# Engineering Privacy and Protest: a Case Study of AdNauseam

Daniel C. Howe
School of Creative Media
City University, Hong Kong
Email: daniel@rednoise.org

Helen Nissenbaum
Cornell Tech
New York University
Email: hfn1@nyu.edu

like DP

*Abstract*—The strategy of obfuscation has been broadly applied—in search, location tracking, private communication, anonymity—and has thus been recognized as an important element of the privacy engineer's toolbox. However, there remains a need for clearly articulated case studies describing not only the engineering of obfuscation mechanisms but, further, providing a critical appraisal of obfuscation's fit for specific socio-technical applications. This is the aim of our paper, which presents our experiences designing, implementing, and distributing AdNauseam, an open-source browser extension that leverages obfuscation to frustrate tracking by online advertisers.

At its core, AdNauseam works like a list-based blocker, hiding or blocking ads and trackers. However, it provides two additional features. First, it collects the ads it finds in its 'Vault', allowing users to interactively explore the ads they have been served, and providing insight into the algorithmic profiles created by advertising networks. Second, AdNauseam simulates clicks on ads in order to confuse trackers and diminish the value of aggregated tracking data.

A critic might ask: why click? Why not simply hide ads from users and hide users from trackers? The twofold answer reveals what may be distinctive elements of the AdNauseam approach. To begin, we conceptualize privacy as a societal value. Whereas many privacy tools offer solutions only for individual users, AdNauseam is built on the assumption that, often, effective privacy protection must be infused throughout a system. This assumption presents different and interesting engineering challenges. Second, AdNauseam seeks to concurrently achieve the goal of resistance through protest. And since protest frequently involves being vocal, AdNauseam's core design conflicts at times with conceptions of privacy based on secrecy or concealment. While such tensions, and the tradeoffs that result, are not uncommon in privacy engineering, the process of designing and building AdNauseam demanded their systematic consideration.

In this paper we present challenges faced in attempting to apply obfuscation to a new domain, that of online tracking by advertisers. We begin with the goals of the project and the implemented features to which they map. We then present our engineering approach, the set of tensions that arose during implementation, and the ways in which these tensions were addressed. We discuss our initial evaluation efforts on both technical and ethical dimensions, and some of the challenges that remain. We conclude with thoughts on the broader issues facing privacy tools that must operate within complex socio-technical contexts—especially those dominated by actors openly resistant to them—informed by our experience with AdNauseam's ban from Google's Chrome store.

## I. Introduction

The ad blocking wars [34] reflect wide ranging resistance to the online advertising landscape. AdNauseam, an open-source, cross-platform browser extension, contributes to the growing arsenal of

点击所有的广告，污染和混淆数据跟踪

tools addressing this issue, by leveraging obfuscation in an attempt to frustrate profiling based on interests revealed through ad clicks. Specifically AdNauseam enables users to click ads behind the scenes in order to register visits in ad network databases. The aim of the software is to pollute the data gathered by trackers and render their efforts to profile less effective and less profitable. At the same time, the software allows users an avenue of proactive expression, by actively disrupting the economic system that drives the system, and by sowing mistrust (advertisers generally pay ad networks for clicks) within it. Additionally, AdNauseam makes the ads it collects available to users to explore via interactive real-time displays.

### A. Engineering Philosophy

Our approach builds on prior work that explicitly takes social values into account during tool design [14], [12], [26]. In planning, development, and testing phases, we have integrated values-oriented concerns as first-order "constraints" together with more typical metrics such as efficiency, speed, and robustness. Specific instances of values-oriented constraints include *transparency* in interface, function, code, process, and strategy; *personal autonomy*, where users need not rely on third parties; *social privacy* with distributed/community-oriented action; *minimal resource consumption* (cognitive, bandwidth, client and server processing); and *usability* (size, speed, configurability, and ease-of-use). Enumerating values-oriented constraints early in the design process enables us to iteratively revisit and refine them in the context of specific technical decisions. Where relevant in the following sections, we discuss ways in which AdNauseam benefited from this values-oriented approach, as well as tensions between design goals that emerged. We have also followed strategies from privacy-by-design [19], [24], [20], [22], [5], including *Data Minimization*, *Legitimacy Analysis* and *Socially-informed Risk Analysis* as elements of our design process.

### B. Design Goals and Constraints

The AdNauseam extension attempts to realize three tangible goals for the context of online tracking via advertising. The first is to offer *protection*; protection against for users against malware and "malvertising" (malicious software that leverages advertising mechanisms to gain access to users' systems [31]), as well as protection against data aggregation and profiling (either for individual users, the aggregate set of users, or both) via clicks on advertisements. The second goal is to provide a means of proactive engagement, allowing users an avenue for *expression* of their dissatisfaction with current advertising mechanisms to those in control of such systems. In the case of AdNauseam, this expression has an interventional aspect, as the software actively attempts to disrupt the economic model that drives advertising surveillance. The third goal is to facilitate *transparency* regarding the advertising ecosystem—and the profiling

on which it operates—by providing users with the ability to view the ads they are served in real-time, and later, to explore interactive visualizations of the ads collected over time, providing a rare glimpse of how advertisers view them.

### C. Social-Technical Context

AdNauseam applies obfuscation to the context of tracking by online advertisers. If we compare this context to others where obfuscation has been applied (e.g., search), we notice similarities and differences. In both cases users are confronted with large corporate entities (search engines and advertising networks) whose business models depend on the bulk collection of personal data. And in both cases users have little say in shaping these interactions, except to take-it-or leave it. One difference is that although "leaving it" may be feasible in search, using instead a non-tracking alternative such as DuckDuckGo, it is unclear what alternative exists for those wanting to opt-out of advertising surveillance – cease using the Web? [1] A second difference is the degree to which users want the service provided by the trackers. In search we can assume most users do in fact want (or need) the service offered by the search engine, which also happens to be the tracker. Advertising, by contrast, is not as clear. Tracking aside, some users may find ads useful; while others prefer not to see ads at all, while still others might tolerate non-tracking ads in order to avoid subscription fees. A third, structural difference, is that in search there is a single adversary with full knowledge of both search and meta data, including prior queries, click frequency, results clicked, timing data, prior interest profiles, etc. By contrast, the advertising ecosystem is fractured into multiple actors, including ad-hosting web sites, advertisers, trackers, advertising networks, ad servers, analytics companies, etc., each of which interacts with the user in different ways, and is privy to different aspects of those interactions.

