The Invisible Hand:

Algorithmic Curation Within the Facebook News Feed

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Jonah E. Kim

Spring 2024

On my honor as a University Student, I have neither given nor received unauthorized aid on this

assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

William J. Davis, Department of Engineering and Society

…he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention.

Adam Smith, 1776

Introduction

      Almost two decades ago in 2006, Facebook introduced the News Feed to help manage the platform's growing userbase. The algorithm now computes on posts from over 3 billion users determining which messages to block, promote, and show. Facebook willingly admits the integral role that the recommendation algorithm plays within their system saying that the News Feed helps connect "people to meaningful posts from their friends and family" (Mosseri, 2018, para. 1). However, while recommendation algorithms operate more uniformly and quickly than any human, the resulting systems are not free from bias and partiality (O'Neil, 2016). Curative algorithms have the potential to censor and distort the truth (Edelson et al., 2021; Pariser, 2011; Pathak et al., 2023). Controversies—such as the misinformation campaigns during America's 2016 presidential election, the Facebook-originated revolution in Egypt, and the Rohingya genocide in Myanmar—all illustrate the News Feed's ability to rapidly disseminate information and misinformation across its platform (Jacoby, 2018a, 16:21-17:44; Jacoby, 2018b, 9:06-12:00; Mozur, 2018; Ortutay, 2022b). The News Feed allows malefactors to spread content aimed to destabilize elections and target minorities. Additionally, studies have corroborated the concerns that interactions on Facebook tend to be more polarized, segregated, and ideologically homogenous (Bakshy et al., 2015; Del Vicario et al., 2016). Political theorists worry that this uniformity of beliefs and partitioning of user audiences artificially privatizes content among social groups (Lazer, 2015; Riemer & Peter, 2021). This amounts to targeted censorship and decreases free speech since people cannot exchange ideas without algorithmic mediation.

      While current academic studies analyze *how* the News Feed impacts misinformation and polarization (Cinelli et al., 2021; Del Vicario et al., 2016; Edelson et al., 2021; Guess et al., 2019; Hargreaves et al., 2018, 2020; Pathak et al., 2023; Pierri et al., 2023), few attempt to determine

why the News Feed recommends misinformation, *why* it lends itself to the formation of ideologically homogenous communities, or *what* in the News Feed's design creates these effects. This presents issues for legislators and designers seeking to improve the social integrity of Facebook. One cannot resolve these issues without first understanding what causes them. Present studies point out flaws without providing actionable solutions. This project argues that, in order to understand and address the present consequences of the News Feed, one must first grasp the artifact's history. Facebook's rapid growth in the 2000s led the company to rely on engagement metrics to filter its content. This decision has created and perpetuated a curative system which promotes misinformation and increases polarization. Understanding this history places one in a better position to form potential solutions as one can address the underlying causes and avoid the architectural flaws inherent within present designs. This work relies on recently leaked internal Facebook documents—collectively known as *The Facebook Papers*—to provide new evidence and recontextualize current research. Technological Momentum and its variations structure this analysis. While other frameworks observe an artifact at a particular snapshot, Technological Momentum's historical and inertial perspective enables the analysis of the News Feed through time, revealing how engagement became embedded within Facebook's platform.

Gaps Within Present Understanding

Facebook enables the rapid transferal of ideas between users sharing similar beliefs, changing how both truths and lies spread. These effects are well reported but not fully explained by media or academia. Wael Ghonim—an activist behind the 2011 protests that ousted the regime of former Egyptian President Mubarak—cited Facebook as a key enabler of the event. Facebook allowed Ghonim to connect with people that he otherwise could not have (Jacoby, 2018a, 16:29-17:40). The platform enabled the social coordination of geographically distributed

but ideologically similar individuals. While activists have used Facebook as a tool for social change, nation-state actors and governments have exploited the platform to destabilize elections and promote extremist ideologies. For example, the Burmese regime in Myanmar abused Facebook to spread propaganda against the local Rohingya Muslim minority. The military posed as public figures, created false reports, and incited a genocide against the Rohingya which displaced hundreds of thousands (Mozur, 2018; Ortutay, 2022b). As Facebook provided many with their sole source of news, the regime effectively altered what was considered 'truth' in the country (Gowen & Bearak, 2017, para. 5). The propaganda united the country underneath a manufactured lie, creating uniform beliefs which the Burmese regime exploited to enact an ethnic cleansing. When reporting on the News Feed, media sources held Facebook accountable for misinformation (Alba, 2021; Ortutay, 2022a; Silverman et al., 2022) but did not provide a cohesive analysis of how or why misinformation spreads. Those that attempted to pinpoint the causes of misinformation pointed to failures in policy, tools, or algorithms (Dwoskin, 2021; Hindman et al., 2022; Roose et al., 2021) but did not explain why these flaws might exist.

