December 18, 2024

COMMENTS IN RESPONSE to Call for Consultation and Evidence on Social Media, Misinformation and Harmful Algorithms by UK Parliament[1]

The Association for Computing Machinery (ACM) is the world's largest and longest established professional society of individuals involved in all aspects of computing. It annually bestows the ACM A.M. Turing Award, often popularly called the "Nobel Prize of computing." ACM's Europe Technology Policy Committee ("Europe TPC") is charged with and committed to providing objective technical information to policymakers and the general public in the service of sound public policymaking.[2] Europe TPC has responded to the European Union (EU) stakeholder's consultations in the past in the context of the AI Ac[3], the Data Act,[4] the Digital Services Act[5], the Digital Citizen Principles[6], and the Cyber Resilience Act[7], amongst others[8]. ACM and Europe TPC are non-profit, non-political, and non-lobbying organisations.

Europe TPC is pleased to provide feedback on the UK parliament's consultation call on social media, misinformation, and harmful algorithms.

[1] The author of the document is: Ahmed Nagy. Additional assistance was provided by: Andrew McGettrick, Michel Beaudouin-Lafon, Advait Deshpande, Tom Romanoff

[2] https://www.acm.org/public-policy/europe-tpc

[3] https://www.acm.org/binaries/content/assets/public-policy/europe-tpc-comments-ai-consultation.pdf

[4] https://www.acm.org/binaries/content/assets/public-policy/acm-eur-tpc-data-act-comments-13may22a.pdf

[5] https://www.acm.org/binaries/content/assets/public-policy/europetpc-digital-services-act-comments.pdf

[6] https://www.acm.org/binaries/content/assets/public-policy/acm-europe-tpc-dsa-comments.pdf

[7] https://www.acm.org/binaries/content/assets/public-policy/europetpc-comments-digital-principles.pdf

[8] https://www.acm.org/binaries/content/assets/public-policy/acm-europe-tpc-cyber-reslience-comments-pdf

Summary

Riot events and the participation of social media platforms in ways that can disrupt society call for new types of social media and AI-powered systems. However, that should not be used as a reason to ban the usage of large language models or machine-generated content. It should be used as an opportunity to educate the stakeholders and put the proper guard rails. That will require an evolution of systems that could keep up with society's needs and consider the socio-technological interaction between the different parts to bring it to harmony. Early observations include a need for new generations of social media platforms that can bring algorithm transparency, prioritise content diversity, create explicit user control, make diversified revenue streams that can bring harmony to society, implement comprehensive data governance and avail full data ownership for users. Further observations include the need to couple social impact monetisation through content value monetisation. In addition, observations include the need for de-centralised architectures for social media platforms that rely on federated learning instead of centralised systems. Further, it is necessary to gradually add models and compliance audits for companies to avoid departing from markets and unpleasant economic drops. Such aspects would require consistent long-term investment and might need to attract socially conscious investors. While that does not come easily, resistance from existing corporate structures should be expected where short-term profitability concerns can dominate the sentiment of investors. Educating users about their data rights and how to use the platforms best goes hand in hand to ensure a harmonised society with properly functional stakeholders. It is necessary to point out that a holistic solution is necessary to address the shortcomings, and it is not a one-stakeholder effort.

Overview

Context: The main incident that triggered the request for consultation took place between 30 July and 7 August 2024. Anti-immigration demonstrations and riots swept across the UK, with mosques and hotels housing asylum seekers being targeted. The unrest was partly fueled by false claims on social media linking asylum seekers to the killing of three children in Southport. Ofcom highlighted the rapid spread of illegal content and disinformation during the crisis, pointing to the role of algorithmic recommendations in amplifying divisive narratives. The response from social media platforms to this harmful content has been described as inconsistent. The Online Safety Act 2023 mandates stricter measures against disinformation, requiring platforms to minimise illegal activity and remove unlawful content. In response to these events, the Science, Innovation and Technology Committee has launched an inquiry into the

links between algorithm-driven content ranking, generative AI, and the dissemination of harmful content. The coming sections will answer the inquiry systematically to address the existing and proposed regulations, including the Online Safety Act, and explore additional measures to address social harms linked to the aforementioned technologies (Dawes, 2024). The response structure will address the questions based on the consultation request.

To what extent do the business models of social media companies, search engines and others encourage the spread of harmful content, and contribute to wider social harms?

