they consume. This user is concerned about the potential biases in the information they encounter on social media and desires a tool that provides a comprehensive analysis of an article or video's credibility. This includes students, professionals, and anyone looking for a trustworthy source of information.

Target users all individuals or demographics with access to the social media and news platforms young adults to adults. With the aim to encourage communities as a collective to actively check the validity of their sources of information before believing it. Intend to have most impact on the working and voting population of a country.

### **Technology Stack:**

The platform will leverage machine learning algorithms for natural language processing (NLP) to analyse text and video content for sentiment and potential biases. As well as an algorithm trained on news data able to detect inaccuracies by fact checking with a databased of fact checked results. It will integrate with external APIs for fact-checking databases and social media platforms for real time updates.

### **User Interface (UI) Design:**

The user interface will feature a simple and intuitive design, allowing users to submit articles or video links easily. The platform will display a summary of the analysis, including affiliations, ties, and fact-check results.

### **Functionality:**

- **Background Search:** Utilize NLP to extract key information about the author, including affiliations, previous work, and associations.
- **Bias Analysis:** InfoGuard utilizes NLP to assess the potential biases present in the content.
- **Fact-Check Search:** Integrate with reputable fact-checking databases to verify the accuracy of the content and provide a confidence score.
- **User Dashboard:** Users can view a summary of the analysis, empowering them to make informed decisions about the content's reliability.

### **Integration:**

The platform will seamlessly integrate with major social media platforms like TikTok, Instagram meta, allowing users to share and analyse content directly from their feeds.

### **Technical Implementation:**

A cloud-based architecture will support the platform, with serverless functions handling data processing. Machine learning models will be deployed to analyse text and video content, and API integrations will facilitate real-time data retrieval.

### **Challenges and Mitigations:**

**Privacy Concerns:** Implement strict data privacy measures to protect user information.

**Algorithm Bias:** Regularly update algorithms to minimize biases and ensure fair evaluations.

### **Market Research:**

A comprehensive market analysis reveals a gap in current solutions, with InfoGuard offering a unique combination of bias analysis, affiliation checks, and fact-checking capabilities.

Benefits:

### **Challenges and Mitigations:**

Potential challenges include the accuracy of NLP algorithms and the availability of real-time data for fact-checking. Continuous refinement of algorithms and partnerships with reliable fact-checking organizations will address these challenges.

### **Market Research:**

While fact-checking tools exist, none currently offer a comprehensive analysis of an author's affiliations and ties. This unique feature sets our platform apart, offering users a more thorough understanding of the context surrounding the content.

### **Proposed Product Lifecycle process:**

- Conduct a feasibility study to assess technical requirements.
- Develop prototype using agile process, incorporating user feedback at each stage to refine the platform further.
- Focus solutions on each iteration on reducing user problems.
- Explore partnerships with fact-checking organizations to establish data-sharing agreements and enhance the platform's credibility.
- Seeking funding opportunities to support the platform's development and deployment.

### **Conclusion:**

The proposed platform addresses a critical need in today's media landscape, providing users with a tool to assess bias and make informed decisions. By leveraging technology to offer a holistic analysis, this platform stands as a beacon for media literacy and responsible content consumption.

### **References**

1. Robins-Early N. Disinformation reimagined: how AI could erode democracy in the 2024 US elections. The guardian [Internet]. 2023 Jul 19 [cited 2024 Jan 12]; Available from: <https://www.theguardian.com/us-news/2023/jul/19/ai-generated-disinformation-us-elections>
2. Coronavirus fake news actions in the UK 2021 [Internet]. Statista. [cited 2024 Jan 12]. Available from: <https://www.statista.com/statistics/1113700/coronavirus-fake-news-actions-uk/>
3. David Klepper, Associated Press, Ali Swenson, Associated Press. AI-generated disinformation poses threat of misleading voters in 2024 election [Internet]. PBS NewsHour. 2023 [cited 2024 Jan 12]. Available from: <https://www.pbs.org/newshour/politics/ai-generated-disinformation-poses-threat-of-misleading-voters-in-2024-election>
4. Adair B. The lessons of Squash, Duke's automated fact-checking platform [Internet]. Poynter. The Poynter Institute; 2021 [cited 2024 Jan 12]. Available from: <https://www.poynter.org/fact-checking/2021/the-lessons-of-squash-the-first-automated-fact-checking-platform/>
5. Bhuiyan J, Robins-Early N. The EU is leading the way on AI laws. The US is still playing catch-up. The guardian [Internet]. 2023 Jun 13 [cited 2024 Jan 12]; Available from: <https://www.theguardian.com/technology/2023/jun/13/artificial-intelligence-us-regulation>