ASPEN TECH POLICY HUB

Draft Regulations for the Implementation of the Honest Ads Act and the Filter Bubble Transparency Act

Matthew Volk
Cecilia Donnelly Krum
Matt Sievers

Summary

Online advertisers exploit outdated political advertising laws to bypass the Federal Election Commission's (FEC's) existing disclaimer and reporting rules. This makes it significantly easier for political propagandists to spread disinformation in online ads without accountability.

Online political ads already have a massive reach, and that reach is growing exponentially. In the summer of 2018, online political ads generated tens of billions of views, and total online political ad spending for 2020 is expected to be over three times what it was in 2016.

Two pieces of legislation recently introduced in Congress – the Honest Ads Act and the Filter Bubble Transparency Act – would help correct the deficiencies in the existing political advertising system, but they leave certain definitions unclear. In this document, we review each piece of legislation and propose regulations that the executive branch could implement were the legislation to be passed. Sample regulatory text can be found in the Appendices.

Honest Ads Act

Senator Amy Klobuchar introduced the Honest Ads Act (HAA) on behalf of a bipartisan group of Senators in 2019. The stated intention is "to enhance the integrity of American democracy and national security by improving disclosure requirements for online political advertisements." The legislation attempts to do so in three ways. First, it changes the rules for political advertising in traditional media to include online and digital advertising. Second, it requires that online platforms publish a record of all requests to purchase political ads. Finally, it requires advertisers to ensure they do not sell ads that are paid for by foreign nationals.

However, **the bill as written leaves significant room for willful misinterpretation and does not cover several misleading online advertising practices.** The Federal Election Commission (FEC) will need to adopt new regulations and update existing ones in order to implement the provisions of this bill. We make three key recommendations that the FEC can implement to ensure maximum applicability of this bill:

1. Clarify that advertisers must report total ad impressions, not just paid views;
2. Require all published variants of online advertisements to appear in the FEC record; and
3. Clarify that paid direct messages sent via platforms count as advertisements.

CLARIFY VIEW COUNTS FOR ONLINE POLITICAL AD REPORTING

First, because online advertisements can be shared, and so often reach many more screens than initially paid for, the FEC should ensure that every view of a political advertisement counts toward the total reported number, and not just the paid subset.

Sec 8(a) of the HAA amends existing disclosure rules for political ads by creating a new requirement for online platforms:

> An online platform shall maintain, and make available for online public inspection in machine readable format, a complete record of any request to purchase on such online platform a qualified political advertisement which is made by a person whose aggregate requests to purchase qualified political advertisements on such online platform during the calendar year exceeds $500.

Sec 8(a) also describes the contents of these records:

> (2) CONTENTS OF RECORD.—A record maintained under paragraph (1)(A) shall contain—
>
> (A) a digital copy of the qualified political advertisement;
>
> (B) a description of the audience targeted by the advertisement, **the number of views generated from the advertisement**, and the date and time that the advertisement is first displayed and last displayed;

There are several methods of counting ad viewership, some of which are better than others.

Facebook defines several metrics for measuring the reach of advertisements, two of which are used as standards in the industry:

- **Post reach** is the total number of unique people who saw an ad at least once.
- **Post impressions** are the number of times an ad was seen.

Further, Facebook breaks down "reach" into three subcategories:

- **Paid reach** is the number of people who had the post placed on their screens for pay.
- **Organic reach** is the number of people who had the post placed on their screens without influence from advertising money.
- **Viral reach**, a type of organic reach, is the number of users who had the post placed on their screens without influence from advertising money because a friend or group shared or liked the post.

Because users engage with and share advertisements originally delivered via paid reach, buying advertisements on social networks has a multiplier effect that causes further organic and viral reach. If paid reach results in influential political pages or personas interacting with the advertisement, others may see the result organically. The corresponding viral reach can dramatically increase the total reach.

