The Ethics of Internet Tracking

Anna L. Villani
Department of Computer Science, University of Colorado Boulder Boulder, CO 80303 USA

Behavioral internet tracking is the practice of tracking web and mobile users on the internet, gathering things like browser history, email interactions, buttons clicked, saved pages, and Facebook likes. One of the main uses of behavioral internet tracking is for marketing and advertising purposes. Internet tracking enables a company to see what their customers are doing online, to develop better marketing strategies for their individual customers. Behavioral tracking can be done at the first-party level, third-party level, and now even across a user's multiple devices. The biggest ethical issue that arises around internet tracking is a person's right to knowledge of who is accessing his or her personal information, and what is being done with that information. There is currently no government policy covering the entirety internet tracking, or even at least for all companies within the United States. Because of the varying nature of opinions on internet tracking, different layers of internet tracking itself, and possible harms it can cause I think that there should be some kind of policy that covers all of these issues.

With first-party and third-party cookies, data is tracked on the internet using cookies. A cookie is a small script placed on the hard drive of a user's computer by the server of a website that user visits. The cookie is placed so that the website recognizes the browser/computer combination of the user when returned to the site. The word 'party' refers to the website that is placing the cookie. If it is a first-party cookie (first-party tracking) the cookie is being placed by the website the user visits. Third-party cookies are placed by another website associated with first-party site. Third-party internet tracking is when a first-party website authorizes a third-party site to learn about its users. Generally how this works is a piece of software from the third-party site is embedded into the first-party site. When users connect to the first-party site, the third-party site is also gains access to the users' information. Trackers can collect information on Internet users based on their browsing behavior. Many first-party sites contain privacy and data policies in their Terms of Service. However, many of their users still don't know the extent they are being tracked because they do not read or understand the Terms of Service.

Facebook's Terms of Service Data Policy (https://www.facebook.com/about/privacy/)

What kinds of information do we collect?

Depending on which Services you use, we collect different kinds of informatio from or about you.

Things you do and information you provide.
We collect the content and other information you provide when you use our Services, including when you sign up for an account, create or share, and message or communicate with others. This can include information in or about the content you provide, such as the location of a photo or the date a file was created. We also collect information about how you use our Services, such as the types of content you view or engage with or the frequency and duration of your activities.

Things others do and information they provide.
We also collect content and information that other people provide when they use our Services, including information about you, such as when they share a photo of you, send a message to you, or upload, sync or import your contact information.

Your networks and connections.
We collect information about the people and groups you are connected to and how you interact with them, such as the people you communicate with the most or the groups you like to share with. We also collect contact information you provide if you upload, sync or import this information (such as an address book) from a device.

Information about payments.
If you use our Services for purchases or financial transactions (like when you buy something on Facebook, make a purchase in a game, or make a donation), we collect information about the purchase or transaction. This includes your payment information, such as your credit or debit card number and other card information, and other account and authentication information, as well as billing, shipping and contact details.

Device information.
We collect information from or about the computers, phones, or other devices where you install or access our Services, depending on the permissions you've granted. We may associate the information we collect from your different devices, which helps us provide consistent Services across your devices. Here are some examples of the device information we collect:

- Attributes such as the operating system, hardware version, device settings, file and software names and types, battery and signal strength, and device identifiers.
- Device locations, including specific geographic locations, such as through GPS, Bluetooth, or WiFi signals.
- Connection information such as the name of your mobile operator or ISP, browser type, language and time zone, mobile phone number and IP address.

Information from websites and apps that use our Services.
We collect information when you visit or use third-party websites and apps that use our Services (like when they offer our Like button or Facebook Log In or use our measurement and advertising services). This includes information about the websites and apps you visit, your use of our Services on those websites and apps, as well as information the developer or publisher of the app or website provides to you or us.

Information from third-party partners.
We receive information about you and your activities on and off Facebook from third-party partners, such as information from a partner when we jointly offer services or from an advertiser about your experiences or interactions with them.

Facebook companies.
We receive information about you from companies that are owned or operated by Facebook, in accordance with their terms and policies. Learn more about these companies and their privacy policies.

The paper, *Reality and Perception of Copyright Terms of Service for Online Content Creation*, gave evidence that in most cases users do not read the Terms of Service because of its long and confusing jargon. The paper states that "considering that reading only the privacy policy of every site visited would take the average Internet user over 200 hours per year, it is not surprising that many do not take the time to read often complicated terms of service. " [1] In cases of third-party tracking, even more users are unaware that they're being tracked. First-party sites may usually explain if they're tracking data themselves, but they don't always tell users if there are third-party sites embedded in their systems. The news article Facebook begins tracking non-users around the internet talks explains how Facebook is now tracking non-users' data. According to the article, "Facebook will now display ads to web users who are not members of its social network, the company announced Thursday, in a bid to significantly expand its online ad network. As The Wall Street Journal reports, Facebook will use cookies, 'like' buttons, and other plug-ins embedded on third-party sites to track members and non-members alike" [2]. An important ethical question here is whether is is right for first-party sites to allow third-party sites access to their users' data without the users' consents. Surveys have consistently shown that most users oppose third parties collecting and using their browsing activity. In their paper Third-Party Web Tracking: Policy and Technology, Mayer and Mitchell offer some data to support this claim. A 2009 representative U.S. phone survey by Turow et al. [4] found that 87% of respondents would not want advertising based on tracking. In an 2010 survey of Amazon Mechanical Turk users by McDonald and Cranor [5], only 45% of respondents wanted to be shown any ads that had been tailored to their interests. A December 2010 USA Today/Gallup poll [3] reported 67% of respondents thought behavioral targeting should be outright illegal.

