

Article

Not peer-reviewed version

Fear of Artificial Intelligence? NLP, ML and LLMs Based Discovery of AI-Phobia and Fear Sentiment Propagation by AI News.

[Jim Samuel](#) , Tanya Khanna , [Srinivasaraghavan Sundar](#) *

Posted Date: 13 March 2024

doi: 10.20944/preprints202403.0704.v1

Keywords: AI phobia; Artificial Intelligence; AI; natural language processing; large language models; machine learning; topics; sentiment; emotion; fear; risk; threat; news; headlines; ML; NLP; LLMs



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Fear of Artificial Intelligence? NLP, ML and LLMs Based Discovery of AI-Phobia and Fear Sentiment Propagation by AI News

Jim Samuel, Tanya Khanna and Srinivasaraghavan Sundar *

Rutgers University; jim.samuel@rutgers.edu

* Correspondence: srinivassundar98@gmail.com

Abstract: Confusion, fear and mixed sentiments prevail in the minds of people towards what is arguably one of the most important of dynamics of modern human society: Artificial Intelligence (AI). This study aims to explore the contributions of news media towards this phenomenon - we analyze nearly seventy thousand recent news headlines on AI, using natural language processing (NLP) informatics methods, machine learning (ML) and large language models (LLMs) to draw insights and discover dominant themes. Our theoretical framework was derived from extant literature which posits the power of fear producing articles and news headlines which produce significant impacts on public behavior even when available in small quantities. We applied extensive textual informatics methods using word and phrase frequency analytics, sentiment analysis and human experts based thematic analysis to discover insights on AI phobia inducing news headlines. Our rigorous analysis of nearly seventy thousand headlines using multiple validation methods in NLP (exploratory informatics including BERT, Llama 2 and Mistral [1,2] based topic identification), ML (supervised informatics) and LLMs (neural nets for sentiment classification, with BERT, Llama 2 and Mistral) demonstrates the presence of an unreasonable level of emotional negativity and fear inducing verbiage in AI news headlines. The framing of AI as being dangerous or as being an existential threat to humanity can have a profound impact on public perception, and the resulting AI phobia and confusion in public perceptions are inherently detrimental to the science of AI. Furthermore, this can also impact AI policy and regulations, and harm society. We conclude with a discussion deducing implications for society and make recommendations for education and policies that could support human identity and dignity.

Keywords: AI phobia; Artificial Intelligence; AI; natural language processing; large language models; machine learning; topics; sentiment; emotion; fear; risk; threat; news; headlines; ML; NLP; LLMs

1. Introduction

"Fear defeats more people than any other one thing in the world." Ralph Waldo Emerson

Artificial Intelligence (AI) is a powerful and valuable science, with tremendous potential for value creation across domains and disciplines [3]. However, along with the recognition of its value, another dominant social narrative has emerged stating that AI is dangerous to humans and unreliable, leading to fear of AI [4,5]. Masses of people around the world are confused about AI, and much of the current public perception is shaped by divergent and opinionated news media reports on AI, as seen in Figures 1 and 2. Extant research has shown that news media play a major role in "shaping, viewing, and addressing concerns", and news media have an increasing propensity to create fear and 'construct' crises [6,7]. Misplaced public perceptions on AI can have significant societal consequences, influencing areas such as public engagement with AI, AI education, AI regulations, AI policy and advancement of AI as a science. Furthermore, given that many non-computer science and non-AI professionals and government officials are also swayed by repetitive themes on AI in news media - especially fear of AI. Business and governance decisions can thus be adversely impacted leading to negative consequences for society. Such anticipation of consequential outcomes from persistent fear and risk messaging on multiple topics is supported by extant research, and as aptly stated by Furedi on deteriorating sociability "*lack of clarity has led to the weakening of trust, which in turn has had profoundly destructive consequences for society*" [8]. More specifically, "narratives of fear" are framed,

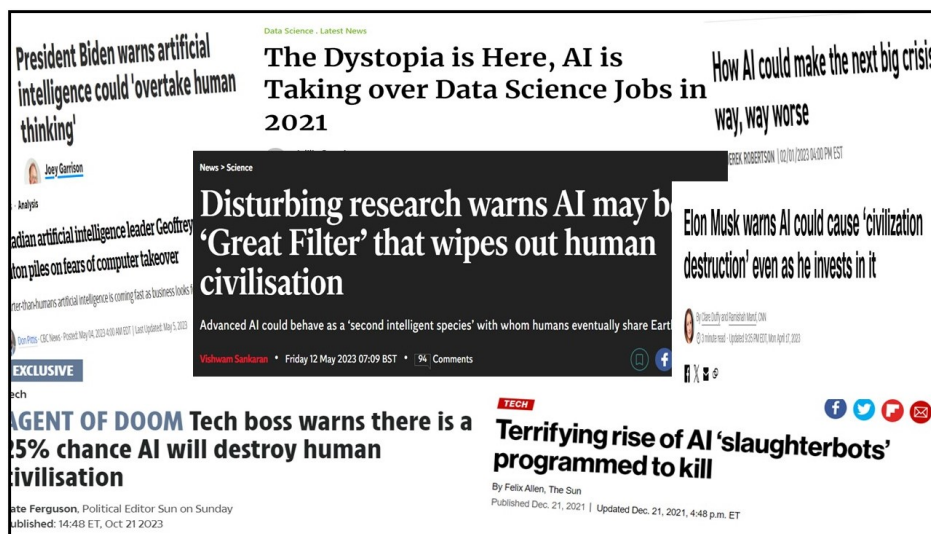


Figure 2. Fear inducing AI news headlines - examples

1.1. What Is AI and Should We Be Afraid?

AI has been defined as being a cluster “of technologies that mimic the functions and expressions of human intelligence, specifically cognition, logic, learning, adaptivity and creativity”, and essentially consists of combinations of math, statistics, data, software and hardware [3,9]. AI is no more dangerous than a math or statistics textbook - it depends on how humans use it: it could be used for education and subsequent noble purposes, or it could be used to cause damage; the knowledge therein could be used to develop advanced therapeutic solutions to benefit humans or to create weapons of mass destruction. It would be inadvisable to call the book dangerous - the book is just a book, it is not by itself dangerous. Similarly, AI as a cluster of technologies is powerful, but it is inherently of itself not dangerous. It is only dangerous to the extent humans misuse AI to develop risky and harmful applications. There must be a separation between AI as a science and persons who use or misuse AI applications and technologies. This simple concept has significant implications for AI law and AI policy - AI regulations must not target, block, limit or stymie the development of AI as a science and instead, AI policy must foster, support and promote advancements in the science of AI. The force of all regulatory initiatives must be focused on targeting persons, companies and governments who indulge in abusive of AI for profiteering, unfair business and governance practices, control, manipulation and other activities causing specific harms. Policies must be proactively directed to ensure shared empowerment of AI capabilities and value creation through transparency (example - open source AI), education and democratization of AI capabilities [9–11].

1.2. Fear of AI around Us

Public perception of AI today is deeply entrenched in fear and concern over moral and ethical quandaries regarding its usage. While it is natural human tendency to be apprehensive of the unknown, the ubiquity of AI-Phobic messages prevalent in society plays a vital role in defining public opinion on the topic. AI News headlines frame information with varying degrees of bias and prejudice, often containing strong emotional undercurrents in order to appeal to the readers imagination as opposed to informing them about the impact of the technology. These headlines often leave a lasting impression on the reader [12]. The fear of AI has been studied and documented in extant research - all the noise surrounding concerns over AI often drowns out the fact that it has rapidly grown into an important science for humanity [4,5]. AI can also be viewed as being a cluster of technologies that is catalyzing innovation and development across a variety of socioeconomic domains. In this context, the propagation of AI Phobia would only serve to impede the progress that has been made

as the acceptance and use of AI applications have been observed to be inversely proportional to the apprehension that permeates in the minds of the public [5].

1.3. Are News Headlines Consequential?

News headlines are consequential because they employ linguistic features "that activate our epistemic and emotional resources, and frame our understanding of covered issues" [12]. News headlines are particularly effective in influencing people on issues of strong importance, such as AI, where systematic and validated factual information is scarce, fragmented information abounds and formal education mechanisms are yet to catch up. News headlines are deliberately designed to be sensational to maximize attention, and fear sentiment is one of the strongest emotional baits as "news items are selected for reporting to engage audiences emotionally rather than intellectually" [12]. Furthermore, 'fear-arousing' news patterns have a multiplicative effect, as it affects individuals and communities, amplifies their risk perceptions and increases the likelihood of them communicating their fears and risk perceptions with other [13]. The growing trend of displaying headlines and placing content behind paywalls implies that many readers will quite likely only read the headlines of AI news articles from news aggregators and social media posts, and not have an opportunity to read the content at all. Hence it is fair to expect news headlines to have a significant impact as they frame news-content, and they influence viewers perceptions on causality of the phenomenon covered by the headlines [14].

1.4. News Impact—Emotions and Sensationalism

Consider the case of a person consuming a pinch of cyanide in a large bowl, mixed with a large quantity of rice - the presence of Cyanide by volume would be statistically non-significant compared to the presence of rice by volume. However, the impacts of this statistically non-significant component would be severe and consequential. Amidst the crowded overdose of daily news, it is important to distinguish between news that impacts human thinking and behavior, and news that merely passes by. On the basis of this inductive logic, we argue in this subsection in favor of quality, impact and effectiveness of emotionally sensitive and sensational news articles and headlines. It is also necessary to distinguish between news media being "creators" of fear-of-AI news versus news media serving as amplifiers and multipliers of fear-of-AI news - news media "*It is important to remember that the media amplify or attenuate but do not cause society's sense of risk. There exists a disposition towards the expectation of adverse outcomes, which is then engaged by the mass media*" A small quantity of emotionally presented sensitive information may lead to tremendous impact on our society regardless of the validity of the news - extant research has highlighted the dramatic impacts of fear and other emotion inducing news [15–19] Misrepresentations of AI as dangerous and as being or as going-to-be more intelligent than humans is consequential - these incite fear and influence imagination and thinking, leading to negative perceptions and resistant behavior towards AI technologies and initiatives [20].

1.5. Motivation and Objectives

In the context of the fact that AI is reshaping and revolutionizing society, it has become critical to understand and address the role of news media in framing AI information for public consumption. The above discussion evidences the widespread presence of fear of AI, also known as AI-phobia [5,21]. There are important and time-sensitive global issues that motivate this study: 1) There is an urgent need to create social awareness about the role of the media in amplifying and multiplying fear of AI, some of which may be justified but is in need of further attention. Society needs to know the magnitude and the consequences of a distorted messaging on AI. 2) Governments needs to become explicitly aware of the extent of AI phobia inducing news. Public leaders and officials need to ensure that the policy and laws enacted are not driven by the forces of media amplified fear and hyped up risk perceptions. 3) AI scientists and business leaders need to become aware of how sensationalism plays out in AI news, and contemplate on the long term damage to public perception.

have also been naively idealistic frames that do not adequately address the risks and potential pitfalls of the technology. Terms such as "AI-Super-intelligence" portray AI and its benefits in a quixotic manner. Between the two extremes, there have also been a number of headlines that present some degree of ambiguity and uncertainty regarding the nature of the technology. Terms such as "black-box" and "enigma" have been more prominent since 2015 [23]. To add to the clutter, News Media has also resorted to terms such as "Godlike AI", "Rogue/Zombie AI" and other expressions seemingly straight out of a science fiction novel [25]. While the intent is often to inform and educate a wider audience about the subject, headlines such as these often have the opposite effect. Often confusing and confounding the reader, muddling discussions about the topic and obscuring the actual impact of the technology. While headlines frequently make bold, attention-grabbing predictions, research indicates that the contents within the headlines often contradict or temper these sensational claims [22,23]. The notion that AI is a threat to humanity was observed to be a recurring theme in several news headlines. This statement would often be accompanied by a quotation from a significant voice within the industry. However, this claim would almost always be refuted by the author. [22]. Concerns pertaining to AI ethics and data privacy was another recurring theme in news discourse. Articles would often raise general, superficial concerns regarding the misuse of AI in the introduction or conclusion of a given article without examining and scrutinizing the topic in detail [22,26]. While AI has also been covered in a effectual, pragmatic way, particularly when covering the applications of AI in finance and healthcare [25], one could argue that the more dramatic headlines frequently overshadow these articles. This pattern indicates a lack of comprehensive understanding in dealing with nuanced concerns [26].

These factors have led to a sense of uncertainty and unease in the minds of the general public. [23]

2.2. *Impact of News Media on Public Perception*

The pervasive nature of media in this day and age coupled with the persuasive effect of news articles fosters an environment where news media influences public opinion [27]. While news media does not drastically change or alter an individual's perception of a given topic, it does determine an individual's perception about public opinion and social climate, thereby influencing popular opinion in a more subtle, indirect way [27,28]. Research in the field of media studies has delved into the dynamics of how media content influences audience perceptions. To tackle the inherent challenges psychology faces in analyzing media content, a new approach called Media Framing Analysis has been introduced. This method aims to provide deeper insights into the ways media narratives shape public opinion and discourse [29]. Additionally, studies on the topic of news media trust have examined its effect on individual media consumption habits, especially in contexts where consumers are presented with a multitude of media choices. These investigations have also shed light on the declining trust in news media within the US [30]. While analyzing the impact of news media in investor sentiment, one study found that in declining markets, investors are influenced by pessimistic news articles while making decisions [31]. Another study found that the coverage of public sentiment in news articles often often caused emotional reactions in the investor, influencing the decision making process [32].

2.3. *Sentiment Analysis of News Headlines*

Sentiment analysis is a widely used NLP technique to identify the emotional tone of a given text [33,34]. Generally, sentiment analysis classifies text into positive, negative or neutral categories, and continuous scores, for example from around -1 to 1. However, more recent tools classify a range of emotions (fear, joy, sadness, etc.). Since the advent of Twitter and other prominent social media platforms, there has been a steep increase in NLP research developing innovative models and methods [35]. Another fascinating facet of deriving sentiment insights is the ability to track public sentiment towards a particular topic over a period of time [36]. This analysis would help in validating the timing of certain decisions and aid in understanding how certain decisions will be perceived by the general public [37].

