Abortion Data Privacy: Analysis of Third-Party Tracking

Bolor-Erdene Jagdagdorj
Carnegie Mellon University

Lujo Bauer
Carnegie Mellon University

Abstract

Third-party tracking is common across the Internet and is used to target advertisements toward individuals. These practices threaten people's privacy, especially those seeking an abortion. The tracking of abortion-related data is especially important in the aftermath of the Supreme Court's ruling in *Dobbs v Jackson Women Health Organization*, which abolished the federal right to abortion. Many states have severely limited or completely banned abortions. Privacy is an increasing concern for women seeking abortions and individuals assisting in or providing abortions. This research examines third-party tracking across abortion-related websites by analyzing advertisement content after visiting such sites. We found that visiting abortion-related sites increased the incidence of health- and pregnancy-related ads shown afterward, suggesting that people's visits to abortion-related sites are tracked by third parties.

1 Introduction and Background

In 2022, the US Supreme Court decision overturned a constitutional right to abortion, allowing each state to decide abortion regulations [5]. With changing laws and restrictions, privacy becomes an increasing concern for women seeking abortions as well as medical professionals and individuals assisting in or providing abortions [10].

Prior work shows third-party requests occur on 99.1% of abortion-clinic websites [6] and 98.6% of hospital websites [7]. Across hospital websites, the most common third-party cookies were found to provide information to Google

and Facebook [9]. These studies show the transfer of medical and abortion data to third parties, but did not determine which requests were tracking requests. Our study focuses on measuring tracking using advertisement content differences.

2 Approach

In this experiment, we explored if third-party tracking is present on abortion-related websites by measuring the difference in advertisement content. We used an approach similar to Balebako et al., creating several topically different browsing profiles and then collecting ads on subsequent webpage visits [4]. Our profiles simulated abortion-seeking individuals and generic Internet users. We analyzed the content of ads collected afterward to determine whether they were likely to have been shown because of third-party tracking on abortion-related websites that were visited during profile creation. We split our research process into two phases: a training phase to build the user profiles and a testing phase to collect and analyze advertisements. Specifically, we created five user profiles: two controls and three experimental groups.

Training: For each condition, a Python script and the Selenium library were used to automate visiting and interacting with 110 total websites. We used the Google Chrome web browser (version 112.0.5615.137) running on the iOS operating system (version 12.0.1). Two groups were used as a control, visiting the same generic webpages. The three experimental groups visited websites related to abortion pills and either abortion-clinic or crisis-pregnancy-center websites (Table 1).

Typical users were mimicked by visiting the top 110 websites from the Tranco list that were able to be visited using our automated browser [3]. We imitated abortion-seeking users who were likely to either look up abortion pills or abortion clinics. For abortion pills, we selected the top 10 pages not linked to any specific clinic when querying two abortion pills, mifepristone and misoprostol. The National Abortion Founda-

Condition name	Websites visited
Control A	110 generic
Control B	110 generic
Experiment A	10 abortion pill, 100 abortion clinic
Experiment B	10 abortion pill, 25 generic, 100 abortion clinic
Experiment C	10 abortion pill, 100 crisis pregnancy center

Table 1: Overview of website types that user profiles were built upon

tion list of clinics was used to choose clinics with a valid and unique URL [8]. Crisis pregnancy centers (CPC) often mask themselves as abortion clinics so people may inadvertently turn to a CPC rather than an actual abortion clinic. These clinics were compiled from the CPC Map [2] by selecting clinics split between various geographic locations and abortion regulating states.

Testing: In the testing phase, ads were collected and analyzed for each previously created profile. Twenty websites were selected from which to collect and analyze ads. Ten of these sites were general news websites from the AllSlides media bias news chart split between left, right, and central news sources [1]. Ten health-related websites were selected from the SimilarWeb list of top health pages as well as health misinformation pages found on the Statista list of top disinformation sites [11]. Each webpage was manually viewed to select a sublink that had collectable ads (collected via our Ad-Trap script, using NodeJS). The tool visited a given webpage, waited for it to fully load, scrolled fully, and collected advertisements shown. For each ad we aimed to store a screenshot, associated link, and the final redirect URL.

We loaded each page three times before capturing ads. This reload gave us a better chance of viewing personalized ads rather than default campaigns. A script was written to synchronize each testing page visit so the time would be the same between each experimental condition. To determine if the ad content differed between conditions, we categorized ads as directly related to abortion or pregnancy, general healthcare, or in categories related to abortions or pregnancy but not directly on those topics.