As an example, the Russian-backed Internet Research Agency (IRA) infamously used online political advertisements to influence the 2016 U.S. election. Facebook identified that the IRA had purchased $100,000 worth of ads, which are estimated to have reached 126 million people. However, based on the average cost per advertisement in the U.S. at the time (0.849 cents per view), $100,000 would buy a paid reach of only approximately 11.8 million. Paid reach misrepresented the total reach substantially, because the paid ads were used primarily to establish audiences that were then reached

organically. Further, many individuals likely saw the same content more than once, and these duplicate views are reflected only by impressions, not reach.

FEC regulations should make explicit that "the number of views generated from the advertisement" represents the total post *impressions* for each advertisement, not the paid reach.

Proposed regulatory language follows in Appendix A.

REQUIRE REPORTING ALL VARIANTS OF POLITICAL ADS IN THE FEC RECORD

Second, because online advertisements often take advantage of testing frameworks that allow them to run thousands of variants of the same advertisements, the FEC should make explicit that all variants of an advertisement must be included in the submitted FEC record.

As proposed in Sec 8(a) of the HAA, the required contents of a digital record include the following:

> (2) CONTENTS OF RECORD.—A record maintained under paragraph (1)(A) shall contain—
>
> (A) **a digital copy of the qualified political advertisement**;
>
> (B) **a description of the audience targeted by the advertisement**, the number of views generated from the advertisement, and the date and time that the advertisement is first displayed and last displayed;

This language is insufficient, as many variants of the same ad might be targeted to different audiences under the same advertising campaign. For example, "dark ads" are online advertisements shown to some target demographics but kept invisible to others. An advertiser can create a dark ad online and show hundreds or thousands of variants to different target demographics without ever publishing the ad itself. In practice, this means that a political campaign might pay an ad network to run 1,000 versions of an ad, each with slightly different wording in order to present a politicized issue from multiple perspectives. This allows an advertiser to appeal to many political ideologies at once without transparency.

Figure 1 (below) provides a mock example of what this might look like in practice. From top to bottom, the three variants of this ad are about health care, student loans, and retirement, respectively. Each would be micro-targeted to a different portion of the electorate and remain unseen by the rest of the population. If used more maliciously, the ad could tell certain demographics to vote while discouraging others from doing so.

In practice, there are often thousands of variants of a single ad. The former director of advertising at the Republican National Committee (RNC) has stated publicly that their organization routinely ran 40,000 to 50,000 variants of ads, testing to see which worked best among different constituencies. On the day of the third presidential debate in October 2016, this team ran 175,000 variations of their ads on Facebook, without disclosing what they were testing or which constituencies saw which messages.

The Honest Ads Act provides the legal grounds to regulate political dark ads and force them to be published. The regulation should make clear that all micro-variants of the same ad should, in line with Sec. 8(a) of the bill, be published along with each variant's target demographic. This will force advertisers to disclose all variants of dark ads.

Proposed regulatory language follows in Appendix B.

