Public Organization Review
https://doi.org/10.1007/s11115-025-00815-2

Artificial Intelligence Tools in Misinformation Management during Natural Disasters

Nadejda Komendantova[1] · **Dmitry Erokhin**[1]

Accepted: 10 February 2025

Abstract

Ensuring accurate information during natural disasters is vital for effective emergency response and public safety. Disasters like earthquakes and hurricanes often trigger misinformation, complicating response efforts and endangering lives. Historical events, such as Hurricane Katrina and the COVID-19 pandemic, illustrate the harmful impact of false information. Artificial intelligence (AI), with technologies like natural language processing and machine learning, offers promising solutions for detecting and mitigating misinformation. This paper explores AI's role in managing misinformation during disasters, highlighting its potential to improve disaster response, enhance public trust, and strengthen community resilience.

Keywords Misinformation · Natural disasters · Artificial intelligence · Natural language processing · Emergency response · Public trust

Introduction

Natural disasters, such as earthquakes, hurricanes, floods, and wildfires, have profound impacts on societies worldwide. These events can cause significant loss of life, displacement of populations, and extensive damage to infrastructure (Teh & Khan, 2021). In the immediate aftermath of such disasters, the need for accurate and timely information becomes critical for effective response and recovery efforts (Natural Research Council, 1999). However, the chaotic nature of these events often leads to the spread of misinformation, which can exacerbate the challenges faced by emergency responders and affected communities (Tran et al., 2020).

✉ Nadejda Komendantova
komendan@iiasa.ac.at

Dmitry Erokhin
erokhin@iiasa.ac.at

1 International Institute for Applied Systems Analysis, Laxenburg, Austria

Published online: 28 February 2025

Misinformation during natural disasters can take many forms, including rumors, false claims, and deliberate disinformation (see Table 1 for a selection of real-world examples). This misinformation can spread rapidly through various channels, particularly social media, and create confusion and panic among the public. For instance, during Hurricane Katrina in 2005, false reports of violence and chaos hampered rescue operations and contributed to a breakdown of trust between the public and authorities (Miller, 2016; Brezina & Phipps Jr, 2009). Similarly, during the COVID-19 pandemic, misinformation regarding safety measures and treatments spread widely and complicated public health efforts (Caceres et al., 2022).

Recognizing the detrimental impact of misinformation, previous research has explored various strategies for its detection and mitigation. Advancements in artificial intelligence (AI) have emerged as promising tools in this regard, offering advanced capabilities for real-time analysis and intervention. Specifically, AI technologies such as natural language processing (NLP) and machine learning algorithms have shown significant potential in identifying and countering false information (Kumar et al., 2021). AI algorithms can achieve an impressive 97% accuracy in classifying news articles, effectively determining whether they are genuine or false (Hashmi et al., 2024). Moreover, AI tools can be highly effective in helping people reconsider their beliefs in conspiracy theories (Costello et al., 2024). In an experiment involving over 2000 participants, researchers used a version of ChatGPT (GPT-4 Turbo) to challenge conspiracy theories. Surprisingly, people's belief in false conspiracy theories dropped by an average of 20%. About 25% of participants reduced their confidence in these theories from above to below 50%. On TikTok, flagging unsubstantiated videos decreased the rate at which they were shared by 24%, while likes on such unsubstantiated content declined by 7% (Hernandez, 2021), whereas hiding content with fact-checking labels on Facebook could reduce content views by up to 95% (Rosen, 2020). In addition, AI tools on Facebook detect almost 100% of spam and remove 99.5% of terrorist-related content, 98.5% of fake accounts, 96% of adult nudity and sexual activity, and 86% of graphic violence (Koebler & Kox, 2018). Endorsement of technology companies' efforts to combat misinformation is also important. A recent survey shows growing public support for technology companies to take steps in restricting false information online, with 65% of respondents in 2023 expressing support, up from 56% in 2018 (PEW Research Center, 2023).

Despite these advancements, effectively managing misinformation during natural disasters remains a significant challenge. The unique characteristics of disaster scenarios– rapidly evolving situations, diverse and unverified information sources, and heightened emotional states among affected populations– pose distinct obstacles that have not been fully addressed in existing research. Challenges include the need for AI systems to process large volumes of data in real-time, understand context-specific language and terminology related to disasters, and deal with the scarcity of labeled datasets for training models in these scenarios. Moreover, there is a gap in research regarding the deployment of AI tools that can adapt to the unpredictable nature of disasters and provide timely interventions without causing additional confusion or

Table 1 Forms of misinformation during natural disasters

Form of misinformation	Description	Real-world examples
Rumors	Unverified information that spreads rapidly due to fear or uncertainty.	- Hurricanes Harvey and Irma: False information spread on social media claiming that immigration status would be checked at evacuation shelters. This caused fear and confusion among residents, potentially affecting their decision to evacuate[a]. - New Year Noto earthquake: After the earthquake on the Noto Peninsula in Japan, rumors spread on social media about the severity of the damage and potential aftershocks. Prime Minister Kishida Fumio had to warn the public about the dangers of spreading misinformation[b].
False claims	Incorrect statements presented as facts.	- Hurricane Milton: False claims suggesting that Hurricane Milton was “engineered” and that the weather in Florida was being “manipulated” spread on social media. These claims were debunked by experts, but they caused significant confusion and fear among the public[c]. - Climate change misinformation: A viral video falsely claimed that decreasing natural disaster-related death counts show climate change is not real. This misinformation spread widely, despite evidence to the contrary[d].
Conspiracy theories	Explanations that attribute events to secret, often malevolent plots by groups.	- Hurricane Milton: Discussions surrounding Hurricane Milton have been clouded by false conspiracy theories, particularly those suggesting that the storm is a result of geoengineering or political manipulation[e]. - Earthquake in Syria and Turkey: The HAARP (High-Frequency Active Auroral Research Program) conspiracy theory suggests that the U.S. government uses this research program as a weapon to intentionally trigger massive earthquakes around the world. Some conspiracy theorists believe that natural disasters, such as the devastating earthquakes in Turkey and Syria, are not accidental but rather the result of human manipulation through HAARP or similar technologies[f].
Misinterpretation of official statements	Incorrect understanding or reporting of information from authorities.	- Oroville Dam crisis: The National Weather Service Sacramento tweeted a map showing expected flooding, which encompassed all of Sacramento County, including the area near Oroville Dam, leading many to mistakenly believe that the entire county was under an evacuation order[g]. - Hurricane Cone Graphic: The National Hurricane Center’s Tropical Cyclone Track Forecast Cone has been widely misinterpreted by the public, leading many to believe that areas outside the cone are safe from storm threats[h].
Hoaxes and pranks	False information spread as jokes or to cause disruption.	- Shooting at the hurricane: A satirical Facebook event encouraged people to shoot guns at the storm to scare it away. Despite its humorous intent, law enforcement had to warn against such dangerous behavior[i]. - Storing valuables in dishwashers: A viral post advised people to store valuables in dishwashers during floods, falsely claiming that they are waterproof. This myth has been debunked by experts, as dishwashers are not watertight during floods[j].

Table 1 (continued)

Form of misinformation	Description	Real-world examples
Scams and fraudulent offers	Deceptive practices aiming to exploit individuals during crises.	- Fake hurricane donation posts: Schemers are creating fake images of hurricane victims and asking people for donations. An example is an image of a girl holding a puppy and being evacuated from a flood area[k]. - Fake earthquake donation posts: Scammers claim to raise money for survivors, left without heat or water. But instead of helping those in need, scammers are channeling donations away from real charities, and into their own accounts and cryptocurrency wallets[l].
False images and videos	Sharing doctored or unrelated media as current disaster events.	- Hurricane Harvey: Following Hurricane Harvey, there were numerous fake images and videos circulating online, including doctored photos and videos from other disasters. These misleading visuals made it difficult for people to discern accurate information[m]. - Hurricane Milton: After Hurricane Milton, numerous false images and videos spread on social media, including AI-generated footage and visuals from unrelated events. These misleading visuals caused confusion and fear among the public[n].

[a]https://new.nsf.gov/news/how-rumors-spread-social-media-during-weather

[b]https://www.nippon.com/en/in-depth/d00987/

[c]https://www.bbc.com/news/articles/cx2lyzw7xwxo

[d]https://eu.usatoday.com/story/news/factcheck/2023/11/27/false-claim-disaster-deaths-show-climate-change-not-real-fact-check/71249882007/

[e]https://www.nbcnews.com/tech/internet/hurricane-milton-conspiracy-theory-government-storm-biden-rcna174558

[f]https://www.codastory.com/newsletters/turkey-earthquake-haarp-conspiracy/

[g]https://www.dhs.gov/sites/default/files/publications/SMWG_Countering-False-Info-Social-Media-Disasters-Emergencies_Mar2018-508.pdf

[h]https://news.miami.edu/stories/2024/02/cone-of-uncertainty-graphic-to-feature-more-information.html

[i]https://www.seattletimes.com/life/what-not-to-believe-viral-hoaxes-about-hurricane-florence/

[j]https://www.seattletimes.com/life/what-not-to-believe-viral-hoaxes-about-hurricane-florence/

[k]https://www.yahoo.com/news/chicopee-police-warn-residents-fake-094259948.html

[l]https://www.bbc.com/news/world-europe-64599553

[m]https://www.artsy.net/article/artsy-editorial-age-fake-photos-videos-spot

[n]https://www.cbsnews.com/news/hurricane-viral-video-how-to-spot-old-fabricated-ai-footage/

panic. This underscores a pressing need to explore how AI can be effectively leveraged to detect and mitigate misinformation in these critical situations.