### D. Feature Mapping

The mapping of goals to features (and to the system modules described below) was performed as follows: The goal of *protection* was implemented at a basic level by the clicking of collected ads, via the visitation module; and by the blocking of non-visual trackers and other malware, via the detection module. The former attempts to protect the user from data profiling via clicks on advertisements, and the latter from non-visual tracking and potential malware. *Expression* was realized through clicks, again via the visitation module, and also in our implementation of the EFF's Do Not Track (DNT) mechanism [11]. With DNT enabled (the default setting), the DNT header is sent with all requests, and ads on DNT sites remain visible. Ads on these DNT pages are also not visited by AdNauseam. The goal of increased *transparency* is realized through the visualization module, specifically via the real-time menu interface, where users can watch as new ads are discovered, then visited; the vault interface (described below), and a range of explanatory links embedded throughout AdNauseam's settings pages. Additionally, an in-depth Frequently-Asked-Questions (FAQ) list is linked from multiple locations within the interface.

### E. Data Minimization

Following a growing body of literature on privacy-by-design [19], [24], [20], [22], [5], our design and implementation process followed principles of data minimization. Thus AdNauseam was designed to function without ever communicating to a "home server" or sending user-data to any other entity, for any reason. For developers, this meant we were unable to access usage patterns and related data, which may have yielded important insights. Yet this both clarified our own position in regard to data collection, and also enabled us to sidestep potential problems of data leakage or interception in transit.

> *[D]ata minimization does not necessarily imply anonymity, but may also be achieved by means of concealing information related to identifiable individuals* [19].

Additionally we have applied the principle of data minimization to our ad export feature, which allows users to export their aggregate ad collection (as JSON) in order to sync or migrate between machines, to backup, or to share. From our experiences user-testing this feature, we noted that such exports contained potentially sensitive data, specifically users' click trails (stored as the locations for found ads), possibly over months or even years. When considering how to handle this data we noted that it also existed in the browser's local storage, which could potentially be accessed by a malicious actor. Thus we subsequently implemented encryption for this data, both in memory and in storage, as well as offering users the option, before each export, to redact ad locations if desired.

### F. Legitimacy Analysis

> *Before any privacy-by-design activities are embarked upon, a discussion needs to take place with respect to the "legitimacy" of the desired system given its burden on privacy."* [28] [20]

A critic might ask: why click? Why not simply hide ads from users and hide users from trackers? There are two reasons. First, AdNauseam is inspired by the path-breaking work of Priscilla Regan, who argued that beyond the protection of individual interests, privacy may serve social ends, similar to collective goods such as clean air or national defense [38]. This notion of privacy as a collective good presents interesting engineering and evaluation challenges, which, in our view, warrant close attention. Thus AdNauseam may stimulate deliberation not only on its particular features, but may draw attention to the conception of privacy it seeks to promote. A second reason for clicking, as opposed to simply blocking, is that AdNauseam seeks concurrently to achieve the goal of expressive resistance to tracking through protest. And since protest generally involves being vocal, AdNauseam's design seeks to give voice to users. Rather than enacting privacy as concealment, AdNauseam provides a means for users to express, in plain sight, their dissent by disrupting the dominant model of commercial surveillance. This approach embodies a principle drawn from the theory of contextual integrity, namely, privacy as appropriate flow of information [36]. Thus, AdNauseam does not hide deliberate clicks from trackers but rather, by surrounding these clicks with decoy clicks, obfuscates inferences from clicks to users' interests, which may be manipulated in various ways, including via behavioral advertising. AdNauseam does not block clicks; instead it blocks inappropriate access to interest profiles that trackers may infer from them.

Some have argued that simply using a quality ad blocker offers similar degrees of protection and expression. Although basic ad blocking may protect individual users, its scope of impact is limited to those users. There is also a need for tools whose impacts reach beyond those individuals who know they exist and possess the sufficient technical competence and confidence to install them. AdNauseam's aim of polluting aggregate data has the potential to reduce its value to profilers and, more generally, to draw attention to the problematic practices of behavioral advertisers. Although blocking may also realize expressive goals, for example, via industry studies and media

[1] From this perspective, obfuscation may be even more legitimate for advertising than for search, due to the lack of viable alternative options.

reports, the expressed message differs from that of AdNauseam's. Ad blocker use is generally interpreted by the advertising industry as a rejection of problematic *aesthetic* aspects of the ad experience, while AdNauseam's expressive intent specifically targets the industry's unethical surveillance practices [2]. Anecdotal reports from tools users, to which we return briefly below, also suggest qualitative differences of intent in their use of AdNauseam.

Finally, critics have claimed that AdNauseam harms independent content producers who can no longer support their sites. As this critique touches a broad array of tools, including standard ad blockers, it will take us too far afield to address it fully here. However, setting aside the rejoinder which points out that these sites are enabling surveillance, or more harshly, "selling out" their visitors, the hope is that loosening the chokehold of tracking over web and mobile domains will allow other business models to flourish. Toward this end we have enabled support in AdNauseam for the EFF's DNT mechanism, a machine-verifiable, and potentially legally-binding, assertion on the part of sites that commit to privacy-respecting behavior [11]. For sites that make this commitment, AdNauseam does not (by default) hide, block, or click their ads.