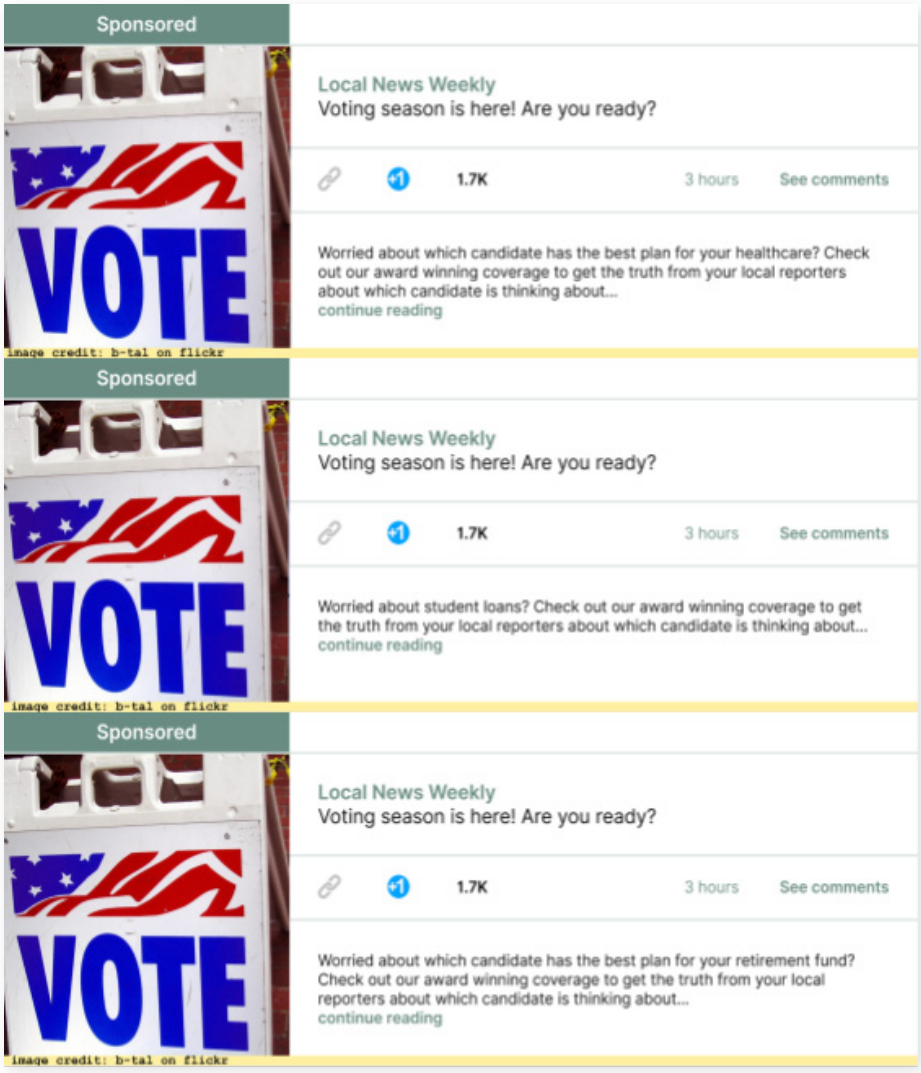

Figure 1: Mocked Dark Ad Example

CLARIFY THAT PAID DIRECT MESSAGES ARE ADVERTISEMENTS

Third, the FEC should make clear that paid direct messaging campaigns are subject to the same record maintenance rules as all other qualified political advertisements. In paid direct messaging campaigns, advertisers purchase the ability to message users of messaging platforms directly, similar to unsolicited emailing. This includes both text messages sent to phones via online texting platforms and direct messages sent via online platforms like WhatsApp and Snapchat.

The reach of online direct messages and text messages is already on par with that of email, and such messages are dramatically more effective at garnering a response. For example, the Democratic National Committee was able to buy 94 million registered cell phone numbers of voters from key demographics, and campaigns and organizations sent 350 million direct messages in 2018 (a six-fold increase from the previous year). Further, it is easy to send text and direct messages at little to no cost: Eleni Kounalakis's campaign for California Lieutenant Governor used the online mass messaging platform Hustle to target 10,000 text messages to specific voters across California, all in under an hour with a dozen volunteers. Unfortunately, this technology has already been used as a tactic to dissuade key demographics from voting. A senior Trump campaign official told *Bloomberg Businessweek* in 2016 that "We have three major voter suppression operations underway . . . . They're aimed at three groups Clinton needs to win overwhelmingly: idealistic white liberals, young women, and African Americans." The last of these "operations" targeted black voters in swing states just before Election Day, with the text "Hillary Thinks African-Americans Are Super Predators."

Because advertisers can purchase the ability to message users of messaging platforms directly, paid direct messages should constitute a form of "paid digital communication" under Sec. 5(a) of the HAA. As a result, these messages should also qualify as online advertisements. Concretely, the regulations should clarify that, under Sec. 8(a) of the bill, if aggregate direct-messaging fees exceed $500 in a calendar year for a political advertiser, this purchasing pattern must be disclosed, along with targeted demographics and template messages.

Proposed regulatory language follows in Appendix C.

Filter Bubble Transparency Act

The Filter Bubble Transparency Act (FBTA) is a bill introduced by Senator John Thune on behalf of a bipartisan group of Senators in the 116th Congress. The stated intention is "to require that internet platforms give users the option to engage with a platform without being manipulated by algorithms driven by user-specific data." The bill first introduces several definitions, including *covered internet platforms*, *user-specific data*, *input-transparent algorithms*, and *opaque algorithms*. It then requires all covered internet platforms to provide users with an option to use an *input-transparent algorithm*, one that only uses data explicitly provided to the platform to return results.