Social media companies and search engines are often profit-driven entities that thrive on user engagement, which in turn generates advertising revenue. Algorithms are designed to prioritise content that captures attention, leading to the amplification of sensational, emotionally charged, and often polarising material. Further, the existing laws that govern the relationship between users and companies have not made it easy for users to understand the implications of being lenient about their privacy and consistently targeted through mining their emotional state and their privacy-sensitive data items. As a result, users need to be empowered by understanding the impact of engaging content and other types of information that are balanced, authentic and non-polarising need to be shown to users. However, the business models existing for a wide range of social media platforms reward “engaging content.” a thematically engaging content can be news, which might be truthful yet excessively promoted and, at worst, completely fabricated evidence with supporting lies (O'neil 2017). Research has shown that:

- Sensational content gains more engagement: posts with divisive, shocking, or emotional language tend to garner higher interaction rates. As algorithms amplify content based on engagement, harmful content often rises to prominence (Bessi & Ferrara, 2016).
- Echo chambers and filter bubbles: platforms promote content that aligns with a user's preferences, reinforcing biases and exacerbating social polarisation (Pariser, 2011).
- Misinformation’s economic model: disinformation spreads quickly as a low-cost method of generating clicks and ad revenue, supported by fake news sites or coordinated campaigns. (Acemoglu 2024 , Aridor 2024)

While several companies occasionally claim to prioritise content moderation, the economic incentives of engaging content directly or indirectly fuel more engagement and act as a cornerstone in the business model for many companies. It is hard to depart from the algorithmic behaviour wired in the model for monetising and ensuring that users are rewarded for engaging content irrespective of the level of factual accuracy, social polarization, misinformation, and disinformation it can bring. Thus, this creates a conflict of interest in fighting misinformation and disinformation with the current business models for most companies. As a result, systematic

fighting of information and disinformation requires a holistic approach that will be discussed in two sections “**Additional Measures**” and “**Actionable Prioritised Recommendations**”.

How do social media companies and search engines use algorithms to rank content, how does this reflect their business models, and how does it play into the spread of misinformation, disinformation and harmful content?

Algorithms are the backbone of how content is ranked and presented. A lot of the models have ways to personalise the recommendations and maximise time spent and interaction. In order to do that, algorithms collect a wide range of analytics that are user-related, network-related, and general. The user-related analytics include interests, location, interaction speed, sentiment level, and polarity for topics for the user. While network-related analytics include centrality in the network connectedness and speed of dissemination for information and advertisement, and reputation among other members. Further, general analytics include time spent in the network, speed of building acquaintances and engaging with the content (data, advertisement, revenue generation), or content generation. A more comprehensive set of analytics could help drive more personalised content and ensure the network is utilised to the maximum. Following are some of the aspects that rely on the details of the algorithms used:

- Personalised recommendations: Algorithms like Facebook’s EdgeRank, YouTube's recommendation system (based on PageRank), or TikTok recommender systems use metrics such as clicks, likes, shares, and watch times to predict user preferences (Vombatkere 2024, Narayanan 2023).
- Business model integration: Their primary goal is to maximise user time spent on platforms, increasing ad views. This approach sometimes conflicts with content moderation goals, as clickbait or sensational material drives higher engagement (Zanker 2019).
- Misinformation amplification: Algorithms struggle to differentiate credible sources from misinformation, as both can perform similarly in terms of virality. Platforms like Twitter (now X) and YouTube have faced backlash for unintentionally spreading conspiracy theories or hate speech (Papakyriakopoulos 2022, Boeker 2022, Lee 2022).

Efforts like Twitter’s Community Notes or Facebook's third-party fact-checking partnerships are steps towards addressing the spread of harmful content. However, critics point to their limited scope and the underlying contradiction between these measures and the business model. As a result, there needs to be a new generation of platforms and systems capable of generating sustainable cash flow yet, at the same time, do not support and promote misinformation and disinformation items. There have been some steps from the industry in that regard, but still not enough; such efforts aim to inform the user about the data that is collected from them, and use

implicit feedback to turn off specific distribution efforts or feed-sharing items. Give users a chance to have full ownership of their data and run a lean server where the users' data are not stored on the company's server but rely on peer storage or third parties. Thus, such efforts could falsely convince the user that privacy is fully protected as full data ownership. However, that alone might not be enough to ensure a sustainable business model; the architecture of the data processing systems requires changes which should include federated learning, distributed processing, unlearning for machine learning algorithms, pseudo anonymity, and collective aggregation instead of individual reporting of clicks or interactions for users.