### G. Socially-informed Risk Analysis

Given the goals we hoped to achieve and the set of features to which these mapped, we set out to identify risks to which users might be exposed. For each such risk, we considered the degree to which the user would be exposed when browsing the web using an unmodified browser, in comparison to the degree of exposure while using AdNauseam. Finally we considered their exposure using existing alternatives, ad-blockers like AdBlock Plus [1] or wide-spectrum blockers like uBlock [17](see, for example, Figure 3 below). The following risks were identified:

- Increased tracking by advertisers and other data-gatherers
- Personal data leakage (via clicks, hiding or export)
- Harms via malware or "malvertising"

To establish a lower-bound on exposure, we imposed a constraint that exposure with AdNauseam must be strictly lower on all dimensions than with an unmodified browser. Conversely, we hypothesized that the current performance of uBlock, the open-source blocker with the best performance metrics, would provide an upper-bound on exposure. As AdNauseam must interact, at least minimally, with advertising servers in order to fulfill its functions, it would necessarily expose users to more risk than the current state-of-the art blocker. For all cases (see *Comparative Evaluation* below) we were able to verify that risk to users was diminished with AdNauseam, both in comparison with the no-blocker case, *and* to AdBlock Plus, the most commonly installed blocker [37].

## II. Architecture

The AdNauseam software is comprised of four modules, each responsible for one of its primary functions: detection, extraction, visualization, and visitation.

### A. Detection

This module is responsible for the analysis and categorization of requests following a page view. Such requests, most often to third-parties, are first classified according to the type of elements they realize; whether advertisements, analytics, beacons, social-media, or functional widgets. The largest proportion of such requests (40-50%) are made to the first group, on which this module focuses, which includes ad and ad-tracking services [43]. This module determines which requests to block and which to allow, and distinguishes, in the latter category, between those that yield visual elements and those used only for tracking.

In order to categorize such requests, we leverage the capabilities of the open-source uBlock-Origin [17] project, a configurable, list-based "blocker" that is effective and efficient [43]. Like other blockers, uBlock allows users to specify publicly accessible lists of resources which contain syntactic matching rules for the retrieval of web resources. Based on these lists, we first determine whether a request should be blocked or allowed, and then, if allowed, whether it should be visible or hidden. If hidden, the element is downloaded and included in the page, but made invisible to the user via a content-script. Both blocking and hiding are specified via rules that may include the serving domain, the type of resource (e.g., images or video), and/or properties of the DOM container (for example, a DIV with a specific id or class). Rules are included from widely distributed lists that are updated and maintained by individuals and communities (e.g, "EasyList" [8]. Additionally, users can augment these lists with custom rules they create, either to block or hide new content, or to whitelist a site, page, or element.

Requests marked as blockable in AdNauseam are disallowed at the network level, mimicking the behavior of most other blockers, including uBlock, AdBlock Plus, Adblock, and Adguard, which perform blocking on some percentage of requests, and hiding on the remainder. The difference for AdNauseam is that a subset of requests which might be blocked in other blockers must be allowed in AdNauseam; specifically those that result in visual advertisements.[3] At the element hiding level, the detection module is invoked incrementally, via content-scripts, as page elements are loaded (or dynamically generated) and inserted into the DOM. Elements marked for hiding are assigned a CSS class that sets their display to invisible, and the surrounding DOM is collapsed so as not to leave blank space on the page. Each hidden element (generally a visual ad) is then passed to the *Extraction* module.

### B. Extraction

Once a visual element has been detected and hidden, we must then determine whether it is in fact an advertisement. If so, the extraction module of the system must extract the properties needed by the *Visualization* and *Visitation* modules. These properties include timestamp, size, content-URL, target-URL, page-detected-on, etc. Text-only ads, as often found on search engines, present a different challenge, as these are generally served inline along with page content rather than requested from a 3rd-party server. In these non-image cases, several additional fields are aggregated to form the content payload (title, description, tagline) and there is no content-URL linking to an external resource. To enable extraction of such data, AdNauseam includes a custom set of CSS selectors used to parse specific DOM attributes from text-ad sites (Google, Ask, Bing, etc.). Such filters run only on specific domains where text-ads have been previously discovered.

### C. Visualization

In order to facilitate transparency regarding tracking and profiling by advertisers, AdNauseam provides users with interactive visualizations of their collected ad data. These visualizations provide both

[2]This is our intent at least; Google's recent ban of the software may imply that this intent is understood.

[3]Interestingly, it is exactly this standard combination of functions—hiding and blocking—that Google cites as being in violation of its Terms of Service, a claim discussed below in the *Distribution* section.

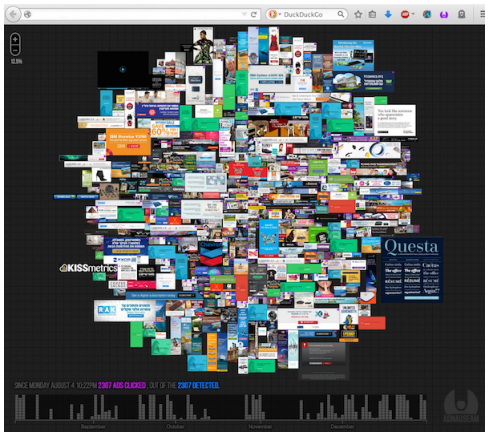


Fig. 1. AdNauseam's AdVault visualization.

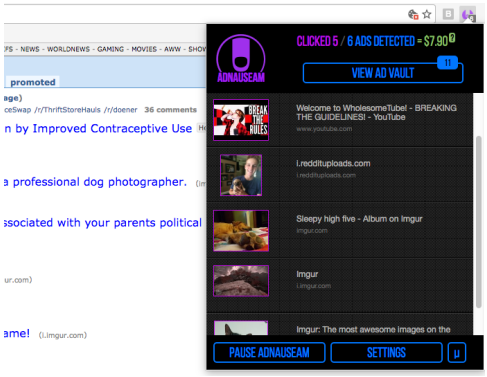


Fig. 2. Estimated cost to advertising networks.

high-level displays of aggregate data (see Figure 1), as well as the option to inspect individual ads for a range of data. A number of derived functions provide additional metrics (i.e., the total estimated charge to advertising networks for the ads visited for a page, site or time-period, as in Figure 2). Ads may be filtered and sorted by date, topic-category, ad-network, page-category, etc. The visualization module is a distinct contribution of AdNauseam that attempts to a) provide users with greater insight concerning their interactions with advertisers, and b) enable interested users and researchers to study the ad data collected. To facilitate the latter, we include mechanisms for importing and exporting ad data sets (as JSON) from within the extension. The use of this data for further research, with appropriate mechanisms for user consent, is an area of future work.

