Harvard Journal of Law & Technology Digest
Online-only publication

Fake News and the Looming "State Action" Problem

Hilary Hurd[*]

Published March 21, 2019
Original link: https://jolt.law.harvard.edu/digest/fake-news-and-the-looming-state-action-problem

Recommended Citation

Hilary Hurd, Note, *Fake News and the Looming "State Action" Problem*, Harv. J.L. & Tech. Dig. (2019), https://jolt.law.harvard.edu/digest/fake-news-and-the-looming-state-action-problem.

[*]Hilary Hurd is a 2L at Harvard Law School. She previously studied international security in the UK on a Marshall Scholarship. Special thanks to Professor Jack Goldsmith and Professor Martha Minow for their helpful comments on this paper.

Abstract

With over 2.34 billion active users,[1] Facebook is the world's largest social media platform.[2] People use it to share intimate information about their lives, but also to share "fake news," either for a devious purpose or, perhaps more troublingly, because they assume the information is true. The 2016 "Pizzagate" scandal epitomized how a false Facebook post could torpedo across the internet, leading a gullible person to take up arms;[3] meanwhile, many books and articles evaluating Russia's impact on the 2016 election underscore just how pervasive false advertisements can be.[4]

This paper explores the "fake news" problem and the challenges it poses for society. It explains what steps Facebook has taken to minimize the effects of fake news on its platform and the technological and philosophical challenges involved. Rather than explore policy proposals to minimize fake news, it is principally concerned with one US legal doctrine which could thwart most policy proposals outright: "state action" theory. It explains how state action theory might be revitalized and why such a revitalization would exacerbate fake news.

[1] Defined as a user who has logged on within the last 30 days. *Number of Monthly Active Facebook Users Worldwide as of 4th Quarter 2018 (in Millions)*, Statista.com, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ (last visited Mar. 10, 2019).

[2] *See Most Famous Social Network Sites Worldwide as of January 2019, Ranked by Number of Active Users (in Millions)*, Statista.com, https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ (last visited Mar. 10, 2019).

[3] *See* Amanda Robb, *Anatomy of a Fake News Scandal*, Rolling Stone (Nov. 16, 2017), https://www.rollingstone.com/politics/politics-news/anatomy-of-a-fake-news-scandal-125877.

[4] *See, e.g.*, Yochai Benkler, Robert Faris & Hal Roberts, Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics (2018).

Table of Contents

I. Understanding the "Fake News" Problem

Facebook insists that it's not a media company—yet a majority of subscribers use Facebook for news.[5] A 2016 study by Pew Research Service found that 66% of people use Facebook for news, compared with just 21% of users on YouTube and 59% on Twitter.[6] Of those who use Facebook for news, an estimated 64% are unlikely to get news from another social media site.[7]

Having a topical News Feed provides many benefits, but the consequences can be disastrous when the information is false. "Fake news" on Facebook is especially pernicious. Not only is the "real identity" of any given Facebook account easier to camouflage than a printed source,

[5] *See* Michelle Castillo, *Zuckerberg Tells Congress Facebook is not a Media Company: 'I Consider Us to be a Technology Company,'* CNBC News (Apr. 11, 2018), https://www.cnbc.com/2018/04/11/mark-zuckerberg-facebook-is-a-technology-company-not-media-company.html.

[6] *See* Jeffrey Gottfried & Elisa Shearer, *News Use Across Social Media Platforms 2016*, Pew Research Center (July 7, 2016), http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/.

[7] *See* Amy Michell et al., *How Americans Get Their News*, Pew Research Center (July 7, 2016), http://www.journalism.org/2016/07/07/pathways-to-news/ (surveying where most people get their news in and finding 50% of individuals between the ages of 18-49 receive their information online).

the prospect of making money through online advertising creates a powerful financial incentive to deceive.[8] Though not all Facebook users are inclined to believe, or share, false stories they see on the internet, stories gain credibility when shared online by a friend.[9]

"Fake news" means different things to different people. Clair Wardle and Hossein Derakhshan break the term "fake news" into three categories: "mis-information" (false information shared without harmful intent); "dis-information" (false information shared with harmful intent); and lastly "mal-information" (genuine information shared to cause harm).[10] Facebook seems to apply related distinctions. In Facebook's 2018 video, "Facing Facts," data science manager, Eduardo Ariño de La Rubia, classifies online content into a four-box matrix.[11] On the x-axis, information increases by truthfulness, moving from less to more true. On the y-axis, the intention underlying the post moves from innocent to devious. Facebook's fake news work is mostly focused on information that is in the upper left quadrant: namely, information that is "less true" but also "devious" by intention. This is what Facebook is traditionally referring to when it talks about "disinformation" or "hoaxes."[12] Nonetheless, Facebook is also concerned with propaganda (especially when coordinated by foreign governments), which falls in the upper right quadrant: "devious" but not necessarily untrue.[13]

---

[8] In 2016, Buzzfeed identified more than 100 pro-Trump news sites with American-sounding domain names operated out of Macedonia. Their scheme was simple: because Google's automated advertising engine, "AdSense," paid them for every click on their websites, they began publishing intentionally salacious political stories that would prompt Facebook users to read sites. *See* Craig Silverman & Lawrence Alexander, *How Teens In The Balkans Are Duping Trump Supporters With Fake News*, Buzzfeed News (Nov. 3, 2016), https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinformation.

