Galvanize:
Improving access to Privacy Preferences and GPDR Requests for Internet Trackers

Sam Boger CS2390, Brown University, 2020

1 Abstract

Privacy controls can be difficult to discover and access. Despite regulations like the GDPR and the CCPA requiring many organizations to implement these controls and provide them to users, major obstacles including discoverability and usability diminish their impact. This project builds some proof-of-concept features that address these obstacles by adding functionality to EFF's Privacy Badger browser extension. These features demonstrate how moving these controls into the user's domain changes the power dynamic between users and trackers. Several limitations like scalability, robustness, and limited privacy controls remain unsolved in this initial version. The report concludes with consolidating the lessons learned into an ambitious vision of an improved online privacy ecosystem.

2 Introduction

Digital advertising is deeply embedded to the open internet as it exists today. Free services and information are frequently funded by selling advertisements. Advertisers are willing to pay more per advertisement to place ads if they can be tailored to a user's demographics or behavior. Therefore, the advertising industry has an incentive to both collect information on users over the long-term to learn their general habits and identities and build user profiles. Then they want to recognize the user just before serving ads to them so the ads can be tailored to the particular user's profile. This practice is known as Interest Based Advertising (IBA).

IBA raises numerous privacy concerns for individuals. One of these is the collecting and building of user profiles. Even though these are often pseudononymous, based on an opaque identifier for each user like their browser cookie or device identifier, these identifiers can plausibly be linked to their personal identity. For example, a sufficiently detailed user profile may in fact be unique to an individual even if their personally identifying information (PII) is not stored. Also, it is plausible to link multiple pseudo identifiers together and then any one such profile that contains PII would reveal the true identity for all such profiles. Other risks relating to these personal profiles include data leakage, sabotage, theft, or legally permissible government access. Therefore users may object to the existence of these profiles built in the service of IBA even before any advertising actually occurs.

Once an ad is actually targeted to a user, a new set of possible objections arise. Advertisements are meant to influence behavior, which a user may object to outright. The website serving the ad benefits from IBA since it makes more money from serving the ad, the advertiser gains effectiveness by serving an ad more likely to give a better return on their investment, but it is more difficult to determine whether the user benefits or is harmed by this arrangement. Oppositions may be based on diminishing personal autonomy, leaking privacy sensitive information by placing these targeted ads in a semi-public display like a computer screen visible to other people and perhaps by other code on the page, or by increasing the possibility for advertisers disseminating misinformation or radicalization.

Requiring websites to collect affirmative consent by users to engage in IBA, or at least having an option for users to opt out of IBA, is a major feature of modern privacy regulations like the GDPR. Despite widespread implementation of options for users to exercise, obstacles persist for users including discovering who has collected data about them and then navigating the process of updating their privacy settings and making data access or deletion requests.

This work improves the situation by creating opportunities for privacy-conscious users who want to exercise their data privacy rights. Existing privacy protection tools like Privacy Badger give users information on who tracks them during their browsing with options to completely block those trackers if desired. To implement these improvements the extension is augmented to include one-click access to engage with tracker's privacy settings and user-centric guidance on using GDPR-related resources.

Several lessons learned while building these proof-of-concept features invite some theorizing on what longer-term

efforts could lead to a substantially improved privacy ecosystem. In particular, this report highlights how shifting privacy controls away from the tracker's direct control would change the power dynamic of individuals in relation to data controllers.

3 Related Work

The primary motivation for this work was the paper on classifying and defining trackers [6]. The authors demonstrate that users prefer explanations about advertising written by third parties rather than by the advertising companies themselves. This work continues that theme by highlighting how privacy controls can be more effective and accessible when implemented on behalf of users rather than by the trackers themselves.

EFF's Privacy Badger extension serves as an invaluable building block for implementing this project since it already detects and classifies trackers on any given page and presents this information to users [4]. The above paper also explains why on-page detection is so important given that trackers form ecosystems that are focused by language and region, so users should prioritize addressing trackers they encounter rather than a global or statically defined "top" tracker list.

Users have access to other privacy tools that relate to trackers and GDPR rights. An ad industry self-regulatory group has developed a bulk opt out tool along with information on digital advertising, focusing on why it is beneficial to users, publishers, and advertisers [5]. Mainstream browsers themselves have added privacy controls including the ability to block all third party cookies or all trackers (as detected and determined by the browser) by choice and even by default in some [2]. Other relevant extensions available for free to users include ad blockers as well as a tool route GDPR requests by email to certain trackers along with templates for drafting the requests [3].

4 Design

4.1 IBA Opt-Out

Thanks to the aforementioned privacy legislation, and perhaps related societal norms and pressures, many trackers provide the option for users to "Opt Out" of IBA. All of these opt outs address the concerns stemming from being served targeted ads as long as the opt out functionality is working as intended and described. The other concern about the building of user profiles at all, however, is only satisfied by some implementations of opting out.

The tracker may implement the opt out by setting the user's cookie to a static non-tracking value like -1 or *null*. In this case, building profiles based on the user's cookie is impossible and both concerns are addressed, with the possible exception of more invasive practices like browser fingerprinting. This is the observed behavior for one of the trackers examined in this project: Scorecard Research.

