AI and Fake News: Unpacking the Relationships within the Media Ecosystem

Communication & Journalism Research
12 (1) pp 15-32

Reprints and Permissions:
masscomhod@uoc. ac. in
ISSN 2348 – 5663

Deepak P *
Queen's University Belfast, UK

Abstract

The field of AI has been rapidly evolving in recent years. We are in an era where AI is being applied within each and every sector of human activity. The relationship between AI and fake news is one of monumental importance, since such dynamics have significant consequences for the foundations of democratic societies. In this paper, we consider the relationship between AI and fake news in with particular reference to various segments of the media ecosystem that relate to the life cycle of fake news. Our analysis focuses on the segments of generation, propagation and mitigation of fake news. Within each of these segments, we unpack the relationships between AI and fake news, attempting to lay bare the intricacies of the connections. Our analysis reveals that the emergence of AI could potentially turbocharge fake news of various kinds, and this could have profound consequences for the sustainability of the media ecosystem. We hope that our analysis will inform deliberations on how the media ecosystem would be re-imagined to function meaningfully and sustainably in the AI era.

Keywords:

Fake news, AI, Media ecosystem, Attention economy

Introduction

The emergence of online media and social media in the early 2000s was a significant disruptive force within the media ecosystem. The disruption was caused by the diversification of the agency of content creation from just a few media outlets to encompass a wide variety of online media sources, and even the whole user base of social media platforms. This facilitated the emergence of a barrage of fake news (aka disinformation), with significant implications for democracy as illustrated during the time of the 2016 US presidential elections , and numerous other elections beyond. We are in the middle of another radical disruption in this space, that by AI. The complexity of AI, and its anthropomorphic presentation (as popular in systems like ChatGPT[3], has led to a murky and ambiguous picture making it difficult to analyze how AI would disrupt the media ecosystem within the context of fake news.

---

* Correspondence: deepaksp@acm.org

[3] https://openai.com/product/chatgpt

In this paper, we consider three segments of the media ecosystem that are particularly relevant to the context of fake news viz., generation, propagation and mitigation. Within each of these segments, we analyze the nature of the emerging influence of AI, and analyze what AI portends within each of those segments.

The Emergence of Generative AI

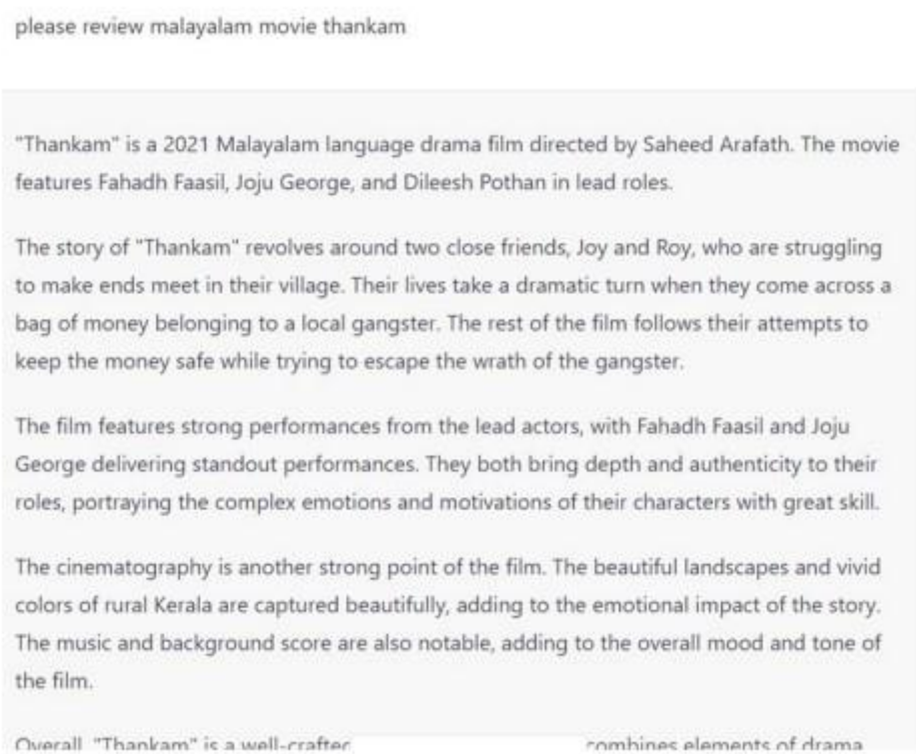
While generation has probably been the least attended AI area within news until recently, the emergence of generally available generative AI, pioneered by the public release of ChatGPT[4], has changed that significantly. This has been followed by a flurry of activity in the space of generative AI, resulting in the emergence of other large language models such as LLaMA[5] (from Meta/Facebook) and Bard[6] (from Google). Large language models are presented as general purpose AI by their creators, wherein they can be queried on any topic under the sun, to elicit responses. These language models were preceded by the emergence of image generation AI, such as DALL.E[7] which are capable of generating images from natural language descriptions. While the specific technological underpinnings are not of relevance to us at this stage, we now look at their defining functional character and the implications of such generative AI, and their general availability, to the news media ecosystem.