The social media platform Twitter makes clear the sort of viewing option the FBTA would require. As Senator Thune described in a press release, "consumers [of Twitter] have the option of viewing the timeline that Twitter has curated for them – which pushes the posts Twitter thinks they want to see to the top of their feed – or viewing an unfiltered timeline that features all posts in chronological order." Further, there is a prominently displayed icon that allows users to switch instantaneously between the two timelines with a single click. If the FBTA is adopted, other platforms would be required to provide a similarly prominent icon that, when clicked, removes customization based upon data not explicitly provided by the user for the given interaction.

The FBTA, if adopted, will require the Federal Trade Commission (FTC) to adopt new regulations in order to enforce its provisions on large online platforms. We make three key recommendations that the FTC can implement to ensure maximum applicability of this bill:

1. Clarify that advertisements are subject to *input transparency*;
2. Explicitly define *data inference*; and
3. Clarify *user-specific data* by providing explicit covered data categories.

CLARIFY THAT ADVERTISEMENTS ARE SUBJECT TO INPUT TRANSPARENCY

First, the regulation should make explicit that the FBTA's input transparency requirement applies to online advertisements. The proposed legislation does not make this connection explicit, leaving it vulnerable to willful misinterpretation. As referenced above, the stated intention of the bill is to

prevent manipulation of users based on their user-specific data, and online advertisements are an increasingly powerful tool whose explicit purpose is to manipulate user behavior.

The market for online advertisements is exploding. In fact, digital ads are so effective that, in 2019, U.S. advertisers spent over $129 billion on digital advertising, surpassing traditional ad sales ($109 billion) for the first time. Further, eMarketer predicts that digital ad spending is growing so much faster that it will exceed two-thirds of total media spending by 2023. Such advertisements can create real-world harm, particularly when it comes to influencing elections. Joan Donovan, who researches online manipulation at Data & Society, summarized the situation well:

> If we were looking for a digital revolution, it happened in advertising online. . . . Political strategists understood this new opportunity and capitalized on it by serving up digital disinformation using ads as the delivery system. No politician can campaign ethically under these conditions because they are just out gunned by those who are willing to use these systems to do damage.

Limiting the nonconsensual manipulation of users, as the FBTA would require, would dramatically curtail the power of such systems.

However, the proposed legislation does not currently make explicit that its rules on input transparency would apply to advertisements. We argue that advertisements already meet its given definitions and that regulation should make this inclusion explicit.

Section 3(b)(1) of the FBTA requires that platforms provide an input-transparent version of the platform to users:

> (1) IN GENERAL.—The requirements of this subsection with respect to a person that operates a covered internet platform that uses an opaque algorithm are the following:
>
> (A) The person provides notice to users of the platform that the platform uses an opaque algorithm that makes inferences based on user-specific data **to select the content the user sees**. . . .
>
> (B) The person **makes available a version of the platform that uses an input-transparent algorithm** and enables users to easily switch between the version of the platform that uses an opaque algorithm and the erosion of the platform that uses the input-transparent algorithm by selecting a prominently placed icon. . . .

The FBTA defines *input-transparent algorithms* in section 2(5) as the following:

> (A) IN GENERAL.—The term "input-transparent algorithm" means an algorithm ranking system **that does not use the user-specific data of a user to determine the order or manner that information is furnished to such user** on a covered internet platform, **unless the user-specific data is expressly provided to the platform by the user for such purpose.**

Advertisements are, by definition, information furnished to users, particularly when they appear as part of a feed or search results. FTC regulations should explicitly clarify that advertisements must rely on input-transparent algorithms as well.

Proposed regulatory language follows in Appendix D.

EXPLICITLY DEFINE DATA INFERENCE

Second, because data inferences often provide the mathematical basis for opaque algorithms, the FTC should explicitly define this term in regulation to avoid willful misinterpretation.

