Chapter 2

Key Technologies with Implications for Privacy

Encryption, Analytics, and Advertising Tracking

Need for Encryption of Websites

To be consistent with library privacy policies and values, the interactions of how persons use library-provided resources must be protected from access by any third party. When communication takes place over a network, especially the internet, it is possible for unknown parties to intercept the data. Interception or eavesdropping can take place on wireless or wired networks and can be opportunistic, targeted, or widespread. Tools for intercepting communications on local networks are readily available and inexpensive. More sophisticated surveillance equipment may be inserted into internet infrastructure to gain more widespread access. This vulnerability to capture of network communications is well known and addressed through well-established encryption techniques. Encryption of web traffic, implemented through the HTTPS protocol, ensures that the contents of the transmission cannot be viewed even if the communication stream is captured.

Given the possibilities for interception and eavesdropping, it must be assumed today that any information transmitted on the web can be captured. The contents of captured communications can be easily accessed when they are transmitted without additional protection. Only with strong encryption technologies can information transmitted across networks be considered private. Encryption does not prevent others from intercepting communications, but it ensures that no one other than the sender and receiver can view the contents and that the contents have not been altered.

The "Policy to Require Secure Connections across Federal Websites and Web" issued by the Chief Information Officer of the Office of Management and Budget of the US federal government mandates the use of HTTPS on government websites and provides a concise summary of the dangers of using HTTP: "The American people expect government websites to be secure and their interactions with those websites to be private." And later: "The unencrypted HTTP protocol does not protect data from interception or alteration, which can subject users to eavesdropping, tracking, and the modification of received data. The majority of Federal websites use HTTP as the as primary protocol to communicate over the public internet. Unencrypted HTTP connections create a privacy vulnerability and expose potentially sensitive information about users of unencrypted Federal websites and services. Data sent over HTTP is susceptible to interception, manipulation, and impersonation. This data can include browser identity, website content, search terms, and other user-submitted information."[1]

HTTPS for Identity Validation

The use of HTTPS also confirms the identity of the website. It is essential that visitors be able to confirm that any website is legitimate and is not being spoofed. The digital certificates used to encrypt the transmission from the site also include authoritative information on the organization to which the certificate was issued. Digital certificates are issued by trusted certificate authorities that validate the ownership of the certificate. To establish a secure connection, a valid certificate must be installed in the web server, and the ownership embedded in the certificate must match its domain. Any mismatch will produce an error and the page will not be secured. Visitors to the website can inspect the certificate used for an

HTTPS site to confirm that the site belongs to the expected organization.

Low Threshold of Difficulty and Expense

The means to protect communications on the web are readily available and inexpensive. Any reasonably current web server software can be configured to encrypt the content it publishes. Once the website has been configured to deliver pages with HTTPS instead of HTTP, it uses a suite of protocols for encryption technologies, including TLS or Transport Layer Security, that cannot be decrypted while the data traverses the internet.

In order to enable HTTPS on a web server, the organization must obtain a digital certificate. These certificates are issued through a "certificate authority" and come in different categories. These certificates differ in the level of validation performed for the organization and its right to use the domain:

- **Extended validation:** The certificate confirms the organization's exclusive right to use the domain and performs an extensive review of the organization details relative to official business records. Sites with this type of certificate will present the name of the organization in the URL bar of most browsers along with the indicator that the site is encrypted using HTTPS.
- **Organization validated:** The certificate authority confirms the organization's right to use the domain. If properly validated, the organization's name will be shown when the user views the details of the certificate in the browser. For sites with this type of certificate, the URL bar of the browser indicates that the site is encrypted using HTTPS.
- **Domain validated:** The certificate authority confirms the organization's right to use the domain but does not require extensive documentation regarding the organization. For sites with this type of certificate, the URL bar of the browser indicates that the site is encrypted using HTTPS.

Certificate authorities will charge higher fees for certificates requiring more extensive organizational vetting and validation. These costs currently are about $25 per year for domain validated certificates; $75 for organization validated; and $400 per year for extended validation. Wild card certificates that support multiple subdomains will also involve additional fees.[2]

The nonprofit initiative Let's Encrypt provides free digital certificates to any organization. Let's Encrypt has developed a method to automatically install, configure, and renew certificates with minimum expertise or effort. While these certificates enable encryption, they do not provide the higher level of organizational validation available through traditional certificates.