The Character of Generative AI

The key aspect about generative AI is that they are trained over very large datasets and assimilate patterns embedded within such datasets within their statistical machinery. This is then used to generate natural language text (or images, in the case of image-generative AI). The patterns they have assimilated make them capable of judging whether *'i had fried rice with chilli chicken'* is a more plausible sentence than *'i had fried rice with pizza'* or *'i had fried rice using chilli chicken'*. In this case, the first would be preferred since it likely aligns better with the patterns found in the large datasets. The second and third sentences are less likely when judged against the patterns in the training data, as far as generative AI is concerned. However, for a human, the second and third sentences are less preferred due to other reasons. While common-sense knowledge that fried rice does not make a good combination with pizza informs the human that the second sentence is not a good choice, it is the grammatical knowledge of english language that makes the third

[4] https://openai.com/product/chatgpt
[5] https://en.wikipedia.org/wiki/LLaMA
[6] https://blog.google/technology/ai/try-bard/
[7] https://openai.com/product/dall-e-2

sentence less preferable for the human. In other words, generative AI has preferences and lack of preferences for distinctly different reasons than humans. The absence of cognitive ability or capacity, common-sense knowledge and grammatical knowledge, are all offset by the abundance of training data. In this particular example, the human may be able to accept that somebody could eat fried rice along with pizza, they would still not consider the third sentence as a good example. However, this distinguishing capacity is not available to the large language model. In other words, it's judgment of language is limited to determining what aligns with the training data.



please review malayalam movie thankam

"Thankam" is a 2021 Malayalam language drama film directed by Saheed Arafath. The movie features Fahadh Faasil, Joju George, and Dileesh Pothan in lead roles.

The story of "Thankam" revolves around two close friends, Joy and Roy, who are struggling to make ends meet in their village. Their lives take a dramatic turn when they come across a bag of money belonging to a local gangster. The rest of the film follows their attempts to keep the money safe while trying to escape the wrath of the gangster.

The film features strong performances from the lead actors, with Fahadh Faasil and Joju George delivering standout performances. They both bring depth and authenticity to their roles, portraying the complex emotions and motivations of their characters with great skill.

The cinematography is another strong point of the film. The beautiful landscapes and vivid colors of rural Kerala are captured beautifully, adding to the emotional impact of the story. The music and background score are also notable, adding to the overall mood and tone of the film.

Fig. 1. ChatGPT's review of Malayalam movie *Thankam*

Being guided by the training data makes the generative AI capable of generating well-formed text that is aligned with the structures in the training data. This makes it a particularly useful tool in creating text with the semblance and syntax of well-authored text. However, the absence of common-sense knowledge, knowledge about the world, or cognitive capability, makes it a potential tool for creation of textual content that is highly divergent from what is plausible in the real world. Consider the example of a ChatGPT interaction pictured in Fig (1). This relates to the Malayalam-language movie Thankam[8]. It may be seen that ChatGPT gets the cast as well as the storyline, two critical aspects, totally wrong within its review. This indicates how it is able to generate plausible sounding text by putting together patterns that are abundant within the training data, but with little or no connection to reality.

Implications for News Ecosystem and Disinformation

Disinformation or fake news, may be generally defined as news that is divergent from the reality. As seen in the previous section, the character of generative AI makes it a potent force for generation of fake news, as also illustrated in Fig (1). This has several implications for the news ecosystem, of which we detail some here.

First, the general availability of generative AI makes it possible for a large profusion of disinformation to be generated freely by various actors in the web. This increases the impetus on various actors in news media organizations to undertake significant efforts at fact-checking, undermining the time and effort they have to attend to other aspects of the news ecosystem.

Second, the fact that generative AI methods are able to generate content that is grammatically well-structured, and bear semblance to realistic content, makes the fact-checking process harder. As an example, in Fig (1), ChatGPT suggests that the movie *Thankam* portrays two characters, played by Fahadh Fasil and Joju George, engaged in a thriller type plot. While these two actors have not been featured in the movie in question, they have been featured together in several thriller-type plots in other movies; such patterns are what likely guided ChatGPT towards that combination of actors to be included in its review. Those familiar with Malayalam movies and the names of these actors would find themselves able to believe in the review in the sense that it is entirely plausible, to a very nuanced level, that these two actors be featured in a plot as narrated. This indicates that the fact-checking process at news media organizations are significantly harder in the face of generative AI, which is another challenge.