Ever since its inception, lexical based methods for sentiment analysis have proven to be robust, reliable, and performant across multiple domains [38]. This technique assigns positive and negative polarity values to the words in a given sentence and combining them using an aggregate function. Several studies utilize this method to perform sentiment analysis on news reports with decent accuracy scores [39]. One study constructed a sentiment dictionary to understand and identify the emotional distribution of a news article [40]. This algorithm often forms the baseline for more complex models to be built on top of it. To tackle more complex datasets with longer texts, a different approach rooted in machine learning is required. The sentences are initially preprocessed through either the Term Frequency - Inverse Document Frequency algorithm [41], or the Bag of Words algorithm. Once this is done, a classification model is used in order to accurately determine the sentiment of a given text. With this method, it is possible to obtain exceedingly high accuracy scores [42]. In sentences comprising of multiple keywords, accurately predicting the sentiment can be a challenge. In order to address this limitation, one study introduced an efficient feature vector to the dataset prior to building the classification model [43]. While a vast array of algorithms have been developed to perform sentiment analysis, in the recent past, LLM based models have become popular. Several studies have demonstrated that LLM based models obtain far better accuracy scores when compared to more traditional ML based Sentiment Analysis models for a wide range of applications. Furthermore, LLM based models are capable of tackling more nuanced problem statements [44,45] Ranging from the prediction of stock prices based on twitter sentiment [46] to accurately predicting the sentiment of a given news headline [47], studies have shown that LLMs have performed exceedingly well in all these areas.

2.4. Other NLP Methods

There exists a rich array of established and recent NLP methods. These include Exploratory Data Analysis (EDA), word and phrase frequency and N-Gram analyses, word and phrase distances, text summarization (abstractive, extractive and hybrid), topic identification, Named Entity Recognition (NER), syntactic and semantics analyses and aspect mining among others, including framed methods such as public sentiment scenario analyses [37]. Text summarization primarily focuses on deriving the essence of a long piece of text and representing it as a coherent, fluent summary [48]. More recently, several deep learning methods have been used to perform extractive text summarization.[49]. In the abstractive approach, a given text is first interpreted in an intermediate form and then the summary is generated with sentences that are not a part of the original document. Extractive summarization method is far more popular than the abstractive summarization approach [50]. Topic modeling is a powerful NLP technique used to discover thematic clusters and often non-obvious areas of interest within large text corpora. This method enables researchers and analysts to identify and categorize otherwise difficult to identify themes or subjects in text data [51]. NER is a method utilized in NLP to identify persons, places and other objects such as names of vehicles, products, companies and government agencies or offices that could be defined using custom algorithmic approaches [52].

2.5. Machine Learning and LLMs

NLP tools use a broad range of AI methods. for example sentiment analyses can be performed using rules based methods, machine learning and LLMs. Extant research has shown that the machine learning methods generate better sentiment analysis accuracy in many domains [53]. Though a nuanced problem to solve, machine learning methods have been used extensively to great effect in emotion recognition [54], often employing a combination of multiple algorithms to achieve high accuracy scores [55]. In addition, identifying and selecting relevant features prior to the classification process has shown a marked improvement in performance [56]. Studies have shown that Hidden Markov Models have outperformed more traditional rule based approaches to Intrusion Detection [57]. The advent of Deep learning opened the doors to solve more complex problems. Deep learning models have also proven to be more robust and resilient to handling adversarial texts [58]. While Convolutional

Neural Networks have been utilized to perform sentiment analyses [59] and text summarization [60], Recurrent Neural Networks with attention mechanisms have found widespread application to solve a variety of problems across multiple domains [61]. Recent advances in LLMs have rapidly established them as a favored solution across various domains. Studies have proven their capabilities in knowledge intensive NLU and NLG tasks [62]. LLMs have been used to accelerate the annotation processes, thereby reducing the cost and time required to annotate texts for NLP algorithms [63,64]. LLMs have also proven to be implicit topic models capable of surmising the latent variable from a given task [65]. Several innovative applications have been developed using LLMs that continue to revolutionize the field of NLP [66,67].

3. Data

Our study leveraged the Google News RSS [68] feed as a primary data source for collecting headlines of news articles which contained AI keywords. The Google News RSS feed provides a broad array of news stories across various categories and regions in an RSS (Really Simple Syndication) format, allowing users to receive updates on the latest news articles. This service was pivotal in enabling access to a wide range of AI-related news content, reflecting a broad range of casual, scientific, global, regional and opinion news-frames on AI.

3.1. Data Collection Methods

We employed a multifaceted approach and used a set of predefined parameters for querying the Google News RSS feed. These parameters, which are delineated in Table 1, include language specifications, the time frame of article publication ensuring contemporary relevance, and the relevant search terms used to filter the content. We employed ScrapingBee selectively when deeper data extraction was necessary. This combined strategy, leveraging both direct RSS feed queries and targeted web scraping, yielded a dataset of 69,080 AI-related news headlines, providing a robust foundation for this research.

Table 1. Google news RSS feed search conditions.

Data Source	Google News RSS Feed
Search Date	November 1, 2020 to February 16, 2024
Search Terms	AI; A.I.; Artificial Intelligence
Retrieval Tool	Custom URL generation
Additional Tool	ScrapingBee API (conditional use)

3.2. Data Processing and Preparation

Once the news headlines were collected from the Google News RSS feed, we prepared the dataset for analysis. The initial step involved the elimination of duplicate entries based on headlines to maintain the uniqueness of our dataset. Next, we employed language detection to identify and translate non-English headlines, creating a unified English-language dataset for consistent analysis. We extracted information from the downloaded data and developed features and temporal attributes, including day of week, month, year, quarter, and weekend indicator.

Lastly, we dissected the URLs to extract the final destination after redirects, the primary domain, any present subdomains, the depth of the URL indicated by the number of slashes, the top-level domain (TLD), and the length of the URL. This process resulted in a refined dataset of 67,091 English-language news headlines, with multiple extracted attributes representing basic Information such as publication title, publication date, publication link, publication source, country (US) and derived Features such as translated title, language of the title, day of week, month, year, quarter, weekend status, holiday status, final redirected URL, domain of the URL, subdomain of the URL, URL depth, TLD, and URL length.

4. Methodology

In this section we outline our methods, applying NLP to generate insights from our data, and applying machine learning and LLM methods to model the data to provide an in-depth analyses of fear of AI spread by news headlines. The exploratory analysis follows standard processes and emphasizes word and phrase frequencies bases analyses. For the core part of the analyses, we employ a 'public-impact' or 'public-influence' perspective. This implies an approach where the data is modeled and evaluated with in the context of the vast masses of people, and not the specialized sections those who are AI specialists or researchers or associated with the AI domain. For example, a hypothetical news headline which says "Meta Leadership Happy as Llama 2 with RAG Gladly Outperforms Llama 2 with Finetuning", would be a happy sentiment for Meta leadership and would be classified as a positive-sentiment by many models. However, we would identify it as a neutral sentiment from a public-impact sentiment perspective. We would classify another hypothetical news headline which says "Meta Leadership Happy as Meta Releases a Free Tool for Students Worldwide" as a positive-sentiment from a public-impact sentiment perspective. Similarly, in studying fear, we identify, analyze and model fear of AI and AI phobia in news headlines from public-impact and public-influence perspectives. In all figures, the character "k" represents the quantity of one thousand.

4.1. Exploratory Data Analysis

As the first step in our methodology, in order to gain insights into the linguistic patterns weaved into AI news, we perform exploratory data analysis (EDA) our data news headlines dataset. Our EDA focuses on examining and summarizing word frequencies, key phrases, n-grams analysis, distribution of words and characters, summary of publication sources, preliminary sentiment analyses for positive, neutral and negative sentiments, temporal and geographic distributions, and qualitative identification of notable themes.

4.1.1. Temporal Distribution of Articles

Our dataset included headlines ranging from the 1st of November, 2020 to the 16th of February 2024. Our temporal distribution assessment indicated a progressive annual increase in the volume of published articles on AI, beginning with 1,761 articles in the last few weeks of 2020, 11,208 in 2021, and 13,952 articles in 2022, indicating an growing engagement with AI topics. In 2023, the numbers jumped to 34,527 articles - this is most likely due to a variety of factors such as the rise of generative AIs, launch of multiple foundation models including numerous LLMs, breakthroughs in AI technology, widespread adoption of AI by industry and government, public policy debates, and significant incidents related to AI that captured public attention. For the first few weeks of 2024, our dataset included 5,643 headlines. A monthly breakdown revealed that October of 2023 experienced the highest frequency of publications, while April 2021 saw the least. Day of week analysis of number of articles published pointed to a pattern where Tuesdays and Wednesdays experienced the most article publications on average, in contrast to lower activity during the weekends. Though our dataset contains only 3 full years (2021-2023) and 2 part-years (2020 and 2024), the yearly trend as seen in Figure 4, indicates a year-over-year increase. Overall, the findings suggest a growing interest in AI topics, with publication trends potentially reflecting cycles of industry, academic, government and public engagement with the rapidly evolving AI domain.

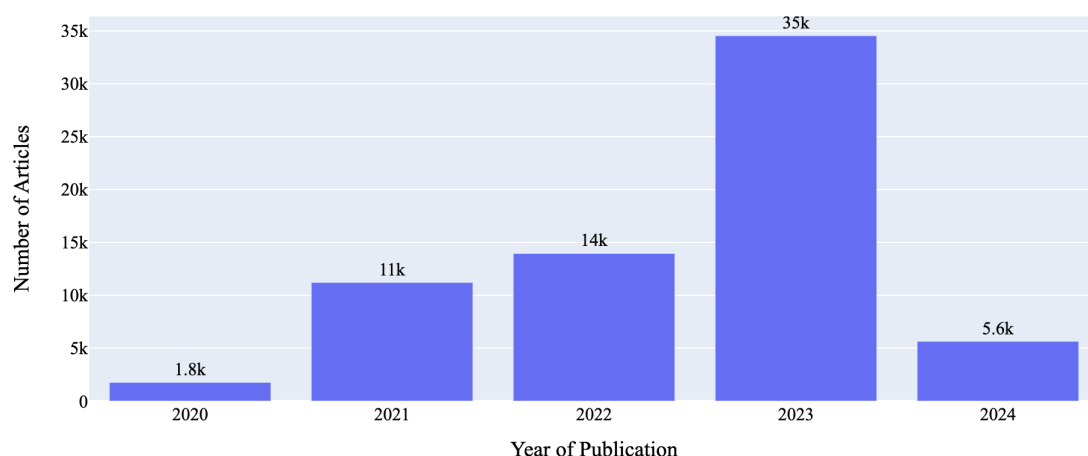


Figure 4. Yearly distribution of articles published in our dataset

4.1.2. Linguistic and Geographic Features

Language Diversity: Our dataset, with English translations from multiple original language news headlines, was tagged for source language. Our analysis revealed the extent of language variations, displaying linguistic diversity in the data. Figure 5 illustrates the distribution of the top non-English languages represented in the headlines. Korean (ko) led with over 5,000 headlines, followed by Japanese (ja) and Spanish (es) exceeding 1,000 each. Other languages like French (fr), German (de), and Indonesian (id) had moderate representation, while Vietnamese (vi), Portuguese (pt), Dutch (nl), Chinese (zh-cn), Russian (ru), and Hindi (hi) were less prevalent. English language dominated, constituting 82.3% of the corpus, highlighting a potential bias and mirroring the internet's linguistic landscape, where English is dominant. Hence our data though global, is skewed towards English language implying that our analyses on understanding the fear of AI and AI phobia propagated by news headlines is largely associated with English speaking people and nations. However, given the global dependence on the English language and widespread reliance on the English language for global news on AI, our dataset provides a fair representation of global news headlines on AI.

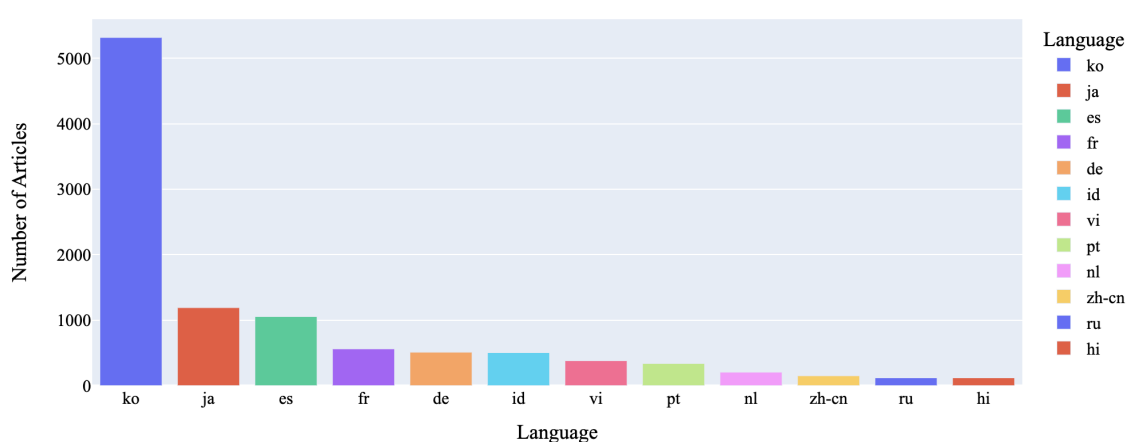


Figure 5. Non english languages distribution of articles in our dataset

Geographic Distribution: To visualize the geographic spread of AI news, we created a choropleth map (see Figure 6) based on identifiable publisher locations. The analysis revealed that most of the articles were from North America (32%) and Europe (29%), potentially reflecting their higher investments into and engagement with AI. The United States, leading the cluster within North America, contributed a the most with 32% of all articles, followed closely by Europe's 29%. This implies a greater

impact and influence of the AI narratives and policy initiatives from these regions in shaping the future of AI perception globally. Asia demonstrated fair presence, contributing 22%, with East Asian nations like China and South Korea providing most of the headlines. Notably, our data revealed an interesting trend of an increasing number of articles from regions traditionally underrepresented in technology discussions. Africa, Latin America, and the Middle East collectively contributed 17% of the publications - this potentially signals a rising global interest in AI conversations, reflecting the universality of AI's importance.

