Cookie Jar: An Experiment in Accessible Privacy Management

Cody Bohlman, Emma Freeman, Joy Hill, Sahree Kasper, Robert Krossa
Carleton College

March 14, 2016

1 Introduction

Our Internet experiences are specifically catered to our interests. Our web browsers know us and companies gather information about online customers. Our society is increasingly dependent on the Internet and more and more people are putting their information online. However, general societal education about how to control our online profile is lacking.

Privacy is complicated. Our comps aims to make it easier to understand how information is spread. In order to improve accessibility for everyone online, we want to make the information contained in cookies more clear. This will allow people to control their privacy. Which sites should be avoided? Which cookies should be deleted? We want to help anyone who uses our comps make informed decisions about their Internet presence. Academics and nonacademics alike have done a significant amount of work in privacy and cookies. Trepte et al. (2015) found that most users are unaware of which privacy settings to use and why. Turow et al. (2015) found in their survey that marketers are misrepresenting a large majority of Americans when it comes to privacy. According to the survey, marketers believe that people are willingly giving out information about themselves for the benefits they receive by using certain websites. Americans, on the other hand, do not believe that this ‘data for discounts’ tradeoff is a square deal. Nissenbaum (2011) portrays online privacy as the wild west, and we are just beginning to understand what is and isn’t acceptable. She compares online privacy to medical surgeries, in that most patients are not educated enough to understand all of the potential risks, but they still get on the operating table because they trust the doctors and the institutions behind them. Nissenbaum argues that privacy should be context specific, and it is up to the companies to know what should and shouldn't be shared based on the context.

Outside of academia, there is a large variety of sources that people can use to learn more about privacy and cookies. HowToGeek (Heddings 2015) wrote an article warning people how browser extensions can collect your metadata and sell your information to corporations. TechCrunch (Essaid 2015) writes that, just like many of the academic papers, the best way to handle Internet tracking is increasing company transparency about what information is collected and shared. VentureBeat (Reilly 2014) writes about how as more Internet traffic moves mobile, companies are coming up with alternatives to cookies to track their users. All of these articles come from media sources meant for readers with above average technical knowledge. Media sources with wider audiences such as The New York Times (Singer 2015) and the Guardian (Kobie 2015) have published some work about privacy and cookies, but not very much. There are several browser extensions that have been made to help people manage and visualize their cookies. By far the most advanced is Lightbeam from Mozilla. Available for the Firefox web browser, this webapp allows users to visualize third party sites that their cookies have connected to. For Google Chrome, there is a plethora of web extensions that allow users to view, edit, and delete their cookies, but there is nothing like Lightbeam. One example, EditThisCookie, allows users to see the cookies set by the current webpage and edit their values. However, you cannot see the third party cookies set by the webpage, nor can you see how everything is connected.

We want to make the language of online privacy more accessible to all levels of technical experience. We have chosen to make a Chrome extension because Google Chrome is used by beginning and advanced computer users. We chose to write a corresponding paper to provide a more in-depth understanding of why we made the choices we made in our extension. We want to get the chance to explain the research we made, how we determined our goals, and make sure that our app is something that people can build on in open source in the future.

2 Cookies and Privacy

Have you ever visited an e-commerce website, say Amazon.com, and searched for an item? If you returned to that site at a later date, you might have noticed that the “suggested items” category contains items either identical or very similar to the thing you had searched for earlier. Some may find this useful, others, not so much. Our goal is to educate people on the use of cookies, and to allow them to manage these pieces of information in one location, easily and efficiently. So, what is a cookie? An HTTP cookie, or just “cookie”, is a bit of text, usually no bigger than 4kb, that is stored on the machine of a user using a web browser. This information typically consists of login information, user location, search preferences, or any of a myriad of individualized bits of data that the web server can read in order to personalize the session of someone browsing their website.

In our research, we focused on the following types of cookies and sought to address the challenge that each of them posed. Session cookies do not have expiration dates and only exist in temporary memory while you browse a particular website. Persistent cookies expire at a specific date or after a period of time. That means for as long as the creators of the cookie want to, every time you visit a website the cookie will transmit data about your browsing history to that website. Persistent cookies are sometimes called tracking cookies for this reason. Secure cookies can only be sent over a secure or encrypted connection, so they reduce the risk of cookie theft by those who may be listening on your connection. Zombie cookies are cookies that even when deleted will be reset by a locally stored script.

First-party cookies are cookies set by the website you are currently browsing. Third-party cookies are set by other domains and typically show up on websites that show advertising content (such as Facebook, Twitter, etc.). They open up the potential for tracking your browsing history in order to figure out what ads to show you.