Third, the availability of generative AI technologies for both text and images, opens up the possibility of generation of efficient fake news that combines both capabilities. Actors within the malicious fake news generation sphere may be able to put together such separate technologies in meaningful ways to generate very vivid fake news. While we are yet to see such amalgamated

[8] https://en.wikipedia.org/wiki/Thankam_(film)

capabilities emerge as yet, these could pose interesting new challenges to the media ecosystem.

Fig. 2. Image editing using NLP interfaces

Fourth, there is an emerging literature within the broader ambit of generative AI that is focused on image editing using natural language interfaces. This makes use of a technology called stable diffusion, and is illustrated in Fig (2). News reporters who were at the scene of an actual event would now be able to take real pictures, and edit it slightly to carry their intended narrative. These kinds of fake news could be significantly harder to fact-check even for another reporter who has been at the event scene.

As one can observe, the emergence of generative AI poses significant challenges to news media organizations. In particular, news media organizations would be faced with a significant mounting challenge of engaging in high-bandwith fact checking at the expense of time that could be spent on other aspects of the news process, to retain and enhance their credibility.

News Propagation and AI

The propagation and distribution of news has changed significantly with the emergence of digital technologies and the web. An important aspect of news propagation is also the modality of consumption of news. The unmistakable radical change in news consumption is there for us to see: news consumption moved from print media to online media and social media. Online news sources have interleaved ad slots which are filled in a personalized manner for the reader using AI technologies. On the other hand, AI technologies have a far greater role in social media in that the ordering of the social media feed, and thus, the determination of what the user should read, is itself left to AI technologies. These questions are particularly political, in the sense, these technologies could improve visibility for some news than others, and such preferences are not based on quality of the news, but by other factors as we will see shortly.

The Attention Economy

The attention economy is a term that is used to denote the economy of the digital world that is dominated by services that are offered free - bereft of any payment - to consumers. The operational model of these services may be characterized by the quote: *If you're not paying for the product, then you are the product*[9]. These platforms work on the model that user attention be captured through their service offerings, so that a part of such attention can be channelized to advertisements in exchange for payments from advertisers. The attention economy has spawned huge corporations, insofar that services that use this model such as Google and Facebook have become among the most valuable companies in terms of market capitalization[10]. The modalities of the attention economy significantly affect the prioritization used by AI algorithms within platforms that use the model.

It has been argued that AI has heralded the ressurection of technological determinism , the idea that the technological progress within a society is determined by the internal logic of efficiency, and such progress has significant consequences for the evolution of societies. While these ideas have roots in Marxist literature, these have also been explored significantly within media studies, most prominently by Neil Postman , an eminent media scholar. Within the attention economy, the internal logic of efficiency translates to the

---

[9] https://blogs.cornell.edu/rosescholarsfall2020/2020/12/18/if-youre-not-paying-for-the-product-then-you-are-the-product/

[10] https://companiesmarketcap.com/

urge to maximize user attention, so that the platforms can create *'value'* by redirecting such attention to advertisers.

Fig. 3. Two Images of Coronavirus Vaccines

Maximizing attention implicitly leads to an ethos of engagement maximization; engagement has intricate connections with disinformation . As a simple example, consider the two Coronavirus vaccines pictured in Fig (3). The left-side picture indicates a vaporizer vaccine, one that doesn't exist (at least, did not exist when this picture was captured), whereas the right-side illustrates a regular vaccine. It is implicitly more engaging to see the left-side vaccine, whereas a user may not even pause scrolling on seeing the run-of-the-mill right-side image. Disinformation, as not constrained by conformance to the real-world, can produce such highly interesting, and thus engaging stories. Thus, the engagement-maximizing logics of the attention economy are implicitly aligned towards disinformation.

Confirmation Bias and Affect

Apart from interestingness as a channel through which the attention economy privileges disinformation, there are at least two other distinct channels. These are based on a juxtaposition of insights from psychology on engagement maximization; we further detail them herein.