[9] *See* Regina Regi, *Fake News and Partisan Epistemology*, 27 Kennedy Inst. of Ethics J. 43 (2017).

[10] *See* Claire Wardle & Hossein Derakhshan, *Information Disorder: Toward and Interdisciplinary Framework for Research and Policymaking*, Council of Eur. (Sept. 27, 2017), https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c.

[11] *See* Facebook, *Facing Facts: An Inside Look at Facebook's Fight against Misinformation*, YouTube (May 23, 2018), https://www.youtube.com/watch?v=zgkF23nFIBw.

[12] *See id.*

[13] *See id.*

Facebook's categories illustrate how intention interacts with truth, but they obfuscate the difficulty in determining "intention." For example, satirical headlines from *the Onion* or *The New Yorker*'s "Borowitz Report" often contain false information with the goal of making an overarching "true" point, but not every reader necessarily regards those stories as satirical. In a story by Christian magazine, *Babylon Bee*, CNN reportedly purchased a giant washing machine in which to "spin" its stories.[14] While many readers might have understood the story was a satirical jab at the liberal media, one of the five fact-checking organizations employed by Facebook, Snopes, flagged the story as "false."[15]

II. WHAT IS FACEBOOK DOING ABOUT FAKE NEWS?

The 2016 Presidential election brought our attention to the digital "fake news" problem. One month later, Facebook revealed its plans to address the issue of "fakes and hoaxes" online.[16] Since then, Facebook has introduced new products and techniques to 1) identify false content, mostly falling within the "hoaxes" category though sometimes extending to "propaganda;" 2) remove fake accounts; and 3) tighten regulation of advertisements.[17] Because much fake news is a financial effort to make money through advertising, Facebook believes the solution lies in removing the underlying financial incentives.

---

[14] *See CNN Purchases Industrial-Sized Washing Machine To Spin News Before Publication*, BABYLON BEE (Mar. 1, 2018), https://babylonbee.com/news/cnn-purchases-industrial-sized-washing-machine-spin-news-publication.

[15] *See* Daniel Funke, *Should Satire be Flagged on Facebook? A Snopes Debunk Sparks Controversy*, POYNTER (Mar. 2, 2018), https://www.poynter.org/news/should-satire-be-flagged-facebook-snopes-debunk-sparks-controversy. Meanwhile, some sites post a "satire" defense to avoid liability. For example, a network of websites run by an infamous hoaxer, Christopher Blair, frequently uses salacious headlines to increase the number of "hits" and advertising revenue generated, even though each article includes a footnoted "satire" disclaimer. *See* Daniel Funke, *Satirical Fake News Site Apologized for Making a Story Too Real*, POYNTER (Nov. 30, 2017), https://www.poynter.org/news/satirical-fake-news-site-apologized-making-story-too-real.

[16] *See* Adam Mosseri, *Addressing Hoaxes and Fake News*, FACEBOOK (Dec. 15, 2016), https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/.

[17] *See* Rob Goldman, *Changes We Made to Ads in 2018*, FACEBOOK (Dec. 21, 2018), https://www.facebook.com/business/news/changes-we-made-to-ads-in-2018.

A. False Content

Facebook's "Community Standards" expressly prohibit hate speech and credible incitements to violence,[18] but they don't expressly forbid false content.[19] For example, when Facebook removed pages run by Alex Jones in 2018, they cited his hateful and bullying speech, not falsehood, as the rationale.[20] Nonetheless, while Facebook does not "remove" false content,[21] it tries to identify it and minimize its spread. To do so, Facebook combines self-reporting by users and third-party identification by independent fact-checkers, including *Factcheck*, *Snopes*, the *Associated Press*, *Politifact*, and the *Weekly Standard*.[22] Initially, Facebook flagged any story identified as false in users' News Feeds with a "disputed" signal and a corresponding article that explained the decision.[23] These stories subsequently appeared lower in the News Feed, reducing views by over 80%.[24] No story flagged as "disputed" could be converted into an advertisement.[25] Facebook later found that the disputed signal actually entrenched deeply held beliefs and was thus counterproductive in discouraging certain media consumption.[26] In 2017, Facebook replaced the

[18] *See Community Standards*, Facebook, https://www.facebook.com/communitystandards (last visited Mar. 11, 2019).

[19] *See Enforcing Our Community Standards*, Facebook (Aug. 6, 2018), https://newsroom.fb.com/news/2018/08/enforcing-our-community-standards/.

[20] *See id.* In an August 2018 post, Facebook went so far as to analogize its rules about false content to the International Covenant on Civil and Political Rights (ICCPR), Article 19, stating that "[h]uman rights law extends the same right to expression to those who wish to claim that the world is flat as to those who state that it's round and so does Facebook. *See* Richard Allan, *Hard Questions: Where Do We Draw the Line on Free Expression*, Facebook (Aug. 9, 2018), https://newsroom.fb.com/news/2018/08/hard-questions-free-expression/.