### D. Visitation

This module simulates clicks (or visits) on collected ads, with the intention of appearing to the serving website (as well as to advertisers and ad networks) as if the ad had been manually clicked. Currently, these clicks are implemented via AJAX, which simulates requests (matching headers, referer, etc.) that the browser would normally send. This provides users with protection against potential malware in ad payloads, as responses are not executed in the browser, and JavaScript, Flash, and other client-side scripting mechanisms are not executed. Similarly, AdNauseam blocks incoming cookies in responses to ad visits. The likelihood that a particular ad will be clicked depends on the user-configurable click-probability setting described further below.

What are the expected results of visiting some percentage of each user's collected ads? First, the data profiles of these users stored by advertising networks and data brokers may be polluted, as users' actual interests are hidden by generated clicks. This both protects individual users (assuming they have clicked, or may click, some ad in the future) as well as the larger user community, as aggregate statistics are less accurate and thus less valuable. Second, as advertisers must now potentially pay publishers for decoy clicks, a degree of mistrust is introduced into the economic model that drives the system. This is perhaps the most compelling argument for this strategy, as it could, given adequate adoption, force advertisers to change their behavior, either by developing new algorithms to filter such clicks, and/or by adopting more privacy-friendly policies (e.g., the EFF's Do Not Track mechanism).

### E. Distribution

Although not often discussed in an engineering context, the distribution issues we experienced highlight concerns we imagine will be only more relevant with the growing influence of corporate players of the software ecosystem.

The prototype for AdNauseam was initially developed as a Firefox-only extension available in Mozilla's addon store. In our production release, we added Opera and Chrome support and made the extension available in the Opera and Chrome stores respectively. We distributed upwards of 50,000 copies of the software over the subsequent six months, with the majority via Google's Chrome store. In January of 2017 however, we learned that Google had banned AdNauseam from the store, and further, had begun disallowing even manual installation or updates, effectively locking users out of their own saved data, all without prior notice or warning.

Google responded to our requests for justification by saying that AdNauseam had violated the following clause of the Store's Terms of Service: "an extension should have a single purpose that is clear to users."[4] The single purpose of AdNauseam, we would argue, is quite clear—namely to resist the non-consensual surveillance conducted by advertising networks, of which Google is a prime example. We do recognize that Google might prefer users not to install AdNauseam, as it opposes their core business model, but the Terms of Service do not (at least thus far) require extensions to endorse Google's business model. Moreover, this is not the justification cited for the ban. Whether or not one is an advocate of obfuscation, it is disconcerting to know that Google can make a privacy extension, along with stored data and preferences, disappear without warning. Here it is a counter-surveillance tool that is banned; perhaps tomorrow it will be a secure chat app, or password manager. For developers, who, incidentally, must pay a fee to post items in the Chrome store, this is cause for concern. Not only can one's software be banned without warning, but comments, ratings, reviews, releases and statistics are removed as well.

## III. Design Tensions

### A. Indistinguishability and Protection

For obfuscation to function effectively as a means of counter-surveillance, the noise generated must exhibit a high degree of *indistinguishability* with regards to data the system intends to capture; that is, it must be difficult for an adversary to distinguish injected

[4]In the one subsequent email we received, a it was stated that a single extension should not perform "both blocking and hiding," a claim that is difficult to accept at face value, as most blockers (including uBlock, AdBlock Plus, Adguard, etc.) perform both blocking and hiding, and have not been banned.

noise from the data it is attempting to collect [15]. However, there are times when this goal comes into tension with other aims of the software, specifically that of protection, e.g., from malware.

For example, following a software-generated ad click, we must decide whether the DOM for the response should be parsed and executed, and whether scripts should be allowed to run. In current AdNauseam versions, visits are implemented via AJAX requests, which means that no DOM is constructed from the response, and scripts are not executed. While protection is maximized here (against malicious code embedded in ads), obfuscatory power may be diminished. For example, one attack we have noted is from an adversary who, upon receiving a click request, sends a preliminary response containing code that executes, within the DOM, the actual request for the ad's target. If the code in the preliminary response never runs, then, from the advertising network's perspective, the click is never executed. We have experimented with solutions that address this issue (including executing clicks in sandboxed invisible tabs), but have yet to settle on a cross-platform solution that adequately protects user from potential malware/malvertising. For now we leave this as future work.

### B. Expression, Detectability, and Social Privacy

We have spoken of the expressive capabilities of data obfuscation generally, and of AdNauseam specifically. But how does this design goal relate to *detectability* (the degree to which an adversary can detect usage of the tool). Abstractly conceived, expression and detectability appear to lie at opposite ends of a spectrum; that is, if a tool is undetectable to an adversary, its expressive capability is minimal, at least in relation to the adversary. Thus, if expression is a goal of an obfuscation tool, designers may wish, perhaps counter-intuitively, to make its use detectable. If a goal of the tool is social privacy—the pollution of aggregate data collected by trackers—then, one might argue, the tool should be undetectable, so that the adversary cannot simply discard all data from those discovered to be using the tool. It appears, at least in a simplistic analysis that a tool cannot simultaneously achieve expressivity and protect social privacy[5].