While academic studies have verified activists' claims that Facebook increases misinformation and polarization, the studies fall into similar explanatory issues as news sources. Two separate inter-university studies conducted by Vicario et al. (2015) and Cinelli et al. (2021) found that both information and misinformation on social media spread through ideologically homogenous networks called "echo chambers" (Cinelli et al., 2021, p. 1; Del Vicario et al., 2016, p. 554). According to the studies, people "tend to select and share content related to a specific narrative and to ignore the rest" (Del Vicario et al., 2016, p. 558). Users are drawn towards communities that reinforce and *echo* their beliefs. This risks increasing polarization and misinformation by segregating speech within homogenous groups. Political philosopher John

3

Stuart Mill argued in his treatise *On Liberty* that disputation and argument keeps truth alive by continually challenging premises and assumptions. Diverse opinions root out false beliefs and strengthen truths (Mill, 1859). If interactions within Facebook tend towards homogeneity, then discourse suffers as people fail to engage with opposing views.

Neither study presented a convincing account of echo chamber's origins. Within either their text or citations, both Cinelli et al. (2021) and Del Vicario et al. (2016) utilized the sociological concept of homophily which states "similarity breeds connection" (McPherson et al., 2001, p. 1). This concept follows the common adage that 'birds of a feather flock together' and frames echo chambers as psychological phenomenon. Homophily has gained traction within academic research (Aiello et al., 2012; Bisgin et al., 2012; Ertug et al., 2022; Fu et al., 2012; Kossinets & Watts, 2009). This perhaps justifies Del Vicario et al.'s (2016) approach of reporting the effects of echo chambers but not explaining their underlying causes. The study implicitly relies on homophily to ground its analysis and simply provides additional evidence in support of the phenomenon. However, homophily alone fails to explain the audience segregation seen on Facebook. Cinelli et al.'s (2021) study found that Facebook experienced higher degrees of user segregation than Reddit. If the two platforms were equivalent, similar patterns of homogenization should arise. A separate factor must exist within Facebook's social or technical structures to account for the gap. Although the researchers indicated possible algorithmic differences, they left analysis to future works. Both studies reported effects cleaved from their causes.

Similar patterns hold across the technical literature. A joint study conducted by American and French universities on misinformation in Facebook ended by saying, "We hope that future research will be able to investigate *why* misinformation generates more engagement" (original

work's emphasis) (Edelson et al., 2021, p. 12). The authors admit that they have only exposed the external effects rather than the internal causes of misinformation. This trend continues in other studies (Bakshy et al., 2015; Guess et al., 2019; Hargreaves et al., 2018; Pathak et al., 2023; Rader, 2017). Although some authors suggested components of Facebook's systems that may be responsible for increased misinformation or polarization (algorithms, policies, etc.), they do not provide holistic accounts of how the technology and its corresponding effects developed.

Technological activist Eli Pariser gives a fuller explanation within his book, *The Filter Bubble*. He argues that the employment of curative systems creates tailored and invisible information boundaries which entrap their occupants in homogenized content (Pariser, 2011, p. 6). Users of platforms such as Facebook implicitly agree to participate in a curated bubble of stories and ads, leading to the external effects seen by the studies such as "echo chambers" and polarized content. Pariser contextualizes Facebook's and other companies' moves towards increased personalization and tailoring as a "race for relevance" in which they attempted to maximize profit through the exploitation of users' data and contextual advertising (Pariser, 2011, Chapter 1). However, this explanation does not appear charitable and does not consider the history of Facebook's development.

Facebook started as a small platform connecting people across college campuses. Its motto has always related to connecting people and building relationships ("Facebook's mission," 2022; Kelly, 2017; Meta, 2024; Reagan, 2009). Pariser's view seems pessimistic; it treats Facebook's move towards personalization as solely motivated by monetary gains. However, observing Facebook's history reveals another explanation, one where technical requirements led to personalization. Facebook created the News Feed in 2006 to manage its growing userbase (Mosseri, 2016, para. 2). The algorithm provided and still provides a curated experience for each

person, filtering through millions (and now billions) of posts and messages (Dixon, 2023; Lapides et al., 2015, p. 163). Former Facebook Executive Adam Mosseri explained that without the News Feed users would be presented with "far too much information for any one person" (Mosseri, 2016, para. 2). While the algorithm undoubtedly provides Facebook with valuable advertising data, the algorithm also plays a critical role in mediating interactions between users and the content hosted on Facebook. Viewing the development of the News Feed solely as a monetary tool neglects the original context in which the technology arose.

In summary, current studies and reports focus on the effects of the News Feed rather than the underlying causes. This paints a false picture of the problem, showing an incomplete context; the *what* is given but not the *why*. Other works attempting to explain the system frame it as a monetary device and neglect to account for its developmental background. This project argues that one must understand the algorithm's history as past actions inform present outcomes. One cannot aim to improve the platform through changes in policies, algorithms, or regulations without understanding what the problems are and how they arose. Neglecting the causes risks repeating the effects.