[21] Barring, of course, any false content that incites violence. *See* Mark Zuckerberg, *Preparing for Elections*, Facebook (Sept. 13, 2018), https://www.facebook.com/notes/mark-zuckerberg/preparing-for-elections/10156300047606634/.

[22] For full list of current fact-checkers see *Verified Signatories of the IFCN Code of Principles*, Poynter, https://ifcncodeofprinciples.poynter.org/signatories (last visited Mar. 11, 2019).

[23] *See* Adam Mosseri, *Addressing Hoaxes and Fake News*, Facebook (Dec. 15, 2016), https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/.

[24] *See* Tessa Lyons, *Replacing Disputed Flags With Related Articles*, Facebook (Dec. 20, 2017), https://newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation/.

[25] *See* Mosseri, *supra* note 23.

[26] *See* Lyons, *supra* note 24.

disputed signal with links to "related articles" whenever a user went to share a disputed post.[27] The "related articles" contained similar, presumably more accurate, information on the same topic. Facebook nonetheless continued to demote disputed stories and prevent disputed stories from being used for advertisements. Publishers whose stories are flagged "false" can contact fact-checkers for challenge or correction.[28] While it's not public how many "disputed" posts external organizations identify,[29] some reports indicate that Facebook doesn't provide its fact-checkers sufficient financial or strategic support.[30] Facebook acknowledges some ongoing challenges associated with external partners, including 1) Facebook currently lacks fact-checkers in some countries; 2) different countries have different journalism standards; and 3) "it can take hours or even days to review a single claim."[31] Nonetheless, Facebook has sought to amplify fact-checkers' work through machine learning techniques, which can identify duplicates of previously debunked stories and to flag posts that replicate those messages across the internet.[32]

---

[27] *See* Mosseri, *supra* note 23.

[28] *See Third-Party Fact-Checking on Facebook*, Facebook, https://www.facebook.com/help/publisher/182222309230722 (last visited Mar. 11, 2019).

[29] *See* Sam Levin, *'They Don't Care': Facebook Factchecking in Disarray as Journalists Push to Cut Ties*, The Guardian (Dec. 13, 2018), https://www.theguardian.com/technology/2018/dec/13/they-dont-care-facebook-fact-checking-in-disarray-as-journalists-push-to-cut-ties.

[30] For example, of the eight employees working at one partnership, Factcheck, only two people are expressly devoted to Facebook-related activities for which Facebook pays the non-profit a total of $189,000 per year. *See* Georgia Wells & Lukas I. Alpert, *Facebook's Effort to Fight Fake News, Human Fact-Checkers Struggle to Keep Up*, Wall Street J. (Oct. 18, 2018), https://www.wsj.com/articles/in-facebooks-effort-to-fight-fake-news-human-fact-checkers-play-a-supporting-role-1539856800; *see also* Daniel Funke & Alexios Mantzarlis, *We Asked 19 Fact-Checkers What They Think of Their Partnership With Facebook. Here's What They Told Us*, Poynter (Dec. 14, 2018), https://www.poynter.org/fact-checking/2018/we-asked-19-fact-checkers-what-they-think-of-their-partnership-with-facebook-heres-what-they-told-us/; Casey Newton, *The Trauma Floor: The Secret Lives of Facebook Moderators in America*, The Verge (Feb. 25, 2019), https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.

[31] *See* Tessa Lyons, *Hard Questions: How Is Facebook's Fact-Checking Program Working*, Facebook (June 14, 2018), https://newsroom.fb.com/news/2018/06/hard-questions-fact-checking/.

[32] *See* Tessa Lyons, *Increasing Our Efforts to Fight False News*, Facebook (Jun. 21, 2018), https://newsroom.fb.com/news/2018/06/increasing-our-efforts-to-fight-false-news/. For example, a fact-checker in France flagged a story which said you can rescue stroke victims by pricking their fingers with a needle. Using machine learning, Facebook identified 20 domains and over 1,400 links spreading the identical claim

B. Fake Accounts

While Facebook's standards don't prohibit fake news, they do require "authenticity"—meaning that users cannot misrepresent their identities online by using a fake account.[33] When Facebook identifies these misleading accounts, it deletes them. For example, Facebook took down a network of more than 270 pages and accounts associated with the Russian Internet Research Agency (IRA) because the IRA "repeatedly used complex networks of inauthentic accounts to deceive and manipulate people who use Facebook, including before, during and after the 2016 US presidential elections."[34] Facebook acknowledged some pages did not contain false content, but emphasized the pages' inauthentic creation.[35] Drawing upon the Facebook's four-part framework, these "inauthentic" campaigns resemble "propaganda" more closely than "hoaxes." In May 2018, Facebook released its first "Community Standards Enforcement Report," which tracks Facebook's progress in minimizing fake accounts—along with spam, violence and graphic content, adult nudity, hate speech, and terrorist propaganda.[36] In the report, Facebook estimated that approximately 3-4% of its active users have "fake accounts."[37] Facebook disabled 1.3 billion of those accounts in early 2018, reporting that it identified almost all of these fake accounts using its software al-

---

and diminished the story's reach. *See id.* Facebook also uses machine learning to identify and prevent "cloaking," a technique whereby bad actors disguise a post's or advertisement's ultimate web-destination, for things like "diet pills, pornography, and muscle building scams," to avoid review. *See* Rob Leathern & Bobbie Chang, *Addressing Cloaking So People See More Authentic Posts*, Facebook (Aug. 9, 2017), https://newsroom.fb.com/news/2017/08/news-feed-fyi-addressing-cloaking-so-people-see-more-authentic-posts/.