This study aims to fill this gap by providing a comprehensive examination of the role of AI in detecting and mitigating misinformation during natural disasters. Specifically, it investigates the types and sources of misinformation unique to disaster contexts, analyzes the capabilities of AI technologies in addressing these issues, assesses the impact of AI-driven interventions on public trust and decision-making, and identifies the challenges and limitations associated with deploying AI tools in such scenarios. By integrating theoretical insights with empirical evidence– including case studies of past and recent natural disasters– the research contributes new knowledge to the field and offers practical strategies for enhancing disaster response efforts.

The remainder of the paper is organized as follows. Section 2 provides a literature review on misinformation during natural disasters, AI tools for misinformation detection, and the relationship between information accuracy and public trust. Section 3 outlines the research methodology. Section 4 delves into AI detection and mitigation techniques. Section 5 presents case studies of natural disasters to illustrate the practical application of AI tools in misinformation management. Section 6 examines the impact of AI-driven interventions on public trust and decision-making during natural disasters. Section 7 addresses the challenges and limitations of using AI for misinformation management. Finally, Sect. 8 concludes by summarizing key findings and discussing their implications.

Literature Review

Misinformation during natural disasters is a pervasive issue that significantly hampers emergency response efforts and public safety (Omar & Van Belle, 2024). It encompasses various forms, including rumors, false claims, and deliberate disinformation, all of which can spread rapidly through communication channels, particularly social media platforms (Aïmeur et al., 2023). The widespread use and immediacy of social media amplify the speed at which misinformation is disseminated, creating challenges for authorities trying to ensure the public receives accurate information.

Several key types of misinformation are commonly observed in these situations. Incorrect safety advice, exaggerated or false information about the scale of the disaster, and rumors regarding resource availability or government response are frequent (Méndez-Muros et al., 2024; Omar & Van Belle, 2024; Zhou et al., 2021). For instance, during the 2011 Tōhoku earthquake and tsunami, widespread rumors about radiation levels caused unnecessary panic (Svendsen, 2013), while false tweets during Hurricane Sandy in 2012 about flooding in the New York Stock Exchange building caused confusion and concern (Hunt et al., 2020a).

The sources of misinformation vary widely, from well-meaning individuals sharing unverified information to malicious actors deliberately spreading falsehoods for political or financial gain. Even reputable news organizations can unintentionally contribute to the problem through reporting errors (Molina et al., 2021; Silverman, 2015). The impact of such misinformation is profound, leading to misinformed decisions, the misallocation of resources, and a general erosion of trust in official communication channels (Cook et al., 2015; Rubin, 2022; Lovari, 2020).

AI has emerged as a crucial tool for addressing this growing problem (Demartini et al., 2020). Advanced AI techniques, such as NLP, machine learning, and data mining, are being developed to detect and mitigate the spread of false information (Meesad, 2021). NLP algorithms, for example, can analyze large volumes of text from social media and news outlets to identify misinformation patterns (Berrondo-Otermin & Sarasa-Cabezuelo, 2023), while machine learning models can be trained on datasets of known misinformation to recognize and flag similar content in real-time.

Several AI tools have been developed with promising results. Google's Fact Check Tools and Facebook's AI-driven misinformation detection systems use machine learning to identify and limit the spread of false information on their platforms (Google, 2024; Meta, 2020). However, there are notable challenges to these technologies. Misinformation evolves rapidly, with new false narratives emerging regularly, and AI systems often struggle with context, sarcasm, and cultural nuances, leading to false positives or negatives (Adriani, 2019; Hershcovich et al., 2022; Muaad et al., 2022).

In the context of natural disasters, misinformation undermines public trust in authorities and official sources of information. When official sources are slow to address or counter misinformation, it creates a vacuum that is often filled by rumors and false claims (Carlson et al., 2018). This can lead to a reluctance to follow official guidance, thereby increasing the risks to public safety during emergencies (Freeman et al., 2022). Accurate, timely communication is critical for effective decision-making, as demonstrated by case studies like the 2010 Haiti earthquake, where reliable information was crucial for coordinating rescue efforts (Van de Walle & Dugdale, 2012). Conversely, the spread of misinformation during the 2017 Las Vegas shooting caused widespread confusion, illustrating the dangers of unchecked false information (Blankenship, 2020).

Despite the advances in AI, there are several gaps in current research. The first is the need for more robust datasets that reflect the evolving nature of misinformation (Hu et al., 2022). Many AI models are limited by the quality and scope of the data on which they are trained. Secondly, interdisciplinary collaboration is essential for developing more effective solutions, incorporating insights from fields such as computer science, communications, psychology, and emergency management (Leitner et al., 2021). Finally, ethical and privacy concerns surrounding the use of AI in misinformation detection remain underexplored. Issues such as surveillance, data privacy, and potential misuse of these technologies need further research (Huriye, 2023).

Methodology

This study employs a qualitative research design centered on a narrative literature review, supplemented by case study analyses. Narrative literature reviews serve a vital scientific function in academia, allowing researchers to examine broader questions and engage in post hoc theorizing (Baumeister & Leary, 1997). These reviews facilitate a comprehensive examination of existing knowledge, theories, and practices through an iterative, non-structured, and multi-layered process (Juntunen & Lehenkari, 2021). They provide flexibility in analyzing and interpreting literature. This approach allows for an in-depth exploration of the complexities, challenges, and opportunities within the AI domain.

The literature review draws from a diverse range of sources, including peer-reviewed journal articles, conference papers, academic books, and reputable online publications. The key thematic areas addressed in the review include investigating the common forms of misinformation that emerge during natural disasters and the channels through which they spread; examining the current AI tools and algorithms used to identify and mitigate misinformation; exploring how accurate information influences public perception and actions during disaster scenarios; and identifying the shortcomings and obstacles faced by existing AI solutions in effectively managing misinformation. By synthesizing findings across these areas, the review aims to identify patterns, gaps, and emerging trends that inform the understanding of AI's role in misinformation management.

To ground the theoretical insights in practical application, the study incorporates case studies of real-world disaster scenarios where AI tools have been employed to manage misinformation. The selection of case studies is based on relevance, ensuring that each case involves the application of AI technologies specifically aimed at managing misinformation during a natural disaster. Diversity is also a factor, with cases selected from different types of natural disasters and various geographical regions to provide a comprehensive analysis. Additionally, impact is considered by choosing cases that demonstrate significant outcomes or offer critical insights into the effectiveness and limitations of AI tools in real-world settings. These case studies serve to illustrate the practical challenges and successes of implementing AI solutions, thereby enriching the findings from the literature review with tangible examples.

While this study does not follow the protocols of a systematic literature review, it maintains scientific rigor through critical analysis and synthesis of the available literature. The narrative approach offers the flexibility to explore a wide range of sources and perspectives, which is essential given the rapidly evolving nature of AI technologies and misinformation tactics.

Several limitations are acknowledged in this study. Dependence on existing literature and case studies means that the findings are contingent upon the availability and quality of prior research, which may be subject to publication bias or incomplete data. The dynamic nature of the field, with advances in AI technologies and changes in misinformation strategies, means that the study's conclusions represent a snapshot in time and may require updates as the field evolves. Additionally, the absence of a systematic methodology may introduce selection bias, although efforts have been made to mitigate this through comprehensive sourcing and critical evaluation. By recognizing these limitations, the study aims to provide a transparent and balanced analysis that contributes meaningful insights to the ongoing discourse on AI applications in misinformation management during natural disasters.

AI Detection and Mitigation Techniques

Natural Language Processing

NLP is a subset of artificial intelligence that focuses on the interaction between computers and human language. It enables machines to understand, interpret, and respond to human language in a way that is both meaningful and useful. NLP is particularly effective in misinformation detection during natural disasters due to its ability to process vast amounts of textual data from various sources such as social media, news articles, and official reports.

One of the primary applications of NLP in misinformation detection is sentiment analysis (Alonso et al., 2021). Sentiment analysis involves determining the sentiment expressed in a piece of text, such as positive, negative, or neutral. By analyzing the sentiment of social media posts and news articles, NLP can help identify potential misinformation. For instance, a sudden surge in negative sentiment around a specific topic might indicate the spread of false information. Additionally, NLP can be used to detect specific keywords and phrases commonly associated with misinformation, such as hoax, fake, or unverified.

Named entity recognition (NER) is another crucial NLP technique used in misinformation detection (Tsai, 2023). NER involves identifying and classifying named entities mentioned in text into predefined categories such as the names of people, organizations, locations, and dates. By cross-referencing these entities with reliable databases, NLP can help verify the accuracy of information and flag potential misinformation. For example, if a social media post claims that a particular government agency issued a warning, NER can help verify whether the agency indeed issued such a warning.

Moreover, NLP can be employed for topic modeling, which involves discovering abstract topics within a collection of documents (John & Keikhosrokiani, 2022). Topic modeling helps in identifying the main themes and trends in large datasets and allows for the detection of emerging misinformation narratives. Techniques such as latent Dirichlet allocation (LDA) are commonly used for this purpose. By continuously monitoring the topics discussed in social media and news articles, NLP can help detect and respond to new misinformation trends as they emerge.

Machine Learning Algorithms

Machine learning algorithms play a vital role in misinformation detection by enabling systems to learn from data and improve their performance over time. These algorithms can be trained on large datasets of known misinformation and legitimate information, which allows them to recognize patterns and classify new information accordingly.