To address this tension, we adapt the design of the user-configurable query-frequency in TrackMeNot [26], to AdNauseam, allowing users to adjust the probability (from 0-100%) that discovered ads will be clicked. As the slider is moved to the left, the likelihood that an ad will be clicked decreases, and vice versa to the right. If we hypothesize that, all other elements being equal, a lower click frequency will be harder to detect, then this setting would represent a mapping between expression, detectability, and social privacy[6]. As the slider moves left, expressivity is reduced as is the likelihood of detection, and the potential for social protection is increased. When moved right, the likelihood of detection increases, as do both expressivity and the potential for (economic) disruption, while the degree of social protection decreases. At the right extreme, when all ads are clicked, there is a higher likelihood the adversary will infer the use of AdNauseam and may choose to discard all click data from the user in question. In this case personal protection may be achieved as the user is no longer profiled, but there is no immediate gain in social protection. (As noted earlier, however, the fractured online advertising ecosystem makes this less obvious than, say, in the domain of search.) At the extreme left, the tool's clicks are undetectable (as there are none), and AdNauseam then functions like a standard blocker, simply blocking and hiding ads.

Detectability by an adversary is not, however, the only measure of a tool's expressive powers, as there may be other audiences that developers seek to impress. Take the case of ScareMail (mentioned in *Related Work* below), an obfuscation tool that appends an algorithmically-generated narrative containing NSA "trigger" words to the end of sent emails. Users are able to express resistance to the recipients of their emails irrespective of whether the adversary, presumably the email provider, is able to detect its use. Whether ScareMail is actually a "privacy tool," or simply a tool for social protest focusing on email privacy, is a question we will not take up here. Our purpose, instead, is to argue that the expressive potential of software need not be mapped only and directly to detectability by the actor identified as the adversary. This would rule out subtle forms of social expression that we are seeing; for example, where users have spontaneously sought ways to share their ad collections online.

## IV. Evaluation

Qualitative evaluation was performed iteratively throughout development, often guided by solicited and unsolicited feedback from various constituencies, including users, developers, reviewers at Mozilla and Opera, and a range of privacy and security advocates. When considering how to evaluate the software, the question of whether AdNauseam in fact "worked" seemed at first to be most obvious and simple to address. We soon realized, however, that the meaning of this question shifted as users' goals, expectations, and perceived risks varied. Evaluating AdNauseam on the basis of feedback from the various constituencies was often a two-part process: first determining user orientations, and then examining feedback in light of their goals, concerns, and priorities. Additionally, beyond the technical issues with which we grappled, a subset of critiques consistently addressed ethical concerns. Thus we have split the discussion below into technical and ethical components.

### A. Technical

Evaluation of obfuscation-based strategies for counter-surveillance is often relatively straightforward. Take search, for example. One can extract query logs from tool users, containing both user-generated and software-generated queries, and then attempt to separate the two, either manually or automatically; in the latter case, by running best-practice machine-learning (or other) algorithms. Although one may not know the exact capabilities of the adversary, evaluators can make educated guesses as to the type of attacks to be considered, whether timing, query content, side-channel, or other means (for details of such evaluations in the search case, see [15]). If we find that the adversary can differentiate true queries with high accuracy, then our generated queries can be filtered and, from a protection standpoint, we must say that the tool fails.[7]

At first glance, evaluating AdNauseam would seem to call for a similar approach in which one measures the difficulty with which an

[5]Real-world domains, like advertising surveillance, are often complicated by a range of socio-economic factors. For example, the analysis above assumes a single adversary with full knowledge of the system, which, as discussed, is not the case here. Further, simply because an adversary *can* filter the data for tool users does not mean they will, especially given high enough adoption rates. Such data is at the core of the business model that drives such collection, and thus ignored profiles have direct economic impact. Clearly there is some number of ignored users after which the practice is no longer economically viable. One must also consider the effort and expense required to initiate such filtering, and the questions which it raises – should, for example, the data of tool users be discarded forever, or are such users to be monitored for tool stoppage as well? A range of social, economic, and cultural factors interact here to influence what is, in the end, a complex business decision.

[6]As a variety of factors influence detectability, the actual assertion of such a linear relationship would require supporting evidence.

[7]We may still argue that the socio-economic cost of filtering is prohibitively high, or that the tool is successful in terms of expression, but these are non-technical concerns which we must bracket for the moment.

adversary, using standard techniques, can distinguish user clicks from generated clicks. However, there are three distinct cases to consider, depending on what ads, if any, are seen by the user. In the first case, where a user enables ad-hiding, disables DNT exceptions, and does not provide a whitelist, no ads are visible to the user, and there are no *true* clicks for an adversary to discover. This is also true for the second case in which the only ads visible are those of DNT-respecting sites (AdNauseam's default settings). As such sites by definition do not track users, there are again no true clicks to discern. The third case applies for users who see non-DNT ads, either because they have disabled hiding entirely, or because they have manually whitelisted sites. Here we must consider the tool's detectability determined in large part by the user-selected click-probability. If this probability is set high enough that detection is possible, the adversary may simply discard all clicks from the user in question; a result similar to that obtained from a successful blocking-only strategy, except with an enhanced expressive component (as the adversary must both recognize the tool and then take action to perform the filtering). While this may be considered a success for some users, as they are no longer profiled via click-data, since the data is discarded there is no net gain in what we have referred to as social privacy. If click-probability is low enough, however, that the tool's actions are not detectable, then in order to evaluate the degree of social protection provided, we need to asses both a) the indistinguishability of the clicks, and b) the impact that successful decoy clicks have on the resulting users profile (a complex question we return to in the *Future Work* below). Of course even if requests themselves are indistinguishable, there may still be side-channels available to adversaries, as discussed above. For the moment we leave the specifics of such evaluations to future work.

*1) Comparative:* To further evaluate performance we compare AdNauseam with commonly used blockers on a range of dimensions, relating both to protection (number of 3rd parties contacted) and usability (page-load speed and memory efficiency). Tests were first run without any extension, then with AdNauseam, Adblock Plus [1], uBlock-Origin [17], and Privacy Badger [10]. Tests were performed with each extension's default settings after resetting the browser to its install state. After visiting the websites in the test set (between 15 and 85 popular URLs, depending on the test) via the Selenium browser automation tool, we evaluated the safety of each extension in terms of the number of 3rd parties contacted (Figure 3), page-load speed (Figure 4), and memory efficiency. As shown in the graphs below, AdNauseam performed better on all dimensions than no blocker and, perhaps surprisingly, better than AdBlock Plus. As expected, AdNauseam performed less well than uBlock, due to the need to allow visual ad resources, rather than blocking them outright.

