Oh, the Places You've Been! User Reactions to Longitudinal Transparency About Third-Party Web Tracking and Inferencing

Ben Weinshel, Miranda Wei, Mainack Mondal†, Euirim Choi,
Shawn Shan, Claire Dolin, Michelle L. Mazurek‡, Blase Ur
University of Chicago, † IIT Kharagpur and University of Chicago, ‡ University of Maryland
{weinshel, weim, mainack, euirim, shansixioing, cdolin, blase}@uchicago.edu, ‡ mmazurek@cs.umd.edu

ABSTRACT

Internet companies track users' online activity to make inferences about their interests, which are then used to target ads and personalize their web experience. Prior work has shown that existing privacy-protective tools give users only a limited understanding and incomplete picture of online tracking. We present Tracking Transparency, a privacy-preserving browser extension that visualizes examples of long-term, longitudinal information that third-party trackers could have inferred from users' browsing. The extension uses a client-side topic modeling algorithm to categorize pages that users visit and combines this with data about the web trackers encountered over time to create these visualizations. We conduct a longitudinal field study in which 425 participants use one of six variants of our extension for a week. We find that, after using the extension, participants have more accurate perceptions of the extent of tracking and also intend to take privacy-protecting actions.

CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**.

KEYWORDS

third-party online tracking, transparency, usable privacy, user study

ACM Reference Format:
Ben Weinshel, Miranda Wei, Mainack Mondal, Euirim Choi, Shawn Shan, Claire Dolin, Michelle L. Mazurek, and Blase Ur. 2019. Oh, the Places You've Been! User Reactions to Longitudinal Transparency About Third-Party Web Tracking and Inferencing. In *2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19), November 11–15, 2019, London, United Kingdom.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3319535.3363200

1 INTRODUCTION

Many websites embed scripts and ads provided by *advertising networks*, or companies that select and display ads on websites. Using data from web requests, persistent cookies, and fingerprints, advertising networks and other companies (collectively termed *trackers*) link records of a user's browsing activity across multiple web pages. These records can be used to make inferences about the user's demographics and interests, enabling *targeted advertising*, in which online ads are tailored for a particular user. Targeted advertising has come under increasing scrutiny as a threat to privacy and a potential enabler of discrimination based on race, age, sexual orientation, and other sensitive categories [5, 20, 39, 87].

Although targeted advertising is ubiquitous on the web and has received substantial news coverage, prior work suggests many people do not understand how they are tracked across websites, nor how their interests are inferred [95]. Users have limited means to learn about the extent and implications of third-party tracking and personalization. Advertising networks provide general explanations of ad targeting, and some companies show interests that have been inferred (Figures 1a, 1b), though they have been shown to provide incomplete and possibly misleading information [4, 89, 99]. User-installed browser extensions (Figures 1c, 1d) highlight the trackers on the web pages a user visits, but provide limited insight into longitudinal tracking and do not explain what companies have learned and inferred about users' interests over the long term [80].

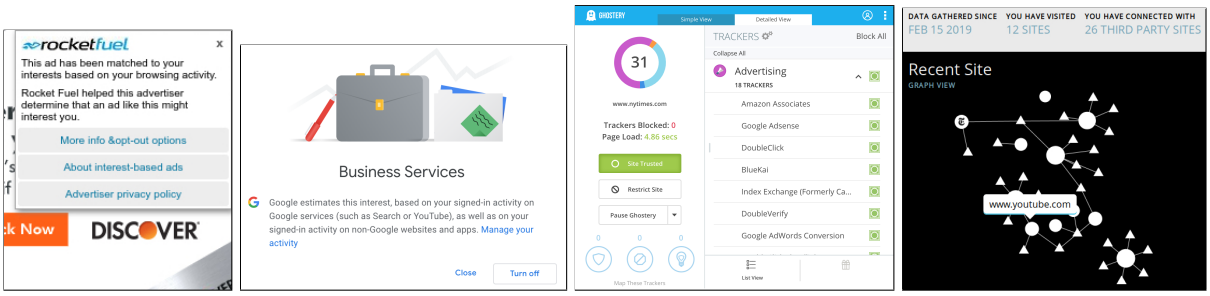
In this work, we seek to understand how different kinds of transparency about tracking—including the longitudinal transparency missing from current tools—affect users' perceptions. To conduct this evaluation, we developed a browser extension, Tracking Transparency, for displaying longitudinal tracking information. We created the extension's user interface iteratively, informed by 13 interviews. The extension's detailed visualizations aim to make tracking, especially its longitudinal aspects, comprehensible to users.

To populate the user interface and visualizations, Tracking Transparency locally stores logs of the pages the user visits and trackers encountered, detected via HTTP requests to known trackers. Further, the extension provides a personalized approximation of the inferences trackers may have made from this accumulated information. For example, a tracker present on a page about dogs and a page about football may infer a user who visited both is interested in pets and sports. To generate this approximation, we apply the TF-IDF topic-modeling algorithm (selected after evaluating several topic-modeling options). Our method for simulating inferences operates client-side, limiting the information shared with the research team, and works without any cooperation from trackers themselves.

The ultimate goal of our work is to understand how visualizing longitudinal and inference-level information about online tracking impacts users' knowledge, perceptions, and attitudes. To this end, the Tracking Transparency tool provides a platform to conduct

CCS '19, November 11–15, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6747-9/19/11.
https://doi.org/10.1145/3319535.3363200



(a) AdChoices text **(b) Google inferred interest** **(c) Ghostery** **(d) Lightbeam**

Figure 1: Examples of the transparency visualizations currently available to users concerning ad targeting and personalization.

studies on how users respond to personalized explanations of tracking using their own browsing history and plausible approximations of what companies could have learned about their interests. The source code for our extension is available on GitHub.[1]

We conducted a longitudinal field study in which 425 participants from Mechanical Turk installed our browser extension and used it for a week. We aimed to understand how particular types of transparency provided either by our tool's visualizations or by visualizations approximating those of existing approaches impact users. Therefore, we randomly assigned each participant to use one of six variants. These variants represented the information currently provided by advertising networks (static explanations), state-of-the-art privacy tools like Ghostery and Lightbeam, and our fully featured Tracking Transparency interface. Participants completed surveys before and after using the extension.

One-third of participants who saw our fully featured interface were surprised by how trackers used their browsing history to infer their interests, and that interests were even inferred in the first place. Most participants who saw any of the variants other than our control condition were surprised by the amount of tracking occurring. The fully featured interface increased participants' self-reported intentions to take privacy-protective actions (e.g., using private browsing) significantly more than static explanations typical of advertising networks' own disclosures, as well as slightly more than state-of-the-art, tracker-centric privacy tools. Tracking Transparency also helped participants more accurately characterize the tracking ecosystem, which they had significantly underestimated.

Currently, privacy tools and advertising networks themselves provide users only blurry snapshots of the online tracking and targeting ecosystem. Tracking Transparency gives richer visibility into the kinds of information that may be inferred. Furthermore, it does so client-side and without cooperation from tracking companies. It is a first step toward helping users recognize and better understand that trackers record what they browsed and make inferences from this knowledge. With improved understanding, users can make more informed privacy decisions and policy demands.

This work's main contribution is to human-subjects research in proposing a new user interaction that presents online tracking through longitudinal and inference-level visualizations, as well as measuring how these visualizations impact users' attitudes and intended behaviors. One key result from our field study is that these new types of visualizations led to increased intention to take privacy-protective actions, especially compared to the static text typical of advertising networks' disclosures. A second key result is that these visualizations increased both participants' knowledge of how inferences are inferred from browsing data and participants' ability to quantify the extent of the tracking ecosystem.

[1] https://github.com/UChicagoSUPERgroup/TrackingTransparencyCCS2019/

2 BACKGROUND AND RELATED WORK

We discuss mechanisms of online tracking, user perceptions of targeted ads, existing privacy tools, and topic modeling.

2.1 Online Tracking and Targeting

Since the first observed third-party web tracker in 1996 [53], the online ecosystem of third-party tracking has grown substantially in magnitude and complexity [102]. Web tracking is an "arms race" between trackers and tracker-blockers [8, 37, 41, 57, 63, 67, 77, 102]. A number of methods can be used for third-party tracking: cookies [28]; fingerprinting [1]; tracking pixels [78]; and more [27].

A growing body of work characterizes user perceptions of online tracking [30, 82, 95, 101]. Users have a wide variety of reactions to web tracking related to targeted ads. These reactions range from positive to negative [79, 90] and from comfortable to creepy [23, 58, 91]. Users often evince a tension between a desire to see relevant ads and their concerns about invasiveness [23, 91]. Further, studies disagree on when users are willing to share information with advertisers [25, 52] or pay for privacy [17, 26, 49]. Users also have generally negative opinions about, or ignore, online tracking disclosures, such as cookie disclaimers [46] and privacy notices [75, 103]. Nevertheless, most of these studies only provide snapshots into users' perceptions of tracking. To our knowledge, how detailed personalized and longitudinal information might qualitatively change user perceptions has not previously been evaluated.

Algorithmic processes assign advertisements to users based on inferred profiles. People do not always understand the output of algorithms [29, 72] despite having many opinions about what should be done if algorithms are biased [31], imperfect [30], or discriminatory [5, 70]. Nonetheless, users can be surprisingly deferential to algorithmic inferences, such as by showing reluctance to make changes to automatically generated profiles [94] or even self-auditing to fit inferences made by an algorithm [86].

2.2 Privacy Tools and Transparency

Growing concern over online tracking has led to an influx of privacy tools. There are many anti-tracking browser extensions [27], and widely-used web browsers now feature cookie blocking and tracking prevention [44, 64, 98]. However, given the rapid evolution and pervasiveness of online tracking, technical privacy tools are unlikely to be the complete solution. Tools to block ads and tracking enjoy meaningful adoption, but users struggle to understand the information they present [56, 59, 80]. Further, with frequent changes in tracking techniques, blocking tracking is an arms race [27].

In this section, we review current privacy tools. In our field study, we compare representatives of the types described below. First, clicking the AdChoices icon that accompanies targeted ads [43] shows text explaining targeting broadly. Users find the icon itself confusing [51]. Furthermore, clicking the icon usually leads to a pop-over in which targeting is explained only in the abstract, without specific or concrete insight into why that ad was chosen. For example, "This ad has been matched to your interests. It was selected for you based on your browsing activity' ' (Figure 1a).

Second, advertising networks (e.g., Google [36], Facebook [32], and Oracle [65]) sometimes provide "privacy dashboards" displaying some of the inferences they have made. For example, Google's Ad Settings (Figure 1b) lists some estimated interests and gives vague explanations of how they were chosen. These dashboards have been shown to be incomplete [89, 99], misleading [4], and potentially inaccurate [9, 21]. They have also been used to show discrimination in advertising [20] and targeting on sensitive topics [48, 99]. Our system greatly expands on this class of transparency (cf. Section 3.3) by simulating principled attribution of inferences.

Third, privacy-focused browser extensions like Ghostery (Figure 1c), Disconnect, and Privacy Badger show the trackers on the current page. They also allow users to block them. Mozilla's Lightbeam displays connections between trackers and websites (Figure 1d) based on the user's browsing history. Although these tools increase awareness of online tracking and privacy issues, users often fail to understand the full impact of tracking and targeting [80]. Tracking Transparency explains to users which inferences trackers could perhaps make about them based on longitudinal information.

Fourth, researchers have evaluated a variety of new techniques for expressing preferences about third-party tracking. For example, TrackMeOrNot gives users the ability to block trackers according to stated privacy preferences [60]. Usability studies of these tools generally find that it is difficult for users to make meaningful changes to their tracking preferences [7, 45, 50, 59].

Researchers have designed interfaces to explain privacy concepts like mobile app permissions [3], app data sharing [6, 92], and web privacy [84]. They have also documented abstract perceptions of online tracking [17, 46, 49, 52, 75, 91]. To our knowledge, no work has focused on explaining the longitudinal and inference-level aspects of web tracking, nor evaluating how such visualizations impact users' attitudes and knowledge. We fill this gap.

2.3 Topic Modeling

The categorization of web pages and attribution of inferences has been approached from multiple angles. Resources like the Open Directory Project [71] have manually categorized many sites. Companies like SimilarWeb provide APIs that categorize sites using proprietary datasets and machine learning [81]. Statistical techniques have been used to attribute inferences in lab settings [7, 20, 48].

For the Tracking Transparency extension, we wanted to categorize arbitrary web pages in a client-side, privacy-preserving manner, which to our knowledge has not been directly explored in the literature. As such, we relied on prior work on topic generation and keyword extraction. Latent Dirichlet Allocation (LDA) [12] uses generative statistical models to group and model documents. The graph-based information retrieval algorithm TextRank [61] is used for keyword extraction, while PageRank [66] ranks the importance of words in texts. Wikipedia is effective at finding topic-related texts [34, 54], especially with graph-based algorithms [16, 100].

3 TRACKING TRANSPARENCY EXTENSION

Our ultimate goal was to contribute an understanding of how longitudinal and inference-focused visualizations impact users' attitudes and awareness about third-party web tracking. However, to our knowledge, no such longitudinal and inference-focused tools existed prior to this study. Thus, we developed the Tracking Transparency extension, which enables personalized visualizations about web tracking by collecting data about a user's browsing client-side. Section 3.1 provides an overview of the user interface, which we refined through interviews (Section 3.2). Section 3.3 describes the topic modeling approach we use to simulate inferencing. Section 3.4 discusses data we collected on the sensitivity of interest categories.