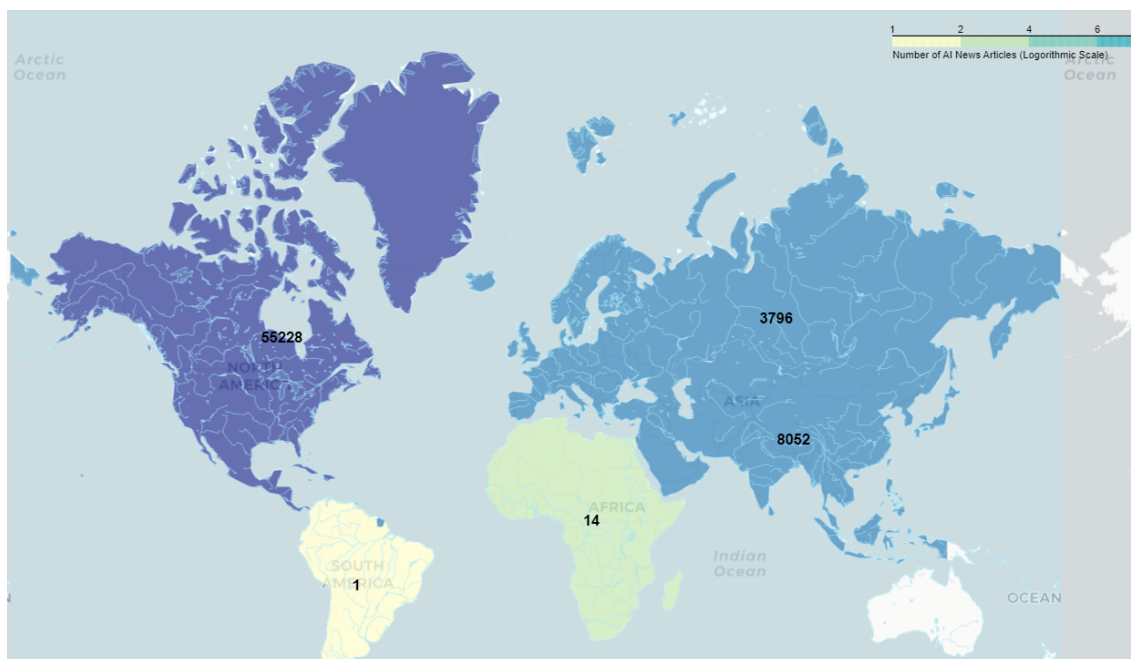


Figure 6. Geographic distribution of AI articles in our dataset

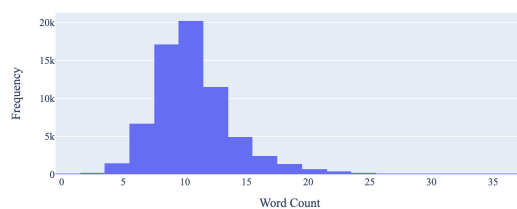
4.1.3. News Headline Textual Analysis

Quantitative Analysis: A quantitative examination of the linguistic components of the AI news headlines reveals that, as an average, the headlines have a character length of almost 68 letters of the alphabet (Figure 7), suggesting a broad range of length of headlines. Word count data as displayed in Figure 7 further supports this trend, with an average of 10.69 words per title, striking a balance between brevity and descriptiveness.

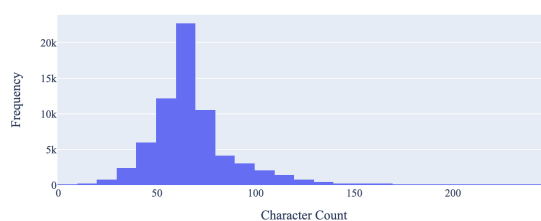
The range of the data depicts variability: titles can be as brief as 8 characters or as expansive as 242, and the number of words can vary from a singular term to 37, reflecting a spectrum of approaches to titling from the succinct to the elaborate. **Text Cleaning and Preprocessing:** The preparation of text and preprocessing steps varied by the type of analysis or model. For example, in creating wordclouds and N-grams, to ensure consistent analysis, all titles were converted to lowercase and stripped of non-alphabetic characters like punctuation and symbols. Titles were then converted to a corpora of individual words from which common stop-words were excluded. Additionally, domain-specific stop words like "AI" and "ML" were removed to focus on more significant terms. We applied standard stop-words available from Python libraries, and we also applied custom stop-words based on irrelevance, commonality and error at multiple points to ensure that the most relevant words and phrases are visible. Terms like "stock" and "buy" were also excluded from the overall n-gram analysis.

Named Entity Recognition (NER) Analysis: NER analysis helped us identify the top individuals, companies, locations and items mentioned in the news headlines. We conducted NER analysis using the state-of-the-art language model "en_core_web_trf" from the SpaCy [69] library. This analysis aimed to extract entities of interest, particularly focusing on individuals (PERSON) and organizations (ORG) mentioned in the data. By identifying the sources frequently referenced in news headlines, we gain

a deeper understanding of the stakeholders influencing public discourse related to AI. The analysis revealed a strong presence of tech giants, such as Google, Meta, Microsoft, and OpenAI, and their products and services such as ChatGPT [70], Bard and chatbots. The presence of regions, like China and the EU (European Union) imply a geopolitical and regulatory dimension to the AI news headlines, and individual tech leaders like Elon Musk, Sam Altman, and Mark Zuckerberg were frequently mentioned, signifying the importance of their actions and opinions on AI.



(a) Distribution of Words



(b) Distribution of Characters

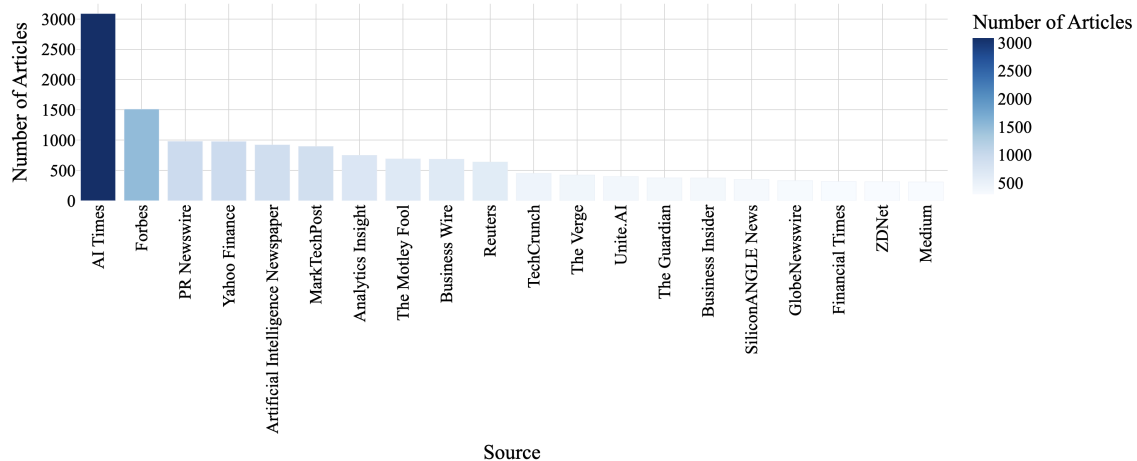
Figure 7. Distribution of words and characters for news headlines in our dataset

4.1.4. Publication Sources

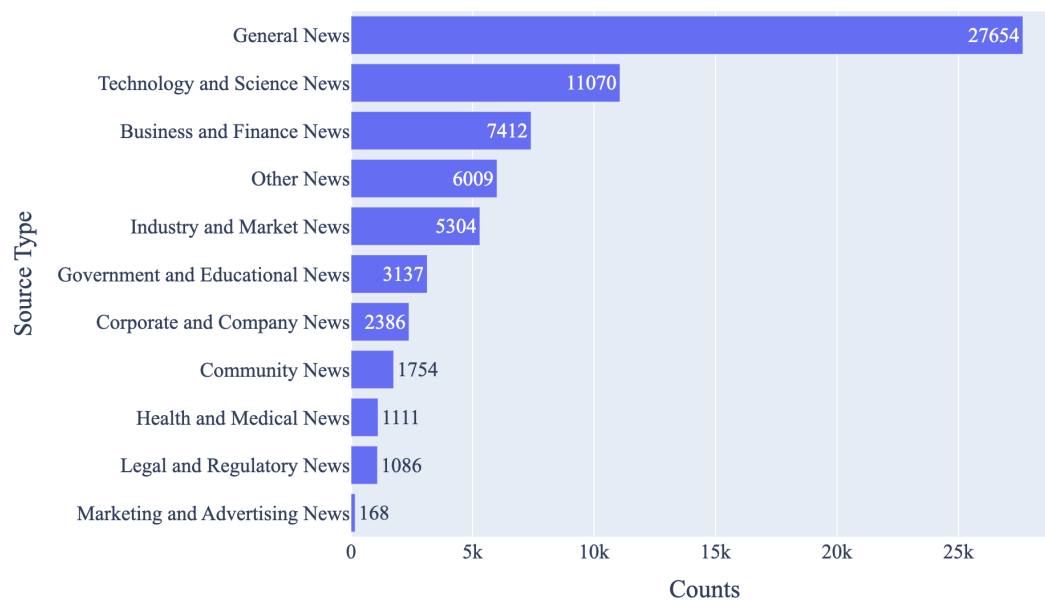
Our analysis summarized the number of AI articles contributed by identifiable sources - we did not check for originality of articles posted by source. Our summary is illustrated in bar chart Figure 8, and surprisingly the biggest source is AI Times from South Korea. Forbes and PR Newswire follow, along with other notable sources which include Yahoo Finance, MarketWatch, and Reuters, each offering hundreds of articles and representing a blend of specialized and general AI news. To better understand the variety of sources reporting on AI, we classified these sources by identifying the most pertinent keywords linked to each type, and this process was augmented by a manual review of the sources as well as by employing the top-level domain attribute, which included designations such as '.org' and '.edu'. The categorization yielded the distribution of publication source categories as illustrated in Figure 8.

An analysis of the top five sources (as illustrated in Figure 9) within each category helped us understand the ways in which sources and source-categories are shaping the AI narrative at an aggregate level. PR Newswire, Reuters, and AI Times in the 'News Outlets' category offer a mix of perspectives highlighting the global impacts of AI. Business Wire and TechCrunch, in the technology and science publications category, provide insights into industry developments and emerging trends. The business and finance category features Forbes with its expert opinions and Yahoo Finance focusing on specific companies and industries impacted by AI. Industry and market insights are offered by Analytics Insight, while government and educational content comes from Government Technology and the Walter Bradley Center, representing both governmental perspectives and specific viewpoints on AI development. Corporate and company blogs like Microsoft and Nvidia offer insights into their own strategies and advancements. Community, AI-topic and AI-methods content features Medium that fosters discussions and blogs from various contributors, while health and medical publications like Medical Xpress and FierceBiotech showcase the intersection of AI with healthcare and biopharmaceutical research. In legal and regulatory news, JD Supra and Lexology provide information

on the evolving legal landscape surrounding AI, and marketing and advertising insights come from Adweek and Marketing Dive, potentially highlighting the niche use of AI in these sectors. These are aggregate level comments, and a number of common AI events and technologies are covered in some form across categories - source analysis underscores the multifaceted nature of the AI narrative, shaped by a range of perspectives from various domains.



(a) Top 20 Publication Sources in our Dataset



(b) Distribution of Source Types in our Dataset

Figure 8. Analysis of publication sources in our dataset

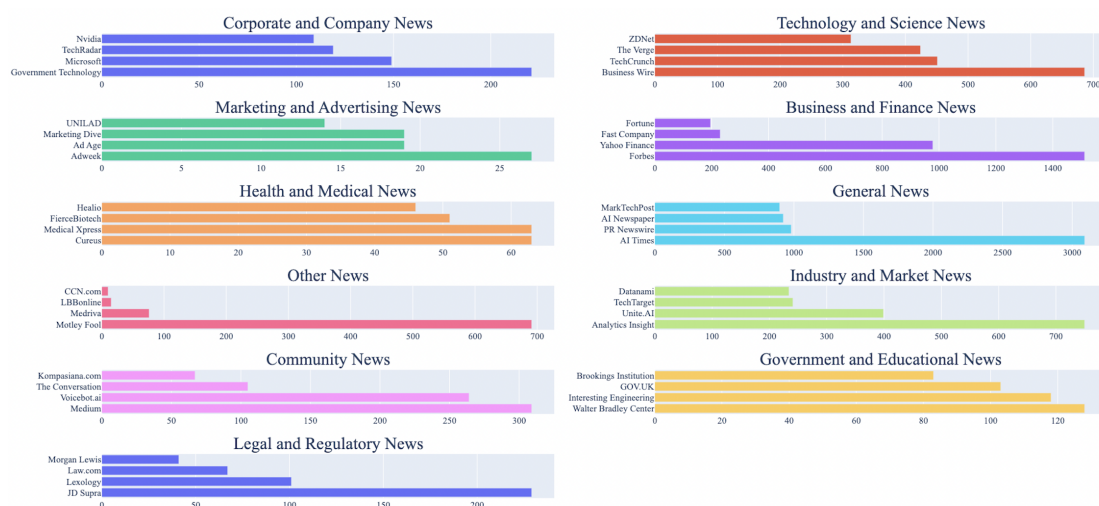
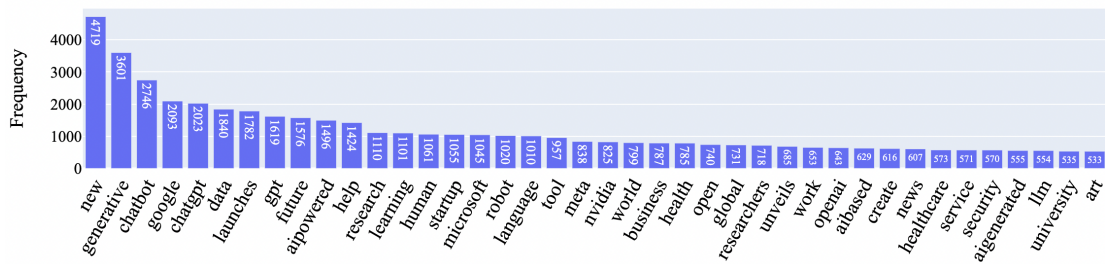


Figure 9. Top source types identified within our dataset

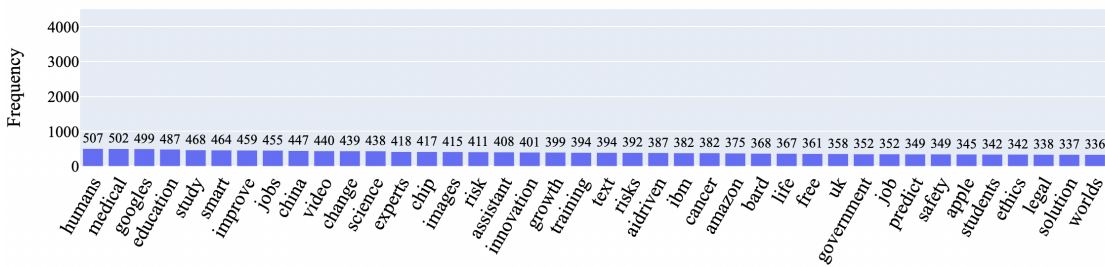
4.1.5. N-Grams Analysis for Full Data

We explored the most frequent single words (unigrams), two-word phrases (bigrams), three-word phrases (trigrams), and four-word sequences (quadgrams) within the titles. To quantify these occurrences, a CountVectorizer [71] transformed the preprocessed titles, creating matrices that reflected the frequency of each word or phrase.

