

## The Black Box: How Toxic Algorithms Keep Social Media in Business

Elyse Jackson

Social Media has evolved a significant amount since its humble beginnings as a simple platform made to connect college students. In February 2004, Facebook launched online founded by Harvard students Mark Zuckerberg, Andrew McCollum, Dustin Moskovitz, Eduardo Saverin, and Chris Hughes.<sup>1</sup> The premise of the social network was to create a literal book of faces for university students to use to connect.<sup>2</sup> While at first intended only for Harvard students, it has since expanded nationwide and then worldwide.<sup>3</sup> What was once a tool to allow communities to connect has become a complex system of algorithms that work to create echo chambers, push annoying (sometimes scam) ads, and promote toxic content.<sup>4</sup> This toxic content can include disinformation, misinformation, hate speech, and downright violence.<sup>5</sup> While mostly hidden behind a black box of mystery, social media algorithms direct user traffic and control the distribution of media on its prescribed platform.<sup>6</sup> As if what we see on our feeds daily isn't bad enough, a data scientist whistleblower has exposed internal documents that demonstrate that Facebook chooses its financial interests over the general good of the public.<sup>7</sup> In other words, if harmful content makes money, then the algorithm will promote it regardless of its effects on society.

In 2021, a Facebook whistleblower by the name of Frances Haugen revealed that Facebook's algorithm promotes dangerous content but chose not to do anything about it to "maximize advertising revenue."<sup>8</sup> Haugen noted that hateful content and fake news drove up engagement both on Facebook and Instagram.<sup>9</sup> Engagement on social media means likes, comments and views.<sup>10</sup> This engagement keeps users on the app and constitutes 95% of Facebook's revenue.<sup>11</sup> The whistleblower filed eight complaints against Facebook with the Security and Exchanges Commission (SEC) backed by tens and thousands of secretly copied

---

<sup>1</sup> Mythili Devarakonda, *'The Social Network': When Was Facebook Created? How Long Did It Take To Create Facebook?*, USA TODAY (July 25, 2022), <https://www.usatoday.com/story/tech/2022/07/25/when-was-facebook-created/10040883002/#:~:text=Facebook%2C%20initially%20titled%20as%20thefacebook,and%20Chris%20Hughes%20founded%20Facebook.>

<sup>2</sup> Sean Michael Kerner, *What is Facebook?*, TECH TARGET (Nov. 2021), <https://www.techtarget.com/whatis/definition/Facebook#:~:text=Facebook%20is%20a%20social%20networking,to%20connect%20and%20share%20information.>

<sup>3</sup> *Id.*

<sup>4</sup> Prabir Purkayastha, *How Facebook Algorithms Promote Hate and Toxic Content*, PEOPLE'S DISPATCH (Oct. 8, 2021), [https://peoplesdispatch.org/2021/10/08/how-facebook-algorithms-promote-hate-and-toxic-content/.](https://peoplesdispatch.org/2021/10/08/how-facebook-algorithms-promote-hate-and-toxic-content/)

<sup>5</sup> *Id.*

<sup>6</sup> *See generally*, Hal Hodson, *The Secret System Controlling Your Facebook News Feed*, NEW SCIENTIST (July 30, 2014), <https://www.newscientist.com/article/mg22329804-200-the-secret-system-controlling-your-facebook-news-feed/#:~:text=Facebook's%20algorithms%20are%20a%20closely,to%20remaining%20non%2Dtransparent.%E2%80%9D.>

<sup>7</sup> Prabir Purkayastham, *supra* note 4.

<sup>8</sup> *Id.*

<sup>9</sup> *Id.*

<sup>10</sup> *Id.*

<sup>11</sup> *Id.*

internal documents.<sup>12</sup> In other words, Facebook is no longer a platform to connect communities as it was in 2004, it is now an intricate business model used to skillfully exploit our subconscious inclinations to sell merchandise and advertising space. Haugen goes on to say that “Facebook’s own research says they cannot adequately identify dangerous content. And as a result, those dangerous algorithms that they admit are picking up the extreme sentiments...”<sup>13</sup>

The question to ask then becomes what is the software code that makes up these unregulated algorithms? Unfortunately for users, again, these algorithms are kept hush-hush due to them being held as trade secrets to gain a competitive advantage over other platforms.<sup>14</sup> However, one loophole to trade secrets is that it is perfectly legal to reverse engineer them to figure out the secret sauce, so to speak.<sup>15</sup> This is the key to understanding how the mysterious Facebook algorithm works and to take steps as users to help make changes in the absence of formal regulations.