Other implementations set one of the cookies for the tracker's domain in the browser to be an opt out value but leave the other cookies intact. In this case, it is more difficult to determine if they continue to collect data on users or if they only stop serving targeted ads and continue to collect information. This is the case for Bing as observed in this project where an IBA opt out cookie is set but other cookies persist. Without much more elaborate and careful study it would be difficult to determine empirically what changes in data collection behaviors.

It's also possible that the opt out only affects a server-side setting for a user, where similarly the user cannot easily determine if the tracker has stopped collecting data about their browsing or simply refrains from using it for targeting. This was the implementation used by AdNexus/Xandr as observed during development.

Regardless of which implementation choices the tracker uses, the design of the IBA opt out feature for this project is intended to work similarly. Privacy Badger already displays the domain assosciated with trackers present on a given webpage along with a slider to allow the user to block cookies or all requests from the domain. The additional features added here are for the user to see whether they have already opted out of IBA from this tracker, and if not, opt out with a single click without leaving the current page. The status check should actively measure this opt out status rather than rely on remembering whether the user opted out earlier, to avoid cases where the tracker resets the users' status or their stated preference otherwise expires.

4.2 GDPR forms

To address another obstacle to accessible privacy options, the extension is modified to guide users through filing data access and deletion requests. The design of this feature is to provide contextual information on the mechanics of filling out these forms. This is motivated by manual examination of these forms and observing that the instructions given by data controllers are often opaque, confusing, and focusing on the data controller's concerns rather than those of the user. However, to not be intrusive or annoying, the guidance should appear only when requested and be easily toggled on and off as desired.

To explain how the intention of the instructions is important, consider the GDPR form on the Rubicon Project which includes questions that ask the user to report their country and their mobile device identifiers. Given the context of the question it is likely that Rubicon wants to know the user's country of "citizenship and/or residence" so they know which privacy laws apply to the user, but they do not tell the user that this is the intention behind the question. A user may be better informed on how to fill in this value if they too are

aware of how it will impact their request. Similarly, a user may not know why Rubicon is asking for their mobile device identifiers. By providing possibilities for how this information will affect how their request is processed, say by allowing Rubicon to look up information associated with this identifier, the user can better decide if and which identifiers to provide. Furthermore the instructions given to users to actually find this information may be insufficiently clear or explained at an overly technical level, which could also be mitigated by these annotations.

4.3 Extensibility

A significant challenge when implementing both opt out functionality and GDPR form annotations is the diversity of implementations by different trackers and data controllers. Some options for addressing this challenge are designing the extension to minimize the effort needed to add each new tracker or form, automating the process, or some hybrid of the two.

Full automation brings issues of addressing false negatives and positives and partial automation with human review raises similar challenges. Partial automation could consist of automatically collected suggestions and possibilities for finding tracker's opt out locations or identifying common elements of a form and which annotation best applies. These suggestions could be accepted, rejected, or modified in manual review. Overall, it's a promising direction, but did fit in the scope of this project and was not pursued.

Instead, the code for this extension was designed such that adding support for additional GDPR forms and tracker opt outs is as simple as possible. To scale this up, it would be possible for a community to incrementally build coverage of more trackers. For instance, the configuration for which GDPR request websites to support and their associated elements and comments could be in a separate configuration file separate from the extension's code.

5 Implementation

Opt out functionality was added by adding buttons to the Privacy Badger popup window that displays when the user clicks the extension icon. The additional buttons are placed below the existing slider for each tracker. The first, when clicked, changes color to green if the user is already opted out of IBA from that tracker and to red if the user is currently opted in, or remaining grey if it cannot be measured. The second, when clicked, performs the opt out on behalf of the user. The user can then click again on the status icon to confirm the opt out was successful.

To check if a user is opted out, either the cookies for that domain must be inspected, a button or rendered text on a webpage can be examined, or a specific request can be made to a server to evaluate the response. For many trackers, all of these signals exist. For the sake of exploration, each tracker added to the extension so far has a different implementation among these listed. Performing the opt out are less varied so far, each requiring opening a page and clicking on a UI element programatically.

Both of these steps have an implementation challenge frequently encountered with web technologies–everything is asynchronous. This necessitates waiting on pages to load before clicking elements, for example by using timeouts. For checking opt outs, the code that depends on the answer needs to be asynchronous which is handled by using JavaScript Promises. Examining cookies for cross-domain sources necessitates that the extension has sufficient permissions for those domains. In this demo version, the extension has permissions to modify content on all domains for simplicity. Mechanically, these checks can be done with XHR requests or more simply by opening tabs in that domain and executing scripts in that context, the latter of which is the approach used here.

The GDPR form submission annotations are toggled using the extensions "action button" located in the URL box. Clicking the action button adds CSS and runs JavaScript that adds on-hover tooltips over specified elements with the desired text.

Having never authored or edited browser extensions before, the Firefox extension demos, including "Beastify" and "apply-css" (which are made available under the Mozilla License [1]) were incredibly helpful. I added code directly from these extensions as a template and modified it to achieve the desired functionality. Of course, everything had to be integrated into the existing Privacy Badger extension.