What role do generative artificial intelligence (AI) and large language models (LLMs) play in the creation and spread of misinformation, disinformation and harmful content?

While generative artificial intelligence techniques and large language models have huge benefits and applications, they come with some risks that need to be addressed. Generative AI and LLMs can play a significant role in both creating and spreading harmful content if misused:

- Creation of fake content: tools like ChatGPT or DALL-E can generate convincing fake news articles, videos, or images that are difficult to discern from authentic content.
- Rapid dissemination: misuse of AI models in automated bot networks can spread false narratives on a massive scale.
- Challenges of moderation: AI tools are trained on vast datasets that may include biased or harmful content, increasing the likelihood of inadvertently propagating harmful ideas. The concept of openness for the data set used in training these models has yet to mature (Liesenfeld 2024).

The rise of deepfake technologies and AI-enhanced bots raises further concerns, as they can be weaponised for propaganda, scams, or reputational damage. As long as there are no guard rails for the usage, that scenario becomes plausible. However, adding the proper compliance mechanisms and auditing the social media companies to ensure that there are sound mechanisms in place that decrease the chances of misuse and could raise flags when there is potential misuse.

What role did social media algorithms play in the riots that took place in the UK in summer 2024?

While the exact role might not be accurately identified, some of the impact can be analysed based on the routine operation of social media algorithms, which have been documented through research and shared by many of the white papers self-reported by the social media companies before the incident. As a result, while concrete evidence is pending, initial investigations indicate that the social media platforms were used to carry out:

- Amplification of violent content: social media algorithms might have highlighted inflammatory posts, making it easier for misinformation to spread.
- Coordination through private channels: encrypted platforms like Telegram likely played a role in organising protests. Public feeds, however, amplified provocative narratives.
- Mob mentality dynamics: viral posts may have fueled emotional responses, encouraging in-person riots; however, conclusive evidence is yet to be verified but initial investigations point to the weaponisation of mob mentality by viral posts that trigger emotional responses served and promoted by social media algorithms.

Historical examples, such as the role of social media in the 2011 UK riots, underline how platforms can catalyse unrest (Bell 2021).

However, it is necessary to take a conservative position from rushing and blaming large language models and generative artificial intelligence since such tools are not known to be integrated into social media platforms. While most of the platforms would not object to machine-created content, it is necessary to analyse how much of the posted content in thiscase were partially or fully created by a machine. A closer analysis should be carried out to understand how much machine-generated posts were promoted by machine bots and whether that dominates the native ranking algorithms of the social media platforms used.

A recent ranking for calculation of AI safety index was produced by the Future of Life Institute, (FLI) where the US GPA system for grade boundaries was constructed to rank the safety of AI used at the different social media companies who rely on AI in their algorithms. This quantified the ranking of the prominent companies; Meta, Facebook's parent company, and developer of the popular Llama series of AI models, was rated the lowest, scoring a F-grade overall. X.AI, an AI company powering X, formerly Twitter, also fared poorly, receiving a D- grade overall, while ChatGPT, OpenAI Google and DeepMind received a D+.. Further, Anthropic, the company behind Claude, which announced safety as a core part of its ethos, received a C grade for its AI engine (FLI AI 2024).

How effective is the UK's regulatory and legislative framework on tackling these issues? How effective will the Online Safety Act be in combatting harmful social media content?

The Online Safety Act is a groundbreaking piece of UK legislation aimed at holding platforms accountable for harmful online content. It brought several contributions to support online safety which include but are not limited to 1) Imposing legal duties on companies to remove illegal or harmful material. 2) Introducing age-verification requirements to protect children from harmful content. 3) Threatens penalties for non-compliance.

While the Act is a significant step, enforcement mechanisms might lag, and companies may find ways to avoid liability. However, without a clear effort to identify and develop alternative business models and support users to know their rights and protect themselves from exploitative behaviours of the algorithms, such acts will have limited effectiveness and the UK, among other countries, will struggle to reduce the promotion of harmful content online. This effort can be achieved through coupling research departments of universities to open source and support the design of models that can promote healthy societies yet ensure that the business does not get harmed. Grants, think tanks, and contests are some of the mechanisms that should be supported to bring that forward to the industry. Further, algorithms that support distributed federated learning and privacy awareness need to be supported in the startup ecosystem to bring in a generation of social media startups that are socially aware and user-centric to protect the privacy of users and to ensure harmony by prompting sound well-founded facts and demoting data items lacking sound evidence. The US Federal Emergency Management Agency, FEMA, built mechanisms to ensure that the data items from social media used for emergency response are well founded in facts. A police unit in Cupertino, California, USA has worked with social media companies to combat school crimes. It is also using social media as one of the sources for building a real-time understanding of riot and mass shooting events to mobilise responses. That means a full integration of the efforts needed to take place to empower policy-forming bodies, police, and response forces to deal with evolving situations using a systematic approach.