Data inferences can allow platforms to access data explicitly not provided by users and use this information in content serving algorithms. This has resulted in significant real-world harms. For example, in 2015 and 2016, web sites commonly used inferences based on device battery life to de-anonymize users. Even if users obfuscated their identities with software like virtual private networks (VPNs), websites could leverage battery life information to provide an almost-unique identifier for each device, de-anonymizing browsing activity and allowing tracking. Though the corresponding API has been removed from most web browsers, similar exploits are commonly used. In a similar vein, Facebook was infamously sued by the Department of Housing and Urban Development for allegedly violating the Fair Housing Act because its advertising infrastructure allowed both humans and algorithms to infer race based on other categories.

Though the FBTA does acknowledge that data inferences should not be used in input-transparent algorithms, it does not define the term. As stated previously, Section 2(5) of the FBTA requires that *input-transparent algorithms* recommend content using only "user-specific data that is expressly provided to the platform by the user for such purpose." It then proceeds to define this data:

> (C) DATA PROVIDED FOR EXPRESS PURPOSE OF INTERACTION WITH PLATFORM.—For purposes of subparagraph (A), user-specific data that is provided by a user for the express purpose of determining the order or manner that information is furnished to a user on a covered internet platform—
>
> (i) shall include user-supplied search terms, filters, speech patterns (if selecting the language in which the user interacts with the platform), saved preferences, and the user's current geographical location;
>
> …
>
> (iv) **shall not include inferences about the user or the user's connected device, without regard to whether such inferences are based on data described in clause (i).**

The FTC should define in regulation that a *data inference* is a deduction about a user based on their data that: (1) links the user with additional unprovided data or (2) defines a feature or category that they would otherwise not fall into. Further, it should be made clear that (3) these inferences can be made by algorithms or by human reviewers. Disallowing inferences that link input data to existing banks of data would prevent platforms from using more data than the user provided for a given interaction. Including new features or categories as an impermissible inference would make it harder for this information to be surfaced to advertisers or ranking algorithms. Lastly, clarifying that both humans and algorithms can make inferences in systematic ways would prevent platforms from *enabling* inferences, even if they do not explicitly make them.

Proposed regulatory language follows in Appendix E.

PROVIDE EXPLICIT CATEGORIES OF COVERED USER-SPECIFIC DATA

Finally, the FTC should provide an explicit, non-exhaustive list of covered categories of user-specific data to ensure maximum compliance with the intention of the bill and to make it easier for platforms to comply with the law.

Section 2(8) of the FBTA defines user-specific data as follows:

> (8) USER-SPECIFIC DATA.—The term "user-specific data" means information relating to an individual or a specific connected device that would not necessarily be true of every individual or device.

This definition on its own is ambiguous. In defining such data in regulations, providing a non-exhaustive list of explicit categories of data that count as user-specific data will help eliminate confusion. We propose five classes of data as a starting point:

- Identifiers, like real name and IP addresses
- Biometric information
- Historical usage information, like browser history
- Protected demographic information, like race or ethnicity
- Inferences based on any other user-specific data

Making explicit that these five categories of data count as user-specific data will reduce the ambiguity around what it means to have information that "would not necessarily be true of every individual or device." It also helps create regulatory certainty for online platforms, removing a market incentive that favors being overly restrictive about what user-specific data is.

Proposed regulatory language follows in Appendix F.

Appendices

APPENDIX A: POST IMPRESSIONS AS A BASIS FOR ADVERTISEMENT REACH

In addition to the definitions set forth in 52 U.S.C. § 30104(j), for purposes of these regulations:

(a) The number of views generated from an advertisement is defined as the number of impressions the ad received. This means the total number of times the advertisement appeared on a person's device in part or in whole. Multiple appearances on the same device count separately toward this number.

APPENDIX B: DARK ADS

In addition to the requirements set forth in 52 U.S.C. § 30104(j), for purposes of these regulations:

(a) In order to provide a complete record of any request to purchase qualified political advertisements pursuant to subsection (j)(1)(A), the record must include all elements described in subsection (j)(2) for every published variant of the advertisement, without regard for their number.