Methods

Previous studies relied on the externally observable effects of the News Feed. Neither the implementation details of the algorithm nor the private research done by Facebook were available for analysis. This changed in 2021. Frances Haugen—a whistleblower from Facebook's Civic Integrity division—leaked thousands of internal company documents, collectively known as *The Facebook Papers* (Mac & Kang, 2023, para. 2). These documents revealed that Facebook knew that its platform increased "hate, misinformation, and political unrest" (Pelley, 2021, para. 1). The papers also included internal experiments and documentation, providing insights into

how the News Feed functions and how Facebook developed it. This information not only reinforces previous studies' findings but also reveals what Facebook believes to be the causes of misinformation and polarization on their platform. For these reasons, this resource proves invaluable for this study to compare and contrast against previous knowledge. It fills gaps within the present understanding of the system, such as what content the algorithm promotes and why.

To help frame the analysis of the News Feed, this project draws upon the Technological Momentum framework. Technological historian Thomas P. Hughes developed the framework to analyze the development of large technological systems, and the framework's ideas have since been extended elsewhere within STS literature (Bijker et al., 2012; Sismondo, 2010). The framework contextualizes technical systems as being inertial artifacts that gain momentum as they develop. This perspective helps bring an artifact's full lifecycle into view. It aids in explaining how large platforms such as Facebook grew and evolved into the monolithic firms seen today. By observing an artifact's development through time rather than at a particular snapshot, the framework provides a wholistic approach to analysis. This helps reveal how initial design decisions cascade into later iterations of a system.

Early in a system's lifecycle, an artifact takes on many roles, interpretations, and designs as stakeholders steer the artifact towards a desired outcome. For instance, Facebook changed its mission statement almost yearly during its initial development, reflecting the platform's growth as it expanded into new contexts and roles (Reagan, 2009). However, as a system matures, the structures built around the artifact start to solidify and inform the actions of people. Developmental bottlenecks within the current design ("reverse salients") attract talent and visionaries to generate solutions. This strengthens and solidifies the system's existing structures—both social and technical—as time and resources are invested (Bijker et al., 2012, p.

223). Ultimately, these systems become "black boxes or facts and artifacts that are taken for granted" (Sismondo, 2010, p. 113). These artifacts become entrenched within a given problem domain and preclude the adoption of other alternatives through their widespread use. In this way, large technological systems gain 'momentum' and become harder to shape and control (Mayntz & Hughes, 1998, p. 13-4).

In the following analysis, these ideas of momentum and inertia help explain several key developmental threads regarding the News Feed. Namely, these concepts clarify why Facebook chose to use recommendation algorithms during the platform's initial development; why engagement became central to the News Feed's operations; and how this reliance on engagement increased polarization and misinformation. The News Feed's development over two decades has made the platform tolerant to outside forces. The system has gained 'momentum' which hinders Facebook's ability to institute new changes and policies. Even when Facebook has repeatedly been presented with evidence that its system harms people and damages the values that the company itself purportedly upholds, the News Feed remains intact.

Analysis

Shifts in Audiences and Technologies

Facebook's growth guided it towards the adoption of recommendation algorithms. Founded in February 2004 as a social tool to connect college students, Facebook gained over 1 million active users by the end of the year (Meta, 2024, timeline event 1 & 4). As the app's social structures and technologies had not yet solidified, the system still had interpretative flexibility allowing it to take on new roles. This enabled the platform to shift its purpose and expand its operations to high-schools and other contexts by 2005 (Greiner et al., 2019, event 5). However,

the company still restricted registration to members of official organizations such as colleges, campuses, and government offices (Facebook, 2006, para. 2). This created tension among the platform's userbase since graduated students wanted to maintain their connections outside of university. Additionally, others not enrolled within recognized institutions wished to join the network and connect with their friends. These pressures led Facebook to expand its registration policies in 2006 to allow anyone with a valid email address to join. Facebook's CEO, Mark Zuckerburg, said in a statement preceding the changes that: "We are expanding to respond to the *requests of millions of people* who want to be part of Facebook" (emphasis added) (Facebook, 2006, para. 3). Facebook had gained enough momentum that people *wanted* it to expand. Users saw the potential of the service and propelled Facebook's growth, shifting the social site from a private to an open platform. Facebook's own popularity fueled its expansion.

This influx of users posed a problem for Facebook: too much data. The News Feed launched twenty days before the open-platform switch (Meta, 2024, timeline event 9) but was not designed for a global context. The feed relied on a simple chronological ranking algorithm ("Why we build feeds," 2019). As the platform continued to grow, the News Feed could not handle the volume of content it experienced, creating a reverse salient in the system. Any of Facebook's millions of users could write a post potentially visible to the entire platform. Furthermore, users tend to 'friend' many people on Facebook while only considering a few connections to be close or intimate in reality. Interpersonal communication studies around this time found that users reported a mean of around 250 friends on Facebook (Boogart, 2006; Tong et al., 2008; Walther et al., 2008). This stands in contrast to the roughly 5 (5.35 STD) close connections which adults typically maintain (Gillespie et al., 2015, calculated from table 1). Sorting posts through simple metrics (such as the chronological ordering Facebook employed)

risks obscuring the few posts which the user cares about while bolstering the visibility of people which the user infrequently engages with. Recommendation algorithms were a natural—perhaps inevitable—solution to this problem.