Each ad was manually analyzed by viewing the screenshot and associated URLs. The ads in each category were summed to give total counts. The ratios of specific ad categories were compared to see if the number of health ads and sensitive ads that could imply an individual was pregnant and/or seeking an abortion were statistically different (using the Fisher Exact Test) for abortion-seeking individuals versus non-abortion-seeking individuals.

3 Results

The overall distribution of advertisements showed increases in health-related ads for abortion-seeking groups when compared to the controls. We were interested in determining if ad content would reflect users visiting abortion websites and thus be tailored to show advertisements indicating a health condition, pregnancy, or abortion specifically. We categorized these as sensitive pregnancy ads, which we defined as ads that were related to

Figure 1: Number of ads related to sensitive health information, pregnancy, and/or abortion

abortion/pregnancies, health, women, kid/family, or a sensitive product/service/provider. The ratio of pregnancy-related advertisements was significantly higher using the Fisher Exact test for the abortion-seeking groups when compared to the control users (Figure 1).
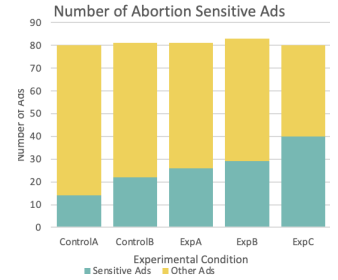
4 Conclusions and Future Directions

Many related topics remain to be investigated. For example, one could analyze the third-party tracking of low-income users and minorities; measure how many abortion-related websites have privacy policies, and how those policies related to organizations' data-sharing policies and observed practices.

As abortion laws are changing, there is an increased urgency for protections for the online data of abortion-seeking individuals. Thus it is imperative for people to be aware of third-party tracking occurring so they can choose what websites or privacy-preserving behaviors to partake in. Previous work found third-party requests are common on hospital and abortion websites. Our study found the ratio of ad content considered sensitive (pregnancy, abortion, health, baby-related, and ads targeting women) was significantly higher for (simulated) abortion-seeking users than control users. These findings indicate third-party tracking occurs across abortion-related and CPC websites. Hence, our work contributes towards understanding privacy risks individuals seeking abortions face. In addition, the policy implications are two-fold: first, for websites that do not fall under HIPAA regulations, new policies should be considered to protect sensitive information from third parties. Second, for websites where users are protected by HIPAA, there should be enforcement to ensure third-party tracking is not violating those regulations.

References

[1] AllSides media bias chart. https://www.allsides.com/media-bias/media-bias-chart. Visited 2023-04-26.

[2] Crisis pregnancy center map & finder. https://crisispregnancycentermap.com/. Visited 2023-04-09.

[3] Tranco: A research-oriented top sites ranking hardened against manipulation. https://tranco-list.eu/. Visited 2023-04-09.

[4] Rebecca Balebako, Pedro G Leon, Richard Shay, Blase Ur, Yang Wang, and Lorrie Faith Cranor. Measuring the effectiveness of privacy tools for limiting behavioral advertising. In *Web 2.0 Security and Privacy Workshop*, 2012.

[5] Center for Reproductive Rights. U.S. Supreme Court takes away the constitutional right to abortion. https://reproductiverights.org/supreme-court-takes-away-right-to-abortion/, June 2022.

[6] Ari B. Friedman, Lujo Bauer, Rachel Gonzales, and Matthew S. McCoy. Prevalence of third-party tracking on abortion clinic web pages. *JAMA Internal Medicine*, September 2022.

[7] Ari B. Friedman, Raina M. Merchant, Amey Maley, Karim Farhat, Kristen Smith, Jackson Felkins, Rachel E. Gonzales, Lujo Bauer, and Matthew S. McCoy. Widespread third-party tracking on hospital websites poses privacy risks for patients and legal liability for hospitals. *Health Affairs (Project Hope)*, 42(4):508–515, April 2023.

[8] National Abortion Federation. Find a provider. https://prochoice.org/patients/find-a-provider/. Visited 2023-04-09.

[9] Joshua D. Niforatos, Alexander R. Zheutlin, and Jeremy B. Sussman. Prevalence of third-party data tracking by US hospital websites. *JAMA Network Open*, 4(9):e2126121, September 2021.

[10] Ian Prasad Philbrick. The end of Roe. *The New York Times*, June 2022.

[11] Statista. Leading health misinformation websites worldwide between may 2019 and may 2020, by estimated views. https://www.statista.com/statistics/266916/top-health-disinformation-websites-visits/. Visited 2023-04-26.