The data the extension generates is stored in a client-side IndexedDB database. For all pages visited, it stores basic metadata (page title, URL, time), detected trackers, and the page's inferred topic. The extension uses web workers to minimize resource usage. We are releasing our extension open-source to bootstrap further human-subjects research into the impact of visualizing tracking.

To detect trackers, we check all outgoing web resource requests to third-party domains against Disconnect's list of known trackers [22]. Mozilla Firefox and DuckDuckGo's privacy extension also use this list [24, 44]. We record a tracker as present for a page if there are any requests to that tracker's domain. Most known types of tracking, including cookies and fingerprinting, generate requests our extension detects. This method lets us both detect trackers without pre-training and associate domains with entities (e.g., mapping both `doubleclick.net` and `google-analytics.com` to "Google"). Blacklist-based detection may not detect all potential trackers, but the Disconnect list is widely used and frequently updated.

3.1 Interface Components

Tracking Transparency provides a dashboard and accompanying interface components to present users with personalized examples of trackers in their online browsing. Note that in our field study (Section 4), some interface components were hidden in some study conditions for purposes of controlled comparison.

Toolbar popup. To access the extension, users first click the extension's icon in the toolbar. This displays the *popup* (Figure 2), which summarizes tracking both on the current page and since installation. It also provides a link to the dashboard home page.

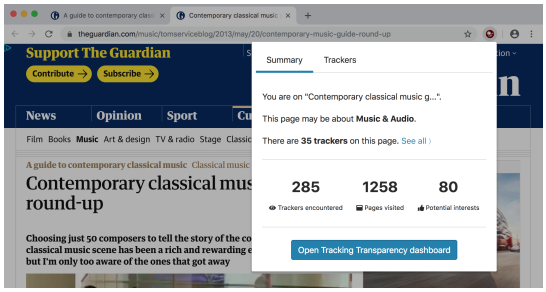


Figure 2: Browser open to a web page with popup visible.

Home page. When users first open the dashboard, they see a *home page* with a static explanation defining trackers and interests (Figure 7 in the appendix). There are also summaries of the top trackers encountered, the top interests inferred, and general statistics about the number of pages visited. The dashboard also defines and explains key concepts like interests, trackers, and sites.

Interests tab. The *Interests* tab (Figure 3) presents simulated interests that trackers may have inferred about that user based on their browsing history. This page's focal point is an interactive sunburst graphic that shows the frequency of interests inferred about them, organized into hierarchical categories. Users can apply filters to the sunburst, displaying interests by recency, popularity, and sensitivity. The popularity filter uses audience-size data scraped from the Google AdWords targeting interface. Sensitivity ratings were determined with a user study (Section 3.4). When a user selects an interest in the sunburst, a sidebar appears summarizing details about that inference. For example, Figure 3 shows an interest in "Home Improvement" may have been inferred on 1 site by 8 trackers.

Interest detail pages. Interest Detail pages feature bar charts to illustrate the relevant site and tracker information (e.g., frequently visited sites where an interest in "Home Improvement" could have been inferred, as well as trackers that could have inferred that interest). A table lists all user-visited pages that our algorithm labeled "Home Improvement," and a bar chart shows how often trackers inferred that interest over time.

Trackers tab. The *Trackers* tab (Figure 8 in the appendix) shows a summary of all observed tracking activity, as well as a bar chart showing how frequently each tracker was observed. As on the *Interests* tab, selecting a tracker displays a sidebar with summary statistics and basic information about the tracker. It also provides a link to the tracker's detail page.

Tracker detail pages. Each *Tracker Detail* page (Figure 10 in the appendix) gives a short description of the tracker and its privacy policy, drawn from Better [11]. A word cloud illustrates interests potentially inferred by the tracker, and a bar chart shows sites most frequently associated with the tracker. A table lists all associated page visits, and a bar chart quantifies tracking over time.

Activity & sites tabs. The *Activity* tab (Figure 9 in the appendix) displays a heatmap of browsing activity and associated tracking for each hour over the past week. The *Sites* tab summarizes all sites the user has visited, highlighting those with many or few trackers. Detail pages for each site visited highlight interests potentially

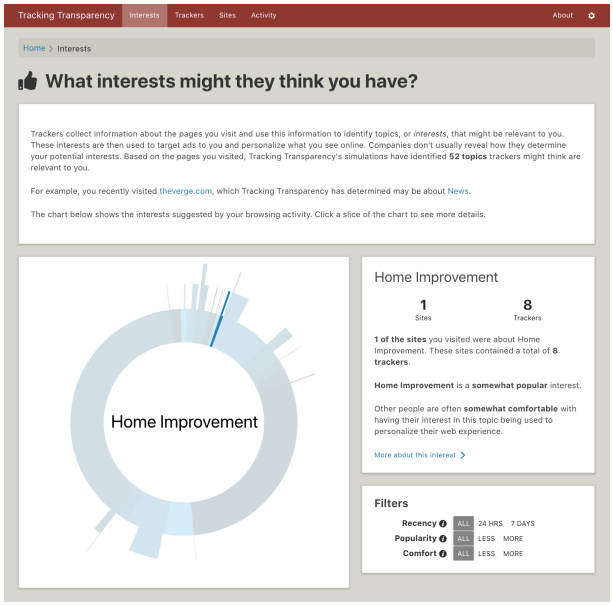


Figure 3: Tracking Transparency's Interests tab.

inferred and trackers encountered on the site, the pages visited within a site, and a bar chart of the visits over time.

Tracker network. As an alternative means of visualizing tracking, one variation tested in our field study included a forked copy of Mozilla's Lightbeam extension [62] styled to match our visual design (Figure 11 in the appendix). Lightbeam illustrates tracking via a network graph of connections between trackers and sites visited.

In-page overlay. Inspired by Ghostery [19], one variation in the user study showed an overlay in the lower-right corner. It listed the trackers observed on each page (Figure 12 in the appendix).

3.2 Interviews

To refine our extension's usability, we conducted 13 interviews with participants who had no prior experience with the extension. These IRB-approved interviews were conducted in the final months of development. Appendix A.1 contains the interview script.

Following library-intercept recruitment models, researchers approached individuals in the the lobby of a large public library to maximize participant diversity. Each interview took about 30 minutes, and compensation was a $10 Amazon gift card. To minimize social desirability bias, participants were told that the moderator had been independently hired to evaluate the tool. Using a provided laptop, participants explored a working prototype of Tracking Transparency populated with simulated data. We asked participants to think aloud as they explored and to suggest improvements [96].

Using these insights, we iteratively improved the extension until participants were no longer providing novel feedback. Because participants obtained the most value from graphics (e.g., bar charts), we converted lists and tables into charts as much as possible. We also used symbols to denote common concepts, added tooltips for clarification, added succinct page headers, and trimmed text. We also turned elements participants expected to be clickable into links.

3.3 Inferring Topics of Web Pages

One of our main goals was to study how users would react to seeing not just the trackers present on each page (the visualization used in most existing privacy tools), but instead the interests (e.g., yoga, classical music) companies could perhaps infer from browsing data. Because, to our knowledge, no current user transparency tools map specific browsing behaviors to potential inferences, we approximated this functionality in Tracking Transparency. We aimed to improve awareness of third-party tracking with meaningful approximations of inferences that advertisers could make about a user. Advertisers' actual inferencing methods are not public, but rather trade secrets. It is an open question whether they use complex multi-step inferences, simply generalize by topic, or do something else entirely. These processes cannot be fully understood or recreated without cooperation from companies. Instead, we strove to make user-intelligible, principled connections between browsing activity and potential inferences.

Other work has found substantial data sharing between trackers [10]. Our tool does not attempt to capture this sharing, instead focusing on user-centered communication. Our design emphasizes the aggregate data collected, rather than specific pages or interests. While our estimates are unlikely to perfectly capture what trackers learn from this data, we believe they provide users with helpful context through specific examples of possible inferences.

For privacy reasons, we wanted all logic to be client-side, with no information about the pages users visit transmitted externally. The topics we assigned needed to reflect one possible way trackers could reasonably assign ad-interest categories to users. As there can be substantial variation in a domain's content across pages, we wanted to assign topics to individual pages, not entire sites.

To produce plausible inferences while also respecting user privacy, we employed a pre-trained, client-side topic-modeling algorithm to determine potential ad-interest categories from each page's content. The extension uses observed tracker activity to link topics together and display inferences a tracker could have made.

3.3.1 Topic Modeling. When a user visits a web page, the extension extracts the visible text and HTML header metadata. These are preprocessed to remove stop words and non-English words, stemming each word in the resulting set [76]. If the page has at least 200 stemmed words, the algorithm uses the extracted text to assign a topic to the page. As the possible output labels, we use 1,932 hierarchical topics (e.g., Games→Board Games→Chess) taken verbatim from Google AdWords categories [38].

Using Wikipedia, we created a training corpus of the top 10 articles for each topic, with articles about specific entities (people, places, products) removed for generalizability. We preprocessed each article as above, but without a keyword threshold. To assign topics, we tested keyword matching and deep learning.

Keyword matching. We identified the 1,000 most relevant words for each hierarchical topic with three algorithms: term frequency-inverse document frequency (*TF-IDF*) [18], *TextRank* [61], and Latent Dirichlet Allocation (*LDA*) [12]. We compute a weighted matching score between the keywords W extracted from a page to the top 1,000 words for each topic T, associating lower weights to lower ranked words: $Score_T = \sum_{i=0}^{1,000} \sum_{K_i^T \in W} \frac{1}{i}$ where K_i^T is the i^{th} most relevant word for topic T. The highest-scoring topic is the output. We also experimented with *Word2Vec* embeddings [69] to assign scores via semantic-similarity-based matching. The topic with highest cosine similarity between embeddings of its top 1,000 words and W was the output.

Deep learning. We trained two LSTM neural network models [40] with each word's Word2Vec embedding. In both cases, we used one layer with 128 cells. In *LSTM*, we kept the 150,000 most frequent words from the *Wiki-corpus* (includes > 99% of all occurrences). In $LSTM_{small}$, to reduce storage size and computation, we kept only the 10,000 most frequent words (> 95% of all occurrences). The storage size of $LSTM_{small}$ (7.5 MB) is half that of *LSTM* (13.0 MB).

3.3.2 Evaluation. To select an inferencing algorithm for the extension, we ran two IRB-approved evaluation studies. We compared nine algorithms: TF-IDF, LDA, Textrank, their Word2Vec variants, two LSTM models, and random topic assignment as a control. We generated our test data using the top 10,000 sites from Alexa [2]. We loaded each domain and clicked two random links so the test set would contain a variety of types of pages. Pages that were not in English, contained under 200 keywords, or took more than 20 seconds to load were programmatically removed, resulting in a list of 5,980 pages. We then manually removed pages that contained only terms of service or privacy policies, contained adult content, or were mostly blank, resulting in a final test set of 2,700 pages.

Accuracy evaluation. We showed 187 MTurk workers a randomly selected page from our test set and the associated topic from a randomly selected inferencing algorithm. Participants rated on a five-point Likert scale whether the topic accurately described the page. Each participant rated 9 topic-page pairings, resulting in a total of 1,683 ratings. This IRB-approved study took about 30 minutes. Compensation was $5.00.

Accuracy ratings differed significantly across algorithms (Kruskal-Wallis, $H = 262.3$, $p < .001$). Dunn's multiple-comparison test with Bonferroni correction found that LDA and its Word2vec variant did not differ significantly from random assignment ($p = .390$ and $p = .330$, respectively). There were few significant differences among the remaining six algorithms, so we focused on the three for which participants' accuracy-agreement ratings were highest: $LSTM_{small}$ (54.0% agreement), LSTM (46.0%), and TF-IDF (45.5%).

Precision and performance evaluation. As AdWords categories are hierarchical, inferences in narrow subcategories present a potential tradeoff between accuracy and precision. To examine this tradeoff, we conducted an additional survey of 54 MTurk workers. Participants rated the accuracy and precision of topic-page pairings assigned using the three finalist algorithms. In addition, we randomly chose one of four display modes: *Top* category only, *TopTwo* categories only, the full hierarchy less one level (*CutOne*), and the *Full* unedited topic hierarchy. Each participant rated 12 pairings, resulting in a total of 648 ratings. The survey length and compensation were the same as previously.

$LSTM_{small}$*-Top*, *LSTM-Top*, and *TF-IDF-CutOne* had the highest participant agreement for both accuracy (66.7%, 61.1%, and 61.1%, respectively) and precision (55.6%, 48.1%, and 50.0%, respectively). As these results are similar, we next considered overhead. We implemented all three algorithms in our browser extension in Chrome

on a machine with a 2.9GHz Intel Core i7 quad-core processor and 16GB RAM. We instrumented the browser to benchmark assigning topics to our 2,700-page test set. Median runtimes were 23.4 seconds and 31.3 seconds for $LSTM_{small}$-*Top* and *LSTM-Top*, respectively, versus 39 ms for *TF-IDF-CutOne*. Because of its comparable accuracy with far less computation, we use *TF-IDF-CutOne* in the extension.