In 2014, two researchers at the University of Illinois and the University of Michigan endeavored to exploit this reverse engineering loophole and conducted an experiment by creating an app called FeedVis.<sup>16</sup> This app created a stream of all the comments, likes and posts of only friends.<sup>17</sup> Compared to the Facebook feed with its carefully curated selection of ads and inflammatory content, users actually saw content within their network that the algorithm hid.<sup>18</sup> In the first study using this app, 62% of people did not know that their network’s posts were being intentionally suppressed and were upset when they found out.<sup>19</sup> During the experiment, the team slowly began to uncover the mysteries by discovering that if you comment on someone’s wall, you’re more likely to see their content.<sup>20</sup> However, the algorithm changes constantly, so what the team discovered in 2014 is likely no longer the exact same.<sup>21</sup>

There is a need for regulation of social media platforms like Facebook, so they will be obligated to take responsibility for the inflammatory and often harmful content that it allows. How much regulation, however, is too much and where should a legislative line be drawn to maintain freedom of expression? In 2021, approximately five bills were introduced to attempt to regulate the software programs that makeup algorithms.<sup>22</sup> Each of these proposals suggests an amendment to Section 230 of the Communications Decency Act.<sup>23</sup> In its current state, Section 230 protects websites and apps from being sued for hosting user-posted content (also called user-generated content).<sup>24</sup> Some bills have proposed removing this protection for certain kinds of harmful content that the algorithm promotes.<sup>25</sup> As a student practically raised on varying forms

---

<sup>12</sup> *Id.*

<sup>13</sup> *Id.*

<sup>14</sup> Hal Hodson, *supra* note 6.

<sup>15</sup> *Reverse engineering*, LEGAL INFORMATION INSTITUTE,

[https://www.law.cornell.edu/wex/reverse\\_engineering#:~:text=Reverse%20engineering%20is%20generally%20legal,engineering%20is%20not%20a%20defense](https://www.law.cornell.edu/wex/reverse_engineering#:~:text=Reverse%20engineering%20is%20generally%20legal,engineering%20is%20not%20a%20defense) (last visited Dec. 14, 2023).

<sup>16</sup> Hal Hodson, *supra* note 6.

<sup>17</sup> *Id.*

<sup>18</sup> *Id.*

<sup>19</sup> *Id.*

<sup>20</sup> *Id.*

<sup>21</sup> *Id.*

<sup>22</sup> Will Oremus, *Lawmakers’ Latest Idea to Fix Facebook: Regulate The Algorithm*, THE WASHINGTON POST (Oct. 12, 2021), <https://www.washingtonpost.com/technology/2021/10/12/congress-regulate-facebook-algorithm/>.

<sup>23</sup> *Id.*

<sup>24</sup> *Id.*

<sup>25</sup> *Id.*

of social media platforms, this would be a significant victory. It would force accountability and increased measures to prevent the widespread dissemination of toxic and harmful content. Another bill proposed by Rep. Tom Malinowski (D-N.J.) and Rep. Anna G. Eshoo (D-Calif.) suggests only removing Section 230 protections when algorithms “promote content related to civil rights violations or international terrorism.”<sup>26</sup> This relatively cautious proposal was intended to narrowly tailor the law to address the problem while preserving the integrity of social media platforms like Facebook.<sup>27</sup> Yet more proposals seek to remove Section 230 protection when the platform promotes medical misinformation or to censor certain political views (a less popular suggestion).<sup>28</sup>

It is very difficult, yet incredibly important to understand where the line should be drawn to prevent over-censorship on the internet. Allie Funk, a senior research analyst at the non-profit Freedom House says that often companies will err on the side of censorship “and removing political, social and religious speech, particularly of those in marginalized communities” when regulations are instituted.<sup>29</sup> While the debate continues to rage over appropriate legislation and over whether TikTok is a tool for Chinese upper-level espionage, algorithmic models continue to wreak havoc in certain areas by promoting content that disrupts communities and harms everyday people. While there are certainly communities of people dedicated to debunking misinformation and educating the masses, there are still some who are not quite so technologically savvy to be able to distinguish the real from the fake.

---

<sup>26</sup> *Id.*

<sup>27</sup> *Id.*

<sup>28</sup> *Id.*

<sup>29</sup> *Id.*