Let's Encrypt
https://letsencrypt.org

Another category of certificates are those issued by the organization itself and not through a certificate authority. These self-signed certificates can be used for basic encryption, but do not provide any assurance that the website is legitimate. These certificates are typically used for testing and will trigger a warning on most web browsers.

Sites without a digital certificate cannot encrypt pages with HTTPS and will be limited to the HTTP protocol, which delivers pages as viewable text. Again, this option does not meet the basic requirement for privacy for a library website.

Advancing to HTTPS Everywhere

The web has been in the process of transition from its initial deployment based on HTTP to universal implementation of HTTPS for more than a decade. In the earlier phases of the web, the HTTPS protocol was available, but its use was targeted to specific tasks involving sensitive information, such as the entry of credit card numbers or passwords. At that time, the process of setting up HTTPS on web servers was more complex and the additional computations needed for encryption were substantial. With current web server hardware and software, the overhead for implementing HTTPS is negligible. Today it is expected that all web traffic should be carried with HTTPS encryption. All major commercial destinations and social networks have switched entirely to HTTPS.

Google has played a major role in the transition to HTTPS. Given its dominance in search, web browsers, and general web services and infrastructure, its policies and practices have a massive impact on the broader sphere. Google Chrome, for example, currently has 63.3 percent of the market share for web browsers, with Firefox a distant second at 9.5 percent.[3]

Google has been exerting increasing pressure to entice websites to make the switch to HTTPS. This pressure comes in the form of warnings issued through its Chrome web browser and through its ranking of search results. All web browsers present some type of indicator when a site has implemented HTTPS. From the earliest phase of the web, users have been aware that they must check for this positive indicator of encryption before entering credit card information, passwords, or other sensitive information.

Pages not encrypted were given a neutral status indicator. Following a generous period of advance notice, Google changed its neutral treatment of non-HTTPS sites to a conspicuous negative indicator. Beginning in July 2018, web pages not encrypted with HTTPS via a valid digital certificate have been flagged as not secure (figure 2.1). Clicking on the information indicator presents this text: "Your connection to this site is not secure. You should not enter any sensitive information on this site (for example, passwords or credit cards) because it could be stolen by attackers."

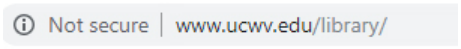

Figure 2.1. Example of Google Chrome unsecure warning

This treatment contrasts with that given pages using HTTPS (figure 2.2). Clicking on the lock icon brings up this text: "Connection is secure. Your information (for example, passwords or credit card numbers) is private when it is sent to this site."

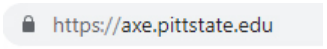

Figure 2.2. Example of secure website on Google Chrome

Meeting the HTTPS Deadline

This date of July 2018 was generally regarded as a deadline by which responsible organizations had to implement HTTPS or face the repercussions of their content being flagged as unsecure.[4] As demonstrated by the data collected in support of this report, most of the public and academic libraries in the United States have met this deadline, though a substantial portion remain not in compliance with this essential requirement.

Scott Helme, an internet security researcher, provides useful information demonstrating the progress made in the transition to HTTPS on the general web. Data describing the use of HTTPS by the top one million websites as tracked by the Alexa Internet service shows a steady climb from 6.71 percent in August 2015 to 58.44 percent in February 2019.[5] The data in this report related to the transition to HTTPS in libraries is roughly on track with the broader web trend seen in the Alexa statistics.

Alexa
https://www.alexa.com/

HTTPS and Only HTTPS

In addition to implementing HTTPS, it is also important to implement mechanisms to ensure that site visitors are not intentionally or accidentally directed to an unsecure version. Websites should be configured to always direct users to the secure version using HTTPS, even if they come to the site with a link coded with HTTP. The website should automatically redirect HTTP to HTTPS, providing protection even if the user types in http:// or comes in through an outdated version of the URL. The "HTTP Strict Transport Security (HSTS)" standard describes a protocol that can be implemented in web servers to implement comprehensive use of HTTPS.[6]

Even if a website has been configured to enable HTTPS, if it allows its pages to be accessed via HTTP, it should be considered vulnerable from the perspective of user privacy. In addition to gathering data on the number of websites for public and academic libraries implementing HTTPS, this report also assesses whether these sites implement the expected redirection behavior to ensure that HTTPS is always used.