Our end goal for topic modeling was to approximate the kinds of inferences trackers might make and thus improve user understanding of the tracking ecosystem. While our final model sometimes assigns an incorrect topic, a model that is correct more often than not is still useful for our purposes of simulating an inferencing algorithm, informing users, and assessing their reactions to this transparency effort. This outcome also aligns with real-world tracking. Prior work has documented the poor accuracy of behavioral profiles built by online advertisers [74, 89], with one study finding only 27% of inferences were strongly relevant [9]. At least 40% of attributes sold by data brokers may be inaccurate [93]. Thus, our approximation algorithm appears comparable in its accuracy to the methods used by companies that invest substantial resources in fine-tuning tracker data collection and inference models.

3.4 Sensitivity of Interest Categories

On the Interests tab, users can filter the chart to highlight topics labeled as more or less sensitive. We quantified sensitivity using an IRB-approved MTurk study in which we asked participants about their comfort with a specific topic being inferred and used for personalization. We based our study on Dolin et al. [23], abbreviating their script and expanding the scope of the study to cover all 1,124 categories used in Google AdWords (excluding world localities).

We obtained 583 responses, each addressing 10 randomly selected categories, from 470 crowdworkers (we permitted participants to take the survey multiple times). Participation took about 15 minutes. We compensated participants $3.00.

Similar to Dolin et al., we found a spectrum of comfort with targeting based on different interest categories. A small number of topics were strongly sensitive or non-sensitive, but most were somewhere in the middle. From this data, we generated a list of 1,124 topics ranked by mean agreement that "I would be comfortable with a company personalizing my web experience based on an inference about my level of interest in [topic]" on a seven-point Likert scale, which formed the basis of sensitivity filtering in the Interests tab.

4 FIELD STUDY METHODOLOGY

We conducted a field study to evaluate how transparency in the form of the Tracking Transparency prototype impacts users' knowledge and attitudes about tracking and inferencing. Participants were randomly assigned to install one of six variants, each with different UI components. At installation, participants completed a pre-usage survey. After one week of normal browsing, we prompted them to explore the extension and complete a post-usage survey.

All participants were recruited through Amazon's Mechanical Turk (MTurk). Participants needed to be located in the U.S., be at least 18 years old, and have a 95% HIT approval rating. Because the extension was built for Google Chrome and Mozilla Firefox, we required participants to regularly use at least one of them. Our IRB approved the study, and the extension itself was reviewed by Google and Mozilla following their standard procedures.

Table 1: A summary of the conditions' key characteristics.

Control:Static	Contains only static text explaining targeted advertising and privacy.
Control:Browsing Only	Provides the dashboard interface with info about browsing history, but no data about tracking.
Current:Trackers	Simulates Ghostery and similar extensions. Contains a list of trackers in the toolbar popup, but no access to the dashboard interface.
Current:Connections	Provides a visually restyled version of Mozilla Lightbeam with no other personalized data.
Longitudinal:Trackers	Contains most interface components. Provides longitudinal info about trackers and browsing, but does not show potential interests inferred.
Longitudinal:Interests	The full interface and data described in Section 3, including potential interests inferred.

4.1 Study Conditions

To gauge the impact of our key transparency features in comparison to state-of-the-art privacy tools' approaches, we randomly assigned participants to one of six versions (*conditions*) of Tracking Transparency. For consistency and comparability, all conditions had the same visual design, branding, text, and UI elements other than the differences being tested, as described below and in Table 1.

Two conditions displayed longitudinal data. The **Longitudinal:Interests** condition is Tracking Transparency as described in Section 3, including longitudinal information about tracking alongside guesses about what interests could have been inferred. To test the impact of displaying these inferencing guesses, **Longitudinal:Trackers** was identical to Longitudinal:Interests except without any inferencing guesses.

Two other conditions replicated UI elements of existing privacy tools for comparison. Similar to tools like Ghostery, Disconnect, and Privacy Badger, **Current:Trackers** showed the trackers on the current page in the toolbar popup, as well as an in-page overlay with the number of trackers. Whereas Longitudinal:Interests was longitudinal, Current:Trackers only provided information about tracking on the current page. **Current:Connections**, based on Mozilla Lightbeam, presented a graph visualization of the connections between websites and trackers, but did not provide Longitudinal:Interests's detailed longitudinal information.

Two final conditions were controls. **Control:Static** provided static text explaining targeted advertising. Comparisons of other conditions to Control:Static thus tested the impact of visualizing personalized data, whether longitudinal or not. To test the impact of focusing on tracking and privacy, **Control:Browsing Only** visualized a user's browsing history without referencing tracking, trackers, or inferences. Appendix A.4 gives additional screenshots.

4.2 Pre-Usage Survey

Participants were asked to install the extension in Chrome or Firefox. Following installation, but before interacting with the extension, participants were directed to the pre-usage survey. We asked about participants' demographics, browsing behaviors, use of relevant browser extensions, and experiences with online shopping and ads. To understand how participants' knowledge and attitude changed after using the extension, we asked a series of questions

in the pre-usage survey that were repeated verbatim a week later in the post-usage survey. These items included seven statements concerning *attitudes* about targeted ads, as well as *knowledge* statements about 15 types of data and 3 broad methods that might possibly be used for targeting. Participants rated their agreement with the former on 7-point Likert scales, and the likelihood of the latter on 7-point likelihood scales ("very unlikely" to "very likely"). The repeated section also included the awareness and collection sub-scales of the Internet Users' Information Privacy Concerns (*IUIPC*) scale [55], as well as questions that asked participants to *quantify tracking* (e.g., the number of trackers they encounter).

Upon completion of this survey, designed to take 15 minutes, we compensated participants $3.00 and reminded them to keep the extension installed for 7 days. On days 4–6, the extension sent browser notifications to encourage participants to explore it.

4.3 Post-Usage Survey

A week after installation, we sent participants a link to the post-usage survey via MTurk. We asked them to "spend a few minutes exploring the extension before beginning the survey," asking two questions about what they saw to encourage them to do so. We then asked four open-ended questions about the information in the extension: "new information," "information you already knew," "surprising information," and what questions they had. We also asked participants to respond to six potential changes in *behavioral intention* (e.g., "Compared to before you used the extension, how likely are you to use a browser's private browsing mode now?") on 7-point scales from "much more likely" to "much less likely." For four potential tradeoffs (e.g., an internet that is free but has tracking versus an internet that costs money but does not have tracking), participants rated which they would choose. As mentioned above, we repeated the batteries of questions concerning attitudes and knowledge, the IUIPC, and quantification of tracking. We also asked the standard System Usability Scale (SUS) [14].

Upon completion of the post-usage survey, designed to take 20 minutes, we compensated participants with a $7.00 bonus payment on MTurk. This larger compensation encompassed both the week of keeping the tool installed and completion of the post-usage survey. Both survey instruments, which are included in Appendices A.2–A.3, were refined through pilot testing and cognitive interviews.

4.4 Participant Privacy

To protect participant privacy, the extension did not report any personally identifiable information. On the participant's own computer, the extension kept a full database of all page visits, trackers encountered, and interest categories, which was used to power the extension's visualizations. This data was stored locally in the browser extension's sandboxed storage and was not accessible to other extensions or web pages. To enable analysis of aggregate data across all users while preserving participant privacy, we collected an anonymized version of the database, with all URLs and page titles hashed with a participant-specific salt generated on the participant's computer and never sent to the researchers. We also collected clickstream data for activity in the dashboard. All data was associated with an anonymous identifier generated by the extension and never associated with the participant's Mechanical Turk ID. The extension did not operate in private browsing mode. Participants were informed about the data collection through both a consent form and a privacy policy. The inclusion of longitudinal visualizations like those in Tracking Transparency in tools intended for wide distribution will require careful communication to users about the potential for privacy leaks on shared devices. To enable longitudinal visualizations, such tools must store a detailed history of a user's web browsing. These extensions should clear their own data when users clear their browser's history, and require additional design considerations around shared devices.

4.5 Analysis Methods and Metrics

For quantitative data, we conducted hypothesis tests with $\alpha = .05$, choosing the test based on the type of data. Questions asked only post-usage, such as behavioral intentions after using Tracking Transparency, elicited responses on scales (e.g., Likert scales). We analyzed this ordinal data with the Kruskal-Wallis H test (KW) for omnibus comparisons. Using the Mann-Whitney U Test (U), KW's analogue for two groups, we ran seven planned contrasts between condition pairs: comparing Longitudinal:Interests to each of the other five conditions, and comparing both Current:Connections and Current:Trackers to Control:Static. To minimize Type II error, we performed Holm correction within each set of contrasts and across each set of omnibus tests. We analyzed SUS data and participants' estimates of tracking similarly, though treating data as continuous.

Many questions asked both pre- and post-usage also elicited responses on scales. To understand how responses in this repeated-measures design changed over time both regardless of condition and by condition, we built repeated-measures ordinal logistic regression models. Responses were the DV for each, and the time period (pre-, post-usage), condition, and interaction between the two were the IVs. We performed Holm correction within each set of questions. Similarly, for the two IUIPC sub-scales, we summed responses across scale items and analyzed these (continuous) sums with a repeated-measures ANOVA.

We analyzed free-response data through qualitative open coding. One member of the research team read responses and created a codebook with thematic codes, iteratively updating as necessary. Each survey question had its own set of 7 or 8 unique, but not mutually exclusive, codes. A second researcher independently coded the full set of data. Inter-coder reliability, measured with Cohen's κ, ranged from 0.76 to 0.82 per question, with a median of 0.80. This level of agreement is "substantial" [47] or "excellent" [33].

4.6 Limitations

To limit self-selection by especially privacy-interested participants, we advertised our study as "evaluating a web browser visualization tool" without mention of privacy, though we did mention tracking as part of the procedures. However, there may also have been contradictory self-selection in which privacy-conscious people may have been unwilling to install an unknown extension and therefore decline participation. Further, MTurk participants are generally younger, more technical, and more privacy-sensitive than the the overall U.S. population [42]. This is evident in our results, which demonstrate high initial levels of knowledge about tracking. We believe these limitations are acceptable, as our tool targets people

with an interest in learning more about online tracking and privacy. Further, while our participants displayed high initial knowledge about tracking and privacy, less-aware populations may stand to benefit even more from visualizations like ours.

As in any online study, participants may not answer carefully, and some may try to participate multiple times. We follow best practices [68], using high-reputation workers and forbidding multiple submissions from one MTurk account. In addition to participants' high initial privacy literacy, the phrasing of our questions is another possible cause of the ceiling effect in some of our results.

We were only able to survey Chrome and Firefox users. These are the two most popular desktop browsers [85], so we consider this reasonably representative. The extension only attempts to detect third-party tracking in desktop browsing; the mobile tracking/ad ecosystem is significantly different. The extension also does not account for cross-browser or cross-device tracking [13, 15, 104].

Our simulation of inferences that could be made based on a user's browsing history is only an approximation of what advertising networks may actually be doing. While the simulated nature of these inferences is a clear limitation of our protocol, advertising networks do not provide consumers or researchers access to actual data mapping precise browsing activities to specific inferences. While imperfect, our methods are one of the only ways for us to evaluate users' reactions to inference-level information.

Further, detecting trackers by using web requests may result in false positives for tracking-unrelated requests, but it captures many types of tracking including cookie storage and access, as well as fingerprinting. Other blocking tools that a user has installed may block requests to trackers and prevent our extension from detecting them, but this would accurately reflect the extent to which the user is actually tracked. We detected whether participants had blocking tools installed and found that there was a slight decrease in the number of trackers detected for those users. Finally, our qualitative results indicated that some participants in the Control:Static condition may have realized they were in a control condition. However, a control was necessary to facilitate comparisons across conditions.

Given ongoing escalations between ad-blockers and advertisers [41], plus the potential of fingerprinting browser extensions [37, 83], it is possible sites could identify and retaliate against future tools like Tracking Transparency. Sites could manipulate the text parsed by the topic modeling algorithm or otherwise try to avoid classification. As we used Tracking Transparency with a small population during a short experiment, it seems unlikely we provoked such retaliation. Any widely deployed tool employing a similar mechanism would need to defend against adversarial scenarios.

5 FIELD STUDY RESULTS

In this section, we present results from our field study assessing how the Tracking Transparency interface affects user attitudes. We begin by characterizing our participants and their usage of the extension (Section 5.1). We then present qualitative analysis of participants' reactions to the information Tracking Transparency presented (Section 5.2). Participants were surprised by the extent of tracking. They newly learned how trackers infer their interests.

Section 5.3 describes how using the extension increased participants' intentions to take privacy-protective actions. Conditions that displayed more information saw larger increases in intentions. We then briefly discuss how the extension did not significantly impact participants' knowledge of targeted advertising (Section 5.4), which was mostly correct to begin with, or their broad attitudes about the practice (Section 5.5). Table 4 in the appendix gives the full statistical results. Finally, Section 5.6 describes how longitudinal information helped participants more accurately quantify tracking.