Recommendation algorithms consume data about a user and a platform's content to produce tailored suggestions (Burbach et al., 2018). Most modern ranking systems—such as those used by Twitter, Netflix, Google, and Facebook—employ Machine-Learning (ML) models (Google, n.d.; Lada et al., 2021; Netflix, n.d.; Twitter, 2023). ML models offer several benefits over hand-crafted solutions. ML models can tackle problems deemed too complex for traditional methods (because of size, scope, data, etc.) or problems which lack known algorithms. Correctly implemented, these models can also adapt to changing data and environments (Géron, 2017, Chapter 1). A move towards recommendation algorithms likely appealed to Facebook since they dealt with millions of users in an emerging problem domain with no widely accepted solution. Although there was not and is not a mathematically *best* way to rank content, a curative system could 'learn' what to recommend. In 2009, Facebook switched from a chronological algorithm to a curative model ("Why we build feeds," 2019). The increasing momentum of Facebook's platform necessitated a shift in technologies to support its rising number of users which now totaled at over 300 million (Carlson, 2009).

This analysis may not convince individuals who believe the decision to move towards personalization was monetarily motivated. However, Facebook turned cash-flow positive *earlier* the same year as the shift to algorithmic ranking, so the decision could not have been solely for profit (Carlson, 2009). Even if one interprets the shift to algorithmic curation as a justification to collect user data for advertising, one must still acknowledge that Facebook needed to earn profit in some manner. With the millions of users dependent upon its system, the company (and likely

the users themselves) wanted a way to maintain the service. The platform had gained too much momentum to simply fade away. Even underneath this interpretation, the momentum behind the platform motivated the creation of the News Feed.

Personalization and Engagement

Machine learning models center around an objective function. An objective function translates a problem's semantics and constraints into a mathematical equation suitable for numerical optimization. It determines the data, the target, and the restrictions of a model. Facebook aims "to personalize… content for more than 2 billion people and show each of them content that is relevant and meaningful for them, every time they come to Facebook" (Lada et al., 2021). However, how does one quantify satisfaction, relevancy, and meaningfulness?

Each time one interacts with Facebook, the platform analyzes one's actions and extracts dozens of signals to feed into algorithmic models. Facebook tracks the stories one likes, the posts one shares, and even the posts visible on one's screen (*Integrity glossary*, n.d.; *MSI useful links*, n.d.; Wodinsky, 2021). Using machine-learning models, Facebook analyzes the captured signals to predict how information flows throughout the platform, what content generates engagement, and what posts contribute to Meaningful Social Interactions (*Information Corridors*, 2021; *integrity tradeoffs*, n.d.; *MSI documentation*, n.d.; *Why signals is a must-win*, n.d.). Meaningful Social Interactions (MSI) provide a quantitative metric for both sorting and evaluating the News Feed (*MSI documentation*, n.d.; *The MSI metrics revisited: Part 4*, 2019). MSI measures positive and active engagement between users wherein a "reciprocal interaction" occurs (*The MSI metrics revisited: Part 2*, 2019). For instance, a user 'hearting' a post may be considered to be a positive interaction (one would presumably not heart a post that one found uninteresting or harmful), an

active interaction (one must choose to interact with the post), and a reciprocal interaction (both parties can see the other's actions).

MSI attempts to reduce meaningful experiences to quantitative and computable values; however, like all models and representations, MSI sacrifices the richness of reality for the simplicity of numbers. A media study conducted by Michigan State University found that Facebook developer blogs frequently ended by saying: "We have learned that the actions people take on Facebook—liking, clicking, commenting or sharing a post don't always tell us the whole story of what is most meaningful to them" (Cotter et al., 2017, p. 1557). Facebook's system ranks with incomplete data about a user's beliefs and values. Nevertheless, Facebook continues to rely on engagement metrics as a predictor of user preferences.

After 15 years of continual development and use, Facebook has grown reliant on engagement and curation. Internal experiments revealed that alternative models, which avoided virality metrics or used unranked feeds, resulted in decreases in engagement, user-satisfaction, and content relevance (*Ranked News Feed*, 2018). If the company were to change technologies, the company would incur financial penalties and worsen users' experiences. However, an anonymous employee commenting on the research said that these results should be expected:

> We've spent a decade perfecting the ranked News Feed, investing person-centuries of time in getting the details right. It would be very weird if a few people working for a few months or a half [year] could make measurably something better—and yet that is often the bar we hold ourselves to when trying alternatives. (*Is ranking good?*, 2018)

The present News Feed was built on a foundation of engagement and virality. Thousands of people have iterated and innovated the design of the News Feed to bring the system to its present

state. Their investment of both time and ability has solidified the importance of engagement within the curative system and has instilled the corresponding metrics with momentum. This inertia prevents the adoption of alternatives, making it seem as if there were only one solution. The system appears like a black box, a taken truth.