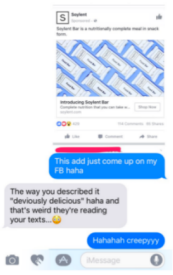
The next big leap in the space of internet tracking is cross-device recognition. Facebook is a company that has had a lot of success in this area already. Although cookies can be used to track a user on the web, cookies are not able to recognize one user as the same person across different devices. They are also not always accurate if multiple people are sharing a computer. The solution of cross-device recognition is to target individuals, not cookies, which can overlap with each

other [6]. I talked to 25 individuals between the ages of 18-26 and asked them some questions about internet tracking. I asked people within this age demographic because I wanted to ensure that most of them were comfortable with the internet, and at least had an idea of what internet tracking is. I made sure that at least half of the participants did not have a proficient technical background of any kind (other than simply using technology). Below is table illustrating the percentage of the 25 people I interviewed who know particular things regarding internet tracking.

Table 1.

Data is tracked on the internet	0.96
What a cookie is	0.68
That sites you don't visit can track your data	0.28
What third-party cookie is	0.36
What cross-device recognition is	0.16

The data shows that almost all of them had a general understanding of internet tracking. Regardless of their opinions on internet tracking, they were not surprised when an ad related to something they had been looking at online showed up on their Facebook pages. However, the percentage dropped drastically with third-party tracking, and only a few people knew what cross-device recognition is. I observed a few people experience ads shows up on their devices through cross-device recognition, and they were a lot more surprised. Some said that since they are used to seeing ads all over their screens it's likely they've seen this happen, but they didn't even notice it as some kind of cross-device recognition. Many thought of this as creepy, especially since it is a whole new level of connection with regards to being tracked, and it went unnoticed for most of them. Two of my friends told me about a conversation they had over text that gives an example of device recognition, and how people react to noticing it for the first time. I asked for some background on the conversation; they were talking about new jobs and how busy they feel all the time. One friend said that he wished preparing food and eating didn't take up so much time, and he wished there was a way to get everything he needed without taking any time. The other friend suggested Soylent. Below is a screenshot of snippet of the text conversation between them (with permission to use).



A few minutes after the messages on the left were sent, one friend went onto Facebook on his computer, and saw the ad for Soylent displayed in the screenshot to the right. The reaction was surprise and concern that technology is now able to process something that is written in a text on a phone, and tie it to a related add that shows up minutes later on that same user's computer. A related ethical question that arises from this is whether a person has a right to knowledge of emerging technologies if they are using that person's personal property (i.e browsing content). It is clear that people now have understanding of 'old' internet tracking, but most users have trouble keeping up with emerging technologies. However, just because something is emerging does not mean it is affecting people, and growing rapidly.

The paper Folk Models of Online Behavioral Advertising performs a qualitative study on. Internet users to determine their actual understandings of online behavioral advertising. From their results they gathered some common misconceptions that users have about internet tracking. One of these misconceptions is that trackers are people with malicious intentions. Many think that trackers want to gain information to break into personal accounts and steal data. Another misconception is that trackers are viruses. This relates to the first misconception of trackers as always bad. Lastly, users often believe that trackers are located on user computers, and can access a users' local files. These misconceptions offer some speculation as to why many users are so averse to being tracked on the internet. In most cases, internet trackers are legitimate companies with the goal of building a business an it being successful. Many times the tracking is aimed with the end goal of improving users' lives in some way.. Most of the internet tracking that goes on is used for these kinds of marketing purposes. In these cases, the marketers do not need to personally identify each consumer, they just need to

be able to aggregate enough information about particular consumers to serve them personalized advertisements. The problem is that enough data can't be gathered so that it is useful to marketers, but that is also not personally identifiable. Personally Identifiable Information is very broadly defined to cover "individually identifiable information about an individual consumer" and includes a consumer's first and last name, home or other physical address, email address, telephone number, and Social Security number [8]. Web browsing history is inexplicably linked to personal information, so for right now, it is practically impossible for trackers to gather useful information about a consumer, while also allowing that consumer to remain anonymous. From the viewpoint that internet tracking is used to gain insights about users, to better understand what they want and need, internet tracking can be a positive contributor to the world.

