Algorithmic curation/exposure

Natalia Umansky[†] & Cornelius Puschmann[‡]

Abstract

Algorithmic curation has emerged as a significant force shaping political communication in the digital era. This process involves actively selecting and organizing information to cater to individual preferences and moderate harmful content, greatly impacting the flow of information and potentially the formation of public opinion. Algorithmic personalization and content moderation are integral components of algorithmic curation. By leveraging vast amounts of data, algorithms aim to deliver personalized experiences to users, aligning content with their interests and past behaviors and to content that violates platform guidelines, such as misinformation, hate speech, explicit material, or harassment. However, concerns have been raised about potential biases and unintended consequences of algorithmic decision-making. Algorithmic curation has profound implications for democratic processes, as it can reinforce existing beliefs,, impede exposure to diverse perspectives and create filter bubbles. Understanding the complexities of algorithmic curation in the realm of political communication is therefore crucial for researchers, policymakers, and society at large. Balancing personalization and the promotion of diverse viewpoints is essential to foster an informed and engaged citizenry in the digital age.

[†]Postdoctoral Researcher, Dep. of Political Science, University of Zurich

[‡]Professor of Media and Communication, Centre for Media, Communication & Information Research, Universität Bremen

Algorithmic curation plays a pivotal role in shaping the landscape of political communication in the digital age (Bandy & Diakopoulos, 2021; Gausen et al., 2022; Jürgens & Stark, 2022). As an increasingly prevalent phenomenon, it significantly influences the information flow and the formation of public opinion. Unlike *algorithmic exposure*, which refers to the mere presentation of content to users, algorithmic curation involves a more active selection and arrangement of information tailored to individual preferences and interests (Bandy & Diakopoulos, 2021; Gausen et al., 2022; Jürgens & Stark, 2022; Swart, 2021). In this context, *algorithmic personalization*, *shadow banning*, and *content moderation* emerge as integral components of the curation process (Cotter, 2021; Duffy & Meisner, 2023). Algorithmic curation operates on the principle of filtering and prioritizing content based on various factors such as relevance, popularity, and user behavior (Bandy & Diakopoulos, 2021; Gausen et al., 2022). By leveraging vast amounts of data, algorithms aim to deliver personalized experiences to users, offering them content that aligns with their preferences and keeping them engaged based on their past behavior, interests, and demographic characteristics. This practice is referred to as algorithmic personalization, and it is one of the integral elements of algorithmic curation. Shadow banning, another facet of algorithmic curation, refers to the practice of limiting or suppressing the visibility of certain content or users without their knowledge (Cotter, 2021). This technique is often employed to mitigate the spread of misinformation, hate speech, or other objectionable content that does not comply with community guidelines and platform policies. Algorithms play a crucial role in this process by automatically detecting and *flagging* potentially inappropriate or harmful content (Duffy & Meisner, 2023; Gillespie, 2020). While content moderation is necessary to maintain a safe and inclusive digital environment, it also raises questions about the potential for biases in algorithmic decision-making. The lack of transparency surrounding shadow banning algorithms raises special concerns about potential biases or unintended consequences, such as the silencing of marginalized voices or the suppression of dissenting opinions (Cotter, 2021; Duffy & Meisner, 2023; Gillespie, 2020).

The impact of algorithmic curation on political communication is profound (Bandy & Diakopoulos, 2021; Jürgens & Stark, 2022). It shapes the information ecosystem, influencing which news articles, opinions, and narratives reach users and how they are presented. By tailoring content to individual preferences, algorithms may reinforce existing beliefs, potentially hinder the exposure to diverse perspectives and contribute to the formation of filter bubbles (Jürgens & Stark, 2022; Ohme, 2021). This has implications for democratic processes, as it may limit the exchange of ideas, hinder critical thinking, and exacerbate societal divisions. Understanding the intricacies of algorithmic curation in the context of political communication is crucial for researchers, policymakers, and society at large (Reviglio & Agosti, 2020). It requires examining both algorithms and the biases they may introduce, as well as the business models that underpin them. Additionally, it necessitates exploring mechanisms to promote transparency, accountability, and the provision of diverse and balanced information. Striking a delicate balance between personalization and exposure to diverse viewpoints is essential to ensure an informed and engaged citizenry in the digital age.