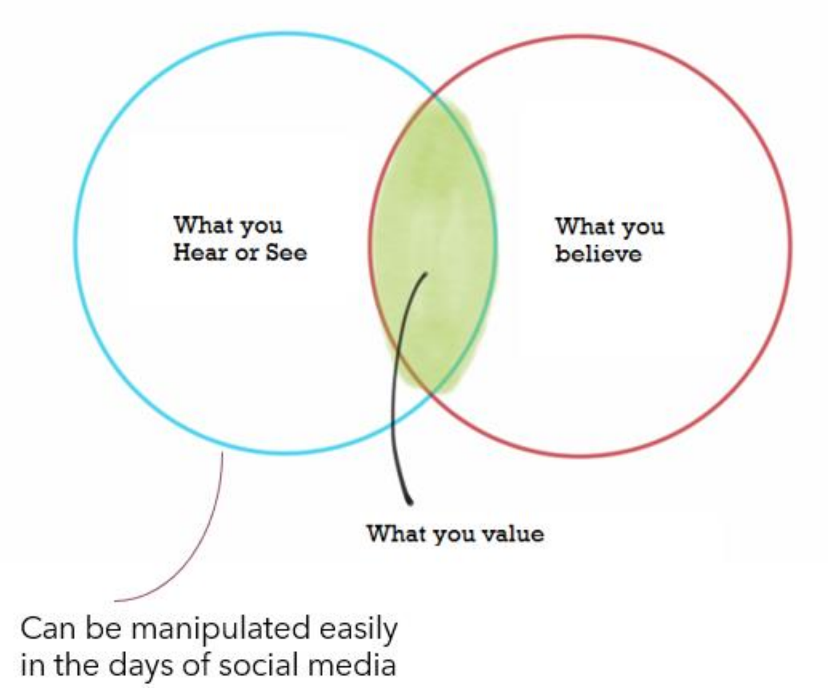


Fig. 4. Illustration of Confirmation Bias

Confirmation Bias:

Confirmation bias is a well-understood cognitive fallacy. This suggests that a user who has some hypothesis in her mind, would find, interpret and remember information in such a way that it systematically impedes the possibility that the hypothesis be rejected. We will illustrate this using an example. In the classical Indian debate relating to marriages viz., *arranged vs. love*, a person who prefers *arranged* is likely to view news about durable arranged marriages as validating her hypothesis, while interpreting failed arranged marriages as one-off events. An attention economy platform that understands this predisposition of the user would benefit from prioritizing news relating to failed love marriages and durable arranged marriages in her feed, since that would cater to her confirmation bias leading to better engagement and more user satisfaction.

Consider Fig (4), which illustrates confirmation bias pictorially. The user is likely to value those that are in the intersection of what she hears or sees, *and* what she believes. In the days of traditional print media, what one hears or sees is quite rigid, in that the news that would go into a print newspapers is determined by the news value and the general interest in the population, rather than being determined by a particular reader's beliefs. However, the rigidity of what is heard or seen is significantly relaxed in the days of AI, since the personalized information delivery paradigm allows the what is heard or seen to be algorithmically and individually determined. This allows an additional degree of flexibility (complementary to interestingness, as we saw earlier) for disinformation actors to leverage, one that is highly individualized.

In fact, such confirmation bias effects in friends recommendation and other aspects of social media have led to what is often called as *echo chambers*, a collection of people who are held together by a strong overlap between their beliefs. Thus, psychological phenomena such as confirmation bias operating in social media have also determined the nature of connections, an interesting fallout of the broader idea of technological determinism. Echo chambers are ways by which confirmation bias is leveraged by disinformation actors by expanding the remit from an individualized setting (as discussed in the prior paragraph) to a social setting. The relationships between echo chambers and disinformation have been subject to much research .

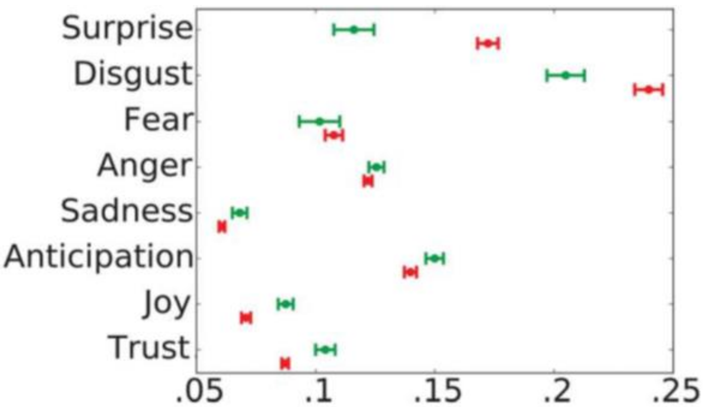


Fig. 5. Emotions and Fake News: Image from [21]

Affect and Fake News:

The relationship between affect (viz., sentiments and emotions) and disinformation is another facet involving significant user engagement gradients. An apparent connection has been between clickbaity headlines, affect and disinformation . In particular, fake news has shown a marked departure in the intensity and kind of emotional content as compared to real news, which could potentially partly explain its virality. Fig (5), is a figure from which indicates the intensity of emotional content between fake and real news. For each emotion, the intensity for fake news is indicated using a *red* range, whereas the intensity for real news is indicated using a *green* range. The figure indicates that fake news has a significantly higher intensity of surprise and disgust, and is shallower on joy and trust.