Misinformation and Polarization

Relying on engagement metrics to determine rankings fosters an environment that promotes the production and perpetuation of viral content that inflames, misleads, and polarizes. The News Feed artificially increases viral content's audiences. Posts that amass attention have a high MSI score and are algorithmically weighted to appear more frequently (*MSI documentation*, n.d.). This allows popular posts to spread rapidly throughout the social network. Viral content gains additional momentum as users interact with it, creating a feedback mechanism which furthers the content's reach. This incentivizes creators to provoke anger and inflame disputes in order to gain wider audiences and to generate ad revenue (Jacoby, 2018b, 9:36-10:52). Inflammatory content is rewarded underneath engagement. Additionally, curative algorithms provide vectors for misinformation that foreign nations can exploit to destabilize elections and to promote propaganda, as in the 2016 presidential election and in the Myanmar genocide (Jacoby, 2018b, 22:22-24:10; The social atrocity, 2022). The issues inherit within engagement metrics extend to genuine content not aimed to explicitly misinform. An internal Facebook document revealed that the company has received complaints from politicians saying that engagement has forced them to take more extreme stances and to directly attack their opponents in order to be seen on the platform (*Political party response*, 2019). Facebook's algorithms incentivize the creation of polarizing content in order to increase a post's reach. Furthermore, the curative system *promotes* the content since it garners more attention. The News Feed's ranking

mechanisms ensure that the system recommends content to persons likely to engage with it. This provides the algorithmic basis for 'echo chambers' by artificially drawing persons of similar interests together (*Do we care?*, 2020). Individuals sharing ideologies or interests are more likely to engage with similar content, thus the News Feed gathers them together and artificially increases homophily.

Facebook knows of the flaws inherent within its systems and attempts to address them; however, these solutions do not correct the underlying issues of engagement metrics. For instance, Facebook developed the High-Risk Early Review Operations (HERO) algorithm to review potentially viral and misleading posts before they can spread across the platform (*Glossary*, n.d.; Wodinsky, 2021). HERO and other systems integrate to perform an 'integrity pass' over potential recommendations within the News Feed, filtering out problematic posts (Lada et al., 2021; *Ranked News Feed*, 2018). Rather than change the metrics causing the issues, Facebook created other algorithmic systems to counterbalance the negative effects.

These integrity systems are imperfect. They misclassify content, erroneously blocking posts from the platform or mistakenly allowing others to circulate (*Facebook and responsibility*, 2020). One internal study estimated that 5-10% of the News Feed's content comprises of borderline or low-quality content (*Negative feedback should be easy*, 2019, p. 5). These machine-learning algorithms miss harmful content and block potentially beneficial posts. Because of the active role that Facebook takes in rankings and filtering its content, one employee stated: "**There is no such thing as inaction on [the] Feed**" (original work's emphasis) (*Facebook and responsibility*, 2020, para. 5). The company is not a passive distributor of information but an active mediator; the company's use of flawed metrics makes it responsible for the outcomes of its system.

Within Facebook, many believe ranking algorithms outweigh their costs. A Facebook employee evaluating the benefits of the News Feed said:

> Even asking [whether ranking is beneficial] feels slightly blasphemous at Facebook. So many experiments and product launches demonstrate the value of ranking that its value is often taken as an article of faith. When I proposed a long-term ranking holdout to a colleague not long ago, we had a serious discussion about whether it would be ethical to deprive users of something so valuable. (*Is ranking good?*, 2018)

Ranking has become an integral part of Facebook's systems. Content tailoring underpins most of the platform's features and provides users with personalized experiences aimed to give them the most value. According to an internal study, limiting algorithmic ranking on the News Feed—all else equal—*increases* harmful content (*Ranked News Feed*, 2018, p. 22). Recommendation algorithms not only *promote* content but also *demote* content. Because of Facebook's reliance on curation, the company cannot safely shift to other technologies. The system has gained momentum which hinders the adoption of alternatives. Political theorist Langdon Winner noted that "once artifacts… have been built and put in operation, the kinds of reasoning that justify the adaptation of social life to technical requirements pop up as spontaneously as flowers in the spring" (Winner, 1980, p. 134). Facebook's reliance on engagement over almost two decades has codified the metrics within the company's systems and algorithms. Rather than reform the platform's foundations, Facebook patched additional curative systems into the News Feed, embedding—instead of removing—engagement from the platform's structures. Despite the continual evidence that virality spreads misinformation and increases polarization, Facebook remains fettered to its curative system.

Conclusion

Algorithmic curation acts as an invisible hand within Facebook. It mediates user interactions, ranking and sorting the words of billions without being seen. While providing a great social utility, the News Feed operates underneath simplifications of reality. There is neither a computable algorithm for determining truth nor for quantifying meaning. Algorithmic suggestions must be made from incomplete and often incorrect data. Facebook's reliance on engagement creates perverse incentives for content creators, enables the spread of misinformation by malefactors, and increases homophily in discourse. Through the compounding investments of both time and talent, Facebook has instilled its curative system with momentum, preventing the company from making meaningful changes to its underlying models. Many argue that Facebook's "relentless pursuit of profits" shaped its algorithms and designs (The Social Atrocity, 2022, p. 7); however, Facebook's history shows a company trying to meet the social and technical requirements demanded by its own curative system. In this way, Facebook is being led by an invisible hand to promote ends which were no part of its original intention.