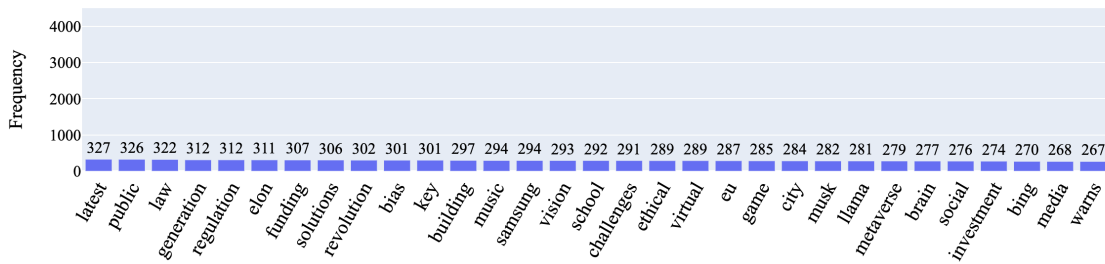
To uncover patterns and trends in the news headlines in our dataset, we conducted an N-gram analysis using the CountVectorizer [71] tool from the scikit-learn library. N-grams are successive sequences of n items from a given text document. By examining these sequences, we can identify commonly used phrases and terms within the dataset, providing insight into recurring themes and topics. For our analysis, we generated unigrams (single words), bigrams (two-word sequences), trigrams (three-word sequences), and quadgrams (four word sequences) to identify key topics and intricate patterns in the news headlines present within our dataset. The CountVectorizer [71] was configured to filter out irrelevant stop words such as “stock” and “buy” to ensure that the focus is on pertinent terms in each headline. A close examination of the unigrams extracted from our dataset reveals recurring keywords such as “new”, “generative”, “chatbot”, google, and “chatgpt” indicating that several discussions about AI in news headlines include the latest applications of the technology and the major players in the market. Moreover, terms such as “business”, “healthcare”, “security”, “art”, and “medical” highlights the ubiquitous impact of the technology across multiple domains. The top 120 Unigrams can be seen in Figure 10. Building on the insights observed in the unigram analysis, a close look at the bigrams highlights terms such as “AI ethics”, “AI regulation”, “responsible AI”, “AI safety”, which collectively suggest a growing concern over the moral and ethical implications of utilizing this technology. Figure 11 presents the top 80 bigram frequencies. An analysis of the trigrams unveils themes such as “artificial intelligence help”, “using artificial intelligence”, and “impact artificial intelligence” underscoring a focus on the usage of the technology and its impact across multiple domains. Delving into the quadrams unravels intricate themes found in our dataset. Phrases such as “things ai missed today”, “artificial intelligence health market”, and “ai ethics ai law” showcase a nuanced discourse where these factors play a crucial role. The top 25 trigrams and quadgrams can be found in Figure 12 and Figure 13 respectively.



(a) Top 40 Unigrams

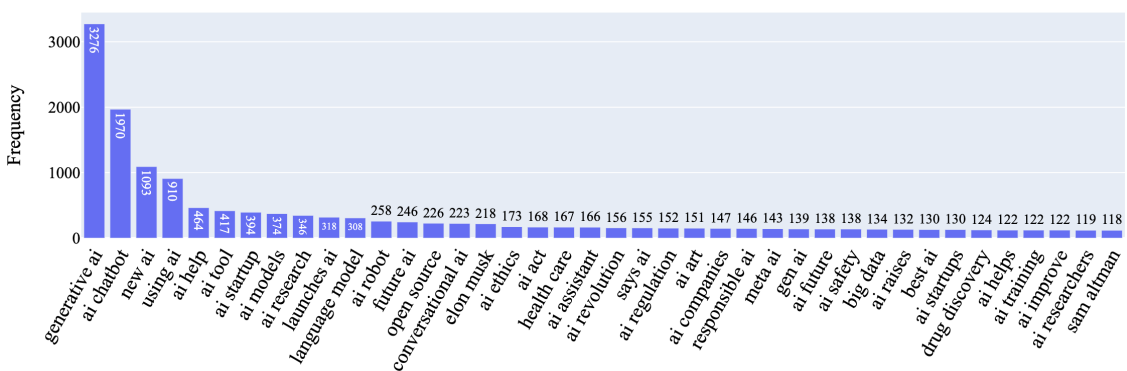


(b) Subsequent 40 Unigrams - 1



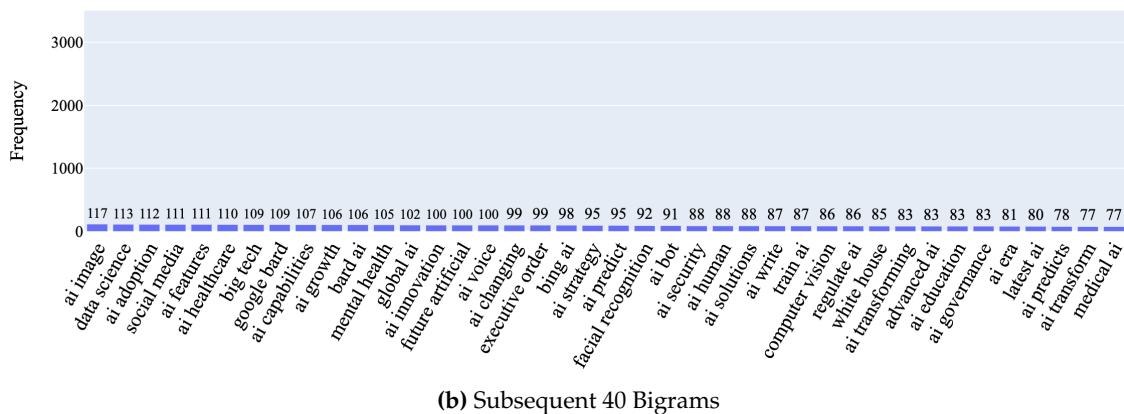
(c) Subsequent 40 Unigrams - 2

Figure 10. Analysis of unigrams present in news headlines in our dataset



(a) Top 40 Bigrams

Figure 11. Cont.



(b) Subsequent 40 Bigrams

Figure 11. Analysis of bigrams present in news headlines in our dataset

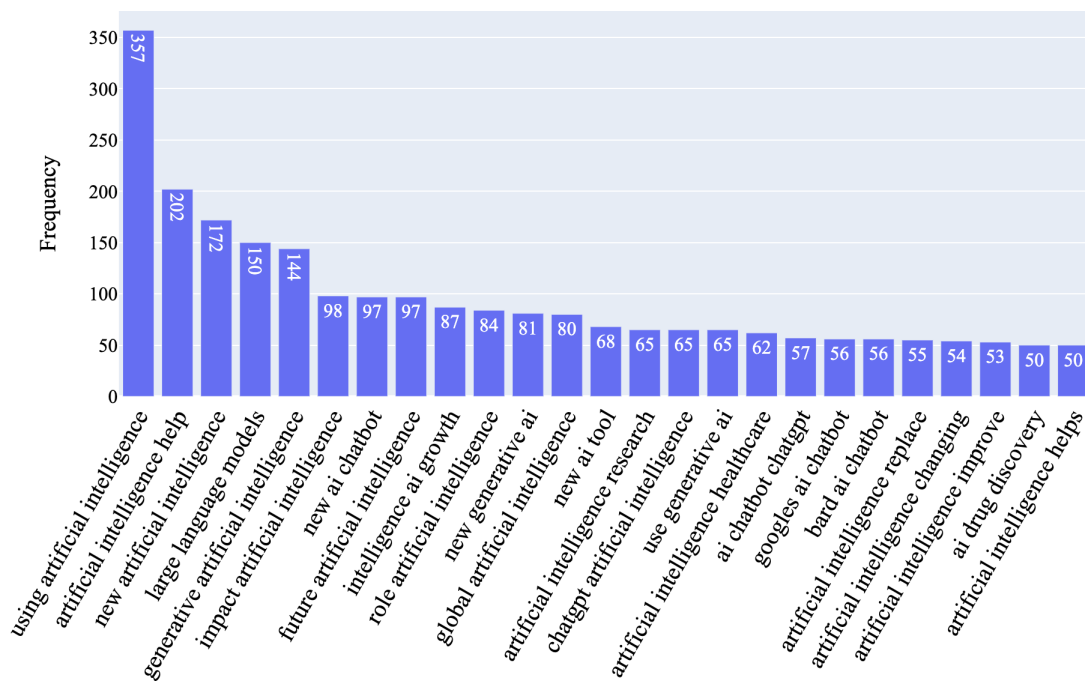


Figure 12. Top 25 trigrams found in our dataset

4.2. Sentiment Classification

We initiated our analysis using three popular sentiment analysis libraries: VADER, AFINN, and TextBlob [72–74]. This step was crucial for establishing a baseline understanding of sentiment distributions within our data. We computed sentiment scores for the entire dataset using all three tools as shown in Table 4. This multi-methods approach provides rich insights and also serves as a robust validation mechanism increasing the rigor and value of the aggregated analyses.

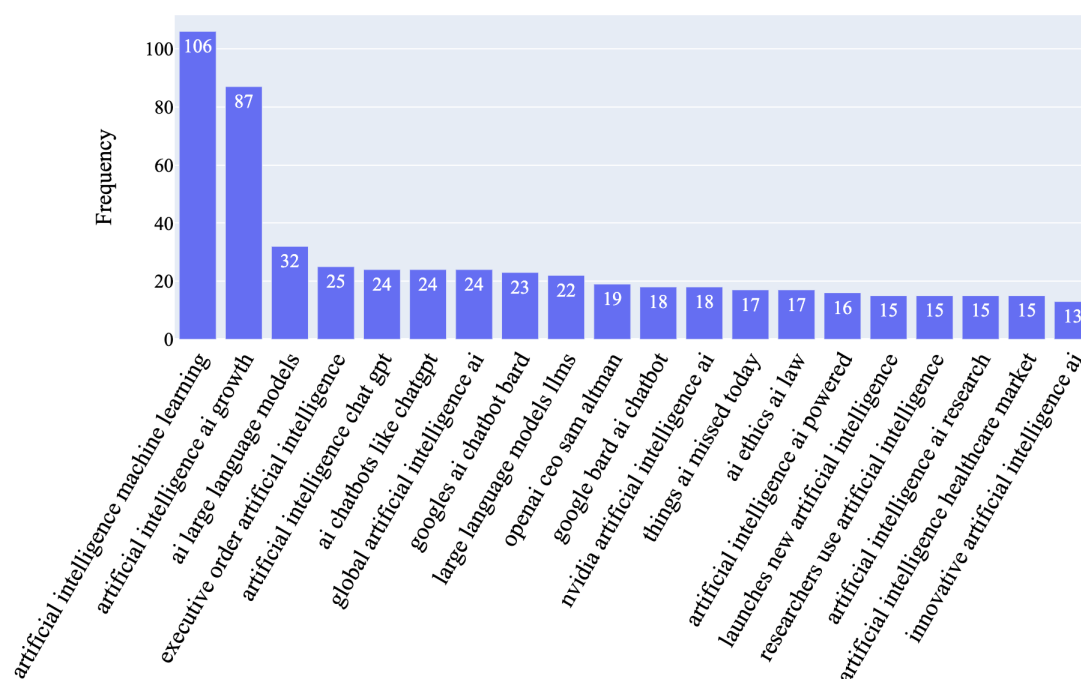


Figure 13. Top 25 quadgrams found in our dataset

4.2.1. Sentiment Analysis

In order to assess the sentiment conveyed in AI news headlines, we employed multiple sentiment analysis methods including sentiment analysis by sentiment word dictionaries and lexicons, LLM based sentiment classification and machine learning based sentiment classification. In this section we focus on lexicon based sentiment analysis, which is the weakest of the methods - especially in the absence of domain specific dictionaries [53]. However, lexicon based sentiment analysis serves as a baseline from which the analysis is built up and for this purpose we report the results from multiple tools, namely TextBlob, Valence Aware Dictionary for Sentiment Reasoning (VADER), Affin and FLAIR [75].

- **TextBlob:** is a library that provides a simple API for diving into common NLP tasks such as part-of-speech tagging, noun phrase extraction, and sentiment analysis. The sentiment function of TextBlob [73] returns a polarity score within the range of -1 to 1, where -1 indicates a negative sentiment, 1 indicates a positive sentiment, and scores around 0 indicate neutrality.
- **VADER,** on the other hand, is a lexicon and rule-based sentiment analysis tool specifically attuned to sentiments expressed in social media. It uses a combination of a sentiment lexicon that is human- and machine-curated and considers factors such as intensity and context. VADER's compound score, which we used, is a normalized, weighted composite score that also ranges from -1 (most extreme negative) to +1 (most extreme positive).
- **Afinn** sentiment analysis tool assigns scores to words based on a predefined list where scores range from -5 to +5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment.
- **FLAIR** We used a pretrained model, FlairNLP, a comprehensive NLP framework. This model leverages sequence labeling in order to detect either positive or negative sentiment in a given text. Table 5 Showcases the results obtained using this model.