Supervised learning algorithms are commonly used in misinformation detection (Poddar & Umadevi, 2019). In supervised learning, algorithms are trained on labeled datasets, where the input data is paired with the correct output. For example, a dataset might consist of social media posts labeled as either true or false. Common supervised learning algorithms used in misinformation detection include support vector machines (SVM), decision trees, and neural networks. These algorithms can classify new posts as true or false based on the patterns they have learned from the training data.

Unsupervised learning algorithms, on the other hand, do not require labeled data (Barve & Saini, 2022). Instead, they identify patterns and relationships within the data itself. Clustering algorithms, such as k-means clustering, are often used in this context. By grouping similar pieces of information together, clustering algorithms can help identify clusters of misinformation. For example, if a large number of social media posts share similar content and are identified as potentially false, this cluster can be flagged for further investigation.

Deep learning has also shown great promise in misinformation detection (Nasir et al., 2021). Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are capable of processing large volumes of data and capturing complex patterns. These models have been used to analyze text, images, and videos for misinformation. For instance, CNNs can be used to detect doctored images (Hamid et al., 2023), while RNNs can analyze the temporal patterns in posts (Shu et al., 2017).

Ensemble methods, which combine the predictions of multiple machine learning models, are also effective in misinformation detection (Ahmad et al., 2020). By aggregating the outputs of different models, ensemble methods can improve the overall accuracy and robustness of the system. Common ensemble methods include random forests, boosting, and bagging. These methods can help mitigate the limitations of individual models and provide more reliable predictions.

Real-time Monitoring Systems

Real-time monitoring systems are essential for the timely detection and mitigation of misinformation during natural disasters. These systems continuously monitor various information sources, such as social media, news websites, and official channels, to identify and respond to misinformation as it arises.

One of the key components of real-time monitoring systems is the use of automated data collection techniques. Web scraping tools can be used to gather data from various online sources in real-time (Rajarathinam et al., 2022). These tools can be programmed to collect specific types of information, such as social media posts containing certain keywords or news articles from specific websites. By continuously collecting data, real-time monitoring systems can ensure that they have up-to-date information on the current state of misinformation.

Once the data is collected, it is processed and analyzed using various AI techniques as also described above. NLP is often used to preprocess the data, to extract relevant information and filter out noise (Al Sharou et al., 2021; Singh, 2018). Machine learning algorithms are then applied to classify the information and identify potential misinformation (Du et al., 2021). The use of AI allows these systems to process large volumes of data quickly and efficiently and makes real-time monitoring feasible.

Real-time monitoring systems also rely on alert mechanisms to notify relevant authorities and stakeholders about detected misinformation (Hussein et al., 2020). These alerts can be generated based on predefined rules or thresholds. For example, if a certain number of social media posts about a false rumor are detected within a short period, an alert can be triggered. These alerts can be sent to emergency responders, government officials, or other relevant parties and allow them to take prompt action to address the misinformation.

In addition to detecting misinformation, real-time monitoring systems can also play a role in mitigating its impact. By providing accurate and timely information, these systems can help counteract false narratives and reassure the public. For instance, real-time monitoring systems can be used to disseminate verified information through official channels, such as government websites and social media accounts. This can help correct false information and prevent it from spreading further.

Real-time monitoring systems can also support the coordination of emergency response efforts. By providing a comprehensive view of the information landscape, these systems can help identify areas where misinformation is causing confusion or hindering response efforts. This allows emergency responders to prioritize their efforts and allocate resources more effectively.

Case Studies

Historical Analysis

Hurricane Katrina (2005)

Hurricane Katrina, one of the most devastating natural disasters in U.S. history, highlighted the significant role misinformation can play during crises (Sun, 2010, 2012; Brezina & Phipps Jr, 2009). In the aftermath of Katrina, misinformation spread rapidly through various channels and exacerbated the chaos. False reports of violence, lawlessness, and rampant looting in New Orleans created a sense of fear and confusion among both the public and emergency responders. These false narratives not only hindered rescue and relief efforts but also strained the relationship between the affected communities and government agencies.

Given that social media was not as prevalent at the time, and some of today's popular networks did not even exist yet (Beigi et al., 2016), most of the analysis done focused on examining the discussion of the hurricane in traditional media (Barnes et al., 2008; Sommers et al., 2006; Tierney et al., 2006).

2010 Haiti Earthquake

The 2010 earthquake in Haiti showcased the potential benefits of using technology to manage misinformation. In the immediate aftermath of the earthquake, there was a significant influx of misinformation on social media (Oh et al., 2010).

The crisis-mapping platform Ushahidi played a crucial role in disaster relief efforts following the earthquake in Haiti (Heinzelman & Waters, 2010). It allowed the international community to access real-time information directly from the Haitian population through text messages and social media sources. By crowdsourcing information, Ushahidi enabled responders to quickly target resources in the rapidly changing disaster environment. It facilitated the collection and mapping of actionable intelligence and helped identify urgent needs and incidents on the ground. Despite challenges in verifying and triaging reports, Ushahidi's deployment in Haiti showcased the potential of leveraging technology for effective disaster response and information gathering.

Recent Examples

Hurricane Harvey (2017)

Hurricane Harvey struck the Gulf Coast of the United States in 2017 and brought unprecedented rainfall and flooding to the Houston area. During this disaster, social media became a double-edged sword. While it was invaluable for disseminating real-time information and coordinating rescue efforts, it also facilitated the rapid spread of misinformation (Hunt et al., 2020b). False reports about mandatory evacuations or the availability of shelters caused confusion and unnecessary panic (Yang & Zhuang, 2020).

Machine learning was applied to track and predict the veracity of tweets related to misinformation spread during the hurricane (Hunt et al., 2020b, 2022; Wang et al., 2020). Researchers were able to analyze and monitor the spread of false rumors on platforms like Twitter. The framework was trained on historical misinformation-related tweets from Hurricane Harvey and other events to predict the accuracy of tweets associated with new cases of misinformation. The machine learning algorithms were capable of accurately predicting the veracity of tweets and categorize them as true, false, or neutral.

COVID-19 Pandemic

The COVID-19 pandemic represents one of the most extensive and prolonged crises where misinformation management has been crucial (Bridgman et al., 2020). The pandemic saw an explosion of misinformation across all media, including social media, news websites, and even traditional media. False claims about the virus's origins, unverified treatments, and misleading information about vaccines created widespread confusion and hindered public health efforts.

AI tools were extensively deployed to combat misinformation during the pandemic. Social media platforms implemented sophisticated AI algorithms to detect and remove false information related to COVID-19 (Shams et al., 2021; Timberg et al., 2020). Additionally, AI-driven chatbots and virtual assistants were deployed by various health organizations to provide accurate information and answer public queries (Agarwal et al., 2024). These AI tools helped in countering misinformation by disseminating verified information and clarifying doubts, thereby improving public understanding and trust (Almalki & Azeez, 2020).

Comparative Analysis

To summarize the above case studies, historically, the management of misinformation during disasters was a significant challenge due to the absence of advanced technological tools. In earlier crises, limited technological infrastructure and the absence of social media meant that false reports often proliferated unchecked through traditional media channels. This could lead to widespread fear, confusion, and strained relationships between affected communities and government agencies. The analysis of misinformation during these periods primarily focused on traditional media's role and highlighted the need for more robust information management systems.

As technology advanced, the potential of AI and machine learning became more apparent in managing misinformation and facilitating disaster response. With the advent of crowdsourced information platforms and real-time data analytics, responders could access situational awareness more effectively. This innovative use of technology allowed for better targeting of resources and demonstrated the benefits of integrating AI-driven platforms in disaster response strategies, despite challenges in report verification.

In recent years, AI has transitioned from a research tool to an active component in managing disasters and combating misinformation. Social media and other digital platforms have leveraged AI-driven machine learning algorithms to track and predict the veracity of information and to enable real-time monitoring and categorization of misinformation. This marked a significant improvement in the ability to manage false information quickly and effectively and showcased AI's growing role in disaster management.

Impact on Public Trust and Decision-making

Building Public Trust

Public trust is a cornerstone of effective disaster management (Oh & Lee, 2022). During natural disasters, the public relies heavily on accurate and timely information to make critical decisions about their safety and well-being. When misinformation spreads, it undermines this trust and may lead to confusion, fear, and potentially harmful behaviors (Korta, 2018). AI tools can play a significant role in building and maintaining public trust by ensuring that accurate information is disseminated, and misinformation is promptly addressed.

AI-driven misinformation detection systems help authorities monitor and analyze large volumes of data from social media, news outlets, and other sources (Ivanov & Petrova, 2024). By identifying and flagging false information, these systems enable rapid response and correction, preventing misinformation from gaining traction. For example, during the COVID-19 pandemic, AI tools were used to identify and correct false claims about the virus, treatments, and vaccines (Ahmad et al., 2022; Chen & See, 2020).

Additionally, AI tools can enhance the transparency of information dissemination processes. By providing clear explanations of how misinformation is detected and why certain information is flagged or corrected, authorities can build trust with the public. Transparency about the use of AI and the steps taken to ensure data accuracy and privacy can also alleviate concerns about surveillance and misuse of technology.

Improving Decision-making

Accurate information is critical for effective decision-making during natural disasters. Both individuals and emergency responders rely on timely and accurate information to make decisions about evacuations, resource allocation, and response strategies. Misinformation can lead to poor decision-making, with potentially life-threatening consequences. AI tools can significantly improve decision-making by ensuring that accurate information is available and accessible.