Those aiming to improve Facebook should start at the foundations of its models, for virality and engagement in their current form have proven to be harmful. Facebook's internal momentum makes it unlikely for the company to shift away from its current metrics by its own volition. Substantial external efforts, through policy and regulation, should be implemented to redirect the course of the system to promote social integrity. Alternatively, research should be done on how to use engagement metrics safely. Cinelli et al. (2021) noted that Reddit experiences less user segregation than Facebook. Making the two companies (and others) exchange information about their systems may reveal what accounts for the difference and provide an actionable solution for reducing misinformation.

16

References

Aiello, L. M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., & Menczer, F. (2012). Friendship prediction and homophily in social media. *ACM Transactions on the Web*, *6*(2), 1–33. https://doi.org/10.1145/2180861.2180866

Alba, D. (2021, March 9). On Facebook, Misinformation is more popular now than in 2016. *The New York Times*. https://www.nytimes.com/2020/10/12/technology/on-facebook-misinformation-is-more-popular-now-than-in-2016.html

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Bijker, W. E., Hughes, T. P., & Pinch, T. J. (Eds.). (2012). *The Social Construction of Technological Systems: New directions in the sociology and history of technology* (Anniversary ed). MIT Press.

Bisgin, H., Agarwal, N., & Xu, X. (2012). A study of homophily on social media. *World Wide Web*, *15*(2), 213–232. https://doi.org/10.1007/s11280-011-0143-3

Boogart, M. R. V. (2006). *Uncovering the social impacts of Facebook on a college campus* [Master's thesis, Kansas State University]. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=94f02f57f73180e4158db06ac5df43b4b85b6e79

Burbach, L., Nakayama, J., Plettenberg, N., Ziefle, M., & Valdez, A. C. (2018). User preferences in recommendation algorithms: The influence of user diversity, trust, and product category on privacy perceptions in recommender algorithms. *Proceedings of the 12th ACM Conference on Recommender Systems*, 306–310. https://doi.org/10.1145/3240323.3240393

Carlson, N. (2009, September 15). Facebook: Cash-flow positive with 300 million users. *Business Insider*. https://www.businessinsider.com/facebook-cash-flow-positive-with-300-million-users-2009-9

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

Cotter, K., Cho, J., & Rader, E. (2017). Explaining the News Feed algorithm: An analysis of the "News Feed FYI" blog. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 1553–1560. https://doi.org/10.1145/3027063.3053114

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the*

National Academy of Sciences, *113*(3), 554–559.
https://doi.org/10.1073/pnas.1517441113

Dixon, S. (2023, August 29). *Facebook: Quarterly number of MAU (monthly active users) worldwide 2008-2023*. Statista. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Dwoskin, E. (2021, September 4). Misinformation on Facebook got six times more clicks than factual news during the 2020 Election, study says. *The Washington Post*. https://www.washingtonpost.com/technology/2021/09/03/facebook-misinformation-nyu-study/

Edelson, L., Nguyen, M.-K., Goldstein, I., Goga, O., McCoy, D., & Lauinger, T. (2021). Understanding engagement with U.S. (mis)information news sources on Facebook. *Proceedings of the 21st ACM Internet Measurement Conference*, 444–463. https://doi.org/10.1145/3487552.3487859

Ertug, G., Brennecke, J., Kovács, B., & Zou, T. (2022). What does homophily do? A review of the consequences of homophily. *Academy of Management Annals*, *16*(1), 38–69. https://doi.org/10.5465/annals.2020.0230

Facebook. (2006, September 26). *Facebook expansion enables more people to connect with friends in a trusted environment*. Meta. https://about.fb.com/news/2006/09/facebook-expansion-enables-more-people-to-connect-with-friends-in-a-trusted-environment/

Facebook and responsibility. (2020, March 9). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21594152-tier2_rank_other_0320

Facebook is creating a big echo chamber for "The government and public officials are lying to us" Narrative—Do we care? (2020, May 18). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21949635-tier2_fa_mi_c19_other_0520

Facebook's mission and vision statement: Analysis and implications. (2022, June 30). *Medium*. https://medium.com/@vberest/facebooks-mission-and-vision-statement-analysis-and-implications-59ab042a8870

Fu, F., Nowak, M. A., Christakis, N. A., & Fowler, J. H. (2012). The evolution of homophily. *Scientific Reports*, *2*(1), 845. https://doi.org/10.1038/srep00845

Géron, A. (2017). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (First edition). O'Reilly Media.