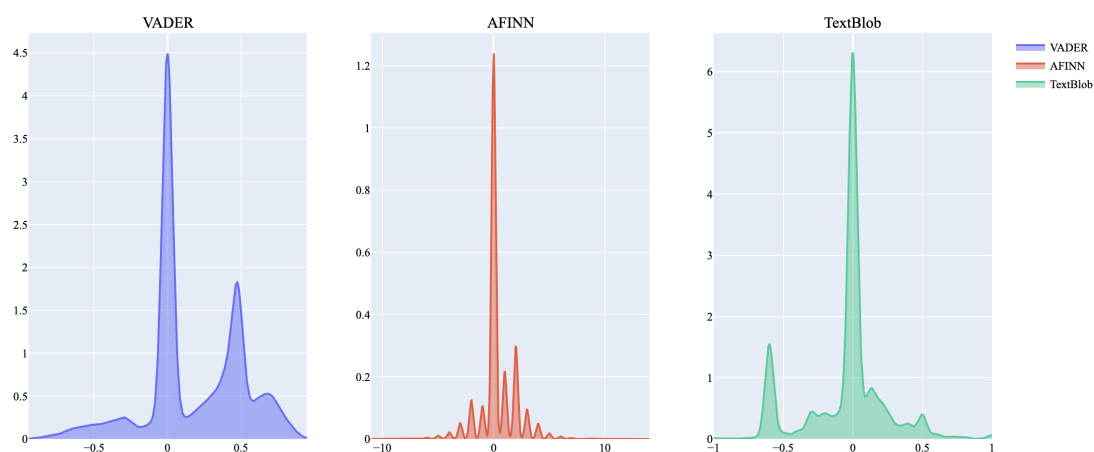


Figure 14. Kernel density estimation of VADER, AFINN, TextBlob

4.3. Large Language Models for Topic Modeling

4.3.1. Topic Modeling with BERT

Topic modeling refers to the statistical and algorithmic techniques used to identify "topics" and themes present within text corpora. It serves as an important approach for discovering hidden semantic similarity clusters and patterns in text data, facilitating the understanding, and sub-grouping of text datasets. By recognizing patterns in word usage across documents and categorizing them into topics, it allows for the examination of a large corpus's thematic framework. In our study, we utilized topic modeling to search for and identify dominant themes from AI-related news headlines. Our methodology involved the use of BERTopic[76], a method that employs BERT(Bidirectional Encoder Representations from Transformers) word embeddings to derive semantically rich sentence embeddings from documents.[77] We preferred BERTopic[76] to Latent Dirichlet Allocation (LDA) because of its superior capability in capturing the contextual meanings of words, a strength supported by previous studies showing BERTopic's effectiveness over LDA and Top2Vec in topic detection from online discussions. [78,79]BERTopic[76] combines advanced NLP techniques for topic modeling, integrating BERT embeddings with dimensionality reduction and clustering algorithms to identify text corpus topics.

We began by converting our text data into embeddings using Sentence Transformers, specifically designed for creating sentence embeddings, thus turning sentences into high-dimensional vectors. To manage the embeddings' high dimensionality, we applied Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction, which maintains the data's local and global structure for easier visualization and clustering. Subsequently, we employed Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) for clustering, an algorithm adept at detecting clusters of various densities and sizes, crucial for accurately defining topics.[80,81] For topic interpretation and representation, we used CountVectorizer[71] to form a bag-of-words matrix, then applied class-based Term Frequency-Inverse Document Frequency [] to refine topic differentiation by highlighting words with higher unique prevalence in specific clusters. This process started with transforming documents into a token count matrix, enabling the exclusion of common stop-words and other preprocessing actions to polish the data. Adjusting the TF-IDF to cluster level, we grouped documents per cluster for term frequency calculations, normalization, and a modified IDF calculation that accounts for term frequency across clusters. We experimented with CountVectorizer[71] settings, like n-grams, to fine-tune topic representations tailored to our needs, aiming for optimal topic granularity. To counteract the dominance of frequently occurring but less informative words, we applied a square root transformation to term frequency, smoothing out term distribution across topics. Additionally, we leveraged Maximal Marginal Relevance to enhance topic keyword diversity and reduce redundancy,

ensuring a balanced representation of terms within topics. Finally, we capitalized on HDBSCAN's capability to pinpoint outliers and amalgamate similar topics via an automatic topic reduction process, refining the topic model by merging closely associated topics for a more cohesive and interpretable set of topics.

4.3.2. Topic Modeling with Llama 2

We leveraged the advancements in LLMs to enhance topic modeling accuracy and efficiency. Specifically, we integrate the capabilities of Llama 2[2], a potent LM with accuracy comparable to OpenAI's GPT-3.5, with BERTopic[76], a modular topic modeling technique. This synergy aims to refine topic representations by distilling information from the topics and clusters generated by BERTopic[76] through fine-tuning with Llama 2[2]. We focus on the 'meta-llama/Llama-2-13b-chat-hf' variant, balancing model complexity with operational feasibility in constrained hardware environments.

- **Model Optimization:** Given the limitation of our hardware, we employed model optimization techniques to facilitate the execution of the 13 billion parameter Llama 2[2] model. The principal optimization technique was 4-bit quantization, significantly reducing the memory footprint by condensing the 64-bit representation to a 4-bit one. This approach is not only efficient but also maintains the model's performance integrity for topic modeling tasks.
- **Prompt Engineering for Llama 2:** To effectively utilize Llama 2 for topic labeling, we designed a structured prompt template, incorporating both system and user prompts. The system prompt positions Llama 2[2] as a specialized assistant for topic labeling, providing a consistent contextual foundation for all interactions. The user prompt, however, is more dynamic, consisting of an example prompt to demonstrate the desired output and a main prompt that includes placeholders for documents and keywords specific to each topic. This design facilitates the generation of concise topic labels, optimizing Llama 2's output for our topic modeling objectives. Within our prompt, two specific tags from BERTopic[76] are of critical importance:

[DOCUMENTS]: This tag encompasses the five most pertinent documents related to the topic.

[KEYWORDS]: This tag includes the ten most crucial keywords associated with the topic, identified via c-TF-IDF. This format is designed to be populated with the relevant information for each topic under investigation.

- **Implementation with BERTopic:** The integration with BERTopic[76] involves a two-step process. Initially, BERTopic[76] generates topics and their corresponding clusters using documents from our dataset. These topics, characterized by their most relevant documents and keywords identified through c-TF-IDF, serve as the input for Llama 2[2], following our prompt template. The template is populated with the top 5 most relevant documents and the top 10 keywords for each topic, guiding Llama 2[2] to generate a short, precise label for the topic.

To enrich our topic representations, we incorporated additional models alongside Llama 2[2]. Specifically, we utilized c-TF-IDF for the primary representation and supplemented it with KeyBERT[82], MMR (Maximal Marginal Relevance), and Llama 2[2] for multi-faceted topic insights. This approach allows for a comprehensive view of each topic from multiple analytical perspectives. With the models and methodologies in place, we proceeded to train our topic model by supplying BERTopic[76] with the designated sub-models. The training process involved fitting the model to our dataset and transforming the data to extract topics. Through careful optimization and prompt engineering, we achieved an efficient and effective topic modeling process, suitable for environments with limited computational resources.

4.4. AI Fear Classification with LLMs

Large Language Models (LLMs) have significantly advanced the field of emotion classification, enabling nuanced detection and analysis of emotional expressions in text. These models, trained on extensive datasets, adeptly identify a vast array of emotions—from joy and sadness to complex states

like anticipation and surprise—with exceptional accuracy. In our research, we harness the pre-trained capabilities of LLMs, specifically DistilBERT, Llama 2, and Mistral, focusing on detecting expressions of fear in AI-related news headlines.[83] Our approach leverages the pre-trained capabilities of these models without additional fine-tuning, relying on their inherent understanding of linguistic nuances to accurately classify fear related to artificial intelligence. This approach allowed us to experiment across different models, exploring their unique strengths and capabilities in the specific context of fear classification without the need for resource-intensive model retraining.

4.4.1. Fear Classification with DistilBert

To enhance the capability of our fear detection framework in processing complex linguistic patterns and understanding contextual nuances, we explored the utilization of the BERT (Bidirectional Encoder Representations from Transformers) architecture. BERT's transformative approach in natural language processing stems from its deep bidirectional representation, which allows it to capture intricate details from the text by considering the context from both directions. We specifically utilized the Falconsai/fear_mongering_detection model[84], which is a fine-tuned variant of DistilBERT from the Hugging Face model repository.[85] DistilBERT is a distilled version of the original BERT architecture, designed to deliver a similar level of performance as BERT while being more lightweight. It achieves this balance by utilizing 6 Transformer layers, as opposed to the 12 layers used in BERT-base, and 12 self-attention heads, totaling around 66 million parameters, which is about 40% fewer than its predecessor. The Falconsai/fear_mongering_detection model had already been specifically fine-tuned for fear mongering detection, suggesting that it would be adept at discerning the subtleties associated with fear-related content. By leveraging this pre-trained model, we could capitalize on its existing specialized knowledge, thus eliminating the need for additional computational expenditure and time that further fine-tuning would entail. This approach also mitigates the risk of overfitting that can come with fine-tuning on a smaller domain-specific dataset.

4.4.2. Fear Classification with Llama 2

Llama 2 [2] is a prominent example of the advancements in Large Language Models (LLMs), equipped to handle a wide range of language processing tasks. Developed by Meta AI, it comes in various sizes spanning from 7 to 65 billion parameters. Utilizing an auto-regressive approach and built upon the transformer decoder architecture, Llama 2 takes sequences of text as input and predicts subsequent tokens through a sliding window mechanism for text generation. In our study, Llama 2 with 7 billion parameters was employed to assess the presence of fear in AI-related news headlines. The model's ability to understand context and nuance was leveraged without further fine-tuning. This approach stems from the model's extensive training on diverse text data, which provides a rich understanding of linguistic patterns and emotional expressions. For the classification task, we utilize a prompt structure to guide the model in evaluating AI-related news headlines. The prompt provides detailed examples of both fear-inducing and non-fear-inducing headlines, aiding the model in making informed classifications. The inclusion of specific instructions within the prompt aims to align the model's generative capabilities with the nuanced objective of detecting fear, worry, anxiety, or concern in the text. These prompts were designed not only to classify the headlines but also to elucidate the rationale behind the model's classification, shedding light on how it interprets the emotional weight of each statement. The setup also involved optimizing the model to efficiently operate with our dataset, ensuring a balance between performance and resource utilization. We apply the BitsAndBytes[86] configuration, enabling 4-bit quantization which significantly reduces memory footprint without compromising model accuracy and enhancing computational efficiency and throughput.

4.4.3. Fear Classification with Mistral

Following our exploration with Llama 2, we extended our methodology to include Mistral 7B v0.1, introduced by Mistral AI. This model, with its 7.3 billion parameters, is crafted to proficiently handle a

vast range of language processing tasks, paralleling the capabilities seen in Llama 2 but with unique architectural enhancements. Just as with Llama 2, Mistral 7B leverages an auto-regressive mechanism and is based on the transformer decoder architecture, enabling it to process sequences of text and predict subsequent tokens effectively. This functionality is essential for our goal of detecting fear in AI-related news headlines, as it relies on the model's inherent capacity to understand context and linguistic nuances without the need for further fine-tuning. For the classification task, we applied the same prompt structure used with Llama 2, to guide Mistral 7B in the evaluation process. This approach was complemented by optimizing Mistral 7B for our dataset, ensuring an optimal balance between computational efficiency and model performance. By employing the BitsAndBytes configuration for 4-bit quantization, we significantly minimized the memory footprint of Mistral 7B, akin to our Llama 2 setup, thereby maintaining model accuracy while enhancing computational throughput.

By adopting a consistent methodology with both Llama 2 and Mistral 7B, our study not only demonstrated the versatile capabilities of these models in emotion detection but also highlighted the nuanced objective of detecting fear in the discourse surrounding artificial intelligence, thereby illustrating the profound potential of LLMs in understanding the complex landscape of human emotions.

4.5. Machine Learning

4.5.1. Data Preparation

In order to perform emotion recognition on our dataset with a focus on identification of inducement of fear of AI, a random subset of 12,000 news headlines, representing around 17.9 % of the total data, were selected for sentiment polarity and fear emotion classification. Polarity and fear classifications of the headlines were identified manually and annotated by human experts. Based on the content, the polarity of the given headline was labeled as positive, negative, or neutral. Similarly, emotion classification was annotated as fear inducing or not, and 1,284 fear inducing news headlines were identified leading to a machine learning dataset of 2,568 AI news headlines. If the AI news headline evoked fear, worry, anxiety, or concern, they were labeled under the "fear" class. Some Headlines were particularly difficult to annotate. For instance, "Can machines invent things without human help?" - this could be classified as not-fear-inducing if treated as a purely hypothetical question, and it could also be classified as fear-inducing since this headline could potentially evoke a sense of unease and unrest, given the possibility of artificially intelligent machines developing risky inventions. Another example of a nuanced headline is "OpenAI aiming to create AI as smart as humans, helped by funds from Microsoft", and though this headline alludes to the creation of a Artificial General Intelligence (AGI), this headline was labeled under "fear" as it raises societal, social, ethical and moral concerns of public consequence. The annotated dataset went through several rounds of review to improve accuracy before it was used to train the machine learning models. Sample headlines and their corresponding sentiment labels from this process can be viewed in Tables 2 and 3.

Table 2. Sentiment classes for sample news headlines from our dataset.

Sentiment	Headline
Positive	4 ways AI can help with climate change, from detecting methane to preventing fires Beneficial AI: Safe, Secure, and Trustworthy Artificial Intelligence for Food Safety How AI can help the education of blind and visually impaired people
Negative	Elon Musk warns AI could cause 'civilization destruction' even as he invests in it Bias in AI is a real problem Stupid Artificial Intelligence
Neutral	3 Artificial Intelligence (AI) Stocks With More Upside Than Nvidia NVIDIA Announces Jetson Platform Expansion for Robotics and Edge Stanford Releases Report on the Current State of AI

Table 3. Emotion class labelling for fear and not-fear for a sample in our dataset.

Emotion	Headlines
Fear	The Dystopia is Here, AI is Taking over Data Science Jobs in 2021
	AI doomsday warnings a distraction from the danger it already poses, warns expert
	The Godfather of Artificial Intelligence warns of a dark future
Absence of Fear	Artificial intelligence could help doctors predict breast cancer risks
	How AI is helping to save the Amazon - Positive News
	AI Predicts Future Pancreatic Cancer Harvard Medical School

4.5.2. Problem Statement

Consider a dataset $H = \{h_1, h_2, \dots, h_N\}$ of N labeled news headlines, each tagged as either inducing fear or not, where N represents the total number of headlines in the dataset. Our objective is to develop a fear detection model $f(H) \rightarrow \hat{y}$, aimed at maximizing the accuracy of identifying headlines that evoke fear. This task is approached as a binary classification problem, where the model predicts a headline as fear-inducing ($\hat{y} = 1$) or not fear-inducing ($\hat{y} = 0$), thereby enabling the nuanced understanding and categorization of news content based on the emotional response it is likely to elicit in readers.

4.5.3. Modeling Process

Our study utilized a manually annotated dataset of 2,568 news headlines, carefully labeled to indicate the presence or absence of fear-inducing content, with an equal number of headlines in each class for balanced representation. The dataset was divided into a training set containing 1,797 headlines and a testing set with 771 headlines, following a 70:30 split. We employed various binary classification algorithms, including Logistic Regression, Support Vector Classifier (SVC), Gaussian Naive Bayes, Random Forest, and XGBoost to evaluate model effectiveness in detecting fear in news headlines.[87–92] Feature extraction from the headlines was performed using Sentence-BERT (SBERT)[93], a variant of the pre-trained BERT model designed to generate semantically meaningful sentence embeddings. SBERT captures contextual relationships within text, making it well-suited for our task, given its ability to generate embeddings that represent emotional cues such as those of fear. These embeddings served as input features for all the classification models. To address the high dimensionality of the feature space and to improve computational efficiency, we applied Principal Component Analysis (PCA) for dimensionality reduction on the training dataset. PCA transforms the data into a new coordinate system, reducing the number of total variables while retaining those that contribute most to the variance. This step was crucial for enhancing model performance by mitigating the curse of dimensionality and focusing on the most informative aspects of the data. For the optimization of model hyperparameters, we utilized Bayesian Optimization, a probabilistic model-based optimization technique. This approach leverages the Bayesian inference framework to model the objective function and identifies the optimal hyperparameters by exploring the parameter space, thus significantly improving the models' predictive accuracy and efficiency.