3 Education and Accessibility

Of course, cookies are just a small part of a much bigger picture when it comes to privacy concerns. When approaching our tool, we wanted to give the user some control over their data privacy without bogging them down in too much confusing detail. At the same time, giving the user easy-to-understand agency over their data privacy might spark an interest in the subject in general.

The application's first priority is to simplify cookies in a way that makes them both readable to the common eye, yet provide enough technical information allow for users with a wide range of technical prowess. In order to account for this range of user skill, we would like to give users access to both the raw cookie information and a visualization of the user's cookies.

We also wanted to create a tool that would be useful to users that have a stronger background in cookie functions and data privacy.

4 Use Cases

When designing this application, we considered the needs of a person concerned with their ad content, a parent concerned about what data about their child is public, and a student majoring in Computer Science who wants to see third-party connections that appear after visiting particular sites. We thought about user goals related to removing cookies (even difficult ones), managing their data privacy by visualizing where standard and third-party connections existed, and better tailoring a user's ad experience to meet their needs or preferences.

Suppose a person sees an advertisement that is set by a third-party cookie that offends them. We wanted the user to be able to open something quickly, ascertain which cookie(s) are potentially generating the ad content, and delete it quickly. If they want to learn more about third-party connections they can easily navigate to visualize their cookies, but it's not necessary. The process should be smooth and simple.

We'd like that, upon first downloading and opening the application, they are immediately directed to a welcome and tutorial page that gives some basic information about different kinds of cookies, their functions, and how they relate to data privacy. Upon learning more about why our tool could be of use to them, they could choose to see their cookie connections and tailor their cookies as they like.

It may also be the case that a user has a stronger base of knowledge and is interested in seeing exactly which cookies are connected to which domains and delete cookies based on the visualization information. We wanted to make sure it was possible to quickly be able to identify the offending cookies and then delete them in a tabular interface.

5 What is CookieJar?

After completing our research about cookies, existing applications that work with cookies, and the current state of public interest and understanding of privacy, we built our Chrome extension. Over the course of several months, we steadily increased functionality and improved our design, resulting in an application that is easy to navigate and reliable for users. We named our application CookieJar not only as a play on words but also because it was an appropriate title for representing the relationship between cookies and their domains.

There are several components to our application that streamline the user's ability to work with their online cookies. When the user clicks the extension button at the top right of their browser window, a popup appears. From this popup, the user is automatically shown cookies that are related to the site they are currently visiting. They can search, delete, and view more information about these cookies right from the popup. There is also a button that allows the viewer to open up a new tab, which immediately shows them an interactive visualization of their cookies. This graph has several buttons and filters to ensure easy navigability. Users can easily see a diagram of the sites that have the most cookies, their newest cookie, and the sites with the most connections and third party connections.

From this page, there are a few tabs giving the user more options for understanding their online profile. We included a Frequently Asked Questions Section for users who want to understand some of the basic information about cookies without having to navigate away from our application to use Google or Wikipedia. There is also a tab that has a table containing all of the user's cookies. This table is easily searchable and has expandable rows that show additional information about each cookie. Through these different functions, we believe we have created an application that services the diverse needs of an online community looking to better manage their online presence.

6 Implementation Choices

As we began development work, we were faced with many choices in our implementation. The first question we were confronted with was in regard to which outside libraries or APIs we wanted to integrate with. We first decided that we wanted to build a Chrome extension. Google Chrome is widely used and extensions are user friendly, even for those who are less technologically experienced. Furthermore, Google Chrome stores cookie data in a format that is readily accessible by Chrome extensions. We decided to take full advantage of the Google Chrome cookies API for these reasons. We also found three other libraries that we deemed extremely useful, which will be discussed later.

The next important decision we made was to create a popup that was easily accessible and would not interfere with browsing, but could still provide relevant information to a user. In this popup, we made a choice to show only cookies from the site a user is visiting. As a user browses the web, it is likely that they'll be interested in the cookies that are relevant to the sites they are visiting in that session. This popup eliminates the extra steps of having to search more carefully through their cookies. It also makes it extremely easy to delete cookies that are undesirable immediately after they are created, freeing up computer disk space. The user will also not have to simply periodically delete all cookies, a process which many people find inconvenient because their favorite sites may have cookies that they prefer to keep. On the popup, there is a button that says "Visualize my Cookies!" which users can click to navigate to the web application.

When the user first installs CookieJar, they will be automatically directed to the Frequently Asked Questions page. We determined from usability testing that this would be the best way to introduce users to our application. This page answers a few common questions about cookies and enhances understanding of the information in the graph and table. We decided to start the user on this page because it is very clear from the animated tabs at the top of the page how to navigate to the graph and table.