6 Evaluation

6.1 What Works Well

The current implementation demonstrates the potential value of making privacy controls more accessible. The experience of exercising control without leaving your browsing context, and with a single click, feels satisfying. GDPR form guidance also seems beneficial although no user studies for either of these have been conducted.

The extensibility also seems promising. Adding support for new trackers now takes much less time (10-15 minutes) and few lines of non-boilerplate code (10 lines of Javascript). Adding support for a new GDPR form requires only boilerplate code and only enumerating element ids and text strings.

6.2 What Could Be Improved

Every aspect of this project could be improved with more time and effort. The highest priority improvements would be:

- The opt-out status icons should run automatically, and perhaps periodically in the background, rather than requiring user action.

- Opting out and checking opt out status should either not open new tabs or automatically close them after the work is complete. This is nontrivial given the current implementation which requires asynchronous responses from executing JavaScript in those new tabs, but is definitely possible. For transparency, perhaps there should be an option to leave the tabs open so users can browse the tracker's settings page if they wish.

- Improve robustness. Finding elements to click, inspect, and annotate by their DOM ids as well as hard coding cookie values that represent opt out status are all somewhat fragile if trackers change their implementations at all. Similarly, hard coded timeouts when loading pages could be improved.

- Partially automate covering new trackers and GDPR forms. This includes a long list of possibilities like crawling sites for their privacy settings page and identifying the opt out button within that page that needs to be pro grammatically accessed.

7 Future Directions

Beyond concrete improvements to this project, there are other promising future directions for improving privacy control accessibility. This is a very open research and engineering area with many possibilities. I will try to pain a picture of one appealing direction.

A long-term goal could be to develop a user-owned "Privacy Console". This would be a web application that acts as a privacy and data protection hub. Likely this would need cooperation from the browser or evolution of web standards to make this possible. Imagine that every time a cookie is set for a user, the domain is automatically registered into the Privacy Console. Such a registration would require that the data controller has functionality that enables the browser to, on behalf of the user, request information and take certain actions with regards to that domain.

For every domain registering a cookie, the user must be able to configure their privacy settings with that domain. The existing options in Privacy Badger are a good starting point: allowing the user to delete the cookie and block all future cookies from the domain, or only allow limited access to the cookie, say for only first-party contexts. Further options could exist like setting retention policies on the cookie, automatically deleting it after a preset number of requests or time duration.

Furthermore, GDPR-like features should be available for every domain registered in the privacy console. The user should, at a minimum, be able to request all their data from that domain and request the deletion of the same. Additional transparency, like automatically populating the data schemas which contain the user data and possibly some sample data rows so the user can get some concrete intuition on how the service stores and uses their data.

Along with setting cookies, other actions by data controllers could be allowed only if the controller supports the Privacy Console features. This could cover sending marketing emails, marketing texts, and implementing customer loyalty cards. All of these use cases share in common that users in today's world are frequently not aware of, and may not actively consent to, their information being stored by the data controller.

Organization and user experience for this Privacy Console would be important to get right. It should be searchable and sortable by different metrics (frequency of access by data controllers, amount of data stored, long retention periods...). Beyond direct user access, having a uniform hub for privacy control would allow the development of additional privacy tools which can use the console to perform more complex tasks on behalf of users.

In the current world, privacy tools like this project must adapt to the implementation details and design decisions of each individual domain/tracker/data controller. The main benefits of this imagined Privacy Console is that it embeds privacy controls in the user's sphere of influence. Having it cover all trackers and forcing consistency in options provided by data controllers are additional benefits that solve discoverability as well as providing simplicity for additional layers of privacy controls to make use of. Data controllers simply do not have sufficient incentive at the moment to solve the problem properly.

Consistency is also important for auditing and investigation. If researchers and auditors have clear and consistent understanding of how privacy controls are supposed to work, it would help them test that functionality effectively. Accountability on top of regulation is important to guard against pyrrhic victories where privacy policies evolve but functionally nothing changes.

References

[1] Mozilla Corporation. Mozilla public license. https://www.mozilla.org/en-US/MPL/.

[2] Mozilla Corporation. Today's firefox blocks third-party tracking cookies and cryptomining by default. 2019. https://blog.mozilla.org/blog/2019/09/03/todays-firefox-blocks-third-party-tracking-cookies-and

[3] Conscious Digital. Your digital rights. https://yourdigitalrights.org/.

[4] Electronic Frontier Foundation. Privacy badger, 2017. https://github.com/EFForg/privacybadger.

[5] National Advertising Initiative. Nai consumer opt out. https://optout.networkadvertising.org/.

[6] Miranda Wei, Madison Stamos, Sophie Veys, Nathan Reitinger, Justin Goodman, Margot Herman, Dorota Filipczuk, Ben Weinshel, Michelle L Mazurek, and Blase Ur. What twitter knows: Characterizing ad targeting practices, user perceptions, and ad explanations through users' own twitter data. In *29th {USENIX} Security Symposium ({USENIX} Security 20)*, pages 145–162, 2020.