Gillespie, B. J., Lever, J., Frederick, D., & Royce, T. (2015). Close adult friendships, gender, and the life cycle. *Journal of Social and Personal Relationships*, *32*(6), 709–736. https://doi.org/10.1177/0265407514546977

Glossary: Glossary of terms commonly used in the Facebook Papers. (n.d.). The Facebook Papers. Retrieved March 24, 2024, from https://facebookpapers.com/glossary/

Google. (n.d.). *How results are automatically generated*. Google Search. Retrieved January 25, 2024, from https://www.google.com/search/howsearchworks/how-search-works/ranking-results/

Gowen, A., & Bearak, M. (2017, December 8). Fake news on Facebook fans the flames of hate against the Rohingya in Burma. *The Washington Post*. https://www.washingtonpost.com/world/asia_pacific/fake-news-on-facebook-fans-the-flames-of-hate-against-the-rohingya-in-burma/2017/12/07/2c1fe830-ca1f-11e7-b506-8a10ed11ecf5_story.html

Greiner, A., Fiegerman, S., Sherman, I., & Baker, T. (2019, February 1). *Facebook at 15: How a college experiment changed the world*. CNN. https://www.cnn.com/interactive/2019/02/business/facebook-history-timeline/index.html

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of Fake News dissemination on Facebook. *Science Advances*, *5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Hargreaves, E., Agosti, C., Menasche, D., Neglia, G., Reiffers-Masson, A., & Altman, E. (2018). Biases in the Facebook News Feed: A case study on the Italian elections. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 806–812. https://doi.org/10.1109/ASONAM.2018.8508659

Hargreaves, E., Mangabeira, E. F., Oliveira, J., Franca, T. C., & Mcnasche, D. S. (2020). Facebook News Feed personalization filter: A case study during the Brazilian elections. *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 615–618. https://doi.org/10.1109/ASONAM49781.2020.9381301

Hindman, M., Lubin, N., & Davis, T. (2022, February 10). Facebook has a superuser-supremacy problem. *The Atlantic*. https://www.theatlantic.com/technology/archive/2022/02/facebook-hate-speech-misinformation-superusers/621617/

Information Corridors: A brief introduction. (2021). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/23775111-tier0_po_exp_undated

Integrity glossary. (n.d.). [Internal Webpage]. The Facebook Papers. Retrieved March 23, 2024, from https://www.scribd.com/document/537183274/FULLY-REDACTED-Glossary-of-Facebook-Integrity-Terms

Integrity tradeoffs. (n.d.). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21600352-tier0_rank_ir_0120

Is ranking good? (2018, May 6). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21600996-tier2_rank_exp_0518

Jacoby, J. (Director). (2018a, October 29). *The Facebook Dilemma, Part one*. PBS. https://www.youtube.com/watch?v=T48KFiHwexM

Jacoby, J. (Director). (2018b, October 29). *The Facebook Dilemma, Part two*. PBS. https://www.youtube.com/watch?v=EuA4qxPbpQE

Kelly, H. (2017, June 22). Mark Zuckerberg explains why he just changed Facebook's mission. *CNN*. https://money.cnn.com/2017/06/22/technology/facebook-zuckerberg-interview/index.html

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, *115*(2), 405–450. https://doi.org/10.1086/599247

Lada, A., Wang, M., & Yan, T. (2021, January 26). *How does News Feed predict what you want to see?* Meta. https://about.fb.com/news/2021/01/how-does-news-feed-predict-what-you-want-to-see/

Lapides, P., Chokshi, A., Carpendale, S., & Greenberg, S. (2015). News Feed: What's in it for me? *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 163–172. https://doi.org/10.1145/2702123.2702554

Lazer, D. (2015). The rise of the social algorithm. *Science*, *348*(6239), 1090–1091. https://doi.org/10.1126/science.aab1422

Mac, R., & Kang, C. (2023, June 23). Whistle-blower says Facebook "Chooses profits over safety." *The New York Times*. https://www.nytimes.com/2021/10/03/technology/whistle-blower-facebook-frances-haugen.html

Mayntz, R., & Hughes, T. P. (1998). *The development of large technical systems*. Westview Press.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meaningful Social Interactions useful links. (n.d.). [Internal Webpage]. The Facebook Papers. Retrieved March 23, 2024, from https://www.documentcloud.org/documents/21601160-tier2_rank_other_undated

Meta. (2024). *Company Information*. Meta. https://about.meta.com/company-info/

Mill, J. S. (1859). *On Liberty*. Dover Publications.

Mosseri, A. (2016). Building a better News Feed for you. *Facebook Newsroom*. https://about.fb.com/news/2016/06/building-a-better-news-feed-for-you/

Mosseri, A. (2018, January 11). Bringing people closer together. *News Feed FYI*. https://www.facebook.com/business/news/news-feed-fyi-bringing-people-closer-together

Mozur, P. (2018, October 15). A genocide incited on Facebook, With posts from Myanmar's military. *The New York Times*. https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

MSI documentation. (n.d.). [Internal Webpage]. The Facebook Papers. Retrieved March 23, 2024, from https://www.documentcloud.org/documents/21601216-tier2_rank_other_undated-1

Netflix. (n.d.). *Recommendations: Figuring out how to bring unique joy to each member*. Netflix Research. Retrieved January 25, 2024, from https://research.netflix.com/research-area/recommendations

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Random House LLC.