[33] *See Community Standards*, *supra* note 17.

[34] *See* Jason Murdock, *What is the Internet Research Agency? Facebook Shuts Hundreds of Accounts Linked to Russian Troll Factory*, Newsweek (Apr. 4, 2018), https://www.newsweek.com/what-internet-research-agency-facebook-shuts-hundreds-accounts-linked-russia-870889.

[35] *See* Alex Stamos, *Authenticity Matters: The IRA Has No Place on Facebook*, Facebook (Apr. 3, 2018), https://newsroom.fb.com/news/2018/04/authenticity-matters/.

[36] *See Community Standards Enforcement Report*, Facebook, https://transparency.facebook.com/community-standards-enforcement (last updated Nov. 15, 2018).

[37] *See id.*

gorithm (98.5% as of Q3), while individual users reported the remaining fake accounts.[38]

C. "Fake News" Advertising

The 2016 election raised new questions about the ability of foreign actors to target Americans with political advertisements.[39] Facebook estimates that 10 million Americans saw Russia-sponsored advertisements shortly before and after the 2016 election.[40] These advertisements were about divisive political and social issues, costing only $6,700 total.[41] In April 2017, Facebook shut down a network of 47 accounts and pages believed to be operated out of Russia with the purpose of spreading fake news. Six months later, Facebook delivered 3,000 advertisements from this network to Congress.[42]

---

[38] *See* Kurt Wagner & Rani Molla, *Facebook Has Disabled Almost 1.3 billion Fake Accounts Over the Past Six Months*, RECODE (May 15, 2018), https://www.recode.net/2018/5/15/17349790/facebook-mark-zuckerberg-fake-accounts-content-policy-update.

[39] *See* Philip Ewing, *Russians Targeted US. Racial Divisions Long Before 2016 And Black Lives Matter*, NPR (Oct. 30, 2017), https://www.npr.org/2017/10/30/560042987/russians-targeted-u-s-racial-divisions-long-before-2016-and-black-lives-matter.

[40] *See* Elliot Schrage, *Hard Questions: Russian Ads Delivered to Congress*, FACEBOOK (Oct. 2, 2017), https://newsroom.fb.com/news/2017/10/hard-questions-russian-ads-delivered-to-congress/.

[41] *See id.*

[42] *See* Mike Isaac & Scott Shane, *Facebook to Deliver 3,000 Russia-Linked Ads to Congress on Monday*, N.Y. TIMES (Oct. 1, 2017), https://www.nytimes.com/2017/10/01/technology/facebook-russia-ads.html. Before 2016, Facebook recognized the impact of its News Feed in shaping political attitudes and associations. In 2012, for example, Facebook had published a report on the 2010 midterms, showing that political mobilization messages on Facebook "directly influenced political self-expression, information seeking and real-world voting behaviour of millions of people." Nonetheless, the notion that foreign actors could infiltrate the US political debate came as a surprise. *See* Robert Bond et al., *A 61-million-Person Experiment in Social Influence and Political Mobilization*, FACEBOOK (Sept. 13, 2012), https://research.fb.com/publications/a-61-million-person-experiment-in-social-influence-and-political-mobilization/?mod=article_inline. In 2017, Facebook announced new protocols to prevent a similar campaign from occurring, focusing principally on political advertisements. *See* Mark Zuckerberg, FACEBOOK (Sept. 21, 2017), https://www.facebook.com/zuck/posts/10104052907253171; *see also*, Rob Goldman, *Update on Our Advertising Transparency and Authenticity Efforts*, FACEBOOK (Oct. 27, 2017), https://newsroom.fb.com/news/2017/10/update-on-our-advertising-transparency-and-authenticity-efforts/.

Unlike the vast majority of Facebook content, which lacks pre-publication review, Facebook analyzes advertisements before publication, using both automated and manual review, to ensure they don't violate Facebook's twenty-nine guidelines for prohibited content.[43] These guidelines range from straight-forward rules to complex standards related to misinformation. For example, guideline thirteen states: "[a]ds, landing pages, and business practices must not contain deceptive, false, or misleading content, including deceptive claims, offers, or methods."[44] Unlike its rules for normal posts, Facebook distinguishes between political and non-political advertisements.[45] In 2017, Facebook enhanced its authenticity requirements for anyone paying to run election-related ads on Facebook, requiring that they disclose their identities and locations.[46] In 2018, Facebook extended the disclosure requirement to anyone who wanted to show politically charged "issue ads,"[47] with a specific list of issues tailored by country.[48] Facebook's policy allegedly applies to ads that take a position on those issues so to influence public debate, promote ballot measures, or elect candidates.[49]

Overall, Facebook's success is mixed. A study by New York University and Stanford found that, after the 2016 election, users' interaction with fake news stories declined by over 50% on Facebook, while they in-

---

[43] *See Advertising Policy*, Facebook (2018), https://www.facebook.com/policies/ads/restricted_content/political (last visited Mar. 11, 2019).