One of the key ways AI improves decision-making is through real-time data analysis and situational awareness. AI systems can process and analyze vast amounts of data quickly and provide real-time insights into the evolving disaster situation. For example, AI-powered dashboards and analytics platforms can visualize data from various sources, such as weather reports, social media, and emergency services, to provide a comprehensive view of the situation (Ramya & Singh, 2024; Reddy et al., 2024; Dias et al., 2023). This allows decision-makers to respond more effectively and allocate resources where they are needed most.

AI tools can also assist in predictive modeling and scenario planning (Varsha et al., 2024). By analyzing historical data and current trends, AI systems can predict the likely impact of a disaster and suggest optimal response strategies. For instance, during hurricane season, AI models can predict the potential path and impact of a storm, enabling authorities to issue timely warnings and plan evacuations more effectively (Hashemi et al., 2016). These predictive capabilities enhance the ability to make informed decisions and reduce the overall impact of natural disasters.

Moreover, AI tools can facilitate better communication and coordination among different stakeholders. During natural disasters, effective communication is essential for coordinating response efforts and ensuring that everyone has the information they need. AI-driven communication platforms can streamline information sharing and enable emergency responders, government agencies, and the public to stay informed and work together more effectively. For example, AI chatbots and virtual assistants can provide real-time updates and answer common questions and free up human resources for more critical tasks (Do et al., 2022; Bonales Daimiel & Martínez Estrella, 2021).

Strategies for Using AI to Enhance Public Trust

To maximize the benefits of AI in building public trust and improving decision-making, several strategies can be implemented. First, it is essential to ensure the transparency and accountability of AI systems. Authorities should clearly communicate how AI tools are used, what data they analyze, and how decisions are made based on AI insights. Transparency builds trust and helps the public understand the value of AI in disaster management.

Second, collaboration between AI developers, emergency responders, and communication experts is crucial. By working together, these stakeholders can design AI tools that are user-friendly, accurate, and effective in addressing misinformation. Collaborative efforts also ensure that AI tools are integrated seamlessly into existing emergency response frameworks and enhance their overall impact.

Third, ongoing training and education are vital for both the public and emergency responders. The public should be educated about the risks of misinformation and the importance of relying on verified sources of information. Emergency responders, on the other hand, should be trained on how to use AI tools effectively and interpret the data provided by these systems. Continuous training ensures that everyone is prepared to leverage AI technologies during disasters.

Finally, it is important to address ethical and privacy concerns related to the use of AI. Authorities should implement robust data protection measures and ensure that AI tools are used responsibly. By prioritizing ethical considerations and safeguarding privacy, authorities can build public confidence in the use of AI for disaster management.

Challenges and Limitations

Technical Challenges

AI technologies, while powerful, face numerous technical challenges when it comes to detecting and mitigating misinformation during natural disasters. One of the primary issues is the constant evolution of misinformation tactics. Misinformation creators frequently adapt their strategies, using new language, images, and dissemination techniques to evade detection by AI systems (Ayetiran & Özgöbek, 2024; Shahid et al., 2022). To address this challenge, implementing automated algorithms that can update in real-time with minimal supervision is essential (Haass, 2022). Utilizing machine learning models that support online learning can help AI systems adapt to new misinformation patterns more efficiently. Additionally, collaboration between organizations to share threat intelligence can lead to the development of more robust models that are better equipped to handle evolving tactics (Preuveneers & Joosen, 2021).

Another technical challenge is the complexity of natural language. Understanding context, sarcasm, and cultural nuances is difficult for AI models (Rahma et al., 2023; Hershcovich et al., 2022). These subtleties often lead to false positives (accurate information being flagged as false) and false negatives (misinformation not being detected). For instance, a sarcastic tweet might be incorrectly classified as genuine information, or a legitimate news article might be flagged due to specific keywords that the model associates with misinformation. Advancements in natural language processing, such as the transition from rule-based systems to sophisticated neural networks, particularly transformer-based models like BERT and GPT, could enhance AI's ability to understand subtle linguistic cues (Zhang, 2024; Chai et al., 2024). These models have improved performance in various NLP applications, including sentiment analysis, machine translation, and content generation. The integration of deep learning techniques has led to more accurate interpretation of contextual relationships and semantic nuances. Multilingual capabilities have been developed to address language diversity (Pritam et al., 2024).

Data quality and availability also pose significant challenges (Velev & Zlateva, 2023). AI systems rely on large datasets to learn and improve. However, obtaining high-quality, labeled data that accurately represents the diverse forms of misinformation can be challenging. During natural disasters, the rapid influx of information can include a lot of noise– irrelevant or redundant data– which makes it harder for AI systems to filter out useful information from misinformation. To mitigate this issue, implementing data preprocessing techniques such as noise reduction algorithms and data augmentation can enhance dataset quality (Maharana et al., 2022). Leveraging semi-supervised and unsupervised learning methods can reduce the reliance on labeled data, allowing AI systems to learn from unlabeled datasets (Nishi et al., 2021; Xie et al., 2020). Additionally, crowdsourcing platforms can be utilized to obtain labeled data more efficiently, engaging volunteers in the data labeling process (Drutsa et al., 2020).

Ethical and Privacy Concerns

The use of AI for monitoring and analyzing public communication raises significant ethical and privacy concerns (Stahl & Wright, 2018). AI systems often require access to vast amounts of personal data, which can include social media posts, location data, and communication patterns. This level of surveillance can lead to privacy infringements if not managed properly. There is a fine balance between collecting enough data to accurately detect misinformation and protecting individual privacy.

One of the most pressing negative consequences is the potential erosion of individual autonomy and freedom of expression. Awareness of surveillance leads to self-censorship and a chilling effect on open discourse (Murray et al., 2024; Stoycheff, 2016). AI systems used in content moderation and personalization on online platforms can create filter bubbles and raise concerns about the legitimacy of content removal decisions (Kolarević, 2022). The deployment of AI tools to manipulate democratic processes and capture personal information poses threats to privacy, anonymity, and autonomy (Manheim & Kaplan, 2019). These effects can erode public confidence in institutions and hinder democratic functioning. The impacts of surveillance are particularly concerning for marginalized communities and can exacerbate existing inequalities. To address these concerns, implementing privacy-preserving techniques such as differential privacy and federated learning can help protect individual data (Rodríguez-Barroso et al., 2020). These methods allow AI models to learn from data without accessing personally identifiable information. Establishing clear data governance policies and obtaining informed consent can also enhance transparency and trust (Janssen et al., 2020; Rossi & Lenzini, 2020).

Transparency about data collection and processing is essential to maintain public trust. However, many AI systems operate as black boxes (Von Eschenbach, 2021), which makes it difficult to explain how decisions are made. This lack of transparency can lead to public skepticism and resistance, particularly if people feel their data is being used without their consent or understanding. Ensuring that AI systems are explainable and that their decision-making processes are transparent is critical to addressing these ethical concerns. Developing explainable AI (XAI) models can help in making AI decision-making processes more transparent (Tiwari, 2023). XAI techniques focus on developing interpretable models and post-hoc explanations to decipher the reasoning behind AI-driven conclusions (Alapati & Valleru, 2023). These approaches include feature importance analysis, model interpretability, and natural language explanations (Tiwari, 2023). By bridging the cognitive gap between complex algorithms and human understanding, XAI empowers stakeholders to evaluate and trust AI systems, ensuring fairness, accountability, and ethical standards (Praveenraj et al., 2023).

Bias in AI models is another significant ethical issue (Gaonkar et al., 2020). AI systems are trained on existing data, which can include biases present in the source material. If not carefully managed, these biases can be perpetuated or even amplified by AI systems. For example, certain communities might be disproportionately targeted by misinformation detection efforts, which leads to unequal treatment. Ensuring fairness and avoiding discrimination requires continuous monitoring and adjustment of AI models. Implementing fairness-aware machine learning practices is crucial for mitigating biases in AI systems. Researchers have developed tools to help data scientists discover and mitigate discrimination in machine learning models (Bantilan, 2018). Strategies such as reweighting, adversarial training, and resampling can be employed to overcome prejudice in algorithms (Dhabliya et al., 2024). Adopting a "fairness-first" approach is essential when developing machine learning models for various applications (Bird et al., 2019). Regular auditing of AI models for bias and incorporating diverse datasets during training can enhance fairness. The Fairway method, which combines pre-processing and in-processing approaches, has been proposed to remove ethical bias from training data and trained models without significantly impacting predictive performance (Chakraborty et al., 2020). Collaborating with ethicists and stakeholders from various communities can provide valuable insights to address potential biases proactively (Bird et al., 2019).

Furthermore, the deployment of AI-based surveillance systems raises serious concerns about data security and the potential for misuse. Large-scale data collection creates attractive targets for cybercriminals, increasing the risk of data breaches that could expose sensitive personal information (Veluru, 2024). There is also the danger that surveillance data could be repurposed for unintended uses, such as commercial profiling or political manipulation, without individuals' knowledge or consent. In the absence of stringent regulations and oversight, these systems might operate with little accountability, making it difficult to challenge abuses of power or seek redress for harms caused (Fontes et al., 2022).

Practical Limitations

The practical implementation of AI systems for misinformation management during natural disasters involves several logistical challenges. One major limitation is the integration of AI tools with existing emergency response frameworks (Damaševičius et al., 2023; Deloitte, 2023). Many emergency response systems are not designed to incorporate advanced AI technologies and retrofitting them can be complex and costly. Coordination between various agencies and stakeholders is essential to ensure that AI tools can be effectively utilized. To facilitate integration, developing modular AI solutions that can easily interface with existing systems is crucial (Agrawal et al., 2021; Su et al., 2019). Establishing standard protocols and APIs can enable seamless communication between AI tools and emergency response platforms. Securing funding and support for technological upgrades can also help in overcoming financial barriers.