Challenges in Implementing HTTPS

Even with the low threshold for the technical implementation, a number of challenges can hinder an organization from making the transition to HTTPS. These challenges often relate to dependencies on external resources that do not support HTTPS. In order to be validated as secure, the page, as well as any links or embedded content, including images, style sheets, and JavaScript libraries, must be delivered via HTTPS. All links to external web pages and services must also be HTTPS. If any HTTP links or content is detected, browsers will issue a conspicuous error message warning of unsecure content mixed into the page.

In the library context, avoiding these mixed content errors means that the library catalog, discovery services, and all information resources linked to from the site must be available via HTTPS links. If any of these vendors cannot conform to this requirement, the library may have to delay its own implementation of HTTPS. Since libraries' websites often exist to provide access to information resources to their patrons, ensuring comprehensive use of HTTPS throughout their portfolio of database and content products can be an extensive process. The switch to HTTPS on the library's main website may also need to be coordinated with similar changes to the online catalog, institutional repositories, blogs, or other local resources.

Some libraries may also opt to make the transition to HTTPS as part of a redesign of the library's website or a move to a new hardware or software platform. When part of a larger project, the implementation of

HTTPS may take longer than if it were an isolated task.

Libraries may also be limited by the technologies implemented by their parent institution. If the library web presence operates within the website of a university or local government, it may not have the means to make this change independently. For some libraries, working with the institutional infrastructure may mean a quicker adoption of more secure technologies.

Mandate for Libraries

Libraries have generally lagged behind the commercial sphere in the transition from HTTP to HTTPS. Despite the values-driven necessity of providing a secure and private environment for accessing library content and services, some libraries may not be well informed regarding these vulnerabilities or may lack the technical expertise or the personnel resources to implement these needed changes.

Not implementing HTTPS places libraries in an unfortunate position of their websites being flagged as not private or secure, despite their role in providing access to trusted and vetted resources. Sites implementing HTTPS will receive no such warnings, regardless of the nature of the content they publish. Although technical security and privacy configurations and the quality of content curated are entirely distinct issues, these distinctions may not be well understood by all persons. The reputation of a library can therefore be diminished if it does not attend to these critical technical details.

Analytics and Advertising Networks

Privacy concerns extend beyond configuring a server to correctly implement HTTPS encryption. Although the content of pages delivered through HTTPS cannot be viewed or altered, many other practices can compromise privacy. Even on encrypted pages, site managers can compromise the privacy of their users by including scripts or widgets that provide data to external entities. These tracking mechanisms may be positioned by the providers as innocuous but need to be well understood by organizations with heightened concerns for privacy such as libraries.

The ALA statement *Privacy: An Interpretation of the Library Bill of Rights* also addresses this topic: "Libraries should not monitor, track, or profile an individual's library use beyond operational needs. Data collected for analytical use should be limited to anonymous or aggregated data and not tied to individuals' personal data."[7]

This report studies two basic categories of tracking agents that might be added to library websites. Those related to analytics pass information regarding the use of the website to an external server, enabling website managers to observe patterns of use. The other category involves making connection to advertising networks, leaving the possibility for intermingling library sites with a presumption of privacy and commercial networks based on extraction and sharing of personal data.

Measuring Website Use through Analytics Services

Libraries, like other types of organizations, have a strong interest in measuring the use of their websites. In addition to gaining a general understanding of a site's level of use, an organization can use sophisticated analytics tools to help identify problems on the site and to inform improvements in design and functionality. Website analytics tools can take two different approaches.

- **Server log analysis:** One category is based on processing the log files produced by web servers that record each resource requested. This approach works without involvement of any external resource but may involve a higher level of difficulty. Log-based analytics require access to the internal system resources of the web server, which may be difficult in some organizations where multiple sites operate through the same server. These products also may involve the installation and configuration of the analytics software. This model of analysis was common during the earlier phase of the web but has declined due to the popularity of Google Analytics. Some organizations will use both server log analysis tools and analytics based on page tagging to get a more complete view of the use of their site. Server log tools, for example, can capture access by search indexing crawlers, which represent a substantial portion of server load, though not actual visitor activity.
- **Page tagging:** The other model relies on sending data to an external analytics service as each page is accessed. The website manager places a snippet of code on each page, usually through a standard inclusion component. The analytics tag would be included in much the same way as headers, navigation, JavaScript libraries, or style sheets to provide consistent branding and layout.