5.1 Participants and Usage

Demographics. A total of 456 participants completed the study. We exclude the 6.8% of participants who visited fewer than 100 web pages, leaving 425 participants. As conditions were randomly assigned, the distribution of participants varied: 71 in Control:Static, 82 in Control:Browsing Only, 63 in Current:Trackers, 70 in Current:Connections, 66 in Longitudinal:Trackers, and 73 in Longitudinal:Interests. In total, 52.2% participants identified as female, 46.8% as male, and 1.0% as non-binary. Most (72.2%) were 25–44 years old; 7.8% were under 25, while 20.0% were 45+. Most had bachelor's degrees (40.5%) or some college (35.3%), while fewer had graduate degrees (9.6%) or high school diplomas (14.6%). Additionally, 23.1% reported holding a degree or job related to IT or CS.

Browser usage. Most (89.9%) participants installed the Tracking Transparency extension on Google Chrome, as opposed to Firefox (10.1%). Participants estimated a median of 80% of their browsing was on the device and browser they installed the extension on.

Just under half (48.5%) of participants reported current use of an ad- or tracker-blocking tool, and an additional 18.6% reported having used such a tool in the past. However, only 8.5% reported current use of a dedicated tracker-blocking tool (Ghostery, Privacy Badger, Firefox Tracking Protection, and Disconnect, in order of frequency). Our extension checked for the presence of other blocking tools by querying whether certain popular extensions were installed, finding that 39% of participants had such a tool. A minority of participants reported having viewed ad preferences pages on Facebook (37.4%) and Google (28.9%), and only 7.5% recognized the AdChoices icon that indicates targeted ads [51].

Over the week-long study, our 425 participants visited a total of 1,068,302 web pages and encountered 533 different trackers. The top trackers observed were Amazon (present on 64.2% of pages), Google (47.0%), Facebook (10.1%), comScore (6.4%), and Microsoft (4.5%). Our extension detected an average of 2.58 trackers per page for users with no other blocking tools installed, and an average of 2.15 trackers for those with a blocking tool installed. Most of the 533 trackers were only observed on a small fraction of pages visited, demonstrating a long-tailed distribution consistent with large-scale measurements by Engelhardt et al.'s OpenWPM tool [27].

Tracking Transparency's inferencing approximation layer (Section 3.3) assigned a total of 230 unique interest categories across participants. The median participant was assigned 59 interest categories (μ = 58.6, σ = 16.8) that the extension guessed might be inferrable from the participant's page visits. "Travel," "News," "Shopping," "Books & Literature," and "Online Communities" were the five most frequent categories, and all 425 participants had at least one page assigned the "Travel" topic. There was a long tail of topics assigned, including relatively obscure and infrequently assigned categories like "Medical Literature & Resources." In total, 58 of the

Table 2: The percentage of participants per condition who organically mentioned different classes of information when describing what was surprising, what was new to them, and what they already knew.

			% of participants mentioning					
	Code	Representative quote	Control:Static	Control:Browsing Only	Current:Connections	Current:Trackers	Longitudinal:Trackers	Longitudinal:Interests
Surprising	Number of trackers	"I would have to say the sheer number of trackers found and how many different pages I actually visited I could not believe it was that many."	4	9	39	49	42	42
	Interests are inferred	"I'm surprised by the depth of the information, such as topics, that are gathered from multiple sites, even my email server."	13	1	4	2	2	22
	Own browsing habits	"I really didn't think that I surfed the web that much."	1	49	3	0	11	14
	Detail of data	"Just how many and how well they track the sites you visit."	1	7	0	3	5	11
	Frequency of tracking	"Just exactly how much of the time that Amazon was tracking me. I mean talk about stalking. I knew that they were suggesting things from my google searches and such but their trackers seem to be on the majority of webpages out there."	0	5	1	11	21	7
	Sites without tracking	"I was surprised at times when nobody was tracking when I expected someone to be."	0	1	1	6	8	7
	Unexpected third parties	"That the information was being sold or shared with so many third party website that I haven't heard of before. I never visited them but they have my information anyways."	0	0	13	14	8	5
	Tracking occurs	"How all of my online activity is tracked and all connected in a virtual world where my fingerprint is all over the place even if I am unaware."	3	1	1	5	5	1
	Connections	"I just didn't know how enmeshed the companies were with each other."	0	1	16	0	0	0
	Nothing	"There was nothing that was very surprising, but it was still interesting to see it all."	73	27	24	19	15	10
New information	Number of trackers	"A lot more services track me than I knew about."	3	7	46	62	52	56
	Frequency of tracking	"I learned that most of the sites were tracking what I was doing."	4	7	17	43	40	27
	How interests inferred	"I learned what information sites are pulling when I'm visiting them."	15	9	7	0	6	21
	Own browsing habits	"I didn't realize how many site/pages I use throughout the day."	1	54	3	0	12	15
	Tracking used to target	"I learned more about how the ads I see when browsing magically appear to be personalized."	14	2	6	9	6	8
	Tracking methods	"I did not know the manner in which trackers tracked my interest."	8	12	6	0	8	7
	Connections	"I learned that there are far more connections between first and third party sites I visit."	0	2	11	3	2	0
	Nothing	"Nothing really. I knew that some sites would track me."	62	21	26	14	8	8
Already knew	Tracking occurs	"I was aware of the presence of trackers…but not to the level that the extension confirmed."	46	33	59	71	74	59
	Tracking used for ads	"I knew some ads were generated based on my browsing and search results."	44	16	20	11	23	21
	Frequency of tracking	"Sites are connected. Google is often at the center of that. Sites are always tracking you."	1	3	11	32	30	19
	Interests are inferred	"Companies would track my activity to pool my interests and then use them to target me with ads going forward. I knew Facebook did this frequently."	3	0	1	0	0	11
	Own browsing habits	"I visit a lot of pages. Most likely a lot of them track my activity."	4	48	9	0	6	11
	Tracking used to target	"I knew that companies were able to see some of the information i search for to input dedicated ads but i did not realize the extent of it."	18	2	3	3	0	4
	Tracking methods	"I knew about cookies, pixels, and browser finger printing."	7	9	11	3	3	4

230 topics (25.2%) were inferred for 5 or fewer of the 425 participants, suggesting Tracking Transparency captures personalized interest profiles for users alongside some common topics.

Our telemetry data indicates participants clicked different parts of the extension's UI a median of 19 times (μ = 24.5, σ = 21.1). Users opened the toolbar popup a median of 5 times (μ = 6.7, σ = 5.3) and the dashboard a median of 3 times (μ = 3.1, σ = 2.5). Most interactions occurred during the post-usage survey.

5.2 General Reactions

System usability. Participants rated Tracking Transparency as highly usable on the ten-item System Usability Scale (SUS), scored from 0–100. The median SUS score for Longitudinal:Interests was 82.5 (μ = 81.5, σ = 13.2). SUS scores differed by condition. The Longitudinal:Interests condition's SUS score was significantly higher than the Current:Connections condition's median SUS score of 72.5 and the Control:Static condition's median SUS score of 70.0. Scores above 68 are typically considered above average [14].

Open-ended questions. Through open-ended questions, participants explained what information the extension presented that was surprising, new to them, or that echoed what they already knew. To minimize coding biases, we mixed together responses from all conditions during coding. We observed significant differences by condition in the codes assigned for information that was surprising ($\chi^2(35) = 374.0$, $p < .001$), new ($\chi^2(35) = 341.5$, $p < .001$), or already known ($\chi^2(35) = 221.4$, $p < .001$). Table 2 summarizes participants' responses to these open-ended questions.

Surprising. Over 90% of participants were surprised by something presented in Longitudinal:Interests, the extension's full version, while 73% of participants who saw Control:Static, the least informative variant, reported they did not find anything surprising.

Participants were surprised by the amount of tracking displayed and the number of trackers. This was echoed by almost half of participants in conditions where participants saw longitudinal information, compared to less than 10% in the control conditions. The extension also revealed surprising information about the interests inferred and the detail of data collected by trackers, according to 22% of participants in Longitudinal:Interests. Interestingly, 13% of participants in Control:Static were surprised by how interests are inferred, potentially because this information was made more salient in this otherwise sparse version of the extension.

Further, and as expected, the condition influenced which classes of information were surprising. For example, 49% of participants in the Control:Browsing Only condition, which focused on browsing

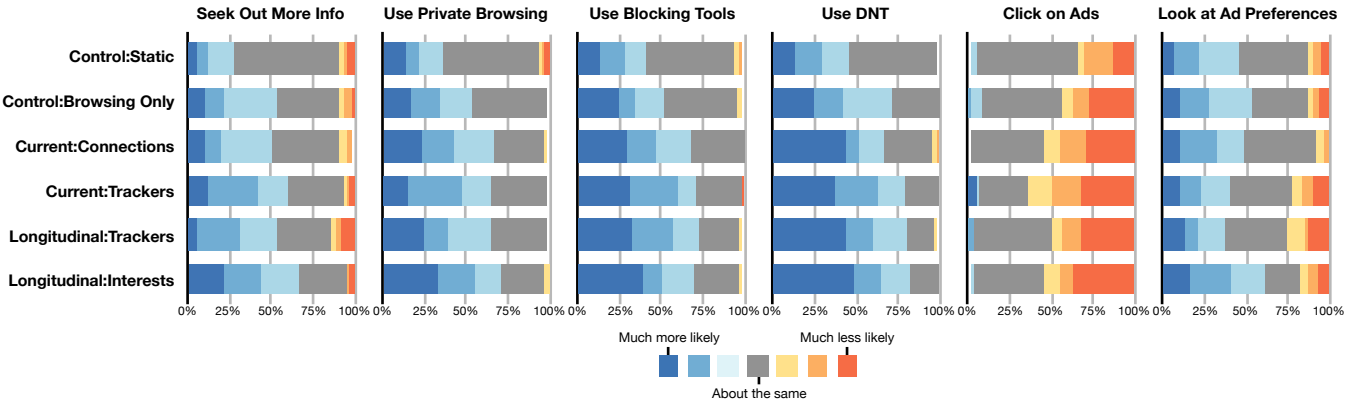


Figure 4: Participants' beliefs about how their intentions to take specific actions changed after using Tracking Transparency. Differences were significant for the four leftmost intentions, and marginally significant for the two rightmost.

information, were surprised by their browsing habits. Similarly, 16% of Current:Connections participants were surprised by connections between trackers.

New information. Tracking Transparency increased participants' awareness of the magnitude of tracking that occurred in their daily browsing. 56% of Longitudinal:Interests participants reported learning about the number of trackers that were tracking their browsing activity. Similarly, for 27% of Longitudinal:Interests participants, the pervasiveness of tracking across sites was also new information. For example, P65 remarked, "The thing I learned is I was being tracked a lot more than I thought originally. I thought some sites weren't big in the tracking game, boy was I wrong!"

The extension provided transparency into the use of inferred interests in tracking for 21% of Longitudinal:Interests participants: "I learned that companies will infer and guess what items I may like" (P232). This was also new for P161: "It shows my top interest is shopping, which i didn't figure that to be true, since i usually hate shopping. but it made me realize that i do a lot of shopping online now. that's new to me. It's also new that I have 75 potential interests." In comparison, few participants in the other conditions reported learning new information about inferences.

Other classes of new information reported reflected information shown only in their condition: 54% of Control:Browsing Only participants learned about their browsing habits, and 11% of Current:Connections participants about tracker connections.

Already knew. Many participants had some familiarity with tracking before using the extension. Across all conditions, 56% said they already knew tracking occurred, and another 16% knew it was prevalent: "I expected a lot of tracking from google, facebook, amazon and other large sites" (P11). For 22% and 5% of participants, it was not new information that tracking was used for ads or targeting, respectively. A minority of participants — 11% in the Longitudinal:Interests condition — already knew that interests can be inferred from browsing. Finally, 6% of participants already knew about tracking methods, frequently mentioning cookies: "I knew about tracking cookies, pixels, and fingerprinting" (P215).

Additional questions. We also asked participants if they had any additional questions about what they saw in the extension. The most common question was how to gain more control of their information or block trackers (asked by 64 participants across conditions). 35 participants had questions about how tracking worked, and another 34 wanted to know more about the information trackers get about them. 19 participants wanted to know more about what trackers do with their information, and 10 even wondered whether tracking was safe or legal. P173 asked, "I would like to know if the sites tracking me are safe or not. Like I can see that Yahoo sometimes tracks, and I will assume that's safe enough, but the comScore one, I'm thinking it's sketchy." P105 wondered, "is this legal? is it ethical for these companies to invade my privacy?"

5.3 Impact on Intended Behaviors

In the post-usage survey, we asked participants to rate how their likelihood to take six different actions changed after using the extension. Figure 4 presents the results. Across conditions, participants overwhelmingly reported increased intention to take privacy-protective actions like using tracker-blocking tools, which may indicate a social desirability bias. Nonetheless, intentions differed significantly by condition for most of these six actions. Conditions richer in information (e.g., Longitudinal:Interests) generally increased these intentions more. These differences show how richer displays contributed to meaningful increases in awareness and interest in behavioral changes. The comparisons between conditions mentioned in this section are all statistically significant. Table 4 in the appendix provides the full statistical results. For four of the six actions, the omnibus test across all conditions was statistically significant. For the remaining two, it was marginally significant.