Algorithmic personalisation

Algorithmic personalization of online services describes "algorithms that tailor information based on what the user needs, wants and who he knows on the social web" (Bozdag, 2013, p. 209). The rapid expansion of user-generated content from the early 2000s onwards called for innovative strategies on the part of online information intermediaries, such as search engines and emerging social media platforms, that would allow users to effectively navigate an evolving information environment. Due to the high choice nature of this new ecosystem, and as a direct result of the advertising-based business model of large platforms, algorithmic personalisation quickly became a popular strategy to disseminate the most relevant content to media users (Bozdag, 2013). Serving personalized information meant increasing the perceived usefulness and relevance content, but has also been characterized as potentially dangerous and detrimental to democracy by limiting users to perspectives congruent with their views (Cinelli et al., 2021; Lury & Day, 2019; Mittelstadt, 2016).

Most personalization systems rely on user profiles to tailor experiences. User profiles typically include demographic information, interests, and preferences, aiming to enhance information access and understand user intentions. The user profiling process consists of three main phases: information collection, profile construction, and utilization of the compiled user profile in the actual web service, usually in conjunction with machine learning. User profiles can be built explicitly or implicitly, with the latter having the advantage of not imposing a burden on the user by having to explicitly formulate preferences. Implicit profiling leverages actions on a service, such as clicking and typing behavior, mouse and trackpad movements, as well as distributed information such as IP address, cookies, and session IDs to infer user interests. Implicit feedback techniques automatically update and continuously refine profiles based on user interactions with the service. User profile information is then matched with the characteristics of content that may subsequently be shown to the user, which is also subject to automated analysis, chiefly using two techniques. Content-based filtering focuses on associations between user choices and properties of new objects, while collaborative filtering utilizes input from other users with similar tastes. Social networks like Facebook have long employed collaborative filtering techniques to customize users' news feeds based on their relationships and preferences, while streaming services such as Netflix or Spotify have more grounds on which to match user profiles with specific content properties. However, both approaches are generally used in combination in increasingly complex and evolving regimes of algorithmic personalization.

It has been argued that algorithmic personalization “create a new type of curated media that can undermine the fairness and quality of political discourse” (Mittelstadt, 2016), while at the same time “[i]t remains an open question how much power algorithms and the design of platforms have over the individual-level visibility of news and political content on social media” (Thorson et al, 2021). Although phenomena such as that of the filter bubble (Pariser, 2012) have potentially far-reaching implications, they have also been widely criticized for being conceptually unclear and overstated in their effects on public opinion (Bruns, 2019; Stark & Jürgens). In any case, concerns surrounding the potential harms (as well as benefits) of algorithmic personalization have been widely discussed in recent years (Magin et al, 2021;

Möller et al, 2018; Stier et al, 2022). While empirical evidence suggests that algorithmic personalisation does not narrow the diversity of users' choices with regard to information sources, it is less clear what long-term impacts, for example on political knowledge or partisanship, it entails. In the past, academic considerations carried little overall weight in relation to applications driven by economic arguments. Couldry and Turow (2014) argue that the application of algorithmic filtering methods used in other areas of the media industry to news inevitably has an adverse effect on public debate. But disentangling the effect of algorithmic nudging towards content that will serve platforms' business objectives by keeping users engaged for longer from ordinary selective exposure remains extremely challenging.

Algorithmic content moderation

Algorithmic content moderation is an integral part of algorithmic curation, enabling online platforms to filter and moderate user-generated content through automated systems driven by algorithms and machine learning techniques (Duffy & Meisner, 2023; Gillespie, 2020). As the volume of user-generated content continues to skyrocket across social media platforms and digital spaces, the challenge of content moderation has grown exponentially. Algorithmic content moderation offers a scalable solution to handle the overwhelming amount of content generated on a daily basis. At its simplest form, algorithmic content moderation relies on keyword filters, flagging specific words or phrases for review or removal (Duffy & Meisner, 2023). This method helps identify and eliminate content that violates platform guidelines, such as hate speech, explicit material, or harassment. However, relying solely on keyword filters presents limitations, including the potential for false positives and false negatives. False positives occur when harmless content is mistakenly flagged as guideline-violating, while false negatives involve the failure to detect and remove problematic content. To address these limitations, more advanced algorithms have been developed, employing techniques like natural language processing (NLP), machine learning, and artificial intelligence. These algorithms analyze content for various attributes, such as offensive language, explicit imagery, spam, and policy violations. By training on extensive datasets containing labeled examples of problematic content, machine learning algorithms learn patterns and make predictions about the nature of unseen content (Duffy & Meisner, 2023; Gillespie, 2020).