One of the topics addressed in this report relates to the use of analytics for library websites. The data collected for this library privacy study demonstrates that a large percentage of libraries use Google Analytics, a free service for measuring use and for optimizing the usability of websites. This service relies on websites

transmitting detailed usage data to Google. Libraries need to assess whether the use of this service falls within what is allowed by professional values and by the privacy policies of each library organization. From a technology perspective, we can observe that the service involves sending data describing patron information-seeking activities to a third party, which must be trusted to limit the way in which that data is used.

Google Analytics and other services from Google are designed to directly or indirectly support the company's business interests. Google earns most of its revenue through advertising. According to Statista.com, in 2018, Google reported total revenue $136 billion; of that, $120 billion came from advertising.[8] The basis of Google Analytics in the commercial advertising ecosystem warrants careful analysis to ensure that its use remains consistent with the library's privacy policies.

Google Analytics has become the dominant tool used for assessing the use of websites. As shown in the data collected for this study, it is used by all types of organizations, including libraries. Although some libraries use other tools for use statistics and analytics, this report focuses on Google Analytics given its widespread use among libraries.

Google Analytics

The implementation of Google Analytics involves two tasks, the creation of an administrative account and the inclusion of a snippet of JavaScript on each page. Each website, or "property," configured through the Google Analytics administrative console is assigned a unique identifier, which must be included in the JavaScript snippet.

Once Google Analytics is activated, each time a page is accessed on the site, information will be transmitted to Google's servers to enable detailed analysis and measurement of use patterns. The data transferred does not necessarily contain personal information about the individuals visiting the website, but it does include detailed information regarding the resources used on the site. In some cases, the data could include information regarding the topics or specific items searched for or accessed on the site. That information can be conveyed on the query string of a URL as one of the elements tracked. All resources accessed within a session are tied together through a unique identifier Google Analytics assigns and records in a browser cookie. This identifier is not associated with a specific individual through the data collected within Google Analytics.

Depending on the circumstances and interpretation, the IP address of a website visitor can be considered a personally identifiable data element. The GDPR (General Data Protection Regulation) framework of the European Union, for example, considers the IP address as personal information in some contexts.[9] Depending on the way that IP addresses are assigned, there can often be a strong correlation between an IP address and a specific device and the individual using that device.

Multiple Tracking Code Options

The code snippets that a site manager places on a web page to enable Google Analytics have changed over time. Each of these options follows the same model of page tracking associated with the site's unique identifier, though with each new version additional features have been added.

The initial Google Analytics snippet (figure 2.3), generally referred to as the Classic version, was introduced prior to HTML version 5 and supported both encrypted and unencrypted transmission of data to the Google Analytics servers. Although this version of the tracking code continues to work, Google recommends that all new sites be configured with the newer Universal analytics code.

```
script type="text/javascript">
 var gaJsHost = (("https:" == document.location.protocol) ? "https://ssl."
: "http://www.");
 document.write(unescape("%3Cscript src='" + gaJsHost + "google-
analytics.com/ga.js' type='text/javascript'%3E%3C/script%3E"));
</script>
<script type="text/javascript">
 try {
 var pageTracker = _gat._getTracker("UA-3203647-3");
 pageTracker._trackPageview();
 } catch(err) {}
</script>
```

Figure 2.3. Original Google Analytics tracking snippet

The Universal version of the Google Analytics tracking snippet uses the analytics.js JavaScript library (figure 2.4). This version always encrypts data as it is transmitted to the Google Analytics servers and includes options for anonymization of IP addresses.

In addition to directly embedding the Google Analytics code snippet into each page, the organization can also use the Google Tag Manager, another free tool from Google. This tool can enable other services that rely on tracking codes in addition to Google Analytics. While it is possible for a site to use the Google Tag Manager and not use Google Analytics, this practice is not common. The presence of the Global Site Tag tracking code for Google Tag Manager is a very strong indicator for the use of Google Analytics for pages where the other Google Analytics tracking snippets are not detected (figure 2.5). It is also possible for both the Google Tag Manager snippet and one of the Google Analytics tracking codes to be present within a web page.