4.5.4. Logistic Regression

We commenced with the Logistic Regression algorithm for our binary classification task. This choice was motivated by Logistic Regression's simplicity, computational efficiency, and its capability to yield highly interpretable models. The optimal hyperparameters for our Logistic Regression model were determined to be solver='saga' and C=4.46. The selection of the solver parameter is critical, as 'saga' is particularly adept at handling the challenges posed by high-dimensional datasets such as ours, providing robust support for both L1 and L2 regularization. This capability is essential for effective feature selection. The regularization strength, denoted by C, was finely tuned to enhance the model's generalization capabilities, striking the right balance between avoiding overfitting and maintaining sensitivity to the underlying patterns in the data. With this optimized configuration, our Logistic

Regression model exhibited strong performance and generalization, making it a robust foundation for our fear detection task in news headlines.

4.5.5. Support Vector Classifier and Gaussian Naive Bayes

Following Logistic Regression, we further experimented with Support Vector Classifier (SVC) and Gaussian Naive Bayes (GaussianNB) algorithms. SVC is valued for its adaptability and efficacy in handling the complex, high-dimensional spaces typical of text data, while Gaussian Naive Bayes, appreciated for its straightforwardness and rapid processing, was also tested. Our optimal SVC model was finely tuned to a C value of 214.3, with gamma set to 0.0005, and utilized a linear kernel. These parameters, chosen via Bayesian optimization, were carefully calibrated to improve the model's capacity to accurately separate classes without sacrificing the ability to generalize. The linear kernel selection aimed to exploit the textual data's intrinsic high dimensionality without unnecessarily complicating the model. For GaussianNB, the variance smoothing parameter was set to 0.1, a decision aimed at bolstering model robustness by fine-tuning the variance in feature likelihood, thereby enhancing prediction consistency across diverse datasets.

4.5.6. Bagging and Boosting

Concluding our exploration of machine learning algorithms for binary text classification, we delved into ensemble methods, specifically focusing on bagging and boosting algorithms to leverage their strengths in our task. For bagging, we employed the Random Forest (RF) algorithm, renowned for its ability to reduce overfitting while maintaining accuracy by aggregating the predictions of numerous decision trees. The most optimal RF model had max-depth=16 and n-estimators=308. For boosting, we selected XGBoost, due to its efficient handling of sparse data and its capability to minimize errors sequentially with each new tree. Our configuration for XGBoost included max-depth=10, and n-estimators=400, among other parameters. Notably, the evaluation metric hyperparameters was set to 'logloss', which focuses the model on improving probability estimates, crucial for the binary classification of text. By incorporating these ensemble techniques, we aimed to harness the collective strength of multiple models, mitigating individual weaknesses and leveraging the diversity of their predictions.

5. Results

5.1. Sentiment Analysis

In our analysis of AI news headlines, as illustrated in Tables 4 and 5, we observe that a substantial 39.74% of headlines convey negative sentiments, potentially indicating fear and apprehension that permeates the public discourse on AI. This marked presence of negative sentiment, particularly in a field as inherently progressive as AI, points to concerns, worries, fears and hyped up risk perceptions concerning AI's encroachment into various aspects of human existence. Further exploration of sentiment analysis through VADER, AFINN, and TextBlob (as detailed in Table 4) offers deeper insights into the sentiment landscape. The mean sentiment scores from VADER (0.175) and AFINN (0.346) indicate a mild positive bias in AI-related headlines. In contrast, TextBlob's mean score (-0.040) highlights a slightly negative sentiment, mildly pointing to the presence of caution and skepticism. The wide range of sentiment scores (from strongly negative to strongly positive) across all tools underscores the polarized nature of AI news headlines, reflecting diverse perspectives ranging from enthusiasm to uncertainty and fear.

Table 4. Summary of VADER, AFINN, and TextBlob methods.

	VADER	AFINN	TextBlob
Mean	0.175	0.346	-0.040
Max	0.949	14	1
Min	-0.944	-11	-1
Variance	0.110	2.640	0.083

Table 5. FLAIR model sentiment analysis

Sentiment	Count	Percentage
Positive	40425	60.25
Negative	26666	39.74

5.2. Results—Topic Modeling with BERTopic

In our study, we applied BERTopic [76] to analyze 67,091 headlines related to AI news. This analysis, enhanced by the c-TF-IDF method for keyword extraction and further enriched through the Nocodefunctions App for semantic network creation, enabled us to visualize and understand the relationships between key concepts across the data. This network (Refer Figure 1 for the semantic network for the most discussed topic covered by news headlines) served as a graphical elucidation of the central and peripheral ideas within each topic, providing an at-a-glance understanding of the thematic structures. Our findings showed that the topic with the largest number of associated headlines was "AI Evolution and Ethical Dilemmas." This topic underscores the growing debate and concern over the swift progress in AI technology and its implications for society and ethics. It highlights the potential risks and the future direction of AI's integration with humanity, mirroring the cautionary stance advocated by many experts in the field.

AI in Education and Workplace topic delves into the nuanced discussions around AI's impact on employment and the educational sector. It reflects concerns and optimism about AI replacing jobs, particularly focusing on the potential shift in the job market that could affect both blue-collar and white-collar professions. The conversation also extends to the educational realm, where there's a push for schools and universities to adapt to this AI-driven future by preparing students with the necessary skills and knowledge. This topic encapsulates the dual narrative of AI as both a disruptor of traditional job roles and an enabler of new opportunities, urging a reevaluation of job skills and educational curricula in anticipation of future demands.

The adoption of AI in Warfare and Security indicates a significant shift towards more sophisticated and intelligent defense mechanisms, whereas its application in Media and Cybersecurity is revolutionizing content creation and bolstering cybersecurity measures. Other topics include open source AI, AI and creativity, AI education, AI and society, AI and art, AI ethics and bias, AI investments and the stock market, intellectual property in AI, legal complexities, and the evolving discourse on ownership and rights in the age of AI innovations.

5.3. Results—Topic Modeling with Llama 2

In conducting our analysis on the extensive range of topics within the domain of AI, we employed the advanced capabilities of Llama 2 [2] for topic modeling. Through this process, Llama 2 [2] identified 189 unique topics, which we then meticulously reviewed to categorize them based on their primary focus and implications. These topics span across various aspects and implications of AI technology, reflecting both its diverse applications and the multifaceted debates surrounding its development, deployment, and regulation.

Given the complexity and the overlapping nature of some topics, we established a set of criteria to ensure each topic was placed in a single category, thus preventing any duplication. These criteria were based on the predominant theme of each topic, considering aspects such as the application domain of AI (e.g., healthcare, education, military), the nature of the discussion (e.g., risks, advancements, ethical

considerations), and the primary stakeholders involved (e.g., businesses, governments, the general public). Following our analysis, we categorized the 189 topics into five primary categories.. Below are examples of topics categorized under each of these headings, reflecting the breadth and depth of AI's impact:

- **AI's Impending Dangers** : This category includes topics that highlight potential risks and challenges posed by AI technology. Examples include "Emerging risks of AI technology," "Generative AI Risks in 2023," and "AI-generated child sexual abuse content."
- **AI Advancements**: Topics under this category are primarily focused on providing insights, explanations, and informative perspectives on various aspects of AI. Examples include "AI technology competition," "Impact of Artificial Intelligence on Education," and "Advancements in AI-assisted image and video editing."
- **Negative Capabilities of AI**: This category encompasses topics that discuss the adverse impacts or capabilities of AI and ChatGPT, shedding light on concerns such as bias, discrimination, and privacy issues. Examples are more nuanced in this category but could include discussions around "Bias and discrimination in AI systems" and "Privacy and security risks in AI-driven data management."
- **Positive Capabilities of AI**: Conversely, this category highlights the beneficial aspects and positive applications of AI and ChatGPT. Topics such as "Artificial Intelligence in Healthcare," "AI in Music Industry," and "Using AI to combat wildfires in California" exemplify the positive impact AI can have across different sectors.
- **Experimental Reporting for AI**: For the category of Experimental Reporting, this encompasses cutting-edge explorations and innovative uses of AI that are at the forefront of technology and research. Examples include the application of AI in predicting natural disasters with greater accuracy and timeliness, such as using machine learning algorithms to forecast earthquakes or volcanic eruptions. Another example is the use of AI in environmental conservation, like deploying AI-driven drones for monitoring wildlife populations or analyzing satellite imagery to track deforestation. Additionally, experimental applications in digital biology and genome editing highlight AI's role in advancing medical science, such as using AI to decipher complex genetic codes or to personalize medicine by predicting an individual's response to certain treatments.

This approach allowed us to highlight the multifaceted nature of AI, encompassing both its challenges and opportunities, and to contribute to a more nuanced understanding of its role in shaping the future.

5.4. Results—Fear Classification with LLMs

5.4.1. Fear Classification with DistilBERT

Upon application of the Falconsai/fear_mongering_detection DistilBERT model to our dataset of 2,568 headlines, the results, as shown in Table 6, revealed a precision of approximately 0.639 for Class 0 (Not Fear-Mongering) and 0.692 for Class 1 (Fear-Mongering). The recall scores were 0.741 for Class 0 and 0.581 for Class 1, leading to F1-scores of 0.686 and 0.632, respectively. The overall accuracy of the model stood at 0.661. Although the precision for Class 1 is moderately high, indicating a good measure of correctly identified fear-mongering headlines, the lower recall suggests that the model missed a significant proportion of fear-inducing content. Conversely, for Class 0, the higher recall indicates a better identification rate of non-fear-inducing headlines.

Table 6. Evaluation metrics for LLMs for fear classification.

Algorithm	Accuracy (%)	Precision	Recall	F1 Score
DistilBERT	66	0.69	0.58	0.63
LLama 2	82.5	0.93	0.71	0.80
Mistral	81.1	0.96	0.63	0.77

Table 7 details examples where the DistilBERT model correctly classified headlines, demonstrating its capability to discern the tone and content effectively in many instances. Table 8, however, illustrates instances of misclassifications, suggesting areas where the model could benefit from further data-specific training to improve its understanding of context and nuance in news headlines. These findings underscore the potential of using pre-trained models like DistilBERT for initial screening in fear detection tasks. However, they also indicate the necessity for fine-tuning on specialized datasets to fully adapt to the nuances of the task at hand.

Table 7. Correctly classified headlines by DistilBERT model.

Headline	Actual	Predicted
How generative AI is boosting the spread of disinformation and propaganda	1	Fear-Mongering
Will definitely replace me': Americans fear artificial intelligence will steal their jobs	1	Fear-Mongering
5 EU FinTechs Using AI to Support Consumers, Businesses	0	Non-Fear-Mongering
Meta's new learning algorithm can teach AI to multi-task	0	Non-Fear-Mongering

Table 8. Misclassified headlines by DistilBERT model.

Headline	Actual	Predicted
Artificial Intelligence Begins to Enter the World of Stock Investment	0	Fear-Mongering
Generative AI-nxiety	1	Non-Fear-Mongering

5.4.2. Fear Classification with Llama 2

Although DistilBERT was specifically trained for the task of detecting fear-mongering content, it exhibits lower performance metrics across all categories when compared to Llama 2. Llama 2 significantly outperforms DistilBERT with an accuracy of 82.5%, alongside higher precision (0.93), recall (0.71), and F1 score (0.80). These results indicate that Llama 2, even without specific fine-tuning, has a superior ability to classify fear-inducing headlines more accurately. The substantial difference in performance between the two models may also reflect Llama 2's advanced understanding of context and emotional expressions, attributable to its larger model size and more sophisticated architecture. Further analysis of the classification performance for each class reveals:

- **Class 0 (Not Fear-Inducing):** Llama 2's performance in identifying headlines that do not induce fear is marked by a precision of 0.7627, suggesting that when it classifies a headline as not fear-inducing, it is correct around 76% of the time. The recall rate of 0.9426 indicates that the model is highly effective, identifying approximately 94% of all not fear-inducing headlines in the dataset. The combination of these metrics leads to an F1 score of 0.8432, reflecting a strong balance between precision and recall for this class. This high recall rate is particularly significant, as it demonstrates Llama 2's ability to conservatively identify content that is unlikely to cause fear, ensuring a cautious approach in marking headlines as fear-inducing.
- **Class 1 (Fear-Inducing):** For headlines that are classified as fear-inducing, Llama 2 shows a precision of 0.9261, meaning that it has a high likelihood of correctly identifying genuine instances of fear-inducing content. The recall of 0.7102, though lower than for Class 0, signifies that the model successfully captures a substantial proportion of fear-inducing headlines. However, there is room for improvement in recognizing every such instance within the dataset. The resulting F1 score of 0.8039 for Class 1 illustrates a robust performance, although the challenge remains to enhance recall without sacrificing the model's high precision.

The disparity in Llama 2's performance between Class 0 and Class 1 highlights its cautious yet effective method in fear detection. The model proficiently reduces false positives by accurately identifying content that does not incite fear. However, enhancing the model's ability to capture all instances of fear-inducing headlines (improving recall) without compromising the precision of its predictions remains crucial for advancing its application in emotion classification within AI news contexts.

5.4.3. Fear Classification with Mistral

Mistral showcases a noteworthy performance with an accuracy of 81.1%, exhibiting the highest precision (0.96) among the three models. Its precision rates for Class 0 (Not Fear-Inducing) and Class 1 (Fear-Inducing) are 0.7364 and 0.9673, respectively, demonstrating exceptional accuracy in identifying fear-inducing content. However, Mistral's recall scores—0.9791 for Class 0 and 0.6382 for Class 1—indicate a strong ability to correctly identify non-fear-inducing headlines but a moderate capacity to capture all fear-inducing headlines. This results in F1 scores of 0.8406 for Class 0 and 0.769 for Class 1, reflecting a balance between precision and recall, especially in distinguishing non-fear-inducing content. This highlights Mistral's strength in precision, particularly in identifying fear-inducing headlines with a high degree of accuracy. However, optimizing recall, particularly for fear-inducing content, remains a critical area for improvement to enhance its utility for fear classification tasks.