Additional Measures

What more should be done to combat potentially harmful social media and AI content?

Additional Measures to combat harmful content include but are not limited to the following:

- Improved transparency: which requires platforms to disclose algorithmic decision-making and content moderation practices partially. A full disclosure is unrealistic and will go against the business rules and fair competition. It is necessary to note that a fully open source is not a realistic target.
- Global collaboration: International standards for content moderation to address jurisdictional challenges are important to help identify issues that can have different reactions from different parts of the world, thus providing a good understanding of what could be considered truthful or manipulative.
- AI oversight: Enhanced auditing of AI systems to ensure they don’t propagate harmful content. Further ensuring that there is a f human in the loop, thus reaching augmented intelligence instead of full automation, might also be necessary, at least at the start for some period.

- Media literacy education: Equipping users to assess the content they consume critically and understand the impact of harmful content on society. The educational efforts should start early and should be included in curricula as early as 5-year-old and 8-year-old pupils in public schools, which should include basic mechanisms and guard rails for using social media.
- Funding of startup and entrepreneurship ecosystems: A critical step is to ensure the existence of sustainable platforms that are well-aligned to the needs of society. This can be covered by supporting the creation of alternative platforms that consider the well-being of society instead of promoting profit. That step needs to be carried out by consistently offering funding opportunities such as UK Innovate to encourage well-aligned homegrown alternatives.

Which bodies should be held accountable for the spread of misinformation, disinformation and harmful content as a result of social media and search engines' use of algorithms and AI?

A holistic stakeholder-centric solution is necessary that looks to the well-being of society, and the users will require the collaboration of companies, users, regulators and developers. Each of the aforementioned stakeholders need to play their role effectively to ensure a harmonised, coordinated, and effective action. Accountability for the spread of misinformation is distributed among several layers of the stakeholders mainly:

1. Social media companies: for designing algorithms that amplify harmful content.
2. Regulators: for enforcing compliance and ensuring safeguards.
3. AI developers: for ensuring ethical use of generative models.

Striking a balance between freedom of speech and content moderation remains a core challenge.

Key Recommendations and Discussion

The intersection of social media algorithms, artificial intelligence, and the propagation of harmful content represents a critical challenge in contemporary digital communication landscapes. The summer 2024 UK riots serve as a poignant case study illustrating the profound societal implications of technological platforms' design and algorithmic mechanisms. According to research by Gillespie (2018) and Zuboff (2019), social media platforms' business models are fundamentally structured around engagement metrics that prioritise user attention over content veracity, creating inherent incentives for algorithmic systems to amplify sensationalist and emotionally charged narratives.

The algorithmic recommendation systems employed by platforms like Facebook, Twitter, and YouTube operate through complex machine-learning models that optimize content visibility based on predicted user engagement. Noble (2018) argues that these algorithms are not neutral technological tools but encode existing social biases and power structures. The recommendation engines utilise sophisticated neural networks that analyse user interaction patterns, learning to prioritise content that generates maximum emotional response and prolonged platform interaction. This dynamic creates dangerous feedback loops where inflammatory, divisive content receives exponential visibility, potentially radicalising user perspectives and fragmenting social discourse.

Generative AI and Large Language Models (LLMs) introduce a new dimension of complexity to misinformation propagation. Researchers like Bender et al. (2021) have highlighted the capacity of these systems to generate contextually sophisticated yet factually unreliable content at unprecedented scales. The ability to produce human-like text that can be emotionally manipulative and strategically targeted represents a significant challenge for traditional fact-checking and content moderation mechanisms. During the UK riots, platforms likely experienced challenges in real-time identification and mitigation of AI-generated disinformation related to the Southport incident.