[44] *Id.*

[45] *See id.* For example, Facebook guideline fourteen says that "[a]ds must not contain content that exploits controversial political or social issues for commercial purposes." *Id.* Facebook has declined to publish further specifics.

[46] *See* Goldman, *supra* note 42.

[47] *See* Rob Goldman & Alex Himel, *Making Ads and Pages More Transparent*, Facebook (Apr. 6, 2018), https://newsroom.fb.com/news/2018/04/transparent-ads-and-pages/.

[48] Facebook has worked with the non-partisan Comparative Agendas Project (CAP) to compile the issue list for the U.S.. Many of the issues included on the list include innocuous words like "values" and "health." *See Ads Related to Politics or Issues of National Importance*, Facebook, https://www.facebook.com/business/help/214754279118974 (last visited Mar. 11, 2019).

[49] *See* Katie Harbath & Steve Satterfield, *Why Doesn't Facebook Just Ban Political Ads*, Facebook (May 24, 2018), https://newsroom.fb.com/news/2018/05/hard-questions-political-ads/.

creased on Twitter.[50] Similarly, another study by the fact-checking arm of French newspaper, *Le Monde*, showed a 50% drop in engagement with fake news sources in France since 2015.[51] Nonetheless, sites repeatedly flagged by Facebook as false still appear on the platform,[52] and Facebook-user engagements with fake news sites still hover around 70 million per month.[53]

[50] *See* Hunt Allcott, Matthew Gentzkow & Chuan Yu, *Trends in the Diffusion of Misinformationon Social Media*, Stan. U. (Oct. 2018), https://web.stanford.edu/g̃entzkow/research/fake-news-trends.pdf.

[51] *See* Adrien Sénécat, *Les Fausses Informations Circulent de Moins en Moins sur Facebook*, Le Monde (Oct. 17, 2018), https://www.lemonde.fr/les-decodeurs/article/2018/10/17/les-fausses-informations-perdent-du-terrain-sur-facebook_5370461_4355770.html.

[52] *See* Daniel Funke, *Fact-Checkers Have Debunked This Fake News Site 80 Times. It's Still Publishing on Facebook*, Poynter (Jul. 20, 2018), https://www.poynter.org/news/fact-checkers-have-debunked-fake-news-site-80-times-its-still-publishing-facebook.

[53] *See* Allcott, *supra* note 50. In the advertising context, recent stories suggest insufficient compliance. As the New York Times reported in June 2018, a Democratic candidate running for Congress in California, Regina Bateson, was explicitly targeted by negative political advertisements by a group that failed to identify itself in keeping with Facebook's rules, not to mention the FEC's updated 2017 guidelines, which similarly require the identity and location of the information's sponsor for online Facebook ads. The man responsible, Paul Smith, said he had successfully placed many political advertisements on Facebook without appropriate labelling. *See* Sheera Frenkel, *Facebook Tried to Rein In Fake Ads. It Fell Short in a California Race*, N.Y. Times (Jun. 3, 2018), https://www.nytimes.com/2018/06/03/technology/california-congressional-race-facebook-election-interference.html. Shortly before the 2018 midterm elections, VICE journalists posed as US Senators to purchase ads from Facebook. Not only were all of the political advertisements approved, they proceeded to share content from fake groups. *See* William Turton, *We Posed as 100 Senators to Run Ads on Facebook. Facebook Approved all of Them*, Vice News (Oct. 30, 2018), https://news.vice.com/en_us/article/xw9n3q/we-posed-as-100-senators-to-run-ads-on-facebook-facebook-approved-all-of-them. The problem is not limited to the US. In October 2018, the New York Times reported that a non-UK based entity "Mainstream Network," had engaged in a massive Brexit campaign, reaching approximately 11 million voters with information about how to send pre-written letters to members of Parliament explaining their opposition to Prime Minister May's negotiation plans. *See* Adam Satariano, *Facebook Ads From Unknown Backer Take Aim at Brexit Plan*, N.Y. Times (Oct. 19, 2018), https://www.nytimes.com/2018/10/19/technology/facebook-brexit-ads.html.

D. Censorship

An outstanding challenge for Facebook is its News Feed algorithm, currently optimized to maximize the "engagement" of its users.[54] Articles not likely to "engage" a user based on their prior conduct (e.g. what they have liked, shared, or posted) are deprioritized in favor of articles that prompt a user to respond, whether because users approve of something and "like" it or because they are outraged. Because fake news stories are frequently designed to be intentionally "outrageous," they receive an algorithmic boost over more reliable content.[55] The aforementioned techniques Facebook uses to minimize "fake news" through diminishing its presence in the News Feed counteract that algorithmic boost but have led some to criticize Facebook for explicit censorship. They are not wrong. After all, the consequence of "diminishing" posts by 80% versus removing them is not substantial.[56]

III. Upending the "Fake News" Project with State Action

Facebook is a private entity unrestrained by the Constitution.[57] So far, efforts to apply the First Amendment to it have failed.[58] Nonetheless, it's wrong to assume that constitutional restraints will never reach

[54] *See* Adam Mosseri, *Bringing People Closer Together*, Facebook (Jan. 11, 2018), https://newsroom.fb.com/news/2018/01/news-feed-fyi-bringing-people-closer-together/.