5.5. Results—Fear Classification with Machine Learning

In evaluating the performance of our algorithms for fear detection in news headlines, we employed a variety of metrics commonly used in classification tasks. This approach allows us to assess the effectiveness of our models in accurately identifying headlines that either evoke fear or do not. True Positives occur when our algorithms correctly identify a headline as fear-inducing, showcasing the model's effectiveness in recognizing fear accurately. True Negatives are instances where headlines are rightly identified as not fear-inducing, reflecting the model's ability to avoid false alarms. On the other hand, False Negatives represent a miss by the model, where it fails to flag a fear-inducing headline. Lastly, False Positives happen when the algorithm mistakenly tags a non-fear-inducing headline as fear-inducing.

Given these definitions, we can express the following metrics to quantify the performance of our models:

- **Precision** is the ratio of correctly predicted fear-inducing headlines to the total predicted as fear-inducing. It is calculated as:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall** (or Sensitivity) measures the proportion of actual fear-inducing headlines that are correctly identified. It is calculated as:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1 Score** provides a balance between Precision and Recall, offering a single metric to assess the model's accuracy. It is especially useful in cases where the class distribution is imbalanced. The F1 Score is calculated as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Accuracy** measures the overall correctness of the model, calculated by dividing the sum of true positives and true negatives by the total number of cases:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

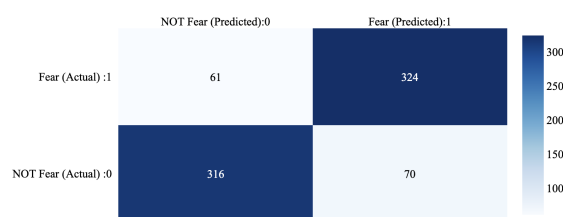
These metrics provide a comprehensive framework to evaluate and compare the performance of our fear detection models, ensuring we accurately capture the nuances of fear-inducing vs. non-fear-inducing classifications in news headlines.

The performance of each classifier in our fear detection task is summarized in Table 9. The classifiers were evaluated based on their accuracy in distinguishing fear-inducing from non-fear-inducing news headlines. Logistic Regression and XGBoost yielded the highest accuracies, 83.0% and 83.3%, respectively, which suggests their potential suitability for the task at hand. The accuracies for SVC and Random Forest, at 81.4% and 81.1%, while lower, still indicate that these models have a degree of effectiveness in the classification of fear-inducing content. Gaussian NB, with an accuracy of 80.5%, performed reasonably well, considering its simplicity and the speed with which it can be implemented. These outcomes suggest that while no single model outperformed the others dramatically, some models may be more adept at capturing the subtleties inherent in textual data related to fear. The findings point to the possibility that ensemble methods such as XGBoost, as well as Logistic Regression, could be favorable choices for text classification. Further investigations could aim to refine these models or potentially combine their strengths to improve the detection of fear in news headlines.

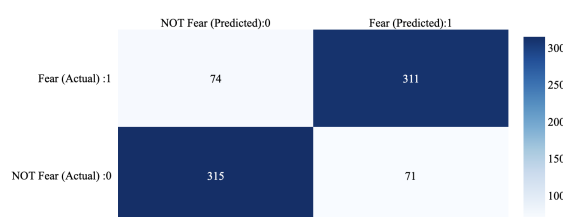
Table 9. Evaluation metrics for Machine Learning classifiers for fear classification

Algorithm	Accuracy (%)	Precision	Recall	F1 Score
Logistic Regression	83.0	0.841	0.822	0.832
SVC	81.4	0.809	0.823	0.816
Gaussian NB	80.5	0.799	0.816	0.807
Random Forest	81.1	0.814	0.808	0.811
XGBoost	83.3	0.823	0.797	0.810

In analyzing the confusion matrices for both Logistic Regression and XGBoost, as depicted in Figure 15 respectively, we obtain a deeper understanding of each model's performance beyond mere accuracy. For Logistic Regression, the confusion matrix reveals that out of the total fear-inducing headlines, 324 were correctly identified (True Positives), while 61 were mistakenly labeled as non-fear-inducing (False Negatives). Conversely, for non-fear headlines, 316 were accurately classified as non-fear-inducing (True Negatives), and 70 were incorrectly identified as fear-inducing (False Positives). This indicates a robust ability to detect fear-inducing headlines while also pointing to areas where the model may benefit from further refinement to reduce misclassification. XGBoost's confusion matrix shows a similar trend with 311 fear-inducing headlines correctly identified (True Positives) and 74 misclassified as non-fear-inducing (False Negatives). In terms of non-fear headlines, 315 were correctly classified (True Negatives), and 71 were wrongly labeled as fear-inducing (False Positives). The slightly higher number of False Negatives compared to Logistic Regression suggests that while XGBoost is generally effective, there might be a need to adjust the model to improve its sensitivity. Both matrices are critical for understanding the trade-offs between sensitivity (Recall) and precision in our models. The proportion of False Positives to False Negatives gives us insight into the type of errors each model is more prone to making, which is valuable information for further tuning and application in real-world scenarios where the cost of different types of errors may not be equal.



(a) Confusion Matrix for Logistic Regression



(b) Confusion Matrix for XG boost

Figure 15. Confusion matrices

Table 10 presents the precision, recall, and F1-score for the optimized Logistic Regression and XGBoost models across both classes of our dataset. For class 0, which represents news headlines that do not induce fear, the Logistic Regression model achieved a precision of 0.84 and a recall of 0.82, resulting in an F1-score of 0.83. The XGBoost model exhibited a slightly lower precision at 0.79 but a higher recall of 0.91, culminating in an F1-score of 0.84 for the same class. This indicates that while the Logistic Regression model is slightly more precise in identifying non-fear-inducing headlines, the XGBoost model is more sensitive, recognizing a greater proportion of the actual non-fear-inducing headlines. In class 1, which comprises fear-inducing headlines, the Logistic Regression model reported a precision of 0.82 and a recall of 0.84, with an F1-score of 0.83. In contrast, the XGBoost model achieved a higher precision of 0.89 but a lower recall of 0.76, leading to an F1-score of 0.82. These results suggest that the Logistic Regression model is more balanced in terms of precision and recall for detecting fear-inducing headlines, whereas the XGBoost model is more precise but less sensitive in identifying fear within the dataset.

Table 10. Class-wise performance metrics for top 2 ML classifiers.

Model	Class	Precision	Recall	F1-Score
Logistic Regression	0	0.84	0.82	0.83
XGBoost	0	0.79	0.91	0.84
Logistic Regression	1	0.82	0.84	0.83
XGBoost	1	0.89	0.76	0.82

6. Discussion and Future Research

Our research addresses the critical area of the framing of AI news headlines for public consumption perception, and highlights the unusual emphasis on fear inducing news headlines by mass media outlets and other sources. Our analysis indicated a significant and potentially unjustified presence of negativity towards AI as indicated by the sentiment analysis using multiple libraries, displayed in Table 4. Furthermore, our EDA revealed the presence of numerous terms and phrases commonly used in news headlines to promote fear and AI phobia which include words and phrases such as 'warns', 'risk', 'bias', 'concern', 'ai ethics', 'ai safety', 'ai regulation', 'responsible ai', 'ai arms race',

We also intend to expand the range of classifier labels to detect other emotions such as mistrust, anger, sadness, trust, confusion, surprise, and joy. Further, we plan to deepen our analyses by examining the subsections within our dataset more closely. To improve the performance of our LLMs, we will fine-tune them and adapt them to our specific use case, employing prompt engineering to enhance model accuracy. An innovative yet limited opportunity lies in analyzing handwritten notes on AI to contrast highly digitized populations with those still reliant on traditional methods. Improved handwriting recognition by AI methods offers a unique research avenue. [95]. Another future research direction involves classifying fear-inducing news articles into justified concerns versus hyped fears that involves the classification of fear inducing news articles into valid items where there is a reasonable basis to warn the public of impending challenges with AI technologies or specific applications such as the proliferation of AI generated deepfakes. This needs to be contrasted with unjustified, exaggerated and hyped fear inducing news headlines and content that sensationalize trivial issues and magnify isolated comments and rare events to create and spread mass AI phobia. Such a classification framework or model, when developed with sufficient accuracy, could be used as an online application to classify and flag the quality and news articles on AI. Finally, it is necessary to encourage AI news reporting in local languages and even those in English could be contextualized to local sensitivities using appropriate linguistic features, as extant research has demonstrated the importance of cultural sensitivity in AI education and AI related communications [96].

7. Recommendations

Based on our analysis we make a few recommendations to improve the counterproductive scenario of negative and fear inducing AI news headlines. It is necessary to develop an ontology on the threats and risks of AI that is relevant to public and news media usage. This will help provide a framework for a more informed news coverage, that will help avoid throwing all of AI under the "fear" factor, for example: Separate AI as science versus AI as applications, contrast use of AI versus abuse of AI and contrast Accidental versus Deliberate harms. Another recommendation is to move towards achieving aggregated honesty - for example, declare a year or prolonged factual reporting on AI. This strategy could involve media outlets collaborating with pragmatic AI experts to provide depth and context and debunk some of the AI fear myths and narratives. This will help balance the unusual amount and extent of negative and AI phobia inducing reporting that has happened over the past few years. Furthermore, it would be worthwhile exploring how news media can be leveraged to facilitate widespread structured knowledge of fundamental AI concepts - this could serve as an avenue for a broad public educational framework on AI concepts and help develop symbiotic and enhanced collaboration between news media, AI experts and AI in academia. This could also lead to news media supported micro-credentials mechanisms and accelerated partnerships and collaborations with AI researchers - colleges and Universities to co-develop reliable and informative AI news content. On a more general level, there are systemic issues such as the need to democratize scientific resources for AI to facilitate public access and use of critical AI infrastructure such as supercomputing resources [11]. It may also be useful for news agencies to provide reporters with training on AI on a voluntary basis. Adopting the spirit of these recommendations will not only improve the quality of AI news coverage but also contribute significantly to building a well-informed society that can engage with advancements in AI critically and constructively.

8. Conclusion

"Ignorance is the parent of fear." Herman Melville

There is a compelling need for AI education, and much of the adverse swaying of public perception and consequential behavior on AI by erroneous AI news headlines, exaggerated content and sensationalism is rooted in the absence of widespread pragmatic education on AI. One of the most urgent and important policies that governments worldwide need to implement is to initiate and facilitate mass public education on the fundamental scientific nature and ontological characteristics

of AI to people of all possible ages and diverse backgrounds. This will equip people to have an objective knowledge based foundational perspective on AI, which will serve as a measure of immunity against the spread of viral negative and fear inducing AI news headlines and content. Furthermore, the constructs need to be clarified: AI is a science like any other science such as mathematics, and AI applications are technology-business. People who use or misuse AI applications need to be held accountable and it is a logical fallacy to mix these constructs. Such a logical fallacy combined with mass illiteracy on AI concepts can have serious consequences for human society. It is therefore critical for news media to assume the responsibility for systematic (individual articles and sources) and aggregate effects (collective societal impact over time) of their reporting. AI phobia inducing news artifacts need to be addressed but not controlled - they can be labelled or tagged using NLP based classification applications. However, comprehensive control could lead to necessary news on the dangers of AI being blocked. AI phobia is not produced by news articles alone, and extant research has highlighted the role of absence of transparency, bias and ethics challenges and socioeconomic systems that promote the oppressive use of AI and a certain level of caution over the potential risks and misuse of AI is justified [10,97]. Furthermore, as a mechanism to both inform the public as well as to serve as an evaluation and performance support mechanism, universities and community organizations could use AI applications trained to detect classes (fear-inducing, fair, etc.) of AI news as conceptually demonstrated in this paper - such an online application could classify and flag the quality of news articles on AI, and even adaptively personalize such information [98]. This research is a call for balance and honesty in AI news reporting and not an exercise aimed at discrediting all reporting on the dangers of AI. Genuine reporting on the factual risks of AI must be encouraged, and sensationalism, along with AI-fear mongering, needs to be avoided.

[Notes: All authors contributed fairly, More details TBL in final version, RAISE URL: <https://sites.rutgers.edu/raise/>]

References

1. Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al. Mistral 7b. *arXiv preprint arXiv:2310.06825*, 2023.
2. Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023.
3. Jim Samuel, Abhishek Tripathi, and Ensela Mema. A new era of artificial intelligence begins—where will it lead us? *Journal of Big Data and Artificial Intelligence*, 2(1), 2024.
4. VASILE GHERHES. Why are we afraid of artificial intelligence (ai). *European Review Of Applied Sociology*, 11 (17):6–15, 2018.
5. Jungmin Kim. Fear of artificial intelligence on people’s attitudinal & behavioral attributes: An exploratory analysis of ai phobia. *GSJ*, 7(10), 2019.
6. David L Altheide. The news media, the problem frame, and the production of fear. *The sociological quarterly*, 38(4):647–668, 1997.
7. David L Altheide. *Creating fear: News and the construction of crisis*. Routledge, 2018.
8. Frank Furedi. *Culture of fear revisited*. A&C Black, 2006.
9. Jim Samuel. A call for proactive policies for informatics and artificial intelligence technologies. *Scholars Strategy Network*, 2021.
10. Jim Samuel. The critical need for transparency and regulation amidst the rise of powerful artificial intelligence models. *Scholars Strategy Network (SSN) Key Findings*. Accessed November, 27:2023, 2023.
11. Jim Samuel, Margaret Brennan-Tonetta, Yana Samuel, Pradeep Subedi, and Jack Smith. Strategies for democratization of supercomputing: Availability, accessibility and usability of high performance computing for education and practice of big data analytics. *Journal of Big Data Theory and Practice*, 2021.
12. Katarzyna Molek-Kozakowska. Towards a pragma-linguistic framework for the study of sensationalism in news headlines. *Discourse & Communication*, 7(2):173–197, 2013.