If indeed there is a causal relationship between emotion profiles and virality (note that virality is positively related to engagement), this adds yet another degree of freedom for disinformation actors to help attention economy platforms maximize engagement. It is to be noted that we do not imply that there is collusion between disinformation actors and attention economy platforms, but simply that disinformation actors could inmplicitly benefit themselves and the platforms by creating a large volume of highly engaging content which could parameterize engagement maximizing AI within the platforms.

Implications for News Ecosystem and Disinformation

Much like in the case of generation, the growing influence of AI within news propagation and distribution, holds out significant implications for the news ecosystem. These are along several facets.

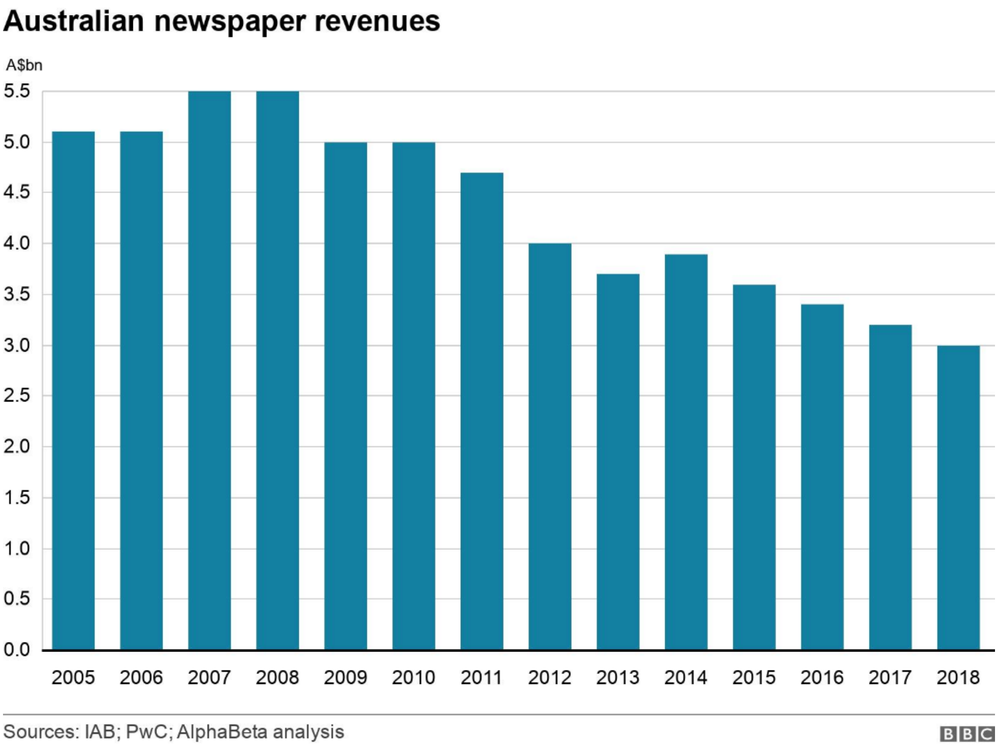


Fig. 6. Australian News Media Revenue Trends. Source: BBC

First, the patterns of news consumption in the online world have shifted the base of consumption from news media outlets (e.g., websites) to attention economy platforms (e.g., social media, and aggregation platforms such as Google News[11]). The aggregation platforms have an apparent advantage in the eyes of news consumers in that they comprise information from across a plurality of media outlets, thus providing a semblance of providing a well-rounded view. This movement of consumption from media outlets to attention economy platforms has had significant ramifications, one of which is that relating to revenue. Fig (6) illustrates the fall of revenues for Australian newspapers over the course of the last two decades. This was the topic of significant controversy amidst the Australian government's move[12] to impose a fairer sharing of revenues between attention economy platforms and news media which create the news stories. The decline of revenue for news media despite producing the news stories that the attention economy platforms use for their own profiteering posits a growing crisis in the media sector.

Second, the advent of AI-based platforms which leverage psychological and other know-how to increase user engagement has led to a situation where news stories that are not predisposed to create user engagement (through having one or more of several factors such as a clickbaity title, a suitable emotion profile, and high interestingness) are automatically given less visibility. Thus, news media houses are now faced with a new challenge in ensuring readership, that of having to pander to the logics of the engagement-focused attention economy. Note that this is different from traditional forms of increasing readership by showcasing adherence to high standards of journalistic practices, and ensuring visibly unbiased and fact-based reporting. This portends a situation where engagement-based reporting could eventually trump fact-based and unbiased reporting, with significant detrimental consequences to the quality of public discourse and prevalance of democracy.