Seeking out more information. 65.8% of Longitudinal:Interests participants reported being between a little more likely and much more likely "to seek out more information about online advertising," and these intentions varied by condition. The Longitudinal:Interests condition displayed the most information about longitudinal tracking. Thus, participants in this condition were more likely to report an intention to seek out more information than those in Current:Connections, which only showed snapshots of tracker connections in their browsing, as well as those in Control:Static, which only showed static descriptions.

Private browsing. Among the Longitudinal:Interests participants, 71.2% reported being at least a little more likely "to use a browser's private browsing mode" after using Tracking Transparency. Participants who saw a representation of current tools were more likely to report an intention to use private browsing than those who only saw static descriptions of third-party tracking. Additionally, participants who saw the fully featured Longitudinal:Interests were also more likely to report intending to use private browsing than those who saw Control:Static or Control:Browsing Only.

Blocking tools and Do Not Track. Participants' reported likelihood "to use browser extensions that block ads and/or online tracking" varied by condition. Reported likelihood to enable their browser's Do Not Track (DNT) setting also varied by condition. Participants who saw our fully featured Longitudinal:Interests or a representation of a current tool were more likely to express an intention to use blocking tools and enable DNT than those who saw static descriptions. While users may not have understood the DNT setting does very little, the responses indicate that participants in more fully featured conditions expressed stronger desire to stop tracking.

Other results. After correcting for multiple testing, differences by condition in participants' responses to "Compared to before you used the extension, how likely are you to click on ads now?" were only marginally significant (KW $\chi^2(5) = 12.663$, $p = .054$). We also asked about changes in likelihood to look at a page provided by advertising companies "to show you what topics they guessed you are interested in," again finding that omnibus differences were again marginally significant (KW $\chi^2(5) = 10.528$, $p = .062$) and were not significant for any pairwise comparisons.

5.4 Users' Knowledge and Attitudes

Knowledge of targeting. To understand how the extension impacted participants' knowledge of targeted advertising, participants rated the likelihood that fifteen types of information and three broad practices are used to target ads. Participants' responses were generally accurate. They did not change significantly between the pre-usage and post-usage surveys, nor did they vary by condition. This may be attributable to a ceiling effect. That is, participants were mostly correct even before the study, capping potential increases.

Participants correctly expected that companies likely targeted ads to them based on their current and past browsing, as well as guesses about their demographics and interests. Between the pre- and post-usage surveys, we observed a significant increase across all conditions in participants' expectation that companies target ads based on guesses about their interests ($\beta = -1.038$, $p < 0.001$).

Impact on attitudes. We asked participants to respond to seven statements measuring their attitudes of tracking and ad targeting. Broadly, participants agreed that transparency is valuable and tracking can be creepy, but expressed divergent and complex opinions regarding the usefulness of relevant ads, inferencing, and third-party tracking (Figure 5). In particular, it may appear as a contradiction that a majority of participants agreed it was "creepy for companies to track websites I visit to show relevant ads," but more than a third of participants also agreed they would be "comfortable with companies guessing my interests based on websites I visit." However, this actually reveals subtle differences between comfort with

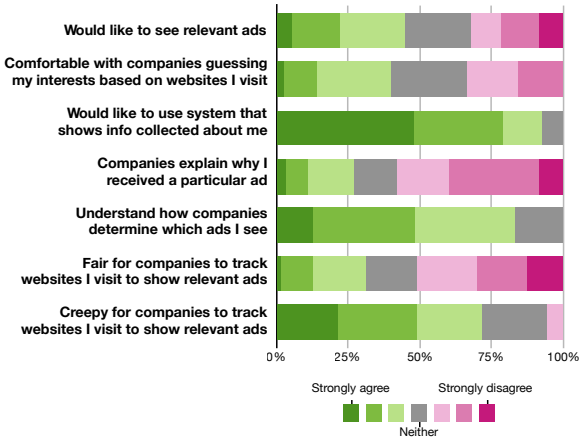


Figure 5: Participants' attitudes about online tracking after using Tracking Transparency for one week.

inferring interests compared to perceiving creepiness in tracking for advertising purposes. Nevertheless, Tracking Transparency did not appear to impact these broader attitudes overall. For six of the seven questions, responses did not significantly change after using the extension, and we did not observe significant differences across conditions in our repeated-measures regression models. Except as noted, we report the distribution of post-usage responses.

Usefulness of ads & tracking. Participants were split regarding the usefulness of personalization. 44.7% agreed they "would like to see ads that are relevant to my interests, as opposed to generic ad," while 32.0% disagreed. Furthermore, whereas 40.0% agreed "I would be comfortable with online advertising companies guessing my interests based on which websites I visit," 44.9% disagreed. While participants overwhelmingly (71.3%) considered it creepy for "for advertising companies to track which websites I visit in order to show me ads that are relevant to my interests," they were split regarding whether the tracking is *fair*, with 30.6% agreeing and 52.9% disagreeing. These results are in line with prior work revealing that some users find personalization useful, but many are uncomfortable with the methods of web tracking [90].

Understanding of tracking. Participants' agreement that they understand tracking increased significantly from 70.1% pre-usage to 82.8% post-usage ($\beta = -0.967$, $p = 0.002$). Further, 48.5% of participants agreed "I would like to use a system that shows me what information has been collected," whereas 64.9% disagreed that ad companies adequately explain why they receive particular ads.

Privacy attitudes. We also studied how Tracking Transparency may have impacted participants' broad privacy attitudes, not observing any effect. Both pre- and post-usage, participants completed the Awareness and Collection subscales of the IUIPC privacy index. They responded to each item on scales from strongly disagree (coded as -3) to strongly agree (coded as 3). Even in the pre-usage survey, participants expressed high privacy concern. The median participant's response, averaged across items, on the Awareness sub-scale was 2.7 ($\mu = 2.3$, $\sigma = 0.8$), between "agree" and "strongly

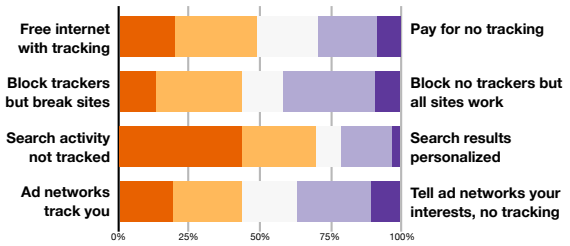


Figure 6: Choices participants preferred in four privacy tradeoffs, which did not vary by condition. Darker colors indicate stronger preferences.

agree." The median participant's average response to items on the Collection sub-scale was 2.0 (μ = 1.7, σ = 1.2).

5.5 Perceptions of Tradeoffs

Targeted advertising, as well as efforts to stop it, manifests as a series of tradeoffs for users. To understand how our extension might have influenced these considerations, we asked participants to pick between four pairs of tradeoffs, choosing whether they would "definitely" or "probably" prefer one of the pair, or whether they were unsure. Making explicit the privacy literature's observations that users have sometimes conflicting or paradoxical attitudes, participants expressed divergent preferences about balancing these tradeoffs (Figure 6). These preferences did not vary by condition.

Whereas 48.7% of participants reported that they preferred the internet be free and have tracking, 29.2% reported preferring to pay for an internet with no tracking. This supports previous work that found some people are willing to pay a premium for privacy, especially if privacy information is made transparent [26, 88].

Similarly, 69.4% preferred that search results not be tracked (and thus not be personalized), while 21.6% preferred the opposite. Tools that block tracking can sometimes "break" web pages. Among participants, 43.5% wanted to block tracking even if it would sometimes break web pages, yet 41.9% preferred that web pages always work.

Currently, tracking is necessary for targeting ads because advertisers otherwise would not know users' interests. A radical alternative to this model would be for users to explicitly tell advertisers their interests. While 43.5% of participants preferred the current system of tracking to learn users' interests, 37.2% would prefer to tell companies their interests and not be tracked.

5.6 Estimates of Browsing and Tracking

In both surveys, participants were asked to numerically estimate how much they browsed the web and how many trackers they encountered. Before using the extension, participants consistently underestimated both, with no variance by condition. However, the extension's longitudinal conditions helped participants better quantify their web browsing (see median estimates by condition, pre- and post-usage, in Table 3 in the appendix). Pre-usage, the median participant per condition estimated visiting 100–200 pages across 22.5–35 domains each week. According to our telemetry data, the median participant actually visited 1,682 web pages on 68 unique domains over the week of the study. Prior to using the extension, the median participant in each condition estimated that they encountered between 10–20 trackers each week. Per our telemetry data, the median participant encountered 148 unique trackers over the week. In conditions that made tracking more transparent, participants' post-usage estimates sharply increased, and these variations across conditions were significant.

The extension's more fully featured conditions helped participants improve their accuracy. Post-usage, all estimates varied by condition, with participants who saw longitudinal data unsurprisingly more closely aligned with the telemetry data. The close alignment between post-usage estimates and our telemetry data for longitudinal conditions is unsurprising because the extension showed them these numbers. More surprising are participants' consistent underestimates of both the number of trackers and the amount they browsed absent this data.

6 DISCUSSION

In visualizing third-party web tracking and the inferences that could be made, we aimed to facilitate conversations about the prevalence of third-party tracking. Advertisers' obscure dashboards and technical knowledge previously formed barriers to retrieving transparent information about online tracking in one's own browsing. Tracking Transparency allows researchers to understand how supplying more information to non-technical users can affect their reactions. Despite previous work that would predict users to be unmotivated [97], our field study indicated that users *are* interested in learning about how they are profiled from their browsing.

In the realm of online privacy, knowledge is power. A better understanding of how online privacy is affected enables better decision making. This parallels security psychology research, which posits that accurate risk perception enables better security decision making [25, 35, 97]. The Tracking Transparency prototype is a step in this direction, as participants who used the fully featured tool were significantly better at quantifying online tracking than those who used a controlled representation of current user interfaces. As P290 explained, "I learned that Google is watching wherever I go and my local news page has more trackers than anyone, which was quite surprising. I knew my ad-blocker stopped a lot of ads there but I had no idea they were still tracking me." Future privacy tools should empower users to learn how such technologies impact them so they can be informed in discussions about tracking and understand the use cases for privacy-preserving measures.

Finally, there is significant room for additional tools and policies to support online privacy. Related work has explored users' contextual preferences regarding web tracking and subsequent technical tools [59, 60]. In this light, future work should explore providing users not only with transparency, but also with greater control over tracking. Additionally, our results highlight the need for companies to provide more transparency about *how* they infer interests and use them for targeting. Some recent initiatives begin to partially support this goal [73]. There has been increasing media attention about the misuse of tracked data, especially regarding discriminatory contexts and political purposes. The Tracking Transparency interface takes an important first step in motivating users to consider behavioral changes, learning, and public policy demands.

7 CONCLUSION

In this paper, we presented Tracking Transparency, a browser extension we created to communicate more information about online tracking to users and to support research into the impact of transparency. Even before using our tool, participants were often aware of the existence of online tracking. However, when confronted with detailed descriptions of tracking in their own browsing, they were often surprised by tracking's extent and prevalence. Further, participants who saw detailed information about potential inferences reported greater intentions to take privacy-protective actions.

Our field study demonstrated the importance of providing detailed, longitudinal tracker data to users. The Tracking Transparency prototype approximates information that advertising companies have little incentive to provide and is otherwise onerous for users to obtain. After completing our study, a number of our participants expressed that they wanted to keep Tracking Transparency installed. This suggests our interface addresses a much-needed intermediate step in the privacy-consciousness spectrum: educating the public about how their own browsing data is collected and used without their explicit permission. Without greater public awareness about the scope and practices of online tracking, advancing privacy-friendly policies or regulatory options is unlikely.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Data Transparency Lab and Mozilla, as well as from a UMIACS contract under the partnership between the University of Maryland and DoD. We thank Lorrie Cranor, Aaron Goldman, Oliver Hahn, Dimitri Vasilkov, Mark Cohen, Juliette Hainline, and Andrew McNutt for their assistance.