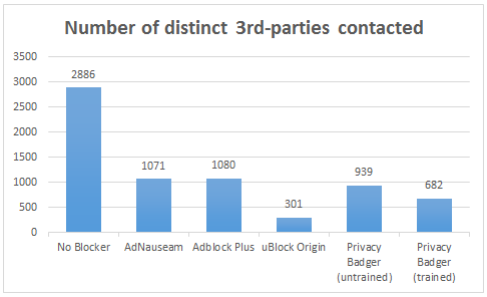


Fig. 3. Number of distinct third-parties contacted.

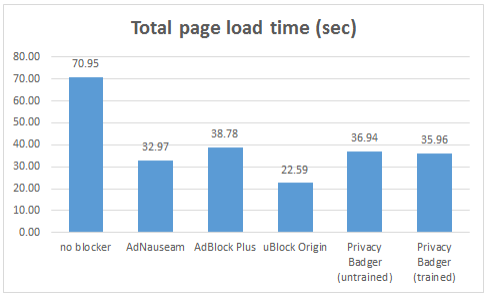


Fig. 4. Total page load time (sec).

## B. Ethical

In adopting the philosophy of data obfuscation AdNauseam seeks to shield users from the inexorable and inappropriate probes of services and third parties. Choosing obfuscation, however, means taking seriously the ethical critiques it has drawn, including charges of dishonesty, wasted resources, and polluted databases. Addressing these issues, Brunton and Nissenbaum [4] ask creators of obfuscating systems to answer two questions: first, whether their aims are laudable; and second, whether alternatives exist that might achieve these aims at lesser cost. Regarding the first charge we begin by saying that ubiquitous online surveillance violates the tenets of a liberal democracy. The troubling nature of this surveillance is exacerbated by its surreptitious operation, its prevarication, and its resistance to the wishes of a majority of users; claims clearly established through systems' analysis, demonstrations and public opinion surveys [42], [16], [41]. Data from this surveillance contributes to the creation of valuable, but often highly problematic profiles that fuel big data industries with uncertain, potentially negative effects on their subjects. Against this backdrop, we judge the aims of AdNauseam, which include the disruption of this process, to be morally defensible.

The second charge asks whether obfuscation imposes a lower collateral costs than alternatives for achieving similar ends. Comparing the purported cost of AdNauseam against alternative approaches involves uncertainties we are unable to tackle here. But, by the same token, this dearth of concrete evidence poses a challenge to critics who accuse ad blockers—and AdNauseam—of harming the web's economy. Even if one holds that the "best" resolution would be societal-level regulation, there has been little progress on this front. As important as seeking credible alternatives, however, is weighing the purported costs of using AdNauseam. Among the latter, the harm of "wasting" network bandwidth or server resources is ironic at best, given the vast amount of bandwidth used by advertisers and trackers, the performance degradation resulting from loading this unwanted content, and the financial toll on those paying for fixed data plans. From an ethical perspective, it is questionable whether the term "waste" is appropriate at all. For those who deliberately choose to use AdNauseam it offers a potential escape from inappropriate profiling. In our view, this is not a worthless endeavor.

One of the most aggressive charges leveled at AdNauseam is that it perpetuates "click fraud." Since obfuscation and fraud both involve forms of lying that disrupt entrenched systems, it is important to evaluate whether the two forms are alike. To carry this out, we consulted various definitions: "[Click] fraud occurs when a person, automated script or computer program imitates a legitimate user of a web browser, clicking on such an ad without having actual interest in the target of the ad's link" [29] comes close to capturing AdNauseam in its notion of clicking without actual interest, but this definition seemed overly broad in that it commits users to click only on ads in which they are interested, and seems an unjustifiable restriction on liberty of action. We also argue that if the automated script is

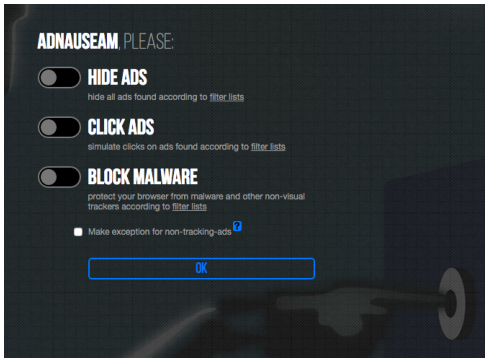


Fig. 5. Opt-in settings on initial install page.

performing as an agent of an individual, through that individual's legitimate choice, then the script is a proxy for the user. John Battelle's account [3], which includes motive and intention, gets closer to the standard meaning of "fraud" in "click fraud": the "'decidedly black hat' practice of publishers illegitimately gaming paid search advertising by employing robots or low-wage workers to repeatedly click on each AdSense ad on their sites, thereby generating money to be paid by the advertiser to the publisher and to Google." While elements of the above definitions overlap with AdNauseam's clicking (without genuine interest in their targets), machine automation is only incidental to click fraud, and may instead involve "low-wage workers." More significant is what AdNauseam does not share with click fraud, namely action on behalf of stakeholders resulting in financial gain. In litigated cases of click fraud the intention to inflate earnings has been critical.

We readily admit that a primary aim of AdNauseam is to disrupt business models that support surreptitious surveillance. It does not follow however that AdNauseam is responsible for the demise of free content on the web. First, it is not, as we make clear on the project page, advertising that is the primary target of the project, but rather the tracking of users without their consent. Contextual advertising that does not involve tracking can certainly support free content just as it has in the past. Second, web content is not actually 'free' as this argument implies. The development of the Internet has been supported largely by government funding (and thus by taxpayers) since its beginning. In fact, vast infrastructure and energy costs are still born in large part by taxpayers, not to mention the potentially species-threatening cost to the environment posed by increasing data traffic [23]. Critics may say that ad blocking users free ride upon those who allow themselves to be tracked, however, in our view this presumes an entitlement on the part of trackers that is indefensible; one may equally charge trackers with destructive exploitation of users [4]. Lastly, in regard to free riding, we wish to point out that the hiding of ads is an optional element of AdNauseam, one that users must explicitly opt into when they install the software (see Figure 5); AdNauseam's visitation and visualization modules work equally well whether the user elects to view ads or to hide them.