Training and capacity-building are also critical practical considerations (Fernando, 2023). Emergency responders and relevant personnel need to be trained on how to use AI tools and interpret the data they provide. This training requires time and resources, and it is essential to ensure that all stakeholders are adequately prepared to leverage AI technologies during disasters. Implementing comprehensive training programs and workshops can enhance the skill sets of emergency responders. Online training modules and simulation exercises can provide practical experience with AI tools. Partnering with educational institutions and technology providers can support ongoing capacity-building efforts.

Resource constraints can limit the deployment and effectiveness of AI systems. Developing and maintaining sophisticated AI models require significant computational power and expertise. In resource-limited settings, such as in many developing countries, the necessary infrastructure and technical expertise may be lacking (Nugraha, 2023; Wahl et al., 2018). This disparity can lead to unequal access to the benefits of AI in misinformation management. Leveraging cloud-based AI services and edge computing can provide access to computational resources for developing regions without significant upfront investment (Vuruma et al., 2024). Open-source AI platforms and collaboration with international organizations can support the deployment of AI solutions in low-resource contexts (Okolo, 2020).

Real-world Constraints

The real-world application of AI systems also faces constraints related to the unpredictable nature of natural disasters (Venkadesh et al., 2024). The rapid onset of disasters means that AI systems must be able to adapt quickly to changing conditions and new types of misinformation. This requires robust and flexible AI models that can operate effectively in real-time, a challenging requirement given the current state of technology. Enhancing the real-time processing capabilities of AI systems through optimization and efficient algorithms can improve responsiveness. Edge computing can improve responsiveness by processing data closer to the source, reducing latency in critical applications (Bertino & Banerjee, 2020). AI-driven edge analytics can analyze aerial footage to create real-time environmental models and detect individuals in distress during disasters (Wagner & Roopaei, 2020). Machine learning algorithms can efficiently analyze and prioritize emergency calls, enabling faster response times and better resource allocation (Reddy et al., 2024). Social media platforms, combined with AI technologies, can provide valuable information during disasters by automatically processing textual messages and imagery data into humanitarian categories (Imran et al., 2017). These advancements in AI and edge computing contribute to improved situational awareness, faster decision-making, and more effective coordination among emergency services, ultimately enhancing disaster management capabilities and societal resilience.

Furthermore, the effectiveness of AI systems can be influenced by the broader information ecosystem. For instance, the presence of multiple platforms with varying levels of content moderation and misinformation control can complicate efforts to manage misinformation. Coordination across different stakeholders (Ramašauskaitė, 2023) is necessary to create a comprehensive approach, but this is often difficult to achieve due to differing policies, technical standards, and interests.

Public perception and behavior also play a crucial role (Kankanamge et al., 2021). Even the most advanced AI systems can be rendered ineffective if the public does not trust or engage with the information provided by these systems. Building and maintaining public trust requires not only technical solutions but also effective communication strategies that emphasize transparency, accountability, and the ethical use of AI. Implementing public awareness campaigns to educate communities about the role and benefits of AI in disaster response can enhance engagement. Providing clear and accessible information about how AI systems operate and the measures taken to protect privacy can build trust. Encouraging public feedback and participation in the development of AI tools can also foster a sense of ownership and acceptance.

Conclusion

Integrating AI into misinformation management during natural disasters addresses key issues in public communication, organizational decision-making, and community resilience. AI technologies such as natural language processing, machine learning algorithms, and real-time monitoring systems offer significant potential for enhancing the accuracy of information dissemination during crises. These advancements support both individual and organizational decision-making by providing accurate, timely information, which is crucial for mitigating the harmful impacts of misinformation on public safety and trust.

However, several technical, ethical, and operational challenges persist. Issues like algorithmic bias, data privacy, and the evolving tactics of those spreading misinformation remain major concerns, requiring ongoing refinement of AI systems. Future efforts must focus on enhancing AI's ability to contextualize misinformation in diverse settings, advocating for a more interdisciplinary approach to AI deployment in crisis management.

The importance of policy and regulatory frameworks cannot be overstated. Effective governance and international collaboration will be critical in standardizing AI use for misinformation management. Additionally, addressing public skepticism through transparent AI operations and educating communities about the role of AI are essential steps in building trust in public organizations.

Lastly, collaboration among AI developers, emergency responders, and policymakers will strengthen the practical application of AI tools. By investing in stakeholder training and interdisciplinary partnerships, current limitations can be overcome and the resilience of both public organizations and communities in disaster recovery efforts can be enhanced.

Funding Open access funding provided by International Institute for Applied Systems Analysis (IIASA). The authors would like to acknowledge following projects which provided funding for this paper:
CORE (sCience and human factOrs for Resilient sociEty) is an EU project funded by the European Commission under Horizon2020 Programme nder grant agreement No. 101021746, CORE (science and human factor for resilient society) as well as the the AGORA project of the European Commission with the grant agreement number 101093921.

Declarations

Ethical Approval The authors declare that they have all needed ethical approvals.

Informed Consent The authors declare that they have all needed consents which are required by the European and International regulation.

Conflict of Interest The authors declare no conflict of interest

References

Adriani, R. (2019). The evolution of fake news and the abuse of emerging technologies. *European Journal of Social Sciences*, *2*(1), 18–24. https://doi.org/10.26417/ejss-2019.v2i1-53

Agarwal, P., Swami, S., & Malhotra, S. K. (2024). Artificial intelligence adoption in the post COVID-19 new-normal and role of smart technologies in transforming business: A review. *Journal of Science and Technology Policy Management*, *15*(3), 506–529. https://doi.org/10.1108/JSTPM-08-2021-0122

Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2021). *AI adoption and system-wide change* (No. w28811). National Bureau of Economic Research. https://doi.org/10.3386/W28811

Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. *Complexity*, *2020*(1), 8885861. https://doi.org/10.1155/2020/8885861

Ahmad, T., Aliaga Lazarte, E. A., & Mirjalili, S. (2022). A systematic literature review on fake news in the COVID-19 pandemic: Can AI propose a solution? *Applied Sciences*, *12*(24), 12727. https://doi.org/10.3390/app122412727

Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, *13*(1), 30. https://doi.org/10.1007/s13278-023-01028-5

Al Sharou, K., Li, Z., & Specia, L. (2021). Towards a better understanding of noise in natural language processing. In *Proceedings of the International conference on recent advances in natural language processing (RANLP 2021)* (pp. 53–62). https://aclanthology.org/2021.ranlp-1.7

Alapati, N. K., & Valleru, V. (2023). The impact of explainable AI on transparent decision-making in financial systems. *Journal of Innovative Technologies*, *6*(1), 1–5. https://www.researchgate.net/publication/384329716_The_Impact_of_Explainable_AI_on_Transparent_Decision-_Making_in_Financial_Systems

Almalki, M., & Azeez, F. (2020). Health chatbots for fighting COVID-19: A scoping review. *Acta Informatica Medica*, *28*(4), 241. https://doi.org/10.5455/aim.2020.28.241-247

Alonso, M. A., Vilares, D., Gómez-Rodríguez, C., & Vilares, J. (2021). Sentiment analysis for fake news detection. *Electronics*, *10*(11), 1348. https://doi.org/10.3390/electronics10111348

Ayetiran, E. F., & Özgöbek, Ö. (2024). A review of deep learning techniques for multimodal fake news and harmful languages detection. *IEEE Access, 12*, 76133–76153. https://doi.org/10.1109/ACCESS.2024.3406258

Bantilan, N. (2018). Themis-ml: A fairness-aware machine learning interface for end-to-end discrimination discovery and mitigation. *Journal of Technology in Human Services*, *36*(1), 15–30. https://doi.org/10.1080/15228835.2017.1416512

Barnes, M. D., Hanson, C. L., Novilla, L. M., Meacham, A. T., McIntyre, E., & Erickson, B. C. (2008). Analysis of media agenda setting during and after Hurricane Katrina: Implications for emergency preparedness, disaster response, and disaster policy. *American Journal of Public Health*, *98*(4), 604–610. https://doi.org/10.2105/AJPH.2007.112235

Barve, Y., & Saini, J. R. (2022). Unsupervised misinformation detection model using incremental k-means algorithm. *International Journal of Intelligent Systems and Applications in Engineering*, *10*(4), 728–735. https://ijisae.org/index.php/IJISAE/article/view/2348

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, *1*(3), 311–320. https://doi.org/10.1037/1089-2680.1.3.311

Beigi, G., Hu, X., Maciejewski, R., & Liu, H. (2016). An overview of sentiment analysis in social media and its applications in disaster relief. *Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence*, 313–340. https://doi.org/10.1007/978-3-319-30319-2_13