Ortutay, B. (2022a, August 15). Facebook misses misinformation in Brazilian election ads. *PBS*. https://www.pbs.org/newshour/world/facebook-misses-misinformation-in-brazilian-election-ads

Ortutay, B. (2022b, September 29). Amnesty report finds Facebook amplified hate ahead of Rohingya massacre in Myanmar. *PBS*. https://www.pbs.org/newshour/world/amnesty-report-finds-facebook-amplified-hate-ahead-of-rohingya-massacre-in-myanmar

Pariser, E. (2011). *The Filter Bubble: How the new personalized web is changing what we read and how we think*. Penguin. https://books.google.com/books?id=wcalrOI1YbQC

Pathak, R., Spezzano, F., & Pera, M. S. (2023). Understanding the contribution of recommendation algorithms on misinformation recommendation and misinformation dissemination on social networks. *ACM Transactions on the Web*, 3616088. https://doi.org/10.1145/3616088

Pelley, S. (2021, October 4). Whistleblower: Facebook is misleading the public on progress against hate speech, violence, misinformation. *CBS News*. https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/

Pierri, F., Luceri, L., Jindal, N., & Ferrara, E. (2023). Propaganda and misinformation on Facebook and Twitter during the Russian invasion of Ukraine. *Proceedings of the 15th ACM Web Science Conference 2023*, 65–74. https://doi.org/10.1145/3578503.3583597

Political party response to '18 algorithm change. (2019, April 1). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21748430-tier2_eu_news_ir_0419

Providing negative feedback should be easy (And why this would be game changing for integrity). (2019, September 26). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21602498-tier1_rank_pr_0919

Rader, E. (2017). Examining user surprise as a symptom of algorithmic filtering. *International Journal of Human-Computer Studies*, *98*, 72–88. https://doi.org/10.1016/j.ijhcs.2016.10.005

Reagan, G. (2009, July 13). *The evolution of Facebook's mission statement*. https://observer.com/2009/07/the-evolution-of-facebooks-mission-statement/

Riemer, K., & Peter, S. (2021). Algorithmic Audiencing: Why we need to rethink free speech on social media. *Journal of Information Technology*, *36*(4), 409–426. https://doi.org/10.1177/02683962211013358

Roose, K., Isaac, M., & Frenkel, S. (2021, January 7). Facebook struggles to balance civility and growth. *The New York Times*. https://www.nytimes.com/2020/11/24/technology/facebook-election-misinformation.html

Silverman, C., Kao, J., & Merill, J. B. (2022, January 4). Facebook groups topped 10,000 daily attacks on election before Jan. 6, Analysis shows. *The Washington Post*. https://www.washingtonpost.com/technology/2022/01/04/facebook-election-misinformation-capitol-riot/

Sismondo, S. (2010). *An introduction to science and technology studies* (2nd ed.). Wiley-Blackwell.

Smith, A. (1776). *The wealth of nations* (R. H. Campbell, A. S. Skinner, & W. B. Todd, Eds.). Liberty Classics.

The meaningful social interactions metrics revisited: Part 2. (2019, November 4). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21602160-tier0_rank_exp_1119-1

The meaningful social interactions metrics revisited: Part 4. (2019, November 6). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21602160-tier0_rank_exp_1119-1

The social atrocity: Meta and the right to remedy for the Rohingya. (2022). Amnesty International. https://www.amnestyusa.org/wp-content/uploads/2022/09/Updated-Final-Report_The-Social-Atrocity-_0609202259.pdf

Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. *Journal of Computer-Mediated Communication*, *13*(3), 531–549. https://doi.org/10.1111/j.1083-6101.2008.00409.x

Twitter. (2023, March 31). *Twitter's recommendation algorithm*. X Blog. https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm

Walther, J. B., Van Der Heide, B., Kim, S.-Y., Westerman, D., & Tong, S. T. (2008). The role of friends' appearance and behavior on evaluations of individuals on Facebook: Are we known by the company we keep?: Facebook friends and individual evaluations. *Human Communication Research*, *34*(1), 28–49. https://doi.org/10.1111/j.1468-2958.2007.00312.x

What happens if we delete ranked News Feed? (2018, February 28). [Internal Webpage]. The Facebook Papers. https://www.documentcloud.org/documents/21748432-tier2_news_ir_0218

Why signals is a must-win. (n.d.). Meta. Retrieved March 23, 2024, from https://www.facebook.com/business/m/HSBSeries/signals

Why we build Feeds. (2019, October 4). [Internal Webpage]. *The Facebook Papers*. https://www.documentcloud.org/documents/21600853-tier1_rank_exp_1019

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

Wodinsky, S. (2021, November 4). *35 internal code words Facebook uses to talk about Its users and tools*. Gizmondo. https://gizmodo.com/35-secret-code-words-facebook-uses-to-talk-about-its-us-1847988690