## V. Related Work

The strategy of obfuscation has been broadly applied—in search [26], location tracking [32], social networks [30], anonymity [7], [39], etc.—and, as such, has been recognized as an important element of the privacy engineer's toolbox. A range of obfuscation-based projects have been described in [4], including FaceCloak [30], for Facebook profiles, BitTorrent Hydra [39], for decoy torrent sites, and CacheCloak [32], for location data. There have also been a number of obfuscation schemes for web search [15].

Other relevant work, described in [27], has come from the art/tech community. "I Like What I See" is a tool that clicks all 'Like' links on Facebook to obscure user interests. "ScareMail" [18] is an extension built atop Gmail that append an algorithmically-generated narrative containing NSA "trigger-words" to the end of each sent email. "Invisible" [21] extends obfuscation to the context of genetic privacy via a spray that obfuscates DNA to frustrate identification.

Two early tools addressing surveillance integrate ad-blocking with some broadly-defined social good: AddArt [2] replaces ads with user-configurable art, while AdLiPo [25] does the same with language art. Lightbeam [33], provides displays of users' connections, including to advertising networks (though not ads themselves). Floodwatch [13] is the one tool we have found that provide visualizations similar to our own, though it requires communication with one or more 3rd-party servers to do so. Privacy Badger [10] blocks third-party requests based on real-time monitoring of the connections they attempt rather than via lists, blocking only those resources engaged in tracking.

## VI. Future Work

AdNauseam provides individuals with the means to express their commitment to online privacy without the need to depend on the good will or intervention of third-parties. Although fully functional, AdNauseam is perhaps best considered as a proof of concept for a particular approach to privacy, that is, privacy through obfuscation. As discussed, AdNauseam's potential lies in its capacity to protect individuals against data profiling, as well as simultaneously providing a proactive means of expressing one's views to monolithic and largely uninterested corporations. Going forward, a scientific approach to evaluating AdNauseam's performance, or the performance of any system adopting obfuscation, needs a rigorous means of measuring success—namely, evidence that decoy clicks have been registered and have an impact on the resulting profile. Such needs are likely to turn not only on the statistical analysis of signal-to-noise ratios, but also on a practical understanding of how click data is actually mined and used, and the extent to which it influences aspects of user profiles. This would allow future iterations of obfuscation-based tools to be both effective and efficient in the noise they produce.

Future work could take several directions. In the near term we hope to better answer the question of how to perform indistinguishable clicks without exposing users to potential harms via downloaded content, as discussed above. Though complex, P2P approaches for the sharing of obfuscation data between users is a ripe area of future work, with users potentially visiting the ads detected by peers as a means of both shielding their data and maximizing indistinguishability. A central challenge here would be meeting functional criteria while not compromising the design constraints discussed early in this paper, e.g., transparency and independence from third-parties. Finally, beyond the technical, work exploring the motivations and qualitative experiences of users who select obfuscation tools could shed light on the unique potential such tools might offer in additional domains.

## VII. Conclusions

AdNauseam operates in a technologically and socially complex environment, one in which user data is perceived to be highly valuable. For individuals, however, recorded patterns potentially open a window into their lives, interests, and ambitions. Thus surveillance via advertising is not only a source of individual vulnerability, but also interferes with the rights to inquiry, association, and expression that are essential to a healthy democratic society. Consequently,

there remain tensions between individual users, collective social and political values, and the economic interests of publishers and advertisers. In a better world, this tension would be resolved in a transparent, trust-based accommodation of respective interests. Instead, concerned users find little transparency and few credible assurances from advertisers that privacy will ever trump the pursuit of profit. Thus trust-based mutual accommodation gives way to an adversarial relationship, one in which we must leverage all the strategies at our disposal. Our success in this endeavor will depend in part on how well we share our experience applying known strategies to new contexts, in concrete and specific detail, according to an evolving set of best practices, as we have attempted above.

We conclude with a philosophical point. In some of the most revealing exchanges we have had with critics, we note a palpable sense of indignation, one that appears to stem from the belief that human users have an *obligation* to remain legible to their systems, a duty to remain trackable. We see things differently; advertisers and service providers are not by default entitled to the externalities of our online activity. Rather, users should control the opacity of their actions, while powerful corporate entities should be held to the highest standards of transparency. Unfortunately this is the opposite of the status quo; our trackers want us to remain machine-readable so that they can exploit our most human endeavors (sharing, learning, searching, socializing) in the pursuit of profit. AdNauseam attempts to represent an alternative position.

## Acknowledgements

The authors wish to thank all those who have helped with the creation and spread of AdNauseam, particularly Sally Chen, Leon Eckert, Cyrus Suen, and Emily Goldsher-Diamond. This paper has been greatly improved by comments from our (anonymous) reviewers, our "shepherd" Ero Balsa, and Seda Gürses; specifically for her substantive guidance throughout the writing process and her unflagging support of productive, cross-disciplinary work, no matter the difficulty. Finally, thanks go to Mushon Zer-Aviv, for his profound contributions to all aspects of the project.