REFERENCES

[1] Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. 2014. The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. In *Proc. CCS*.
[2] Alexa. Fetched on October 5, 2017. Top 1 Million Sites. http://s3.amazonaws.com/alexa-static/top-1m.csv.zip.
[3] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Cranor, and Yuvraj Agarwal. 2015. Your Location has been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging. In *Proc. CHI*.
[4] Athanasios Andreou, Giridhari Venkatadri, Oana Goga, Krishna P. Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations. In *Proc. NDSS*.
[5] Julia Angwin and Terry Parris. 2016. Facebook Lets Advertisers Exclude Users by Race. https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race.
[6] Rebecca Balebako, Jaeyeon Jung, Wei Lu, Lorrie Faith Cranor, and Carolyn Nguyen. 2013. "Little Brothers Watching You:" Raising Awareness of Data Leaks on Smartphones. In *Proc. SOUPS*.
[7] Rebecca Balebako, Pedro Leon, Richard Shay, Blase Ur, Yang Wang, and L. Cranor. 2012. Measuring the Effectiveness of Privacy Tools for Limiting Behavioral Advertising. In *Proc. W2SP*.
[8] Muhammad Ahmad Bashir, Sajjad Arshad, and William Robertson. 2016. Tracing Information Flows Between Ad Exchanges Using Retargeted Ads. In *Proc. USENIX Security*.
[9] Muhammad Ahmad Bashir, Umar Farooq, Maryam Shahid, Muhammad Fareed Zaffar, and Christo Wilson. 2019. Quantity vs. Quality: Evaluating User Interest Profiles Using Ad Preference Managers. In *Proc. NDSS*.
[10] Muhammad Ahmad Bashir and Christo Wilson. 2018. Diffusion of User Tracking Data in the Online Advertising Ecosystem. *PoPETS* 2018, 4 (Oct. 2018), 85–103.
[11] Better. 2019. Trackers Collections. https://better.fyi/trackers/.
[12] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3 (2003), 993–1022.
[13] Károly Boda, Ádám Máté Földes, Gábor György Gulyás, and Sándor Imre. 2011. User Tracking on the Web via Cross-Browser Fingerprinting. In *Proc. NordSec*.
[14] John Brooke. 1996. SUS – A Quick and Dirty Usability Scale. *Usability Evaluation in Industry* 189, 194 (1996), 4–7.
[15] Yinzhi Cao, Song Li, and Erik Wijman. 2017. (Cross-)Browser Fingerprinting via OS and Hardware Level Features. In *Proc. NDSS*.
[16] N. Chaignaud Chahine, C. Abi and J.-Ph. Kotowicz. 2008. Context and Keyword Extraction in Plain Text Using a Graph Representation. In *Proc. SITIS*.
[17] Farah Chanchary and Sonia Chiasson. 2015. User Perceptions of Sharing, Advertising, and Tracking. In *Proc. SOUPS*.
[18] Prabhakar Raghavan Christopher D. Manning and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press.
[19] Cliqz. 2019. Ghostery. https://www.ghostery.com/.
[20] Amit Datta, Michael Carl Tschantz, and Anupam Datta. 2015. Automated Experiments on Ad Privacy Settings. *PoPETS* 2015, 1 (2015), 92–112.
[21] Martin Degeling and Jan Nierhoff. 2018. Tracking and Tricking a Profiler - Automated Measuring and Influencing of Bluekai's Interest Profiling. In *Proc. WPES*.
[22] DisconnectMe. Accessed November 2018. Disconnect-Tracking-Protection. https://github.com/disconnectme/disconnect-tracking-protection.
[23] Claire Dolin, Ben Weinshel, Shawn Shan, Chang Min Hahn, Euirim Choi, Michelle L. Mazurek, and Blase Ur. 2018. Unpacking Privacy Perceptions of Data-Driven Inferences for Online Targeting and Personalization. In *Proc. CHI*.
[24] DuckDuckGo. 2019. Browser Extension. https://duckduckgo.com/app.
[25] Janna Lynn Dupree, Richard DeVries, Daniel M. Berry, and Edward Lank. 2016. Privacy Personas: Clustering Users via Attitudes and Behaviors toward Security Practices. In *Proc. CHI*.
[26] Serge Egelman, Adrienne Porter Felt, and David Wagner. 2012. Choice Architecture and Smartphone Privacy: There's A Price for That. In *Proc. WEIS*.
[27] Steven Englehardt and Arvind Narayanan. 2016. Online Tracking: A 1-million-site Measurement and Analysis. In *Proc. CCS*.
[28] Steven Englehardt, Dillon Reisman, Christian Eubank, Peter Zimmerman, Jonathan Mayer, Arvind Narayanan, and Edward W. Felten. 2015. Cookies That Give You Away: The Surveillance Implications of Web Tracking. In *Proc. WWW*.
[29] Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I like it, then I hide it: Folk Theories of Social Feeds. In *Proc. CHI*.
[30] Mohtarre Eslami, Sneha R Krishna Kumaran, Christian Sandvig, and Karrie Karahalios. 2018. Communicating Algorithmic Process in Online Behavioral Advertising. In *Proc. CHI*.
[31] Motahhare Eslami, Kristen Vaccaro, Karrie Karahalios, and Kevin Hamilton. 2017. "Be careful; things can be worse than they appear": Understanding Biased Algorithms and Users' Behavior around Them in Rating Platforms. In *Proc. AAAI*.
[32] Facebook. 2019. About The Ads You See From Facebook. https://www.facebook.com/ads/settings.
[33] J.L. Fleiss. 1981. *Statistical Methods for Rates and Proportions (2nd ed.)*. John Wiley.
[34] Evgeniy Gabrilovich and Shaul Markovitch. 2007. Computing Semantic Relatedness using Wikipedia-based Explicit Semantic Analysis. In *Proc. AAAI*.
[35] Vaibhav Garg and Jean Camp. 2012. End User Perception of Online Risk Under Uncertainty. In *Proc. HICSS*.
[36] Google. 2019. Ad Settings. https://adssettings.google.com.
[37] Gabor Gyorgy Gulyas, Doliere Francis Some, Nataliia Bielova, and Claude Castelluccia. 2018. To Extend or not to Extend: on the Uniqueness of Browser Extensions and Web Logins. In *Proc. WPES*.
[38] Google Ads Help. Accessed November 2018. Add Topics to Ad Groups. https://support.google.com/adwords/answer/156178.
[39] Alex Hern. 2018. Cambridge Analytica: how did it turn clicks into votes? https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie.
[40] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* 9, 8 (1997), 1735–1780.
[41] Umar Iqbal, Zubair Shafiq, and Zhiyun Qian. 2017. The Ad Wars: Retrospective Measurement and Analysis of Anti-Adblock Filter Lists. In *Proc. IMC*.
[42] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. 2014. Privacy Attitudes of Mechanical Turk Workers and the U.S. Public. In *Proc. USENIX Security*.
[43] Saranga Komanduri, Richard Shay, Greg Norcie, and Blase Ur. 2011. Adchoices? Compliance with Online Behavioral Advertising Notice and Choice Requirements. *I/S: A Journal of Law and Policy for the Information Society* 7 (2011), 603.
[44] Georgios Kontaxis and Monica Chew. 2015. Tracking Protection in Firefox For Privacy and Performance. In *Proc. W2SP*.
[45] Balachander Krishnamurthy, Delfina Malandrino, and Craig E Wills. 2007. Measuring Privacy Loss and the Impact of Privacy Protection in Web Browsing. In *Proc. SOUPS*.

[46] Oksana Kulyk, Annika Hilt, Nina Gerber, and Melanie Volkamer. 2018. "This Website Uses Cookies": Users' Perceptions and Reactions to the Cookie Disclaimer. In *Proc. EuroUSEC*.

[47] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174.

[48] Mathias Lecuyer, Riley Spahn, Yannis Spiliopolous, Augustin Chaintreau, Roxana Geambasu, and Daniel Hsu. 2015. Sunlight: Fine-grained Targeting Detection at Scale with Statistical Confidence. In *Proc. CCS*.

[49] Linda Naeun Lee, Richard Chow, and Al M. Rashid. 2017. User Attitudes Towards Browsing Data Collection. In *Proc. CHI EA*.

[50] Pedro Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako, and Lorrie Cranor. 2012. Why Johnny Can't Opt out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising. In *Proc. CHI*.

[51] Pedro Giovanni Leon, Justin Cranshaw, Lorrie Faith Cranor, Jim Graves, Manoj Hastak, Blase Ur, and Guzi Xu. 2012. What do Online Behavioral Advertising Privacy Disclosures Communicate to Users?. In *Proc. WPES*.

[52] Pedro Giovanni Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor. 2013. What Matters to Users? Factors that Affect Users'Willingness to Share Information with Online Advertisers. In *Proc. SOUPS*.

[53] Adam Lerner, Anna Kornfeld Simpson, Tadayoshi Kohno, and Franziska Roesner. 2016. Internet Jones and the Raiders of the Lost Trackers: An Archaeological Study of Web Tracking from 1996 to 2016. In *Proc. USENIX Security*.

[54] Zhiyuan Liu, Peng Li, Yabin Zheng, and Maosong Sun. 2009. Clustering to Find Exemplar Terms for Keyphrase Extraction. In *Proc. ACL*.

[55] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Casual Model. *Information Systems Research* 15, 4 (2004), 336–355.

[56] Arunesh Mathur and Jessica Vitak. 2018. Characterizing the Use of Browser-Based Blocking Extensions To Prevent Online Tracking. In *Proc. SOUPS*.

[57] Jonathan R Mayer and John C Mitchell. 2012. Third-party Web Tracking: Policy and Technology. In *Proc. IEEE S&P*.

[58] Aleecia M McDonald and Lorrie Faith Cranor. 2010. Americans' Attitudes About Internet Behavioral Advertising Practices. In *Proc. WPES*.

[59] William Melicher, Mahmood Sharif, Joshua Tan, Lujo Bauer, Mihai Christodorescu, and Pedro Giovanni Leon. 2016. (Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking. *PoPETS* 2 (2016), 135–154.

[60] Wei Meng, Byoungyoung Lee, Xinyu Xing, and Wenke Lee. 2016. TrackMeOrNot: Enabling Flexible Control on Web Tracking. In *Proc. WWW*.

[61] Rada Mihalcea and Paul Tarau. 2004. TextRank: Bringing Order into Texts. In *Proc. EMNLP*.

[62] Mozilla. 2019. Lightbeam. https://addons.mozilla.org/en-US/firefox/addon/lightbeam/.

[63] Jason R C Nurse and Oliver Buckley. 2017. Behind the Scenes: a Cross-Country Study into Third-Party Website Referencing and the Online Advertising Ecosystem. *Human-centric Computing and Information Sciences* 7, 1 (2017), 40.

[64] Opera. [n.d.]. New ad blocker - Built into the Opera browser. https://www.opera.com/computer/features/ad-blocker. (accessed November 2018).

[65] Oracle. 2019. Oracle Data Cloud Registry. https://datacloudoptout.oracle.com/registry/.

[66] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank Citation Ranking: Bringing Order to the Web.* Technical Report. Stanford InfoLab.

[67] Xiang Pan, Yinzhi Cao, and Yan Chen. 2015. I Do Not Know What You Visited Last Summer: Protecting Users from Third-party Web Tracking with TrackingFree Browser. In *Proc. NDSS*.

[68] Eyal Peer, Joachim Vosgerau, and Alessandro Acquisti. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods* 46, 4 (2014), 1023–1031.

[69] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proc. EMNLP*.

[70] Angelisa C Plane, Elissa M Redmiles, Michelle L Mazurek, and Michael Carl Tschantz. 2017. Exploring User Perceptions of Discrimination in Online Targeted Advertising. In *Proc. USENIX Security*.

[71] Open Directory Project. 2019. http://www.odp.org/.

[72] Emilee Rader and Rebecca Gray. 2015. Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. In *Proc. CHI*.

[73] Prabhakar Raghavan. 2019. Raising the bar on transparency, choice and control in digital advertising. Google Ads Blog.

[74] Ashwini Rao, Florian Schaub, and Norman Sadeh. 2014. What do they know about me? Contents and Concerns of Online Behavioral Profiles. In *Proc. ASE BigData*.

[75] Ashwini Rao, Florian Schaub, Norman Sadeh, and Alessandro Acquisti. 2016. Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online. In *Proc. SOUPS*.

[76] Mohammad Rezaei, Najlah Gali, and Pasi Fränti. 2015. ClRank: A Method for Keyword Extraction from Web Pages Using Clustering and Distribution of Nouns. In *Proc. WI-IAT*.

[77] Franziska Roesner, Tadayoshi Kohno, and D Wetherall. 2012. Detecting and Defending Against Third-Party Tracking on the Web. In *Proc. USENIX Security*.

[78] Jukka Ruohonen and Ville Leppänen. 2018. Invisible Pixels Are Dead, Long Live Invisible Pixels!. In *Proc. WPES*.

[79] Sonam Samat, Alessandro Acquisti, and Linda Babcock. 2017. Raise the Curtains: The Effect of Awareness About Targeting on Consumer Attitudes and Purchase Intentions. In *Proc. SOUPS*.

[80] Florian Schaub, Aditya Marella, Pranshu Kalvani, Blase Ur, Chao Pan, Emily Forney, and Lorrie Faith Cranor. 2016. Watching Them Watching Me: Browser Extensions' Impact on User Privacy Awareness and Concern. In *Proc. USEC*.

[81] SimilarWeb. 2019. Website Categorization API. https://www.similarweb.com/corp/developer/website_categorization_API.

[82] Lisa Singh, Grace Hui Yang, Micah Sherr, Andrew Hian-Cheong, Kevin Tian, Janet Zhu, and Sicong Zhang. 2015. Public Information Exposure Detection: Helping Users Understand Their Web Footprints. In *Proc. ASONAM*.

[83] Oleksii Starov and Nick Nikiforakis. 2017. XHOUND: Quantifying the Fingerprintability of Browser Extensions. In *Proc. IEEE S&P*.

[84] Oleksii Starov and Nick Nikiforakis. 2018. PrivacyMeter: Designing and Developing a Privacy-Preserving Browser Extension. In *Proc. ESSoS*.

[85] Statcounter. 2018. Desktop Browser market share worldwide: Oct 2018. http://gs.statcounter.com/browser-market-share/desktop/worldwide#monthly-201810-201810-bar.