A notable aspect of algorithmic content moderation lies in its capacity to detect hate speech and harmful behavior. Algorithms can scrutinize text and multimedia content to identify hate speech, discriminatory language, and threats, as well as patterns of cyberbullying or harassment. Automating this process enables platforms to swiftly take action by removing or flagging problematic content, fostering a safer and more inclusive online environment. Another moderation technique relates to shadow banning, which involves the covert restriction of an individual's visibility or reach without their knowledge as a means of deterring problematic behavior without outright removing the user or their content (Cotter, 2021). Nevertheless, algorithmic content moderation does encounter challenges. The algorithms used for moderation are trained on existing data, which can reflect societal biases and norms (Duffy & Meisner, 2023; Gillespie, 2020; Gorwa et al., 2020). Consequently, biased outcomes may arise, disproportionately targeting or protecting certain groups or perspectives. Striking a balance between preventing harmful content and upholding freedom of expression remains a challenge. Furthermore, the intricate nature of language and context presents difficulties for algorithms. Accurately interpreting sarcasm, irony, and cultural nuances proves challenging, leading to errors in content classification (Gillespie, 2020; Gorwa et al., 2020). The ever-evolving landscape of online communication and the emergence of new trends and slang further complicate the task of algorithmic content moderation.

References

Bandy, J., & Diakopoulos, N. (2021). More accounts, fewer links: How algorithmic curation impacts media exposure in Twitter timelines. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), 1-28.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. Ethics and information technology, 15, 209-227.

Bruns, A. (2019). Filter bubble. Internet Policy Review, 8(4). https://policyreview.info/concepts/filter-bubble

Cotter, K. (2021). "Shadowbanning is not a thing": black box gaslighting and the power to independently know and credibly critique algorithms. Information, Communication & Society, 1-18.

Couldry, N., & Turow, J. (2014). Big Data, Big Questions| Advertising, Big Data and the Clearance of the Public Realm: Marketers' New Approaches to the Content Subsidy. International journal of communication, 8, 17.

Duffy, B. E., & Meisner, C. (2023). Platform governance at the margins: Social media creators' experiences with algorithmic (in) visibility. Media, Culture & Society, 45(2), 285-304.

Gausen, A., Luk, W., & Guo, C. (2022). Using agent-based modelling to evaluate the impact of algorithmic curation on social media. ACM Journal of Data and Information Quality, 15(1), 1-24.

Gillespie, T. (2020). Content moderation, AI, and the question of scale. Big Data & Society, 7(2), 2053951720943234.

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. Big Data & Society, 7(1), 2053951719897945.

Jürgens, P., & Stark, B. (2022). Mapping exposure diversity: The divergent effects of algorithmic curation on news consumption. Journal of Communication, 72(3), 322-344.

Lury, C., & Day, S. (2019). Algorithmic personalization as a mode of individuation. Theory, Culture & Society, 36(2), 17-37.

Magin, M., Geiß, S., Stark, B., & Jürgens, P. (2022). Common Core in Danger? Personalized Information and the Fragmentation of the Public Agenda. The International Journal of Press/Politics, 27(4), 887–909. https://doi.org/10.1177/19401612211026595

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.

Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity. Information, Communication & Society, 21(7), 959–977. https://doi.org/10.1080/1369118X.2018.1444076

Ohme, J. (2021). Algorithmic social media use and its relationship to attitude reinforcement and issue-specific political participation–The case of the 2015 European immigration movements. Journal of Information Technology & Politics, 18(1), 36-54.

Reviglio, U., & Agosti, C. (2020). Thinking outside the black-box: The case for “algorithmic sovereignty” in social media. Social Media+ Society, 6(2), 2056305120915613.

Swart, J. (2021). Experiencing algorithms: How young people understand, feel about, and engage with algorithmic news selection on social media. Social media+ society, 7(2), 20563051211008828.

Thorson, K., Cotter, K., Medeiros, M., & Pak, C. (2021). Algorithmic inference, political interest, and exposure to news and politics on Facebook. Information, Communication & Society, 24(2), 183–200. https://doi.org/10.1080/1369118X.2019.1642934

Turow, J. (2011). Media today: An introduction to mass communication. Taylor & Francis.

Zeng, A., Yeung, C. H., Shang, M. S., & Zhang, Y. C. (2012). The reinforcing influence of recommendations on global diversification. Europhysics Letters, 97(1), 18005.

Zhou, T., Kuscsik, Z., Liu, J. G., Medo, M., Wakeling, J. R., & Zhang, Y. C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. Proceedings of the National Academy of Sciences, 107(10), 4511-4515.