Berrondo-Otermin, M., & Sarasa-Cabezuelo, A. (2023). Application of artificial intelligence techniques to detect fake news: A review. *Electronics*, *12*(24), 5041. https://doi.org/10.3390/electronics12245041
Bertino, E., & Banerjee, S. (2020). Artificial intelligence at the edge. arXiv preprint arXiv:2012.05410. https://doi.org/10.48550/arXiv.2012.05410
Bird, S., Kenthapadi, K., Kiciman, E., & Mitchell, M. (2019). Fairness-aware machine learning: Practical challenges and lessons learned. In *Proceedings of the twelfth ACM international conference on web search and data mining* (pp. 834–835). https://doi.org/10.1145/3289600.3291383
Blankenship, M. (2020). How misinformation spreads through twitter. https://digitalscholarship.unlv.edu/brookings_capstone_studentpapers/6
Bonales Daimiel, G., & Martínez Estrella, E. C. (2021). Using virtual assistants and chatbots for crisis communication. *aDResearch ESIC: International Journal of Communication Research/Revista Internacional De Investigación en Comunicación*, *25*(25). https://doi.org/10.7263/adresic-025-04
Brezina, T., & PhippsJr, H. E. (2009). False news reports, folk devils, and the role of public officials: Notes on the social construction of law and order in the aftermath of Hurricane Katrina. *Deviant Behavior*, *31*(1), 97–134. https://doi.org/10.1080/01639620902854803
Bridgman, A., Merkley, E., Loewen, P. J., Owen, T., Ruths, D., Teichmann, L., & Zhilin, O. (2020). The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. *Harvard Kennedy School Misinformation Review*, *1*(3). https://doi.org/10.37016/mr-2020-028
Caceres, M. M. F., Sosa, J. P., Lawrence, J. A., Sestacovschi, C., Tidd-Johnson, A., Rasool, M. H. U.,..., & Fernandez, J. P. (2022). The impact of misinformation on the COVID-19 pandemic. *AIMS Public Health*, *9*(2), 262. https://doi.org/10.3934/Fpublichealth.2022018
Carlson, M., Jakli, L., & Linos, K. (2018). Rumors and refugees: How government-created information vacuums undermine effective crisis management. *International Studies Quarterly*, *62*(3), 671–685. https://doi.org/10.1093/isq/sqy018
Chai, Y., Jin, L., Feng, S., & Xin, Z. (2024). Evolution and advancements in deep learning models for Natural Language Processing. *Applied and Computational Engineering*, *77*, 144–149. https://doi.org/10.54254/2755-2721/77/20240674
Chakraborty, J., Majumder, S., Yu, Z., & Menzies, T. (2020). Fairway: A way to build fair ML software. In *Proceedings of the 28th ACM Joint meeting on european software engineering conference and symposium on the foundations of software engineering* (pp. 654–665). https://doi.org/10.1145/3368089.3409697
Chen, J., & See, K. C. (2020). Artificial intelligence for COVID-19: Rapid review. *Journal of Medical Internet Research*, *22*(10), e21476. https://doi.org/10.2196/21476
Cook, J., Ecker, U., & Lewandowsky, S. (2015). Misinformation and how to correct it. In *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*, 1–17. https://cssn.org/wp-content/uploads/2020/12/Misinformation-and-how-to-correct-it-John-Cook.pdf
Costello, T. H., Pennycook, G., & Rand, D. G. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science*, *385*(6714), eadq1814. https://doi.org/10.1126/science.adq1814
Damaševičius, R., Bacanin, N., & Misra, S. (2023). From sensors to safety: Internet of Emergency Services (IoES) for emergency response and disaster management. *Journal of Sensor and Actuator Networks*, *12*(3), 41. https://doi.org/10.3390/jsan12030041
Deloitte. (2023). Leveraging AI for effective emergency management and crisis response. *Deloitte Insights*. https://www2.deloitte.com/us/en/insights/industry/public-sector/automation-and-generative-ai-in-government/leveraging-ai-in-emergency-management-and-crisis-response.html
Demartini, G., Mizzaro, S., & Spina, D. (2020). Human-in-the-loop artificial intelligence for fighting online misinformation: Challenges and opportunities. *IEEE Data Eng Bull*, *43*(3), 65–74. https://www.damianospina.com/publication/demartini-2020-human/demartini-2020-human.pdf
Dhabliya, D., Dari, S. S., Dhablia, A., Akhila, N., Kachhoria, R., & Khetani, V. (2024). Addressing bias in machine learning algorithms: Promoting fairness and ethical design. *E3S Web of Conferences, 491*, 02040. https://doi.org/10.1051/e3sconf/202449102040
Dias, T., Fonseca, T., Vitorino, J., Martins, A., Malpique, S., & Praça, I. (2023). From data to action: Exploring AI and IoT-driven solutions for smarter cities. In *International symposium on distributed computing and artificial intelligence* (pp. 44–53). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-38333-5_5
Do, V., Huyen, A., Joubert, F. J., Gabriel, M., Yun, K., Lu, T., & Chow, E. (2022). A virtual assistant for first responders using natural language understanding and optical character recognition. In *Pattern Recognition and tracking XXXIII* (Vol. 12101, pp. 94–104). SPIE. https://doi.org/10.1117/12.2620729

Drutsa, A., Fedorova, V., Ustalov, D., Megorskaya, O., Zerminova, E., & Baidakova, D. (2020). Crowdsourcing practice for efficient data labeling: Aggregation, incremental relabeling, and pricing. In *Proceedings of the 2020 ACM SIGMOD international conference on management of data* (pp. 2623–2627). https://doi.org/10.1145/3318464.3383127

Du, J., Preston, S., Sun, H., Shegog, R., Cunningham, R., Boom, J., & Tao, C. (2021). Using machine learning–based approaches for the detection and classification of human papillomavirus vaccine misinformation: Infodemiology study of reddit discussions. *Journal of Medical Internet Research*, *23*(8), e26478. https://doi.org/10.2196/26478

Fernando, Y. S. (2023). Transforming emergency response and disaster management with Ai: Developing intelligent systems for triage, resource allocation, and situational awareness in healthcare crises. *International Journal of Applied Machine Learning and Computational Intelligence*, *13*(12), 11–20. https://neuralslate.com/index.php/Machine-Learning-Computational-I/article/view/114

Fontes, C., Hohma, E., Corrigan, C. C., & Lütge, C. (2022). AI-powered public surveillance systems: Why we (might) need them and how we want them. *Technology in Society*, *71*, 102137. https://doi.org/10.1016/j.techsoc.2022.102137

Freeman, D., Waite, F., Rosebrock, L., Petit, A., Causier, C., East, A., & Lambe, S. (2022). Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological Medicine*, *52*(2), 251–263. https://doi.org/10.1017/S0033291720001890

Gaonkar, B., Cook, K., & Macyszyn, L. (2020). Ethical issues arising due to bias in training AI algorithms in healthcare and data sharing as a potential solution. *The AI Ethics Journal*, *1*(1). https://doi.org/10.47289/AIEJ20200916

Google. (2024). Introduction to AI for journalists. https://newsinitiative.withgoogle.com/resources/trainings/introduction-to-ai-for-journalists/

Haass, J. C. (2022). Cyber threat intelligence and machine learning. In *2022 fourth international conference on transdisciplinary AI (TransAI)* (pp. 156–159). IEEE. https://doi.org/10.1109/TransAI54797.2022.00033

Hamid, Y., Elyassami, S., Gulzar, Y., Balasaraswathi, V. R., Habuza, T., & Wani, S. (2023). An improvised CNN model for fake image detection. *International Journal of Information Technology*, *15*(1), 5–15. https://doi.org/10.1007/s41870-022-01130-5

Hashemi, M. R., Spaulding, M. L., Shaw, A., Farhadi, H., & Lewis, M. (2016). An efficient artificial intelligence model for prediction of tropical storm surge. *Natural Hazards*, *82*, 471–491. https://doi.org/10.1007/s11069-016-2193-4

Hashmi, E., Yayilgan, S. Y., Yamin, M. M., Ali, S., & Abomhara, M. (2024). Advancing fake news detection: Hybrid deep learning with FastText and explainable AI. *Ieee Access: Practical Innovations, Open Solutions*. https://doi.org/10.1109/ACCESS.2024.3381038

Heinzelman, J., & Waters, C. (2010). *Crowdsourcing crisis information in disaster-affected Haiti*. US Institute of Peace. http://www.jstor.com/stable/resrep12220

Hernandez, G. (2021). New prompts to help people consider before they share. https://newsroom.tiktok.com/en-us/new-prompts-to-help-people-consider-before-they-share

Hershcovich, D., Frank, S., Lent, H., de Lhoneux, M., Abdou, M., Brandl, S., & Søgaard, A. (2022). Challenges and strategies in cross-cultural NLP. arXiv preprint arXiv:2203.10020. https://doi.org/10.48550/arXiv.2203.10020

Hu, L., Wei, S., Zhao, Z., & Wu, B. (2022). Deep learning for fake news detection: A comprehensive survey. *AI Open*, *3*, 133–155. https://doi.org/10.1016/j.aiopen.2022.09.001

Hunt, K., Agarwal, P., Al Aziz, R., & Zhuang, J. (2020a). Fighting fake news during disasters. *OR/MS Today*, *47*(1), 34–39. https://www.acsu.buffalo.edu/~kylehunt/ORMS%20Today%20(preprint).pdf

Hunt, K., Wang, B., & Zhuang, J. (2020b). Misinformation debunking and cross-platform information sharing through Twitter during hurricanes Harvey and Irma: A case study on shelters and ID checks. *Natural Hazards*, *103*(1), 861–883. https://doi.org/10.1007/s11069-020-04016-6

Hunt, K., Agarwal, P., & Zhuang, J. (2022). Monitoring misinformation on Twitter during crisis events: A machine learning approach. *Risk Analysis*, *42*(8), 1728–1748. https://doi.org/10.1111/risa.13634

Huriye, A. Z. (2023). The ethics of artificial intelligence: Examining the ethical considerations surrounding the development and use of AI. *American Journal of Technology*, *2*(1), 37–44. https://doi.org/10.58425/ajt.v2i1.142