[86] Christopher A Summers, Robert W Smith, and Rebecca Walker Reczek. 2016. An Audience of One: Behaviorally Targeted Ads as Implied Social Labels. *Journal of Consumer Research* 43, 1 (2016), 156—178.

[87] Latanya Sweeney. 2013. Discrimination in Online Ad Delivery. *CACM* 56, 5 (2013), 44–54.

[88] Janice Tsai, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. 2007. The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. In *Proc. WEIS*.

[89] Michael Carl Tschantz, Serge Egelman, Jaeyoung Choi, Nicholas Weaver, and Gerald Friedland. 2018. The Accuracy of the Demographic Inferences Shown on Google's Ad Settings. In *Proc. WPES*.

[90] Joseph Turow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy. 2009. *Americans Reject Tailored Advertising and Three Activities that Enable It.* Technical Report. Annenberg School for Communication.

[91] Blase Ur, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. 2012. Smart, Useful, Scary, Creepy: Perceptions of Online Behavioral Advertising. In *Proc. SOUPS*.

[92] Max Van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J Weitzner, and Nigel Shadbolt. 2017. Better the Devil You Know: Exposing the Data Sharing Practices of Smartphone Apps. In *Proc. CHI*.

[93] Giridhari Venkatadri, Piotr Sapiezynski, Elissa Redmiles, Alan Mislove, Oana Goga, Michelle Mazurek, and Krishna Gummadi. 2019. Auditing Offline Data Brokers via Facebook's Advertising Platform. In *Proc. WWW*.

[94] Jeffrey Warshaw, Tara Matthews, Steve Whittaker, Chris Kau, Mateo Bengualid, and Barton A Smith. 2015. Can an Algorithm Know the "Real You"?. In *Proc. CHI*.

[95] Jeffrey Warshaw, Nina Taft, and Allison Woodruff. 2016. Intuitions, Analytics, and Killing Ants: Inference Literacy of High School-educated Adults in the US. In *Proc. SOUPS*.

[96] Susanne Weber, Marian Harbach, and Matthew Smith. 2015. Participatory Design for Security-Related User Interfaces. In *Proc. USEC*.

[97] Ryan West. 2008. The Psychology of Security. In *Communications of the ACM*.

[98] John Wilander. 2017. Intelligent Tracking Prevention. https://webkit.org/blog/7675/intelligent-tracking-prevention/.

[99] Craig E Wills and Can Tatar. 2012. Understanding What They Do with What They Know. In *Proc. WPES*.

[100] Shaohui Yang Xu, Songhua and Francis C.M. La. 2010. Keyword Extraction and Headline Generation Using Novel Word Features. In *Proc. AAAI*.

[101] Yaxing Yao, Davide Lo Re, and Yang Wang. 2017. Folk Models of Online Behavioral Advertising. In *Proc. CSCW*.

[102] Zhonghao Yu, Sam Macbeth, Konark Modi, and Josep M. Pujol. 2016. Tracking the Trackers. In *Proc. WWW*.

[103] Jun Zhao, Reuben Binns, Max Van Kleek, and Nigel Shadbolt. 2016. Privacy Languages: Are we there yet to enable user controls?. In *Proc. WWW*.

[104] Sebastian Zimmeck, Jie S. Li, Hyungtae Kim, Steven M. Bellovin, and Tony Jebara. 2017. A Privacy Analysis of Cross-device Tracking. In *Proc. USENIX Security*.

A APPENDIX

A.1 Interview Script

The purpose of this study is to inform the design of an app to help users like you learn more about browsing the internet and online trackers. You are allowed to leave at any time. If you still consent to being part of this study, please say, "Yes." [wait] Are you okay with me recording audio for our session? [wait] Today's study has three parts. First, I have some intro questions about your experiences with online tracking, and I'll also explain what the app is supposed to do. Second is the main part, where we'll go through each page of the app. At the end, we have a few short closing questions.

- On a scale of 1 to 5, with 1 being the least and 5 being the most, how knowledgeable would you say you are about online tracking and how it works?
- On a scale of 1 to 5, how interested would you say that you are in learning more online tracking? Why did you choose that number?
- In your own words, could you explain to me what you know about online tracking?

Today we are testing an app called "Tracking Transparency." It was developed by researchers at the University of Chicago, and we were hired by these researchers to get feedback about the app. The app is a browser extension that shows your browsing history, which trackers you have encountered online, and what inferences they might have made about you, based on web pages you have visited. Now, I'm going to read a short description of online tracking that the researchers want you to know: When you browse the internet, your online activity is tracked by the website you are visiting, as well as by third-party advertising and analytics companies. These third-party companies use logs of your browsing behavior to infer your interests, preferences, and demographics. They can then tailor your internet experience in part based on those inferences, impacting the search results, ads, and social feeds that you see. For example, if you visit a blog about traveling with dogs and a third-party tracker on that site infers that you are interested in dogs. Later, you might encounter an ad that was targeted specifically to dog lovers.

Again, we were hired by the researchers to get feedback on their Tracking Transparency browser extension. I didn't make the app, so please feel free to give me any and all feedback, I won't be offended! There is not one particular design they hope you'll like better than the others; they're most interested in your honest and blunt feedback for everything you see. As you go through the app, I would like you to think-aloud for me as you answer. You know the little voice in your head that sometimes narrates as you answer questions or take surveys? I just want you to vocalize that little voice out loud for me as you go through. Let's get started. We have a copy of the extension on this computer for you to use. As you can see, the browser extension logo is in the upper right corner. If we click on the extension logo, the extension provides an overview of the tracking information, and by clicking on the "Show me more..." button, we will get to the main pages of the browser extension. You can access the individual tabs at the top of the browser extension: trackers, inferences, domain, recent activity, and time. Please think-aloud as you go to each of these tabs.

Researcher will direct participant to each tab and to think aloud with their perceptions. If they are silent, researcher will use the questions listed below to guide their thinking.

- What are you thinking now?
- Why did you do that?
- What is this tab telling you?
- What do you think the graph is showing you?
- What do you think the table is showing you?
- What is confusing on this tab to you? (Change/add/remove to make clearer?)
- What do you think would happen if you clicked this link?
- Do you think the graph is interesting?
- Would you add/change/remove anything to make this tab more interesting?
- Would you go return to this tab after the first time you see it? Why?
- What is the most interesting thing on this page to you?
- Is there anything you would want to know but isn't explained on this tab?
- Do you feel like this tab is telling you new and interesting information, compared to previous tabs?

This is the last section - we just have a three last questions.

- On a scale of 1 to 5, with 1 being the least and 5 being the most, how informative would you say the app was?
- On a scale of 1 to 5, how knowledgeable would you say you are now, after this session, about how online tracking works?
- On a scale of 1 to 5, how interested would you say that you are now, after this session, in learning more online tracking? Why did you choose that number?

And finally, do you have any questions for us? Thank you so much for your participation in today's session. We are very grateful for all your comments today, and will be passing them on to the researchers' for their final design. Here is a $10 Amazon gift card. If you have any questions about this research, you may contact our Principal Investigator or the IRB at the contact info on the consent form. Thank you again!

A.2 Survey 1 Instrument

Internet Usage

Which of the following browsers do you regularly use? Select all that apply. ◯ Chrome ◯ Firefox ◯ Safari ◯ Opera ◯ Internet Explorer/Edge ◯ Epic ◯ Brave Firefox Focus ◯ Tor ◯ Other: ___

What percentage of your online browsing is on the device and browser you are using right now, compared to other devices or other browsers? [slider: 0 … 100]

How often do you make purchases online using a web browser (as opposed to through an app)? ◯ Never ◯ Rarely ◯ Monthly ◯ Weekly ◯ Daily ◯ Multiple times a day ◯ Don't know

Have you ever heard of or used the following software, browser extensions, websites, or tools? (Answer choices for each) ◯ Don't use it and have never heard of it ◯ Don't use it, but have heard of it ◯ Previously used it ◯ Currently use it

- (Matrix-style grid with the following rows): AdBlock Plus; AdBlock; Disconnect; Facebook; Firefox Tracking Protection; Ghostery; Gmail; HTTPS Everywhere; Privacy Badger; uBlock Origin

[AdChoices icon] Have you seen this icon while browsing online? ◯ Yes ◯ No ◯ Don't know

Whether or not you have seen this icon while browsing online, what is your best guess of what this icon indicates?

Experiences with Advertising

(Answer choices for each) ◯ Yes ◯ No ◯ Don't know

- Have you ever looked at your Facebook ad preferences (…shown below)?
- Have you ever looked at your Google ad settings (…shown below)?

The following question was asked with all of the following combinations: A [read, click on], B [on social media, in search results, on all other websites]

To the best of your memory, how often do you [A] advertisements [B] ◯ Never ◯ Rarely ◯ Sometimes ◯ Often ◯ Always ◯ I don't use [B] ◯ Don't know

To the best of your memory, how many times have you bought something as a result of reading or clicking on an online advertisement in the last year? ◯ Never ◯ Once ◯ 2-5 times ◯ 6-10 times ◯ More than 10 times ◯ Don't know

Opinions about Online Advertising and Ads

During the rest of this survey, we use the term "online advertising companies" to refer to companies that show you advertisements online. Note that these companies that select and display advertisements are distinct from the companies whose products are being advertised. Please select the answer choices that best describes your agreement or disagreement with the statements shown below.

Answer choices for all questions in this section: ◯ Strongly agree ◯ Agree ◯ Somewhat agree ◯ Neither agree nor disagree ◯ Somewhat disagree ◯ Disagree

- I would like to see ads that are relevant to my interests, as opposed to generic ads.
- I would be comfortable with online advertising companies guessing my interests based on which websites I visit.
- If it were available, I would like to use a system that shows me what information has been collected about me online.
- I feel that online advertising companies adequately explain why I received a particular ad.
- I feel that I understand how online advertising companies determine which advertisements I see.
- I would consider it fair for advertising companies to track which websites I visit in order to show me ads that are relevant to my interests.
- I would consider it creepy for advertising companies to track which websites I visit in order to show me ads that are relevant to my interests.

Knowledge Certainty and Facts

Imagine a regular Internet user who has many online accounts, including social media, email, and more. This person uses their browser with the default configurations. How likely or unlikely do you expect it is that, while browsing online, this user will see ads that advertising companies targeted to them based on the following types of information? (Answer choices for each) ◯ Very likely ◯ Likely ◯ Somewhat likely ◯ Neither likely nor unlikely ◯ Somewhat unlikely ◯ Unlikely ◯ Very unlikely

- Background audio captured by their microphone
- Their geographic location
- Their Social Security Number
- The brand/model of the device they are using to access the internet
- A company's guess about their race
- A company's guess about their gender
- A company's guess about their age

- A company's guess about their political views
- A company's guess about specific products they might be interested in
- Websites they have visited in the past
- A company's guess about topics they're interested in
- Times of day when they often browse online
- The color of the device they are using to access the internet
- A company's guess about the number of browsers they use
- The website they are currently visiting (and on which the ad is being displayed)

In the last week, about how many different online advertising companies do you think tried to collect information about your browsing history?

In the last week, about how many websites do you think you've visited? For example, if you went to 3 pages on 1 news site, count this as 1.

In the last week, about how many pages do you think you've visited? For example, if you went to 3 pages on 1 news site, count this as 3.

Of the pages you've visited in the last week, about what percentage do you think had online advertising companies on them? [slider: 0 . . . 100]

Answer choices for the following three questions: ◯ Very likely ◯ Likely ◯ Neither likely nor unlikely ◯ Unlikely ◯ Very unlikely ◯ Don't know

- How likely or unlikely do you think it is that an online advertising company would attempt to collect information about users' interests in broad categories (e.g., Arts & Entertainment, Business & Industrial, Sports)?
- How likely or unlikely do you think is it that an online advertising company would attempt to collect information about users' interests in specific categories (e.g., Classical Music, Livestock, Cricket)?
- How likely or unlikely is it that small online advertising companies (e.g., Pubmatic, Taboola, TurnTo) can track which websites you visit?

IUIPC Awareness and Collection Subscales
Answer choices for all questions: ◯ Strongly agree ◯ Agree ◯ Somewhat agree ◯ Neither agree nor disagree ◯ Somewhat disagree ◯ Disagree ◯ Strongly Disagree

- Companies seeking information online should disclose the way the data are collected, processed, and used.
- A good consumer online privacy policy should have a clear and conspicuous disclosure.
- It is very important to me that I am aware and knowledgeable about how my personal information will be used.
- It usually bothers me when online companies ask me for personal information.
- When online companies ask me for personal information, I sometimes think twice before providing it.
- It bothers me to give personal information to so many online companies.
- I'm concerned that online companies are collecting too much personal information about me.

Demographics
With what gender do you identify? ◯ Female ◯ Male ◯ Non-binary ◯ Other ◯ Prefer not to say

What is your age? ◯ 18-24 ◯ 25-34 ◯ 35-44 ◯ 45-54 ◯ 55-64 ◯ 65 or older ◯ Prefer not to say

What is the highest degree or level of school you have completed? ◯ Some high school ◯ High school ◯ Some college ◯ Trade, technical, or vocational training ◯ Associate's degree ◯ Bachelor's degree ◯ Master's degree ◯ Professional degree ◯ Doctorate ◯ Prefer not to say

Which of the following best describes your educational background or job field? ◯ I have an education in, or work in, the field of computer science, engineering, or IT. ◯ I do not have an education in, or work in, the field of computer science, engineering, or IT. ◯ Prefer not to say

A.3 Survey 2 Instrument

This survey, Survey 2, is about Tracking Transparency, the extension you installed about a week ago. This survey will take about 20 minutes to complete.