If the organization has deployed Google Analytics using the Google Tag Manager, it may not be possible to detect the presence of the tracking code when inspecting the source code for the page. The Google support documentation states that only the page owner can see the tags activated through the Google Tag Manger console (see figure 2.6).[10] Browser plugins, such as Ghostery, will be able to detect the use of Google Analytics for these sites.

The default tracking code snippet currently presented through the Google Analytics console takes the form of the Global Site Tag rather than the Universal Analytics previously recommended.

Because of privacy concerns, Google Analytics includes a feature to anonymize IP addresses before they are recorded. This anonymization is essentially a truncation of the address so that it retains some useful information regarding the general location of the user. IP address anonymization can be specified in the Google Analytics JavaScript snippet, or it can also be configured in the administrative console of the Google Tag Manager.[11]

Google Analytics also includes a feature through which specific users can be tracked. This User-ID

```
<script>
(function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function()
{
 (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new
Date();a=s.createElement(o),

m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(
a,m)
 })(window,document,'script','https://www.google-
analytics.com/analytics.js','ga');
 ga('create', 'UA-32191981-1', 'auto');
 ga('set', 'anonymizeIp', true);
 ga('send', 'pageview');
</script>
```

Figure 2.4. Universal version of Google Analytics tracking snippet

```
<!-- Global site tag (gtag.js) - Google Analytics -->
<script async src="https://www.googletagmanager.com/gtag/js?id=UA-
000000000-1"></script>
<script>
 window.dataLayer = window.dataLayer || [];
 function gtag(){dataLayer.push(arguments);}
 gtag('js', new Date());

 gtag('config', 'UA-32191981-1'),{'anonymize_ip': true};
</script>
```

Figure 2.5. Global Site Tag tracking code for Google Tag Manager

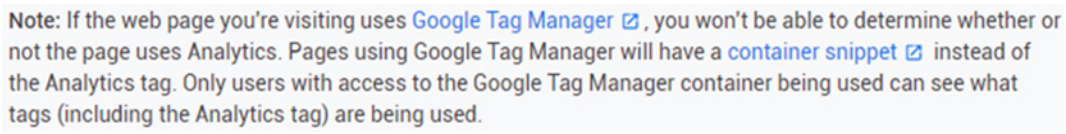

Figure 2.6. Explanation from Google Analytics support page.(Source: "Check if a Web Page Uses Analytics," Google Analytics Help, accessed July 24, 2019, https://support.google.com/analytics/answer/1032399?hl=en)

feature must be specifically configured in the Google Analytics Console, including a step agreeing to the associated privacy policy. The tracking code of the site is also updated to include the unique user identifier for the person accessing the page, which could be provided for those who have logged into the site. Activation of this feature would be inconsistent with the privacy policies of most libraries since it not only creates nonanonymized records of patron information-seeking activities, but also shares that data with Google.

When Google Analytics is used, all data relating to website use is transmitted to Google's servers. That data is used for reporting through the organization's Google Analytics account, but it may also be part of broader analytics or data mining. The Google Analytics console offers options regarding how Google employees may access the organization's data (figure 2.7). Enabling access to either Google's marketing specialists or all its sales personnel would seem inconsistent with general library practices regarding the treatment of patron use data.

Google also includes a variety of features in Google Analytics that allow an organization to enable linking with one or more Google Ads accounts. These features are useful to organizations that subscribe to Google's advertising services but would rarely be used on a library website, which usually does not offer advertising. Enabling these features allows collection of additional data and may also trigger collection of personally identifying information, such as for site visitors who are logged into a Google account. Figures 2.8 and 2.9 show the selections within the Google Analytics console that enable advertising features and extended data collection.

Advertising Networks and Social Media

The intermingling of library websites with advertising networks can introduce concerns for privacy. Analytics services involve transmission of data that may contain information-seeking activities of website visitors. Tracking codes and cookies for ad networks and social media sites represent a larger concern in regard to the privacy of patrons who access library websites. These organizations have strong interest in collecting or using information related to personal identity, interests, and past online interactions for targeting ads. In some cases, the tracking and interactions may be anonymized, and in others any current active logins, previously deposited browser cookies, or other mechanisms enable personal identification.