Hussein, M. R., Shams, A. B., Rahman, A., Raihan, M. S., Mostari, S., Siddika, N.,..., & Apu, E. H. (2020). Real-time credible online health information inquiring: A novel search engine misinformation notifier extension (SEMiNExt) during COVID-19-like disease outbreak. https://doi.org/10.21203/rs.3.rs-60301/v1

Imran, M., Alam, F., Ofli, F., Aupetit, M., & Bin Khalifa, H. (2017). *Enabling rapid disaster response using artificial intelligence and social media*. Qatar Computing Research Institute. https://mimran.me/papers/enabling_rapid_disaster_response_using_artificial_intelligence_and_social_media.pdf

Ivanov, M., & Petrova, E. (2024). Guarding mobile networks: Leveraging AI and ML for enhanced security and misinformation detection. *Innovative Computer Sciences Journal*, *10*(1), 1–9. https://innovatesci-publishers.com/index.php/ICSJ/article/view/14

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, *37*(3), 101493. https://doi.org/10.1016/j.giq.2020.101493

John, S. A., & Keikhosrokiani, P. (2022). COVID-19 fake news analytics from social media using topic modeling and clustering. In *Big data analytics for healthcare* (pp. 221–232). Academic Press. https://doi.org/10.1016/B978-0-323-91907-4.00003-0

Juntunen, M., & Lehenkari, M. (2021). A narrative literature review process for an academic business research thesis. *Studies in Higher Education*, *46*(2), 330–342. https://doi.org/10.1080/03075079.2019.1630813

Kankanamge, N., Yigitcanlar, T., & Goonetilleke, A. (2021). Public perceptions on artificial intelligence driven disaster management: Evidence from Sydney, Melbourne and Brisbane. *Telematics and Informatics*, *65*, 101729. https://doi.org/10.1016/j.tele.2021.101729

Koebler, J., & Cox, J. (2018). The impossible job: Inside facebook's struggle to moderate two billion people. https://www.vice.com/en/article/how-facebook-content-moderation-works/

Kolarević, E. (2022). The influence of artificial intelligence on the right of freedom of expression. *Pravo Teorija i Praksa*, *39*(1), 111–126. https://doi.org/10.5937/ptp2201111k

Korta, S. (2018). Fake news, conspiracy theories, and lies: An information laundering model for homeland security. *Homeland Security Affairs*. https://apps.dtic.mil/sti/citations/AD1052689

Kumar, V., Kumar, A., Singh, A. K., & Pachauri, A. (2021). Fake news detection using machine learning and natural language processing. In *2021 international conference on technological advancements and innovations (ICTAI)* (pp. 547–552). IEEE. https://doi.org/10.1109/ICTAI53825.2021.9673378

Leitner, S., Gula, B., Jannach, D., Krieg-Holz, U., & Wall, F. (2021). Understanding the dynamics emerging from infodemics: A call to action for interdisciplinary research. *SN Business & Economics*, *1*(1), 23. https://doi.org/10.1007/s43546-020-00027-4

Lovari, A. (2020). Spreading (dis) trust: Covid-19 misinformation and government intervention in Italy. *Media and Communication*, *8*(2), 458–461. https://doi.org/10.17645/mac.v8i2.3219

Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, *3*(1), 91–99. https://doi.org/10.1016/j.gltp.2022.04.020

Manheim, K., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy. *Yale JL & Tech*, *21*, 106. https://heinonline.org/HOL/LandingPage?handle=hein.journals/yjolt21&div=4&id=&page=

Meesad, P. (2021). Thai fake news detection based on information retrieval, natural language processing and machine learning. *SN Computer Science*, *2*(6), 425. https://doi.org/10.1007/s42979-021-00775-6

Méndez-Muros, S., Alonso-González, M., & Pérez-Curiel, C. (2024). Disinformation and fact-checking in the face of natural disasters: A case study on Turkey–Syria earthquakes. *Societies*, *14*(4), 43. https://doi.org/10.3390/soc14040043

Meta. (2020). *Here's how we're using AI to help detect misinformation*. https://ai.meta.com/blog/heres-how-were-using-ai-to-help-detect-misinformation/

Miller, D. S. (2016). Public trust in the aftermath of natural and na-technological disasters: Hurricane Katrina and the Fukushima Daiichi nuclear incident. *International Journal of Sociology and Social Policy*, *36*(5/6), 410–431. https://doi.org/10.1108/IJSSP-02-2015-0030

Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021). Fake news is not simply false information: A concept explication and taxonomy of online content. *American Behavioral Scientist*, *65*(2), 180–212. https://doi.org/10.1177/0002764219878224

Muaad, A. Y., Jayappa Davanagere, H., Benifa, J. B., Alabrah, A., Naji Saif, M. A., Pushpa, D.,... & Alfakih, T. M. (2022). Artificial intelligence-based approach for misogyny and sarcasm detection from arabic texts. *Computational Intelligence and Neuroscience*, *2022*(1), 7937667. https://doi.org/10.1155/2022/7937667

Murray, D., Fussey, P., Hove, K., Wakabi, W., Kimumwe, P., Saki, O., & Stevens, A. (2024). The chilling effects of surveillance and human rights: Insights from qualitative research in Uganda and Zimbabwe. *Journal of Human Rights Practice*, *16*(1), 397–412. https://doi.org/10.1093/jhuman/huad020

Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, *1*(1), 100007. https://doi.org/10.1016/j.jjimei.2020.100007

Nishi, K., Ding, Y., Rich, A., & Höllerer, T. (2021). Improving label noise robustness with data augmentation and semi-supervised learning (student abstract). *Proceedings of the AAAI Conference on Artificial Intelligence,* 35(18), 15855–15856. https://doi.org/10.1609/aaai.v35i18.17924

Nugraha, T. (2023). Managing cybersecurity risks related to ai in developing countries: Challenges and strategies part I. *Modern Diplomacy*. https://moderndiplomacy.eu/2023/12/20/managing-cybersecurity-risks-related-to-ai-in-developing-countries-challenges-and-strategies-part-i/

Oh, J., & Lee, D. (2022). Role of trust in government and collaboration in building disaster resilience. *Social Science Quarterly*, *103*(7), 1647–1658. https://doi.org/10.1111/ssqu.13229

Oh, O., Kwon, K. H., & Rao, H. R. (2010). An exploration of social media in extreme events: Rumor theory and Twitter during the Haiti earthquake 2010, ICIS 2010 Proceedings, 231. https://aisel.aisnet.org/icis2010_submissions/231

Omar, S., & Van Belle, J. P. (2024). Disaster misinformation management: Strategies for mitigating the effects of fake news on emergency response. In *International conference on information technology & systems* (pp. 308–318). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54235-0_29

PEW Research Center. (2023). Support for the U.S. government and tech companies restricting false information online has risen steadily in recent years. https://www.pewresearch.org/short-reads/2023/07/20/most-americans-favor-restrictions-on-false-information-violent-content-online/sr_2023-07-20_contentmod_1-png/

Poddar, K., & Umadevi, K. S. (2019). Comparison of various machine learning models for accurate detection of fake news. In *2019 Innovations in power and advanced computing technologies (i-PACT)* (Vol. 1, pp. 1–5). IEEE. https://doi.org/10.1109/i-PACT44901.2019.8960044

Praveenraj, D. D. W., Victor, M., Vennila, C., Alawadi, A. H., Diyora, P., Vasudevan, N., & Avudaiappan, T. (2023). Exploring explainable artificial intelligence for transparent decision making. *E3S Web of Conferences, 399*, 04030. https://doi.org/10.1051/e3sconf/202339904030

Preuveneers, D., & Joosen, W. (2021). Sharing machine learning models as indicators of compromise for cyber threat intelligence. *Journal of Cybersecurity and Privacy*, *1*(1), 140–163. https://doi.org/10.3390/jcp1010008

Pritam, K., Chaubey, A., & Dhole, A. (2024). Advancements and methodologies in natural language processing and machine learning: A comprehensive review. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, *12*(VI), 1495. https://doi.org/10.22214/ijraset.2024.63359

Rahma, A., Azab, S. S., & Mohammed, A. (2023). A comprehensive review on arabic sarcasm detection: Approaches, challenges and future trends. *Ieee Access: Practical Innovations, Open Solutions, 11*, 18261–18280. https://ieeexplore.ieee.org/document/10049545/

Rajarathinam, R., Bhuvaneswari, A., & Jothi, R. (2022). A real-time crisis informatics system by fusing geo-spatial Twitter streams and user-posted images. *Materials Today: Proceedings*, *62*, 4738–4744. https://doi.org/10.1016/j.matpr.2022.03.252

Ramašauskaitė, O. (2023). The role of collaborative networks in combating digital disinformation. In *2. International conference on economics regional development-digital economy: Proceedings book. December 21–23, 2023* (pp. 432–437). Liberty Academic Publishers. https://vb.mruni.eu/object/elaba:184652082/

Ramya, R. V., & Singh, R. K. (2024). The role of artificial intelligence in climate change visualisation. *ldealistic Journal of Advanced Research in Progressive Spectrums (IJARPS)*, *3*(04), 48–54. https://journal.ijarps.org/index.php/IJARPS/article/view/341

Reddy, M. S., Vamsi, C., & Kathambari, P. (2024). Rescue me: AI emergency response and disaster management system. In *2024 2nd international conference on artificial intelligence and machine learning applications theme: Healthcare and internet of things (AIMLA)* (pp. 1–5). IEEE. https://doi.org/10.1109/AIMLA59606.2024.10531386