Introduction
Please spend a few minutes exploring the extension before beginning the survey. Click the icon in your browser toolbar near the top-right corner of the browser window.

If condition 4: Visit a few websites to see the information that Tracking Transparency displays. Look for the overlay in the bottom-right corner of the window.

If not condition 4: Click on the "Open Tracking Transparency dashboard" button. Be sure to view all parts of the dashboard.

In a few sentences, what do you think is the purpose of the Tracking Transparency tool?

If condition 1: In your own words, please explain the information displayed by the Tracking Transparency tool.
If condition 2, 3, 5, 6: In your own words, please list and briefly describe the information displayed on each tab of the Tracking Transparency tool.
If condition 4: In your own words, please briefly describe the information displayed in the popup boxes of the Tracking Transparency tool.

Open-Ended Reactions

- Please list the new information, if any, you learned by using this extension.
- Please list the information you already knew, if any, that the extension told you.
- Please list the surprising information, if any, that the extension told you.
- What questions, if any, do you have about what you saw in the extension?

Post-Extension Intended Behaviors
Answer choices for all questions: ◯ Much more likely ◯ More likely ◯ Somewhat more likely ◯ About the same as before ◯ Somewhat less likely ◯ Less likely ◯ Much less likely ◯ Don't know

- Compared to before you used the extension, how likely are you to seek out more information about online advertising now?
- Compared to before you used the extension, how likely are you to use a browser's private browsing mode now?
- Compared to before you used the extension, how likely are you to click on ads now?
- Compared to before you used the extension, how likely are you to use browser extensions that block ads and/or online tracking now?
- The Do Not Track (DNT) setting is a browser setting to indicate to web pages you visit that you do not want to be tracked online. Compared to before you used the extension, how likely are you to use the DNT setting now?
- Imagine that online advertising companies provided a page to show you what topics they guessed you are interested in. Compared to before you used the extension, how likely are you to spend time looking at such a page now?

Tradeoffs
Imagine that you had a choice between: Option A: [A]; Option B: [B]. Which option would you choose? ◯ Definitely A ◯ Probably A ◯ I'm not sure ◯ Probably B ◯ Definitely B

- [A] free online browsing, but all of your browsing history is collected; [B] a monthly fee to browse online, but none of your browsing history is collected
- [A] block all online trackers, but some web pages or parts of pages don't work; [B] block no online trackers, but all web pages work
- [A] your search results are not personalized or relevant, but none of your searches are tracked; [B] your search results are personalized and more relevant, but all of your searches are tracked
- [A] ad networks collect your browsing history to guess your interests, and use this to show ads relevant to you; [B] ad networks don't collect your browsing history, but every month you are required to fill out an online form about your interests so that they can show ads relevant to you

Opinions about Online Advertising and Ads Repeated from Survey 1

Knowledge Certainty and Facts Repeated from Survey 1

System Usability Scale
This page shows statements about your experiences with the Tracking Transparency tool, called the "system" below. Please select the answer choice that best describes your agreement or disagreement with the statements. (Answer choices for each) ◯ Strongly agree ◯ Agree ◯ Neither agree nor disagree ◯ Disagree ◯ Strongly disagree

- I think that I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think that I would need the support of a technical person to be able to use this system.
- I found the various functions in this system were well-integrated.
- I thought there was too much inconsistency in this system.
- I would imagine that most people would learn to use this system very quickly.
- I found the system very cumbersome to use.
- I felt very confident using the system.
- I needed to learn a lot of things before I could get going with this system.

IUIPC Awareness and Collection Subscales Repeated from Survey 1

A.4 Additional Screenshots and Figures

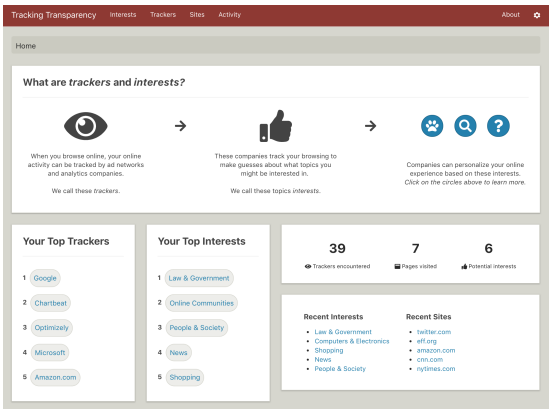


Figure 7: Dashboard homepage (*Longitudinal:Interests*).

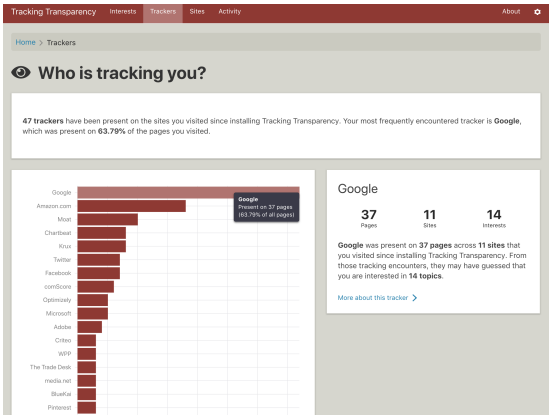


Figure 8: Trackers tab (*Longitudinal:Trackers* and *Longitudinal:Interests*).

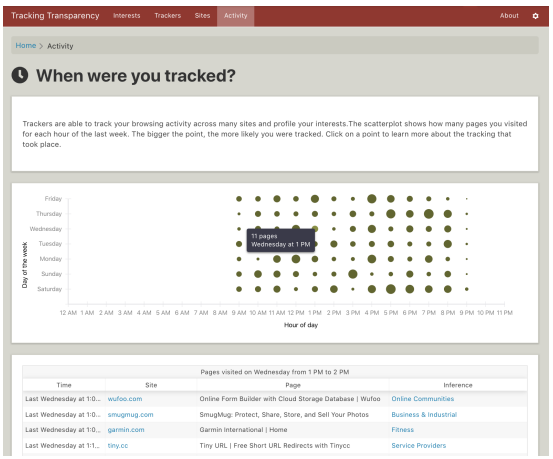


Figure 9: Activity tab (*Control:Browsing Only*, *Longitudinal:Trackers*, and *Longitudinal:Interests*).

Figure 10: A tracker detail page highlighting longitudinal tracking information for a single tracker (in this case, Google). Interest and Site detail pages follow the same structure, but show longitudinal data for a single interest or site. Detail pages were shown only for *Longitudinal:Trackers* and *Longitudinal:Interests*.

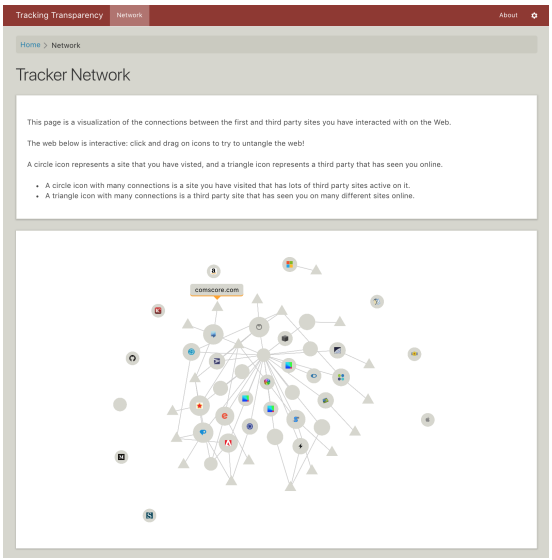


Figure 11: Network tab simulating Mozilla Lightbeam (*Current:Connections*).

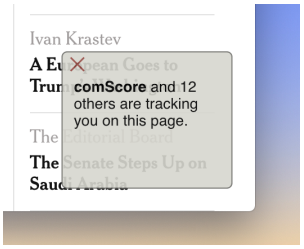


Figure 12: Tracking Transparency's in-page overlay (*Current:Trackers*) simulating tools like Ghostery.

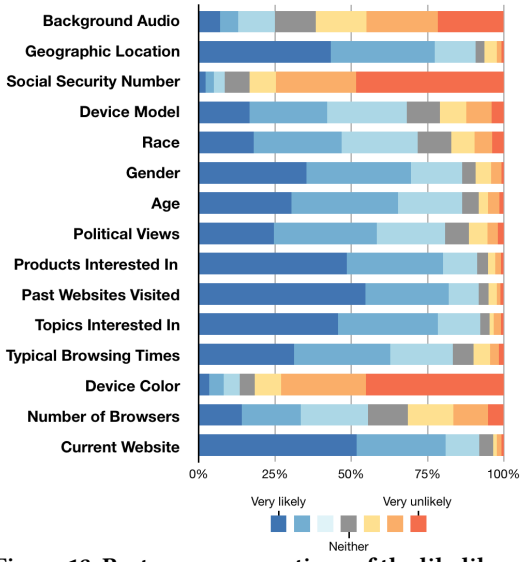


Figure 13: Post-usage perceptions of the likelihood companies use types of information to target ads.

Table 3: Changes in participants' median estimates of tracking and their own browsing across surveys by condition.

Survey	Control:Static	Control:Browsing Only	Current:Connections	Current:Trackers	Longitudinal:Trackers	Longitudinal:Interests
	Estimated # Trackers Encountered					
Pre-Usage	10	10	20	15	10	20
Post-Usage	20	50	50	30	80	100
	Estimated % Browsing Tracked					
Pre-Usage	80%	80%	80%	70%	70%	80%
Post-Usage	70%	70%	80%	80%	80%	80%
	Estimated # Domains Visited Weekly					
Pre-Usage	30	25	30	25	22.5	35
Post-Usage	30	83.5	30	25	70	75
	Estimated # Pages Visited Weekly					
Pre-Usage	200	100	150	100	110	125
Post-Usage	150	2,000	100	100	924	1,500

Table 4: As discussed in Section 4, we performed seven pairwise comparisons of conditions to investigate targeted hypotheses about the impact of different visualizations. Here, we report the significant results of the associated Mann-Whitney U tests in each cell. The leftmost column reports the omnibus Kruskal-Wallis tests.

Question	**Longitudinal:Interests** > Control:Static	**Longitudinal:Interests** > Control:Browsing Only	**Longitudinal:Interests** > Current:Connections	**Longitudinal:Interests** > Current:Trackers
System usability ($\chi^2(5) = 52.6, p < .001$)	$U = 1112, p < .001$	–	$U = 1720, p < 0.001$	–
Seek more info ($\chi^2(5) = 26.059, p < .001$)	$U = 1474, p < .001$	–	$U = 1919, p = .040$	–
Use priv. browsing ($\chi^2(5) = 24.988, p < .001$)	$U = 1527, p < .001$	$U = 2239, p = .028$	–	–
Use blocking tools ($\chi^2(5) = 26.567, p < .001$)	$U = 1665, p = .001$	–	–	–
Use DNT ($\chi^2(5) = 34.602, p < .001$)	$U = 1300, p < .001$	$U = 2176, p = .012$	–	–
Number domains ($\chi^2(5) = 58.298, p < .001$)	$U = 1467, p < .001$	–	$U = 1558, p < .001$	$U = 1116, p < .001$
Number pages ($\chi^2(5) = 132.26, p < .001$)	$U = 856, p < .001$	–	$U = 722, p < .001$	$U = 766, p < 0.01$
% pages tracked ($\chi^2(5) = 11.974, p = .035$)	$U = 1963, p < .001$	$U = 2186, p = .003$	–	–
Number trackers ($\chi^2(5) = 42.207, p < .001$)	$U = 1191, p < .001$	$U = 2352, p = .021$	$U = 1969, p = .018$	$U = 1429, p < .001$

Question	**Longitudinal:Interests** > Longitudinal:Trackers	**Current:Connections** > Control:Static	**Current:Trackers** > Control:Static
System usability ($\chi^2(5) = 52.6, p < .001$)	–	–	$U = 1207, p < .001$
Seek more info ($\chi^2(5) = 26.059, p < .001$)	–	–	$U = 1405, p = .001$
Use priv. browsing ($\chi^2(5) = 24.988, p < .001$)	–	$U = 1684, p = .005$	$U = 1524, p = .006$
Use blocking tools ($\chi^2(5) = 26.567, p < .001$)	–	$U = 1671, p = .003$	$U = 1362, p < .001$
Use DNT ($\chi^2(5) = 34.602, p < .001$)	–	$U = 1725, p = .014$	$U = 1242, p < .001$
Number domains ($\chi^2(5) = 58.298, p < .001$)	–	–	–
Number pages ($\chi^2(5) = 132.26, p < .001$)	–	–	–
% pages tracked ($\chi^2(5) = 11.974, p = .035$)	$U = 1953, p = .047$	–	–
Number trackers ($\chi^2(5) = 42.207, p < .001$)	–	$U = 1614, p < .001$	–