We end this section with a prophetic observation from Neil Postman from his book on the future of the world, contrasting the theories of two eminent scholars. The quote was not within the context of the AI age but for a prior era; yet, it could be argued that it rings much truer in this age.

What Orwell feared were those who would ban books. What Huxley feared was that there would be no reason to ban a book, for there would be no one

[11] https://news.google.com/

[12] https://www.bbc.co.uk/news/world-australia-56107028

who wanted to read one. Orwell feared those who would deprive us of information. Huxley feared those who would give us so much that we would be reduced to passivity and egoism. Orwell feared that the truth would be concealed from us. Huxley feared the truth would be drowned in a sea of irrelevance. Orwell feared we would become a captive culture. Huxley feared we would become a trivial culture, preoccupied with some equivalent of the feelies, the orgy porgy, and the centrifugal bumblepuppy.

AI-based Fake News Mitigation

Unlike the media ecosystem segments of generation and propagation where AI may be broadly seen to pose fresh challenges or deepen existing ones, the segment of mitigation may be seen as one where AI could operate in a positive and constructive manner. Yet, the dynamics of AI within fake news mitigation are more nuanced than what may seem apparent. We first start with outlining the paradigms of AI and the high-level design aspects of AI as used within the space of fake news mitigation. This will be followed by a discussion on the issues arising from potential usage scenarios and disconnects between fact-checking paradigms used by expert humans and AI.

AI for Fake News Detection and Mitigation

The AI community has largely addressed the task of fake news detection by framing it as a conventional binary classification problem; a survey of AI-based methods for fake news mitigation appears at . Within the remit of binary classification, the modality is to gather a large dataset of news articles with each article labelled as either *fake* or *real*. These are then fed into a machine learning model that identifies patterns within news articles that are correlated with either of the labels. Patterns could involve words, word sequences, or other kinds of lexical regularities; for example, within a health news dataset, *X causes cancer* may be a pattern that is correlated with fake news, given that fake news is often used to advance fake claims with respect to cancer, which is often a fatal disease. These patterns are implicitly compiled into a statistical model, which can then be applied to hitherto unseen news articles. The model profiles such news articles by assessing their adherence to patterns that correlate with *fake* or *real* labels, and use such assessments to derive a label that would then be output as the decision for the article. The different methods for classification differ in the kinds of patterns they seek, and the kind of statistical model they use to store the patterns. While the narration so far may have given the impression that it is just the text content in the article that is used for the decision making, current AI models often use a more comprehensive feature representation. For example, feature representations used by AI models may involve the *source* of the article (e.g., media house, or top-level web domain), the *author* of the article, and metrics of engagement that they have generated within social media. This allows classification models to identify and exploit patterns of propagation that are preferentially correlated with one label than the other, and use that in the decision-making process. There are even classification methods that consider presence of emojis, hashtags and URLs as special features than just as part of the article contents.

While classification methods would require a fully labelled dataset where each article is labelled as either *real* or *fake,* there have been recent advances in AI that seek to exploit unlabelled data in building statistical decision-making models. Such models are called unsupervised models, and have been subject to much recent interest; an overview appears at . These often use prior knowledge of differences between real and fake news, and use that to bootstrap a learning paradigm. We will illustrate one such unsupervised learning paradigm using an example. Consider a large dataset of news articles that is devoid of any labeling. If we know that fake news articles tend to be highly oriented towards particular emotions, we could identify the top-k articles that have the highest density of such emotions and label them as *fake*; analogously, the top-k articles that have the least density of such emotions may be labeled as *real*. This gives us a starting point towards understanding the distinguishing character between fake and real news. Now, we could consider each unlabelled article and label it appropriately based on whether they are similar to the articles with seed labels as *real* or *fake*. Such progressive propagation of labels may then be spread across the entire dataset to achieve a full labelling of the dataset. The above paradigm of unsupervised labelling is not meant to typify all unsupervised methods, but instead to illustrate the spirit of unsupervised learning as applied to fake news detection. The nuanced reader would appreciate that the success of such methods depend on the initial heuristic (such as the emotions assumption in the above example) about what distinguishes between real and fake news. These exist several possibilities of designing a seeding heuristic based on understandings of real vis-a-vis fake news. These include the observation of multiple propagation peaks in fake news , more connected diffusion patterns in real news , behavior synchronicity among social media actors in fake news or the propensity of fake news to generate higher level of enquiry posts in social media .