[55] Zuckerberg acknowledged this problem in a November 2018 blog post. He said, "[W]hen left unchecked, people will engage disproportionately with more sensationalist and provocative content . . . no matter where we draw the lines for what is allowed." Therefore, he said that Facebook would start to penalize "borderline content" in the algorithm so that it receives fewer views. Nonetheless, it remains unclear how Facebook will identify "borderline content." *See* Mark Zuckerberg, *A Blueprint for Content Governance and Enforcement*, Facebook (Nov. 15, 2018), https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/.

[56] While Facebook repeatedly champions the right of users to be wrong, it's hard to reconcile that position with the effects of their policies, which cause disputed posts to lose 80% of their views when flagged. *See* Lyons, *supra* note 24.

[57] Only the Thirteenth Amendment applies to private actors. *See, e.g.*, Jones v. Alfred H. Mayer Co., 392 U.S. 409, 438-40 (1968) ("Congress has the power under the Thirteenth Amendment rationally to determine what are the badges and the incidents of slavery, and the authority to translate that determination into effective legislation.") (internal citations and quotations omitted)

[58] *See* Forbes v. Facebook, Inc., 2016 U.S. Dist. LEXIS 19857, 2016 WL 676396, at *2 (E.D.N.Y. Feb. 18, 2016) (finding that Facebook is not a state actor); *see also* Young

it. The Court has imposed constitutional restraints on private actors in the past, and there is reason to believe it could do so again, potentially this term in *Halleck v. Manhattan Community Access*.[59] Should the Court decide that regulating speech in a "public forum" necessarily qualifies as state action under the "public function" test, it would pave the way for transforming Facebook into a state actor with respect to government pages. After all, Facebook's current program of content moderation—from minimizing posts, to deleting accounts, to evaluating ads—certainly qualifies as speech regulation. While free speech advocates are right to demand that Facebook enhance transparency and accountability, treating Facebook as a state actor would exacerbate the fake news problem.

A. State Action Doctrine Today

The Court currently employs a two-part framework to determine whether the Constitution applies against a private actor—often referred to as the *Lugar-Edmondson* framework.[60] First, litigants must prove "the claimed deprivation resulted from the exercise of a right or privilege having its source in state authority."[61] Second, litigants must show that the private parties whose actions caused the deprivation "may be appropriately characterized as 'state actors.' "[62] Where deprivation is caused by a nominally private entity, the Court employs different tests: the "public function test" (when the state delegates a so-called "public function") is the most relevant when considering how state action doctrine might be expanded to include Facebook.[63]

---

v. Facebook, Inc., 2010 U.S. Dist. LEXIS 116530, 2010 WL 4269304, at *3 (N.D. Cal. Oct. 25, 2010) (holding that Facebook is not a state actor); Shulman v. Facebook, Civil Action No. 17-764 (JMV), 2017 U.S. Dist. LEXIS 183110, at *9 (D.N.J. Nov. 6, 2017) (holding that Facebook is not a state actor).

[59] Halleck v. Manhattan Cmty. Access Corp., 882 F.3d 300 (2d Cir.), cert granted, 139 S. Ct. 360 (2018).

[60] *See* Lugar v. Edmondson Oil Co., 457 U.S. 922 (1982).

[61] *Id.* at 939.

[62] *Id.*

[63] The other two tests include the "compulsion test" (when a private entity is effectively controlled by the state); the "joint action test" or "close nexus test" (when the private entity participates in a joint activity with the state). *See, e.g.*, Brentwood Acad. v. Tenn. Secondary Sch. Athletic Ass'n, 531 U.S. 288, 296 (2001).

At present, the public function test is narrowly conceived. In *Jackson v. Metropolitan Edison Co.*, the Court defined a "public function" as the exercise "by a private entity of powers traditionally *exclusively* reserved to the state."[64] These "traditionally *exclusively* reserved" functions include running elections;[65] maintaining a municipal park;[66] and—in the days of company towns—operating an entire town.[67] Despite the fact that many private entities engage in activities serving the public, the Court has avoided classifying more under the "public function" umbrella, for fear of impinging on private actors' freedom.[68] Consequently, the Court engages in awkward analysis trying to distinguish actions that are the *exclusive* prerogative of governments from those that governments *often* perform.[69] As the doctrine stands, Facebook fails the public function test because what Facebook does—from moderating content to "storing, caching, or providing access to content"—is not a traditional function exclusively reserved to the government.[70] Nonetheless, it's possible that the Court might adopt a more expansive notion of "public function" so as to include some Facebook's activities: namely, its moderation of online content, including fake news or hate speech.