This publication has been supported in part by grants from US NSF CNS/NetS 105833, US NSF SATC 1642553, and the Research Grants Council of Hong Kong, China (Project No. CityU 11669616)

## References

[1] AdBlock Plus. "AdBlock Plus." n.d. https://adblockplus.org/.
[2] AddArt. "AddArt." n.d. http://add-art.org/.
[3] Battelle, John. *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. Nicholas Brealey, 2011.
[4] Brunton, Finn, and Helen Nissenbaum. *Obfuscation: A User's Guide for Privacy and Protest*. MIT Press, 2015.
[5] Cavoukian, Ann, and Michelle Chibba. "Cognitive Cities, Big Data and Citizen Participation: The Essentials of Privacy and Security". *Towards Cognitive Cities*. Springer International Publishing, 2016. 61-82.
[6] Click Fraud. (n.d.). In Wikipedia. Retrieved Aug 1, 2016. https://en.wikipedia.org/wiki/Click_fraud
[7] Chakravarty, Sambuddho, et al. "Detecting Traffic Snooping in Anonymity Networks Using Decoys." 2011.
[8] "EasyList." 2016. https://easylist.to/
[9] Englehardt, Steven, and Arvind Narayanan. "Online Tracking: A 1-million-site Measurement and Analysis." *Proceedings of the ACM SIGSAC Conf. on Computer and Communications Security*. ACM, 2016.
[10] Electronic Frontier Foundation. "Privacy Badger." n.d. https://www.eff.org/privacybadger.
[11] Electronic Frontier Foundation. "Do Not Track." n.d. https://www.eff.org/issues/do-not-track.
[12] Flanagan, Mary, Daniel C. Howe, and Helen Nissenbaum. "Embodying Values in Technology: Theory and Practice." *Information technology and moral philosophy*. (2008): 322-353.
[13] Floodwatch. "Floodwatch." n.d. https://floodwatch.o-c-r.org/.
[14] Friedman, Batya, Daniel C. Howe, and Edward Felten. "Informed Consent in the Mozilla Browser: Implementing Value-Sensitive Design". *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*. IEEE, 2002.
[15] Gervais, Arthur, et al. "Quantifying Web-Search Privacy." *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2014.
[16] Goldfarb, Avi, and Catherine Tucker. "Shifts in Privacy Concerns." *The American Economic Review* 102.3 (2012): 349-353.
[17] Gorhill. "uBlock Origin - An efficient blocker for Chromium and Firefox." 2016. https://github.com/gorhill/uBlock
[18] Grosser, Ben. "ScareMail." 2013. Web http://bengrosser.com/projects/scaremail/.
[19] Gürses, Seda, Carmela Troncoso, and Claudia Diaz. "Engineering Privacy by Design." *Computers, Privacy & Data Protection* 14.3, 2011.
[20] Gürses, Seda, Carmela Troncoso, and Claudia Diaz. "Engineering Privacy by Design Reloaded." *Amsterdam Privacy Conference*. 2015.
[21] Dewey-Hagborg, H. "Invisible." 2014. http://www.newmuseumstore.org/browse.cfm/invisible/4,6471.html.
[22] Hansen, Marit, Meiko Jensen, and Martin Rost. "Protection Goals for Privacy Engineering." *Security and Privacy Workshops*. IEEE, 2015.
[23] Hazas, Mike, et al. "Are there limits to growth in data traffic?: On time use, data generation and speed." *Proceedings of the Second Workshop on Computing within Limits*. ACM, 2016.
[24] Hoepman, Jaap-Henk. "Privacy Design Strategies." *IFIP International Information Security Conference*. Springer Berlin Heidelberg, 2014.
[25] Howe, Daniel C. "AdLiPo" 2014. http://rednoise.org/adlipo/.
[26] Howe, Daniel C. and Helen Nissenbaum. "TrackMeNot: Resisting Surveillance in Web Search." *Lessons from the Identity Trail: Anonymity, Privacy and Identity in a Networked Society* 23, 2009: 417-436.
[27] Howe, Daniel C. "Surveillance Countermeasures: Expressive Privacy via Obfuscation". APRJA, A Peer-Reviewed Journal About Datafied Research 4.1, 2015.
[28] Iachello, Giovanni, and Gregory D. Abowd. "Privacy and Proportionality: Adapting Legal Evaluation Techniques to Inform Design In Ubiquitous Computing." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2005.
[29] Liu, De, Jianqing Chen, and Andrew B. Whinston. "Current Issues in Keyword Auctions". *Business Computing (Handbooks in Information Systems, Vol. 3)* 2009: 69-97.
[30] Luo, Wanying, Qi Xie, and Urs Hengartner. "FaceCloak: An Architecture for User Privacy on Social Networking Sites" *International Conference on Computational Science and Engineering*, 2009.
[31] Mansfield-Devine, Steve. "When advertising turns nasty". *Network Security* 2015.11 2015: 5-8.
[32] Meyerowitz, Joseph and R. R. Choudhury. "Hiding stars with fireworks: Location privacy through camouflage." *Proc. of the 15th annual international conference on Mobile computing and networking*. ACM, 2009.
[33] Mozilla. "Lightbeam." 2016. https://www.mozilla.org/en-US/lightbeam/.
[34] Murphy, Kate. "The Ad Blocking Wars." ***New York Times***, 20 Feb. 2016.
[35] Nikiforakis, Nick, et al. "Cookieless monster: Exploring the ecosystem of web-based device fingerprinting." *IEEE symposium on Security and privacy (SP)*. IEEE, 2013.
[36] Nissenbaum, Helen. *Privacy in Context: Technology, Policy and the Integrity of Social Life*. Palo Alto: Stanford University Press, 2010.
[37] PageFair, Adobe "The cost of ad blocking–PageFair and Adobe 2015 Ad Blocking Report", 2015.
[38] Regan, Priscilla M. *Legislating privacy: Technology, social values, and public policy.* Univ of North Carolina Press, 1995.
[39] Schulze, Hendrik, and Klaus Mochalski. "Internet Study 2008/2009." *Ipoque Report* 37 2009: 351-362.
[40] Spiekermann, Sarah, and Lorrie Faith Cranor. "Engineering Privacy." *IEEE Transactions on software engineering* 35.1 2009: 67-82.
[41] Tucker, Catherine E. "Social networks, personalized advertising, and privacy controls." *Journal of Marketing Research* 51.5 2014: 546-562.
[42] Turow, Joseph, et al. "Americans reject tailored advertising and three activities that enable it." 2009.
[43] Wills, Craig E., and Doruk C. Uzunoglu. "What Ad Blockers Are (and Are Not) Doing." *Fourth IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb)*. IEEE, 2016.