The UK's regulatory response, embodied by the Online Safety Act 2023, represents a pioneering legislative attempt to address these technological challenges. Drawing from recommendations by the Carnegie UK Trust and digital policy experts, the Act introduces legal obligations for platforms to proactively moderate content and mitigate potential societal harms. Ofcom's role has been particularly critical, with the regulatory body emphasising the need for algorithmic transparency and accountability. The Science, Innovation and Technology Committee's inquiry signals a sophisticated understanding that technological regulation requires nuanced, adaptive approaches.

Accountability mechanisms remain complex and multifaceted. Van Dijck (2020) argues for a comprehensive approach that distributes responsibility across multiple stakeholders: technology companies, algorithm designers, content creators, and regulatory bodies. The potential for holding platform executives and algorithmic design teams legally accountable represents an emerging regulatory frontier. International collaborations and standardised impact assessment frameworks will be crucial in developing robust, adaptable regulatory mechanisms.

The role of organisations similar to Ofcom and the National Security Online Information Team has become increasingly sophisticated. Such bodies are expected to develop rapid response capabilities, integrate advanced technological monitoring systems, and create flexible guidelines that can adapt to rapidly evolving digital communication landscapes. The summer 2024 riots

demonstrated the urgent need for real-time intervention strategies that can detect and mitigate potentially harmful content propagation.

Technological solutions must be complemented by broader societal interventions. Digital literacy programs, critical thinking education, and transparent algorithmic reporting can empower users to navigate complex information environments more effectively. The potential of AI and machine learning to both create and combat misinformation suggests a nuanced technological ecosystem that requires continuous adaptation and critical reflection.

This analysis demonstrates the intricate relationship between technological platforms, algorithmic systems, regulatory frameworks, and societal dynamics. The summer 2024 UK riots serve as a critical case study illuminating the profound challenges and potential interventions in managing digital communication ecosystems.

Prioritised Actionable Recommendations

Reimagining Social Media Business Models: Balancing Profitability, User Privacy, and Social Responsibility The current social media ecosystem is fundamentally broken, driven by an advertising-based model that commodifies user attention and personal data. A sustainable alternative requires a fundamental redesign of revenue streams, platform governance, and user value proposition.

We included a set of recommendations in the section "**Additional Measures".** The current section focuses on foundational recommendations for business models and ecosystem improvement, which aims to drive sustainable models and align the stakes of interest for the main stakeholders of social media (users, companies, regulators, and AI developers). The current section will further summarise our main recommendations for sustainable social media business models into **five key principles**.

1. User-centric revenue streams: traditional social media platforms rely almost exclusively on targeted advertising, which creates inherent conflicts of interest. A more sustainable approach needs to be built which addresses the shortcomings of existing models by including micro-compensation schemes for content creators and for data sharing.
2. Data governance and privacy: a radical reimagining of data management would include explicit, granular user consent mechanisms, user ownership and control of personal data, and blockchain-based data management systems. That transformation needs to take place gradually, and a shift from the classic to the needed systems should avoid abrupt switches to ensure a convivial, peaceful society.

3. Content ecosystem design: redesigning algorithmic recommendations to prioritise factual, verified content with diverse perspectives, constructive dialogue, and community well-being over engagement metrics.
4. Ethical monetisation strategies: alternative revenue models that align platform interests with user interests that should include subscription tiers with clear value propositions, creator support mechanisms, community-driven funding models and value-added enterprise solutions
5. Technological architecture: implementing architectural solutions that prioritise decentralised platform designs

Conclusion

The intersection of social media platforms, algorithmic systems, and generative AI poses profound challenges in mitigating the spread of harmful content. The UK riots of 2024 exemplify the critical societal risks stemming from disinformation and algorithmic amplification of polarising narratives. While the Online Safety Act 2023 establishes a robust foundation for tackling online harms, the complexity of digital ecosystems necessitates a multi-stakeholder, adaptive approach. Companies, regulators, AI developers, and users must collectively work to balance the priorities of innovation, privacy, free expression, and societal well-being. Without addressing the inherent conflicts in profit-driven business models and ensuring transparency in algorithmic design, technological interventions alone may prove insufficient. A holistic strategy combining robust regulation, technological oversight, societal education, and international collaboration is essential to creating a sustainable and secure digital environment.

To combat harmful content effectively, platforms must transition from engagement-driven revenue models to privacy-focused, sustainable systems, incorporating subscription options and user-owned data. Greater transparency and accountability are needed, with platforms disclosing algorithmic influence on content moderation and undergoing independent audits. Digital literacy programs, spanning all age groups, should build critical thinking skills to counter misinformation. Technological solutions like federated learning, privacy-centric architectures, and human oversight in AI deployment can limit the spread of unverified content. Promoting ethical AI research and global standards for its responsible use is crucial to preventing misuse. Strengthening regulatory enforcement, enhancing international collaboration, and aligning global policies will ensure a coordinated response to cross-border challenges, balancing innovation with societal well-being.