In October, the Supreme Court granted certiorari in *Halleck v. Manhattan Community Access*, a Second Circuit case that applied a new state action test for determining whether a public access television channel qualified as a state actor for purposes of 42 U.S.C. § 1983.[71] The Second Circuit relied on Justice Kennedy's concurrence in *Denver Area Educational Telecommunications Consortium, Inc. v. FCC*, which said that when public access channels are required by law, those channels qualify as

---

[64] Jackson v. Metro. Edison Co., 419 U.S. 345, 352-53 (1974) (emphasis added).

[65] *See* Nixon v. Condon, 286 U.S. 73, 85 (1932) ("Whatever power of exclusion has been exercised by the members of the committee has come to them, therefore, not as the delegates of the party, but as the delegates of the State.").

[66] *See* Evans v. Newton, 86 S. Ct. 486, 487 (1966).

[67] *See* Marsh v. Alabama, 326 U.S. 501, 507 (1946) ("We do not think it makes any significant constitutional difference as to the relationship between the rights of the owner and those of the public that here the State, instead of permitting the corporation to operate a highway, permitted it to use its property as a town . . . .").

[68] *See* Martha Minnow, *Alternatives to the State Action Doctrine in the Era of Privatization*, 52 Harv. C.R.-C.L. L. Rev. 145, 150 (2017).

[69] *See* Am. Mfrs. Mut. Ins. Co. v. Sullivan, 526 U.S. 40 (1999).

[70] *See* Jonathan Peters, *The Sovereigns of Cyberspace*, 32 Berkeley Tech. L.J., 889 (2018).

[71] *See* Halleck v. Manhattan Cmty. Access Corp., 882 F.3d 300 (2d Cir.), *cert granted*, 139 S. Ct. 360 (2018).

per se "designated public forums."[72] From this, the court reasoned that, because federal law authorized setting aside channels for public access and a municipality contracted with a non-profit organization to then run those channels, those channels qualified as "public forums."[73] To determine whether the First Amendment would apply against the non-profit tasked to run those channels, the Second Circuit evaluated the connection between the municipality and the non-profit—concluding that because the president of the municipality designated the Manhattan Neighborhood to run the channels, the non-profit had a sufficient connection to the state to qualify as a "state actor."[74] In a concurring opinion, Judge Lohier endorsed the majority's hybrid state action test.[75] Furthermore, he argued that on the basis of *Lee v. Katz*,[76] a Ninth Circuit decision which held regulating speech in a public forum qualified as a "public function" traditionally within the exclusive domain of the State, the non-profit qualified as a "state actor" under the traditional public function analysis.[77] On appeal, the Court will likely address whether its three state actor tests are exhaustive or, in keeping with the Second Circuit's opinion, whether an alternative state-actor test might be adopted. Either way, the Court will likely confront whether the regulation of speech in a "public forum" necessarily qualifies as a "public function" for purposes of the state actor test, as Judge Lohier argued in his concurrence. If the Court says that regulating speech in a public forum is indeed a public function that qualifies a private entity as a state actor, the outcome would have immediate implications for deciding whether Facebook qualifies as a state actor. After all, Facebook's moderation of content—from diminishing posts to deleting accounts—would certainly count as "regulation." The follow-on inquiry then becomes: which Facebook pages and groups necessarily qualify as "public forums" for purposes of the state action test, if any qualify at all.

---

[72] *See* Denver Area Educ. Telecomms. Consortium v. FCC, 518 U.S. 727, 732 (1996).

[73] *See Halleck*, 882 F.3d at 307.

[74] *See id.*

[75] *See id.* at 308.

[76] Lee v. Katz, 276 F.3d 550 (9th Cir. 2002).

[77] *See Halleck*, 882 F.3d at 308.

B. When Does Facebook Become a "Public Forum?"

At present, the Court broadly defines a public forum as a designated place for the "free exchange of ideas" where speakers "cannot be excluded without a *compelling* governmental interest" that is "narrowly drawn to achieve that interest."[78] When determining whether a space is a public forum, courts examine the "nature of the property" and "its compatibility with expressive activity."[79] As a general rule, public forums involve government property and government control.[80] However, in some instances the Court has treated privately owned property controlled by the government as a public forum.[81]

Unlike the public access channel at issue in *Halleck v. Manhattan Community Access*, the government never delegated to Facebook the job of creating and maintaining its online platform. Instead, Facebook's platform and governing rules fall entirely within Facebook's proprietary domain.[82] Thus, some argue that Facebook is not a public forum, suggesting that any effort to moderate false content on Facebook would *not* qualify as state action under Judge's Lohier's application of *Lee v. Katz*.[83] Nonetheless, the Court's 2017 decision in *Packingham v. North Carolina* suggests a more expansive public forum doctrine that could encompass certain pages on Facebook.[84] Striking down a North Carolina statute prohibiting registered sex offenders from accessing social networking websites, the Court in *Packingham* likened social media

[78] Cornelius v. NAACP Legal Def. & Educ. Fund, 473 U.S. 788, 800 (1985) (emphasis added).

[79] In Widmar v. Vincent, the Court found the meeting facilities at a state university qualified as public forum because the university had an express policy of making those spaces available to registered student. 454 U.S. 263 (1981). Similarly, in *Madison Joint School District v. Wisconsin Employment Relations Comm'n*, the Court held that a state statute providing for open school board meetings created a public forum for citizen involvement. 429 U.S. 167 (1976).