Although the majority of tracking is used for advertising, the existence of it still allows for harm. Since most users do not know how exactly they're being tracked, it is hard for them to protect themselves from tracking. Malicious hackers can easily use tracking software to gather information about a user, identify them, and use it to their advantage or to harm the user. If a first-party site is deemed untrustworthy, a user can just decline to visit it. However, users are often unaware of the very existence of many third-party sites, so they cannot protect

themselves against them if they are malicious. Almost all first-party stakeholders, third-party stakeholders, and consumers alike can agree that consumers should have some degree of control over web tracking. However, different parties disagree greatly on varying specifics. There are arguments over what a consumer should be able to control: the content of the data or just how it is being used. There is also disagreement about what the default setting for users should be. This relates to what the majority of users want, but third-party sites and advertising trade groups can argue that the economic right of their practices outweighs consumer's wishes. Most countries have privacy laws that require sites to include a privacy policy - a statement of your data collection as a disclosing service to your visitors or users - as a website owner or app developer. For the most part in the US, it is up to individual sites to set their own privacy policies. Risk associated with internet tracking can be heightened by the lack of market pressure to exercise good security and privacy practices. Since there are many stakeholders involved in internet tracking (users, first-party sites, third-party sites, advertisers, publishers, etc.), and there is much disagreement on the ethics of internet tracking, I think that putting a government policy in place that extends to all parties involved would have benefits for the majority.

The authors of the paper Folk Models of Online Behavioral Advertising offer some evidence as to what kinds of privacy features the users want based on their data analyses. The three main features users came up with are block tracking, transparency, and effortless to use. Users want the ability to turn ads off if they want, they want to understand what's happening with their personal data, and they want everything to be easy to understand and use. Based on all the information I have gathered about internet tracking, what users want, and what the advertisers want I have come up with a policy I think all parties involved should adhere to. The main components that should be included in the policy are as follows:

-> First-party sites make known if they are using third-party sites to track data, and what those sites are

-> Explain clearly and explicitly what is going to be done with the data, including if/where the data is being to another party

-> Data/Privacy policy has Flesh-Kincaid level > 60 (10th-12th grade or easier)

-> Ensuring that data will be protected, and acceptance of consequences if data is breached by unwanted party

-> Make known if a site supports cross-device recognition

If a company adheres to all of these conditions, it is considered "policy protected". This will benefit the users by making them more comfortable and understanding of what is happening with their personal data. It will also benefit the companies involved because it will make them more comfortable working with each other and help reduce uncomfortable discrepancies, by knowing that everyone is following the same set of policies. For this policy, I included two of the wanted user features: transparency and ease of use. I think that all users have a right to knowledge of who is tracking their data, why, and where it is going. Then, users have the right information to make a correctly informed decision about whether they should use a particular service or not. I did not include the ability to turn off tracking. This is because of the evidence that so many people have the wrong idea of what internet tracking is really used for, so my guess is that a very large number of them would automatically use this feature. I think would negatively effect the economy, as so many companies rely on internet tracking and it has really become a part of the internet system as whole. I presume that if a policy first focuses on transparency and ease of use, many users' feelings about internet tracking will change over time.

Internet tracking involves three levels: first-party tracking, third-party tracking, and cross-device recognition. The deeper these levels go, the more information can be tracked and the less of users understand what's going on with their data. It is my belief that from an ethical standpoint, users have a right to knowing what happens with data that both personally identifies them and contains a lot of information about them. It is wrong for companies to track users' data without them knowing, or take advantage of the their knowledge about users, and intentionally use that to do things in a way that they won't find out. However, because of market pressure and the common user's view of internet tracking, many companies do not want to explicitly spell out what they're doing, or set up protective privacy policies. Thus, a government policy is really needed, to ensure that all companies involved in tracking or at least offering the right level of transparency to their users.

References

[1] Fiesler, Casey, Cliff Lampe, and Amy S. Bruckman. "Reality and Perception of Copyright Terms of Service for Online Content Creation." WordPress. CSCW '16, 2016. Web. 6 Oct. 2016.

[2] Toor, Amar. "Facebook Begins Tracking Non-users around the Internet." The Verge. N.p., 27 May 2016. Web. 13 Dec. 2016.

[3] Gallup. (2010, December) USA Today/Gallup poll. [Online] Available: http:// gallup.com/poll/ File/145334/Internet Ads Dec 21 2010.pdf

[4] J. Turow, J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessy, "Americans reject tailored advertising and three activities that enable it," September 2009.

[5] A. M. McDonald and L. F. Cranor, "Beliefs and behaviors: Internet users' understanding of behavioral advertising, "Proceedings of the 2010 Research Conference on Communication Information and Internet Policy, October 2010.

[6] Reilly, Richard Byrne. "The Cookie Is Dead. Here's How Facebook, Google, and Apple Are Tracking You now." VentureBeat. VB, 6 Oct. 2014. Web. 06 Oct. 2016.

[7] Yao, Yaxing, David Lo Re, and Yang Wang. "SIGCHI Conference Proceedings Format." ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2017) (n.d.): n. pag. Web. 17 Nov. 2016.

[8] Ubenda. "Why I Need a Privacy Policy - for Websites and Apps | Iubenda." Iubenda. Milan Chamber of Commerce, n.d. Web. 9 Nov. 2016.