APPENDIX C: PAID DIRECT MESSAGES

For context, the bill would create 52 U.S.C. § 30104(j), which would state:

> (3) ONLINE PLATFORM.—For purposes of this subsection, the term 'online platform' means any public-facing website, web application, or digital application (including a social network, ad network, or search engine) which—
>
> (A) sells qualified political advertisements; and
>
> (B) has 50,000,000 or more unique monthly United States visitors or users for a majority of months during the preceding 12 months.
>
> (4) QUALIFIED POLITICAL ADVERTISEMENT.—
>
> (A) IN GENERAL.—For purposes of this subsection, the term 'qualified political advertisement' means any advertisement (including search engine marketing, display advertisements, video advertisements, native advertisements, and sponsorships) that—
>
> (i) is made by or on behalf of a candidate; or
>
> (ii) communicates a message relating to any political matter of national importance, including—
>
> (I) a candidate;
> (II) any election to Federal office; or
> (III) a national legislative issue of public importance.

The regulations should provide the following definition:

Definitions.

In addition to the definitions set forth in 52 U.S.C. § 30104(j), for purposes of these regulations:

(a) Direct messages sent via an online platform are qualified political advertisements if they meet the criteria defined in subsection (j)(4)(A).

APPENDIX D: ADVERTISEMENTS ARE SUBJECT TO INPUT TRANSPARENCY

For context, section 2(1) of the FBTA defines an *algorithmic ranking system* as follows:

> (1) ALGORITHMIC RANKING SYSTEM.—The term "algorithmic ranking system" means a computational process, including one derived from algorithmic decision making, machine learning, statistical analysis, or other data processing or artificial intelligence techniques, **used to determine the order or manner that a set of information is provided to a user on a covered internet platform,** including the ranking of search results, the provision of content recommendations, the display of social media posts, or any other method of automated content selection.

The proposed regulatory language below clarifies "set of information provided to a user on a covered platform."

Definitions.

In addition to the definitions set forth in Section 2 of the Filter Bubble Transparency Act, for purposes of these regulations:

> (a) SET OF INFORMATION PROVIDED TO A USER ON A COVERED PLATFORM.—The term "set of information provided to a user on a covered platform" means all data sent to the user by the platform and viewable on said platform. This includes search results, content and recommendations, social media posts, and advertisements.

APPENDIX E: EXPLICITLY DEFINE DATA INFERENCE

Definitions.

In addition to the definitions set forth in Section 2 of the Filter Bubble Transparency Act, for purposes of these regulations:

(2) INFERENCES ABOUT A USER OR A USER'S CONNECTED DEVICE.—The term "inferences about a user or a user's connected device" means any deduction, made by a covered algorithmic ranking system or by any person using the ranking or advertising system, that—

(i) links the user or user's connected device to existing data not explicitly provided by the user for the interaction; or

(ii) defines a feature of the user or user's connected device that was not explicitly provided by the user for the interaction.

APPENDIX F: EXPLICIT CATEGORIES OF USER-SPECIFIC DATA

For context, section 2(8) of the FBTA defines user-specific data as follows:

> (8) USER-SPECIFIC DATA.—The term "user-specific data" means information relating to an individual or a specific connected device that would not necessarily be true of every individual or device.

The proposed regulatory language below provides a non-exhaustive list of user-specific data, and directly uses some language from the California Consumer Privacy Act (CCPA), Section 1798.140(o).

Definitions.

In addition to the definition set forth in Section 2(8) of the Filter Bubble Transparency Act (FBTA), for purposes of these regulations:

> (3) "User-specific data" includes, but is not limited to, the following if it relates to an individual or a specific connected device but would not necessarily be true of every individual or device:
>
> (i) identifiers such as real name, alias, postal address, unique personal identifier, online identifier, internet protocol address, email address, account name, social security number, driver's license number, passport number, or other similar identifiers;
>
> (ii) biometric information;
>
> (iii) internet or other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding a consumer's interaction with an internet website, application, or advertisement;
>
> (iv) protected demographic information, including, but not limited to, race, gender, sexual orientation, or political affiliation; and
>
> (v) inferences drawn from any user-specific data.