13. Hye-Jin Paek, Sang-Hwa Oh, and Thomas Hove. How fear-arousing news messages affect risk perceptions and intention to talk about risk. *Health communication*, 31(9):1051–1062, 2016.
14. Rachel C Adams, Petroc Sumner, Solveiga Vivian-Griffiths, Amy Barrington, Andrew Williams, Jacky Boivin, Christopher D Chambers, and Lewis Bott. How readers understand causal and correlational expressions used in news headlines. *Journal of experimental psychology: applied*, 23(1):1, 2017.
15. Celine Klemm, Tilo Hartmann, and Enny Das. Fear-mongering or fact-driven? illuminating the interplay of objective risk and emotion-evoking form in the response to epidemic news. *Health communication*, 34(1): 74–83, 2019.
16. Valerie Hase and Katherine M Engelke. Emotions in crisis coverage: How uk news media used fear appeals to report on the coronavirus crisis. *Journalism and Media*, 3(4):633–649, 2022.
17. Sushree Panigrahi and Jeet Singh. Deadly combination of fake news and social media. *Rajiv Gandhi Institute for Contemporary Studies*, 4, 2017.
18. Yuwei Chuai and Jichang Zhao. Anger makes fake news viral online. *arXiv preprint arXiv:2004.10399*, 2020.
19. Natascha de Hoog and Peter Verboon. Is the news making us unhappy? the influence of daily news exposure on emotional states. *British Journal of Psychology*, 111(2):157–173, 2020.
20. Stuart WG Derbyshire. Culture of fear: Risk taking and the morality of low expectation. *BMJ*, 315(7111):823, 1997.
21. Junesoo Lee and Jaehyuk Park. Ai as “another i”: Journey map of working with artificial intelligence from ai-phobia to ai-preparedness. *Organizational Dynamics*, 52(3):100994, 2023.
22. Lucía Martín Holguín. Communicating artificial intelligence through newspapers: Where is the real danger, 2018.
23. Hannes Cools, Baldwin Van Gorp, and Michael Opgenhaffen. Where exactly between utopia and dystopia? a framing analysis of ai and automation in us newspapers. *Journalism*, 25(1):3–21, 2024.
24. Dennis Nguyen and Erik Hekman. The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation. *AI & SOCIETY*, pages 1–15, 2022.
25. Shaojing Sun, Yujia Zhai, Bin Shen, and Yibei Chen. Newspaper coverage of artificial intelligence: A perspective of emerging technologies. *Telematics and Informatics*, 53:101433, 2020.
26. Ching-Hua Chuan, Wan-Hsiu Sunny Tsai, and Su Yeon Cho. Framing artificial intelligence in american newspapers. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 339–344, 2019.
27. Albert C Gunther. The persuasive press inference: Effects of mass media on perceived public opinion. *Communication Research*, 25(5):486–504, 1998.
28. Diana C Mutz and Joe Soss. Reading public opinion: The influence of news coverage on perceptions of public sentiment. *Public Opinion Quarterly*, pages 431–451, 1997.
29. David Giles and Rachel L Shaw. The psychology of news influence and the development of media framing analysis. *Social and personality psychology compass*, 3(4):375–393, 2009.
30. Jesper Strömbäck, Yariv Tsfati, Hajo Boomgaarden, Alyt Damstra, Elina Lindgren, Rens Vliegenthart, and Torun Lindholm. News media trust and its impact on media use: Toward a framework for future research. *Annals of the International Communication Association*, 44(2):139–156, 2020.
31. Wen Yang, Dongtong Lin, and Zelong Yi. Impacts of the mass media effect on investor sentiment. *Finance Research Letters*, 22:1–4, 2017.
32. Qing Li, Tiejun Wang, Ping Li, Ling Liu, Qixu Gong, and Yuanzhu Chen. The effect of news and public mood on stock movements. *Information Sciences*, 278:826–840, 2014.
33. Md Mokhlesur Rahman, GG Md Nawaz Ali, Xue Jun Li, Jim Samuel, Kamal Chandra Paul, Peter HJ Chong, and Michael Yakubov. Socioeconomic factors analysis for covid-19 us reopening sentiment with twitter and census data. *Heliyon*, 7(2):e06200, 2021.
34. GG Ali, Md Mokhlesur Rahman, Amjad Hossain, Shahinoor Rahman, Kamal Chandra Paul, Jean-Claude Thill, Jim Samuel, et al. Public perceptions about covid-19 vaccines: Policy implications from us spatiotemporal sentiment analytics. *Available at SSRN 3849138*, 2021.
35. Myles D Garvey, Jim Samuel, and Alexander Pelaez. Would you please like my tweet?! an artificially intelligent, generative probabilistic, and econometric based system design for popularity-driven tweet content generation. *Decision Support Systems*, 144:113497, 2021.
36. Jim Samuel, GG Ali, Md Rahman, Ek Esawi, Yana Samuel, et al. Covid-19 public sentiment insights and machine learning for tweets classification. *Information*, 11(6):314, 2020.

37. Jim Samuel, Md Mokhlesur Rahman, GG Md Nawaz Ali, Yana Samuel, Alexander Pelaez, Peter Han Joo Chong, and Michael Yakubov. Feeling positive about reopening? new normal scenarios from covid-19 us reopen sentiment analytics. *IEEE Access*, 8:142173–142190, 2020.
38. Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307, 2011.
39. Antony Samuels and John Mcgonical. News sentiment analysis. *arXiv preprint arXiv:2007.02238*, 2020.
40. Yanghui Rao, Jingsheng Lei, Liu Wenyin, Qing Li, and Mingliang Chen. Building emotional dictionary for sentiment analysis of online news. *World Wide Web*, 17:723–742, 2014.
41. Lars Buitinck and et al. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.
42. Rajkumar S Jagdale, Vishal S Shirsat, and Sachin N Deshmukh. Sentiment analysis on product reviews using machine learning techniques. In *Cognitive Informatics and Soft Computing: Proceeding of CISC 2017*, pages 639–647. Springer, 2019.
43. MS Neethu and R Rajasree. Sentiment analysis in twitter using machine learning techniques. In *2013 fourth international conference on computing, communications and networking technologies (ICCCNT)*, pages 1–5. IEEE, 2013.
44. Anurag Singh and Goonjan Jain. Sentiment analysis of news headlines using simple transformers. In *2021 Asian Conference on Innovation in Technology (ASIANCON)*, pages 1–6. IEEE, 2021.
45. David Rozado, Ruth Hughes, and Jamin Halberstadt. Longitudinal analysis of sentiment and emotion in news media headlines using automated labelling with transformer language models. *Plos one*, 17(10): e0276367, 2022.
46. László Nemes and Attila Kiss. Prediction of stock values changes using sentiment analysis of stock news headlines. *Journal of Information and Telecommunication*, 5(3):375–394, 2021.
47. Faheem Aslam, Tahir Mumtaz Awan, Jabir Hussain Syed, Aisha Kashif, and Mahwish Parveen. Sentiments and emotions evoked by news headlines of coronavirus disease (covid-19) outbreak. *Humanities and Social Sciences Communications*, 7(1), 2020.
48. Wafaa S El-Kassas, Cherif R Salama, Ahmed A Rafea, and Hoda K Mohamed. Automatic text summarization: A comprehensive survey. *Expert systems with applications*, 165:113679, 2021.
49. Dima Suleiman and Arafat A Awajan. Deep learning based extractive text summarization: approaches, datasets and evaluation measures. In *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pages 204–210. IEEE, 2019.
50. Robiert Sepúlveda-Torres, Marta Vicente, Estela Saquete, Elena Lloret, and Manuel Palomar. Headlines-tancechecker: Exploiting summarization to detect headline disinformation. *Journal of Web Semantics*, 71: 100660, 2021.
51. Ike Vayansky and Sathish AP Kumar. A review of topic modeling methods. *Information Systems*, 94:101582, 2020.
52. Zara Nasar, Syed Waqar Jaffry, and Muhammad Kamran Malik. Named entity recognition and relation extraction: State-of-the-art. *ACM Computing Surveys (CSUR)*, 54(1):1–39, 2021.
53. Jim Samuel, Gavin Rozzi, and Ratnakar Palle. The dark side of sentiment analysis: An exploratory review using lexicons, dictionaries, and a statistical monkey and chimp. *Dictionaries, and a Statistical Monkey and Chimp (January 6, 2022)*, 2022.
54. Nourah Alswaidan and Mohamed El Bachir Menai. A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, 62:2937–2987, 2020.
55. Huimin Xu, Man Lan, and Yuanbin Wu. Ecnu at semeval-2018 task 1: Emotion intensity prediction using effective features and machine learning models. In *Proceedings of the 12th international workshop on semantic evaluation*, pages 231–235, 2018.
56. Lovejit Singh, Sarbjeet Singh, and Naveen Aggarwal. Two-stage text feature selection method for human emotion recognition. In *Proceedings of 2nd International Conference on Communication, Computing and Networking: ICCCN 2018, NITTR Chandigarh, India*, pages 531–538. Springer, 2018.
57. Zarrin Tasnim Sworna, Zahra Mousavi, and Muhammad Ali Babar. Nlp methods in host-based intrusion detection systems: A systematic review and future directions. *Journal of Network and Computer Applications*, page 103761, 2023.

58. Basemah Alshemali and Jugal Kalita. Improving the reliability of deep neural networks in nlp: A review. *Knowledge-Based Systems*, 191:105210, 2020.
59. Kian Long Tan, Chin Poo Lee, Kalaiarasi Sonai Muthu Anbananthen, and Kian Ming Lim. Roberta-ilstm: a hybrid model for sentiment analysis with transformer and recurrent neural network. *IEEE Access*, 10: 21517–21525, 2022.
60. Seunghak Yu, Sathish Reddy Indurthi, Seohyun Back, and Haejun Lee. A multi-stage memory augmented neural network for machine reading comprehension. In *Proceedings of the workshop on machine reading for question answering*, pages 21–30, 2018.
61. Diksha Khurana, Aditya Koli, Kiran Khatter, and Sukhdev Singh. Natural language processing: State of the art, current trends and challenges. *Multimedia tools and applications*, 82(3):3713–3744, 2023.
62. Jingfeng Yang, Hongye Jin, Ruixiang Tang, Xiaotian Han, Qizhang Feng, Haoming Jiang, Shaochen Zhong, Bing Yin, and Xia Hu. Harnessing the power of llms in practice: A survey on chatgpt and beyond. *ACM Transactions on Knowledge Discovery from Data*, 2023.
63. Surendrabikram Thapa, Usman Naseem, and Mehwish Nasim. From humans to machines: can chatgpt-like llms effectively replace human annotators in nlp tasks. In *Workshop Proceedings of the 17th International AAAI Conference on Web and Social Media*, 2023.
64. Akshay Goel, Almog Gueta, Omry Gilon, Chang Liu, Sofia Erell, Lan Huong Nguyen, Xiaohong Hao, Bolous Jaber, Shashir Reddy, Rupesh Kartha, et al. Llms accelerate annotation for medical information extraction. In *Machine Learning for Health (ML4H)*, pages 82–100. PMLR, 2023.
65. Xinyi Wang, Wanrong Zhu, Michael Saxon, Mark Steyvers, and William Yang Wang. Large language models are implicitly topic models: Explaining and finding good demonstrations for in-context learning. In *Workshop on Efficient Systems for Foundation Models@ ICML2023*, 2023.
66. Yasir Gamiieldien. Innovating the study of self-regulated learning: An exploration through nlp, generative ai, and llms. 2023.
67. Lei Li, Yongfeng Zhang, and Li Chen. Prompt distillation for efficient llm-based recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 1348–1357, 2023.
68. Googlenewsrss. <https://news.google.com/>, 2024.
69. Matthew Honnibal and Ines Montani. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear, 2017.
70. Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
71. Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.
72. Clayton Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 216–225, 2014.
73. Steven Loria et al. textblob documentation. *Release 0.15*, 2(8):269, 2018.
74. Finn Årup Nielsen. A new ANEW: evaluation of a word list for sentiment analysis in microblogs. In Matthew Rowe, Milan Stankovic, Aba-Sah Dadzie, and Mariann Hardey, editors, *Proceedings of the ESWC2011 Workshop*, volume 718, pages 93–98, 2011. URL http://ceur-ws.org/Vol-718/paper_16.pdf.
75. Alan Akbik, Duncan Blythe, and Roland Vollgraf. Contextual string embeddings for sequence labeling. In *COLING 2018, 27th International Conference on Computational Linguistics*, pages 1638–1649, 2018.
76. Maarten Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022.
77. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
78. David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
79. Dimo Angelov. Top2vec: Distributed representations of topics. *arXiv preprint arXiv:2008.09470*, 2020.

80. Leland McInnes, John Healy, Steve Astels, et al. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, 2(11):205, 2017.
81. Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
82. Maarten Grootendorst. Keybert: Minimal keyword extraction with bert., 2020. URL <https://doi.org/10.5281/zenodo.4461265>.
83. Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*, 2019.
84. Falconsai. Finetuned distilbert for fear mongering detection. https://huggingface.co/Falconsai/fear_mongering_detection, 2023. Accessed: 2024-03-11.
85. Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Huggingface’s transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*, 2019.
86. Tim Dettmers, Mike Lewis, Younes Belkada, and Luke Zettlemoyer. Llm.int8(): 8-bit matrix multiplication for transformers at scale. *arXiv preprint arXiv:2208.07339*, 2022.
87. David R Cox. The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):215–232, 1958.
88. Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
89. Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16*, pages 785–794, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4232-2. doi:10.1145/2939672.2939785. URL <http://doi.acm.org/10.1145/2939672.2939785>.
90. Tin Kam Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE, 1995.
91. Karl Pearson F.R.S. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901. doi:10.1080/14786440109462720.
92. Peter I Frazier. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.
93. Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on EMNLP-IJCNLP*, pages 3982–3992, Hong Kong, China, November 2019. ACM. doi:10.18653/v1/D19-1410.
94. Richard Anderson, Carmela Scala, Jim Samuel, Vivek Kumar, and Parth Jain. Are emotions conveyed across machine translations? establishing an analytical process for the effectiveness of multilingual sentiment analysis with italian text. *Journal of Big Data and Artificial Intelligence*, 2(1), 2024.
95. Parth Hasmukh Jain, Vivek Kumar, Jim Samuel, Sushmita Singh, Abhinay Mannepalli, and Richard Anderson. Artificially intelligent readers: An adaptive framework for original handwritten numerical digits recognition with ocr methods. *Information*, 14(6):305, 2023.
96. Yana Samuel, Margaret Brennan-Tonetta, Jim Samuel, Rajiv Kashyap, Vivek Kumar, Sri Krishna Kaashyap, Nishitha Chidipothu, Irawati Anand, and Parth Jain. Cultivation of human centered artificial intelligence: culturally adaptive thinking in education (cate) for ai. *Frontiers in Artificial Intelligence*, 6:1198180, 2023.
97. Joseph Jones. Don’t fear artificial intelligence, question the business model: How surveillance capitalists use media to invade privacy, disrupt moral autonomy, and harm democracy. *Journal of Communication Inquiry*, page 01968599241235209, 2024.
98. Jim Samuel, Rajiv Kashyap, Yana Samuel, and Alexander Pelaez. Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. *International journal of information management*, 65:102505, 2022.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.