References

Acemoglu, Daron, Asuman Ozdaglar, and James Siderius. "A model of online misinformation." Review of Economic Studies 91.6 (2024): 3117-3150.

Agarwal, Ashish, Kartik Hosanagar, and Michael D. Smith. "Location, location, location: An analysis of profitability of position in online advertising markets." Journal of marketing research 48.6 (2011): 1057-1073.

Aridor, Guy, et al. "The economics of social media." Journal of Economic Literature 62.4 (2024): 1422-1474.

Balkin, Jack M. "The fiduciary model of privacy." Harv. L. Rev. F. 134 (2020): 11.

Bell, Bethan. BBC online. " Riots 10 years on: The five summer nights when London burned" 6 August 2021 last accessed https://www.bbc.com/news/uk-england-london-58058031. 15 December 2024

Bender, E. M., et al. (2021). "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" FAccT Conference.

Bessi, Alessandro, and Emilio Ferrara. "Social bots distort the 2016 US Presidential election online discussion." First Monday 21.11-7 (2016).

Boeker, Maximilian, and Aleksandra Urman. "An empirical investigation of personalization factors on TikTok." Proceedings of the ACM web conference 2022. 2022.

Dawes, Melanie. Open letter to Secretary of state 22 October 2024. Last accessed 15 December 2024
https://www.ofcom.org.uk/siteassets/resources/documents/about-ofcom/public-correspondence/2024/letter-from-dame-melanie-dawes-to-the-secretary-of-state-22-october-2024.pdf?v=383693

Gillespie, T. (2018). "Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media"

Ko, B. "The Role of User Interactions In Social Media On Recommendation Algorithms: Evaluation of Tiktok’s Personalization Practices From User’s Perspective." 2023,

Lee, Angela Y., et al. "The algorithmic crystal: Conceptualizing the self through algorithmic personalization on TikTok." Proceedings of the ACM on Human-computer Interaction 6.CSCW2 (2022): 1-22.

Liesenfeld, Andreas, and Mark Dingemanse. "Rethinking open source generative AI: open washing and the EU AI Act." The 2024 ACM Conference on Fairness, Accountability, and Transparency. 2024.

Narayanan, Arvind. "Understanding social media recommendation algorithms." (2023).

Noble, S. U. (2018). "Algorithms of Oppression: How Search Engines Reinforce Racism"

O'neil, Cathy. Weapons of math destruction: How big data increases inequality and threatens democracy. Crown, 2017.

Papakyriakopoulos, Orestis, et al. "How algorithms shape the distribution of political advertising: Case studies of Facebook, Google, and TikTok." Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society. 2022.

Pariser, Eli. The filter bubble: What the Internet is hiding from you. penguin UK, 2011.

Rezac, Fabien. "The Role of Privacy Protection in Business Models for Sustainability: A Conceptual Integration from an Ecosystem Perspective." Journal of Business Models 10.2 (2022): 31-57.

Siderius, James. Understanding Social Media: Misinformation, Attention, and Digital Advertising. Diss. Massachusetts Institute of Technology, 2023.

Trusov, Michael, Anand V. Bodapati, and Randolph E. Bucklin. "Determining influential users in internet social networks." Journal of marketing research 47.4 (2010): 643-658.

Van Dijck, J. (2020). "The Platform Society: Public Values in a Connective World"

Vombatkere, Karan, et al. "TikTok and the Art of Personalization: Investigating Exploration and Exploitation on Social Media Feeds." Proceedings of the ACM on Web Conference 2024. 2024.

Wirtz, Bernd W., Oliver Schilke, and Sebastian Ullrich. "Strategic development of business models: implications of the Web 2.0 for creating value on the internet." Long range planning 43.2-3 (2010): 272-290.

Zanker, Markus, Laurens Rook, and Dietmar Jannach. "Measuring the impact of online personalisation: Past, present and future." International Journal of Human-Computer Studies 131 (2019): 160-168.

Zuboff, S. (2019). "The Age of Surveillance Capitalism"

Future of Life, FLI AI Safety Index. Last accessed 15 December 2024. https://futureoflife.org/document/fli-ai-safety-index-2024/