Role of AI based Fake News Detection in Practical Settings

The implicitly intended role of AI based methods for fake news detection is to serve as an aid for news media personnel by providing veracity information on news nuggets to feed into their editorial processes, or to be embedded within news consumption as a veracity signal to the user. However, as highlighted in an article on the ethics of AI-based fake news detection , the framing of fake news detection as a technical problem does little justice to the ethical and normative considerations that are indispensable within the media sector. In particular, the alignment of such AI paradigms for fake news detection to extant recommendations such as those from the EU High-level Expert Group on Disinformation , is very poor.

While focuses on the usage of AI-based fake news detection methods within end-user facing scenarios, there are analogous issues that emanate from their usage within media houses. Consider, for example, the usage of such AI methods as a ’vetting’ process whereby an editor runs a prepared article by an AI method to see whether it would indeed be classified as *real*, prior to publication. If the AI method classifies the document as *fake*, what does it imply for the editor? Would the editor need to re-check the article for veracity even if they are confident about the article being non-fake? Or would the editor need to abandon publishing the article altogether? There are emerging methods in AI-based fake news detection that seek to provide explanations rather than just a binary decision (a survey appears at) which mitigate some high-level issues by providing further context to determine recourse actions. Yet, there are deeper disconnects which emanate from the reductionist view of fake news detection as a technical problem, a perspective that is embedded within extant AI techniques for the same. For instance, while it is convenient to view *fake* and *real* as crisp and distinct labels, most news articles lie somewhere in the real-fake spectrum. This could also differ based on the perspective applied by specific individuals, especially when it comes to political news; a left-wing viewpoint may be viewed by a right-wing sympathizer as closer to fake, and vice versa. Such nuances are hardly considered within extant literature on AI-based fake news detection.

Fake News Detection: AI vs. Human Expert

The complexities of using AI-based methods for fake news detection off-the-shelf are more evident when analyzing the working models of those techniques vis-a-vis the working model adopted by human fact-checking experts, who are employed by several media houses to defend themselves against the onslaught of fake news. While it is well-known that AI methods operate in a significantly different manner to humans across a variety of scenarios, the importance of such divergences in a highly consequential domain such as fake news detection have been highlighted recently .

While AI models employ inductive reasoning (e.g., *have articles of this style previously been found to be fake?*), expert humans employ common-sense reasoning (e.g., *is the described event even plausible?*). AI models lack human sensibilities and sensitivity and are dispassionate in their decision making, whereas expert judgments are influenced by the vulnerabilities expressed within the article (e.g., *who is portrayed in bad light? is this likely to deepen harm for vulnerable?*) and any implicit malicious intent reflected in the article (e.g., *who would profit financially from spreading this news?*). In another avenue of sharp contrast with state-of-the-art AI, humans implicitly leverage the *theory of mind* in understanding and assimilating the stances (e.g., political orientation, regional affiliation, misogyny) within the article; in other words, experts can make judgments on *where the author is coming from*, and factor those inferences meaningfully in making judgments. Most fact checkers, in making decisions, leverage *lateral reading* (understanding the context of the article from its source, authorship etc.) than *vertical reading* (reading the entire article as the primary judgment). In contrast, AI methods seek to form a comprehensive representation of an article, which makes it aligned with vertical reading. In summary, AI methods offset their deficits of sensibility, theory of mind, lateral reading and common sense through the abundance of data that they have at their disposal.

The sharp contrasts between human and AI methods are not without consequences. The *replacement of sensibility in decision making through data abundance* has marked the transition from expert fact checking to AI-driven fact checking and is at the heart of the fragility of the latter. While biases along race, gender, sexual orientation, religion and economic classes have been well studied in AI-based fake news detection, geo-political and cultural biases in fake news AI are starting to be recognized. The information ecosystem has been revolutionized and challenged very recently by the watershed moment that marked abundant end-user availability of generative AI viz., ChatGPT and its ilk, something we considered while discussing the generation segment. These technologies are capable of generating text that confirm to the *structure or semblance* of real news, but embedded with *content* that have no necessary connection with reality. The pattern-based decision making paradigm embodied within AI makes it systemically incapable of addressing the needs of such a rapidly changing media ecosystem. It is notable that the foundational pillars of human expert veracity judgments - viz., common sense, theory of mind, sensitivity and lateral reading - remain potent ammunition against these new developments. These highlight and foreground the need for a new breed of AI for fake news which closely align to human approaches to fact-checking, to ensure robust automation of fake news detection.

In summary, despite the tall promises and hopes raised by AI-based fake news detection, they need to go a long way - and in particular directions - for them to serve as a meaningful aid to curb fake news and to be integrated within the media ecosystem seamlessly.