Rodríguez-Barroso, N., Stipcich, G., Jiménez-López, D., Ruiz-Millán, J. A., Martínez-Cámara, E., González-Seco, G., & Herrera, F. (2020). Federated learning and differential privacy: Software tools analysis, the Sherpa. AI FL framework and methodological guidelines for preserving data privacy. *Information Fusion*, *64*, 270–292. https://doi.org/10.1016/j.inffus.2020.07.009

Rosen, G. (2020). *An update on our work to keep people informed and limit misinformation about COVID-19*. https://about.fb.com/news/2020/04/covid-19-misinfo-update/

Rossi, A., & Lenzini, G. (2020). Transparency by design in data-informed research: A collection of information design patterns. *Computer Law & Security Review*, *37*, 105402. https://doi.org/10.1016/j.clsr.2020.105402
Rubin, V. L. (2022). The problem of misinformation and disinformation online. *Misinformation and disinformation: Detecting fakes with the Eye and AI* (pp. 3–28). Springer International Publishing. https://doi.org/10.1007/978-3-030-95656-1_1
Shahid, W., Jamshidi, B., Hakak, S., Isah, H., Khan, W. Z., Khan, M. K., & Choo, K. K. R (2022). Detecting and mitigating the dissemination of fake news: Challenges and future research opportunities. *IEEE Transactions on Computational Social Systems*. https://doi.org/10.1109/TCSS.2022.3177359
Shams, A. B., Apu, H., Rahman, E., Sarker Raihan, A., Siddika, M. M., Preo, N., & Kabir, R. B. (2021). R. Web search engine misinformation notifier extension (SEMiNExt): A machine learning based approach during COVID-19 pandemic. *Healthcare,* 9(2), 156. https://doi.org/10.3390/healthcare9020156
Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36. https://doi.org/10.1145/3137597.3137600
Silverman, C. (2015). Lies, damn lies and viral content. https://doi.org/10.7916/D8Q81RHH
Singh, S. (2018). Natural language processing for information extraction. arXiv preprint arXiv:1807.02383. https://doi.org/10.48550/arXiv.1807.02383
Sommers, S. R., Apfelbaum, E. P., Dukes, K. N., Toosi, N., & Wang, E. J. (2006). Race and media coverage of Hurricane Katrina: Analysis, implications, and future research questions. *Analyses of Social Issues and Public Policy*, *6*(1), 39–55. https://doi.org/10.1111/j.1530-2415.2006.00103.x
Stahl, B. C., & Wright, D. (2018). Ethics and privacy in AI and big data: Implementing responsible research and innovation. *IEEE Security & Privacy*, *16*(3), 26–33. https://doi.org/10.1109/MSP.2018.2701164
Stoycheff, E. (2016). Under surveillance: Examining Facebook's spiral of silence effects in the wake of NSA internet monitoring. *Journalism & Mass Communication Quarterly*, *93*(2), 296–311. https://doi.org/10.1177/1077699016630255
Su, J., de Prado, M., Dahyot, R., & Saeed, R. (2019). AI Pipeline-bringing AI to you. End-to-end integration of data, algorithms and deployment tools. In *Emerging deep learning accelerators (EDLA) Workshop at HiPEAC 2019* (pp. 1–9). Emerging deep learning accelerators (EDLA) Workshop at HiPEAC 2019. https://mural.maynoothuniversity.ie/15157/1/RD_AI%20pipeline.pdf
Sun, L. G. (2010). Disaster mythology and the law. *Cornell Law Review, 96*, 1131. https://heinonline.org/HOL/LandingPage?handle=hein.journals/clqv96&div=44&id=&page=
Sun, L. G. (2012). Disaster mythology and availability cascades. *Duke Environmental Law & Policy Forum*, *23*, 73. https://heinonline.org/HOL/LandingPage?handle=hein.journals/delp23&div=7&id=&page=
Svendsen, E. R. (2013). A new perspective on radiation risk communication in Fukushima, Japan. *Journal of the National Institute of Public Health*, *62*(2), 196–203. https://www.niph.go.jp/journal/data/62-2/201362020012.pdf
Okolo, C. T. (2020). AI in the real world: Examining the impact of AI deployment in low-resource contexts. *arXiv Preprint arXiv: 2012 01165*. https://doi.org/10.48550/arXiv.2012.01165
Teh, D., & Khan, T. (2021). Types, definition and classification of natural disasters and threat level. *Handbook of disaster risk reduction for Resilience: New Frameworks for Building Resilience to disasters* (pp. 27–56). Springer International Publishing. https://doi.org/10.1007/978-3-030-61278-8_2
Tierney, K., Bevc, C., & Kuligowski, E. (2006). Metaphors matter: Disaster myths, media frames, and their consequences in Hurricane Katrina. *The Annals of the American Academy of Political and Social Science*, *604*(1), 57–81. https://doi.org/10.1177/0002716205285589
Timberg, C., Room, T., & Greene, J. (2020). Tech firms take a hard line against coronavirus myths. But what about other types of misinformation? *The Washington Post*. https://www.washingtonpost.com/technology/2020/02/28/facebook-twitter-amazon-misinformation-coronavirus/
Tiwari, R. (2023). Explainable Ai (Xai) and its applications in building trust and understanding in Ai decision making. *International J Sci Res Eng Manag*, *7*, 1–13. https://doi.org/10.55041/IJSREM17592
Tran, T., Valecha, R., Rad, P., & Rao, H. R. (2020). Misinformation harms: A tale of two humanitarian crises. *IEEE Transactions on Professional Communication*, *63*(4), 386–399. https://doi.org/10.1109/TPC.2020.3029685
Tsai, C. M. (2023). Stylometric fake news detection based on natural language processing using named entity recognition: In-domain and cross-domain analysis. *Electronics*, *12*(17), 3676. https://doi.org/10.3390/electronics12173676

Van de Walle, B., & Dugdale, J. (2012). Information management and humanitarian relief coordination: Findings from the Haiti earthquake response. *International Journal of Business Continuity and Risk Management*, *3*(4), 278–305. https://doi.org/10.1504/IJBCRM.2012.051866

Varsha, V. R., Naganandini, S., & Hariharan, C. (2024). Utilizing AI and machine learning for natural disaster management: Predicting natural disasters with AI and machine learning. In *Internet of things and ai for natural disaster management and prediction* (pp. 279–304). IGI Global. https://doi.org/10.4018/979-8-3693-4284-8.ch013

Velev, D., & Zlateva, P. (2023). Challenges of artificial intelligence application for disaster risk management. *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, *48*, 387–394. https://doi.org/10.5194/isprs-archives-XLVIII-M-1-2023-387-2023

Veluru, C. S. (2024). Responsible artificial intelligence on large scale data to prevent misuse, unethical challenges and security breaches. *Journal of Artificial Intelligence & Cloud Computing. SRC/JAICC-349. DOI: doi.org/10.47363/JAICC/2024 (3)*, *331*, 2–6. https://doi.org/10.47363/jaicc/2024(3)331

Venkadesh, P., Divya, S. V., Marymariyal, P., & Keerthana, S. (2024). Predicting natural disasters with AI and machine learning. In *Utilizing AI and machine learning for natural disaster management* (pp. 39–64). IGI Global. https://www.igi-global.com/chapter/predicting-natural-disasters-with-ai-and-machine-learning/345853

Von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, *34*(4), 1607–1622. https://doi.org/10.1007/s13347-021-00477-0

Vuruma, S. K. R., Margetts, A., Su, J., Ahmed, F., & Srivastava, B. (2024). From cloud to edge: Rethinking generative AI for low-resource design challenges. arXiv preprint arXiv:2402.12702. https://doi.org/10.48550/arXiv.2402.12702

Wagner, J., & Roopaei, M. (2020). Edge based decision making in disaster response systems. In *2020 10th annual computing and communication workshop and conference (CCWC)* (pp. 0469–0473). IEEE. https://doi.org/10.1109/CCWC47524.2020.9031219

Wahl, B., Cossy-Gantner, A., Germann, S., & Schwalbe, N. R. (2018). Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings? *BMJ Global Health*, *3*(4), e000798. https://doi.org/10.1136/bmjgh-2018-000798

Wang, R. Q., Hu, Y., Zhou, Z., & Yang, K. (2020). Tracking flooding phase transitions and establishing a passive hotline with AI-enabled social media data. *Ieee Access: Practical Innovations, Open Solutions*, *8*, 103395–103404. https://doi.org/10.1109/ACCESS.2020.2994187

Xie, Q., Dai, Z., Hovy, E., Luong, T., & Le, Q. (2020). Unsupervised data augmentation for consistency training. *Advances in Neural Information Processing Systems*, *33*, 6256–6268. https://proceedings.neurips.cc/paper/2020/hash/44feb0096faa8326192570788b38c1d1-Abstract.html

Yang, J. Z., & Zhuang, J. (2020). Information seeking and information sharing related to Hurricane Harvey. *Journalism & Mass Communication Quarterly*, *97*(4), 1054–1079. https://doi.org/10.1177/1077699019887675

Zhang, Z. (2024). Advancements and challenges in AI-driven language technologies: From natural language processing to language acquisition. *Applied and Computational Engineering*, *57*, 146–152. https://doi.org/10.54254/2755-2721/57/20241325

Zhou, C., Xiu, H., Wang, Y., & Yu, X. (2021). Characterizing the dissemination of misinformation on social media in health emergencies: An empirical study based on COVID-19. *Information Processing & Management*, *58*(4), 102554. https://doi.org/10.1016/j.ipm.2021.102554