If it is not the user's first time using the extension, the visualization button on the popup will take them directly to a graph view of their cookies. This graph can be a little overwhelming to look at if the user has over a few hundred cookies, so we included a few different ways to make it easier to handle. For this graph, we decided to use a library called Cytoscape, which has a variety of implementation options and user-friendly animation. First of all, the cookies look like cookies and the domains look like cookie jars, a distinction we thought made the graph much easier to understand. By using checkboxes, the user can also decide to show only domains or only cookies, which we believe helps a user navigate a more cluttered graph. Through our research and testing we gained a very clear picture of what users might be interested in seeing on the graph, which informed our decision to include buttons to show the user their newest cookie, domain with the most cookies, and cookies with the most connections. On this page, the user can also use the search bar to find a specific cookie and the graph will zoom in on that domain. Another choice we made on this page was to include a tooltip with additional information that opens with a single click on an object in the graph. We did not want users to have to switch pages and search for a cookie in the table if they wanted more information and this was an incredibly user-friendly way to make information accessible.

For the table, we used a library called DataTables that helped us implement a few exciting features. The table made it easy to sort by name of cookie or by domain. The users could also search for cookies, which was one of our primary goals for the tables usability. It is very easy for users to delete all cookies associated with a certain domain, which we found to be one of the primary goals of our potential users.

Finally, we took our design considerations very seriously. Our application is organized so that users can navigate easily between the graph, table, and FAQ and use the popup at any point. We integrated with a framework called Bootstrap, which is a popular HTML, CSS, and JS framework for web development. After experimentation with color, organization, and a cookie logo, we settled on the application's look. We also included popup bubbles to answer user questions, which are animated and intuitive for all users. As a result, we have created a Chrome extension for easily tracking and visualizing Cookies to streamline user understanding of their online presence.

7 Usability Tests

As discussed in our description of our implementation choices, many of our decisions were results of our extensive usability testing. We acknowledge that our test users were mostly limited to Carleton College students and adults who were directly related to the developers working on this application, which was certainly a limiting factor. However, we worked to get a diverse group of users to try out our application at various points in the process.

We had users answer a few questions about cookies and their experiences with managing their online presence before they worked with our product in order to understand their previous knowledge. Next, we had them complete a few simple tasks. A few examples include opening the popup and deleting a cookie, navigating the graph and finding their newest cookie, and searching the table for a specific domain. The FAQ was in various stages, as it took some time to evolve into a useful resource as we worked to understand what users wanted to know. We then asked the test users questions about cookies to see what new knowledge they had gained about cookies from using the application. Finally, we asked them about individual features to better understand what they could see themselves using regularly if they had this extension on their own browser. These tests taught us a great deal about what people wanted to be able to do with their cookies more easily and drove a great deal of our development work.

References

Essaid, Rami. "Internet Privacy Is The Wrong Conversation." *TechCrunch.* TechCrunch, 26 Apr. 2015. Web. 21 Oct. 2015.

Heddings, L. “Warning: Your Browser Extensions Are Spying On You.” *HowToGeek.* HowToGeek, 20 Jan. 2014. Web. 12 Oct. 2015.

Kobie, Nicole. “Why the Cookies Law Wasn’t Fully Baked and How to Avoid Being Tracked Online.” *The Guardian.* The Guardian, 19 Mar. 2015. Web.

Nissenbaum, H. “A Contextual Approach to Privacy Online.” *American Academy of Arts and Sciences.* 140.4, 32. 2005. Web. 22 Sep. 2015.

Reilly, R.B. “The Cookie Is Dead. Here’s How Facebook, Google, and Apple Are Tracking You now.” *VentureBeat.* N.p., 6 Oct. 2014. Web. 12 Oct. 2015.

Singer, Natasha. ”Sharing Data, but Not Happily.” *The New York Times.* The New York Times, 04 June 2015. Web. 22 Oct. 2015.

Trepte, S., et al. “Do People Know About Privacy and Data Protection Strategies? Towards the ‘Online Privacy Literary Scale.’” *Privacy Online: Perspectives on Privacy and Self-disclosure in the Social Web. Heidelberg: Springer-Verlag*, 2011. 333-65. Web. 23 Sep. 2015.

Turow, J., Hennessy, M., Draper, N. “The Tradeoff Fallacy: How Marketers Are Misrepresenting American Consumers And Opening Them Up to Exploitation.” *Annenberg School for Communication.* N.p., Jun. 2015. Web. 9 Sep.2015.