[80] *Cornelius*, 473 U.S. at 802-03 ("The government does not create a public forum by inaction or by permitting limited discourse." (internal citations omitted)).

[81] *See* Southeastern Promotions, Ltd. v. Conrad, 420 U.S. 546, 555 (1975) (holding that a municipal auditorium and city-leased theatre to be a public forum because it was "designed for and dedicated to expressive activities.")

[82] *See* Thomas Wheatley, Why Social Media is Not a Public Forum, Washington Post (Aug. 4, 2018), https://www.washingtonpost.com/blogs/all-opinions-are-local/wp/2017/08/04/why-social-media-is-not-a-public-forum/?utm_term=.fc3b983c8dc9.

[83] *See* Lee v. Katz, 276 F.3d 550 (9th Cir. 2002).

[84] *See* Packingham v. North Carolina, 137 S. Ct. 1730 (2017).

sites to parks and streets calling social media sites "the modern public square."[85] By "analogizing to public space," some argue, the Court "suggested that the public forum doctrine . . . might extend to all or parts of the internet and social media," despite the dual private-public nature of the site.[86] So far, litigants have not succeeded on this very broad application of the public forum doctrine. Nonetheless, they have succeeded on a narrow reading whereby Facebook pages qualify as "public forums" where the government exercises control over a page or site.[87] For example, in 2017, Twitter users blocked by President Trump filed a lawsuit arguing that because his Twitter Page constituted a "public forum," denying them access to it violated their First Amendment rights; the District Court judge agreed.[88] Similarly, in *Davison v. Loudoun County Board of Supervisors*, the Court held a local official violated the First Amendment by banning an individual from her Facebook page and issued declaratory judgement that the official's "social media page operated as a forum for speech."[89]

Though narrowly tailored to government-created pages, the logic of the aforementioned decisions could remake the state action doctrine. After all, the ability to control content in government-created pages—that is, the ability to remove content or minimize posts on a page—belongs to both the government and Facebook. Should the Court decide that regulating speech in a public forum is necessarily a "public function"—as Judge Lohier argues—it would open the door for litigants to bring First Amendment actions against Facebook, rather than just against the

[85] *Id.* at 1731.

[86] *Packingham v. North Carolina (Note)*, 131 HARV. L. REV. 233 (2017).

[87] *See* Amanda Shanor, *The President's Twitter Account & the First Amendment*, TAKE CARE (Jun. 12, 2017), https://takecareblog.com/blog/the-president-s-twitter-account-and-the-first-amendment. ("[I]t is not enough to say as a categorical matter, 'the First Amendment doesn't apply to private companies like Twitter.' Constitutional principles may apply to spaces or channels of communication that the government controls or uses for official purposes, even if they are owned by a private entity.").

[88] *See* Knight First Amendment Inst. at Columbia Univ. v. Trump, 302 F. Supp. 3d 541, 549 (S.D.N.Y. 2018) (holding that "portions of the @realDonaldTrump account—the 'interactive space' where Twitter users may directly engage with the content of the President's tweets—are properly analyzed under the 'public forum' doctrines set forth by the Supreme Court, that such space is a designated public forum, and that the blocking of the plaintiffs based on their political speech constitutes viewpoint discrimination that violates the First Amendment").

[89] *See* Davison v. Loudoun Cty. Bd. of Supervisors, 267 F. Supp. 3d 702, 706 (E.D. Va. 2017).

government, anytime their content is minimized or removed from a government-created page. This would be a bad result.

Applying the state action doctrine against Facebook for purposes of government-created pages would chill Facebook's incentive to moderate false content on those pages. Considering the government-created pages that currently exist on Facebook, this result could be disastrous. Millions follow the Federal Bureau of Investigation (FBI) on Facebook. Were a false comment to appear on the FBI website showing a made-up video of an FBI-sponsored "attack," it could undermine public trust in the FBI. Neither Facebook nor the government would be able to do anything about it. Similarly, local, state, and national officials make personal accounts to raise awareness about their campaigns and government initiatives. Were Facebook treated as a state actor with respect to government pages, it's possible that Facebook would not moderate any content affecting those pages. Treating Facebook as a state actor would result in a clutter of false, misleading content on important sites. Further, one might argue that bad actors, seeking to thwart any moderation by Facebook, would increasingly try to misrepresent their pages as somehow "government-related."[90] Legal commentators caution against applying the First Amendment to Facebook because it would curtail the ability of the sites to moderate harmful content, creating an internet that "nobody wants."[91] While Facebook users are certainly right to worry about censorship by Facebook, processes can be created to maximize accountability and transparency without resort to the state action doctrine. The 2016 election showed the consequences when Facebook doesn't believe it has a social obligation to act. Hopefully, the Court will avoid that signal. ■

[90] Considering how easy it was for VICE journalists to buy political advertisements posing as US Senators, it's likely these bad actors would succeed. *See* Turton, *supra* note 53.

[91] *See* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 Harv. L. Rev. 1598, 1659 (arguing that it's highly unlikely the Court would consider Facebook to be a state actor for purposes of the first amendment, seeing as it would require a "very expansive interpretation" of *Alabama v. Marsh*).