Summary and Conclusions

In this paper, we considered the impacts that AI has with fake news within the context of the media ecosystem. We analyzed three segments of the media ecosystem as it applies to the fake news life cycle viz., generation, propagation and mitigation. We observed that generative AI has started to pose fresh challenges to the media ecosystem and actors within it, significantly enhancing pressures on the necessary time and depth for fake news detection. Within the propagation segment, we observed that the logics of the attention economy privilege disinformation through various pathways, including those aligned with current understandings on the psychology of the human mind. Coming to mitigation, a segment which is often presented as one where AI can play a meaningful and positive role, we analyzed the current paradigms of AI and their operational modalities. Critically analyzing AI within this setting, we argued that AI usage within disinformation is highly reductionist, and does little justice to the nuances of the task, and the operating models of media houses. We further observed that the divergence in the decision making styles between humans and AI could have epochal consequences within the realm of fake news.

In short, the rapidly evolving AI era poses fundamental challenges to the news media ecosystem within the context of fake news. These encompass both operational and short-term challenges, as well as systemic and long-term challenges. While deliberations on pathways on addressing these challenges are outside the remit of this paper, we hope that our analysis will inform debates on how media could function meaningfully and sustainably in an AI-dominated era.

References

1. Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. Journal of economic perspectives, 31(2):211–236, 2017. Title Suppressed Due to Excessive Length 15

2. AB Athira, SD Madhu Kumar, and Anu Mary Chacko. A systematic survey on explainable ai applied to fake news detection. Engineering Applications of Artificial Intelligence, 122:106087, 2023.

3. Erik Brynjolfsson and JooHee Oh. The attention economy: Measuring the value of free digital services on the internet. 2012.

4. Madeleine de Cock Buning et al. A multidimensional approach to disinformation. EU Expert Group Reports, 2018.

5. P Deepak. Ethical considerations in data-driven fake news detection. 2021.

6. P Deepak. On unsupervised methods for fake news detection. 2021.

7. Adrien Friggeri, Lada Adamic, Dean Eckles, and Justin Cheng. Rumor cascades. In proceedings of the international AAAI conference on web and social media, volume

8. pages 101–110, 2014. 8. Siva Charan Reddy Gangireddy, Cheng Long, and Tanmoy Chakraborty. Unsupervised fake news detection: A graph-based approach. In Proceedings of the 31st ACM conference on hypertext and social media, pages 75–83, 2020.

9. Gerd Gigerenzer. How to stay smart in a smart world: Why human intelligence still beats algorithms. MIT Press, 2022.

10. Mih´aly H´eder. Ai and the resurrection of technological determinism. INFORMACI ´ OS T ´ ARSADALOM: T ´ ARSADALOMTUDOM ´ ANYI FOLY ´ OIRAT ´ , 21(2):119–130, 2021.

11. Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. Prominent features of rumor propagation in online social media. In 2013 IEEE 13th international conference on data mining, pages 1103–1108. IEEE, 2013.

12. Qiwei Li. Clickbait and emotional language in fake news. Department of Computer Science. The University of Texas at Austin. EE. UU. Recuperado de https://bit. ly/2T4msNQ, 2019.

13. Stacy Miller, Philip Menard, David Bourrie, and Scott Sittig. Integrating truth bias and elaboration likelihood to understand how political polarisation impacts disinformation engagement on social media. Information Systems Journal, 2022.

14. C Thi Nguyen. Echo chambers and epistemic bubbles. Episteme, 17(2):141–161, 2020.

15. Margit E Oswald and Stefan Grosjean. Confirmation bias. Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory, 79, 2004.

16. Neil Postman. Technopoly. Bollati Boringhieri, Torino, 2003.

17. Neil Postman. Amusing ourselves to death: Public discourse in the age of show business. Penguin, 2005.

18. Wajiha Shahid, Bahman Jamshidi, Saqib Hakak, Haruna Isah, Wazir Zada Khan, Muhammad Khurram Khan, and Kim-Kwang Raymond Choo. Detecting and mitigating the dissemination of fake news: Challenges and future research opportunities. IEEE Transactions on Computational Social Systems, 2022.

19. Petter T¨ornberg. Echo chambers and viral misinformation: Modeling fake news as complex contagion. PLoS one, 13(9):e0203958, 2018.

20. Siva Vaidhyanathan. Antisocial media: How Facebook disconnects us and undermines democracy. Oxford University Press, 2018.

21. Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. science, 359(6380):1146–1151, 2018.

22. Zhe Zhao, Paul Resnick, and Qiaozhu Mei. Enquiring minds: Early detection of rumors in social media from enquiry posts. In Proceedings of the 24th international conference on world wide web, pages 1395–1405, 2015.