ProPublica has done research on the way that advertising and social networks track personal data. As far back as 2016, ProPublica reported that Google no longer separates information that it has about an individual through Gmail and other accounts and other browser data collected through DoubleClick: "The practical result of the change is that the DoubleClick ads that follow people around on the web may now be customized to them based on your name and other information Google knows about you. It also means that Google could now, if it wished to, build a complete portrait of a user by name, based on everything they write in email, every website they visit and the searches they conduct."[12]

Personal information is widely shared in the advertising ecosystem. This sharing of data across organizations can be easily observed. A search for a product on Amazon.com will cause ads for that product or similar ones to appear on Facebook and other sites. This "retargeting" mechanism is widely used by web destinations to show relevant ads based on browser

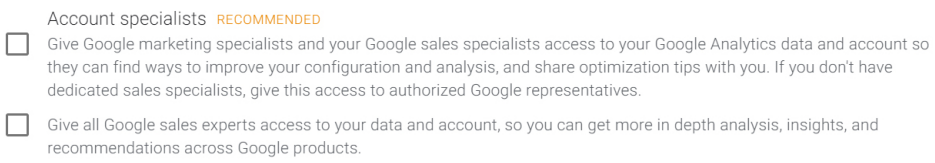

Figure 2.7. Google Analytics options for access to data by its personnel.

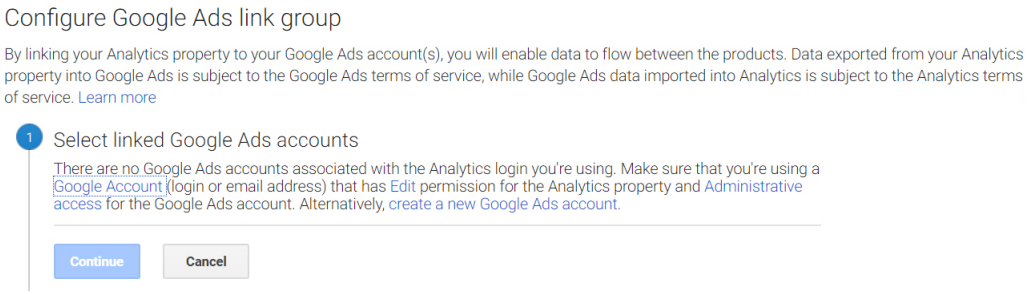

Figure 2.8. Configuring Google Ads link group

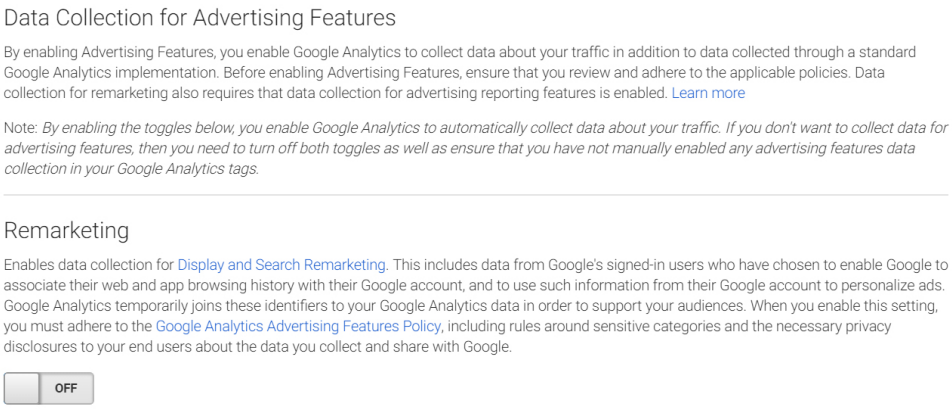

Figure 2.9. Data collection for advertising features and remarketing

history, third-party cookies, and other mechanisms.

The types of data and the mechanisms for sharing it among organizations and websites in the advertising ecosystem are complex and ever-changing. Libraries opting to enable ad-related tracking technologies will want to carefully investigate any possible external exposure of personal information or browsing history as individuals visit their websites and use their resources. Any scenario that allows content items searched for or viewed on a library website to later appear as ad suggestions on another site would not be consistent with library privacy values or most library privacy policies.

The advertising ecosystem continues to evolve toward ever more precise targeting capabilities, extending deeper into the realm of personally identifying information. One recent technique, seen with Google and Facebook, involves the concept of custom audiences. This technique involves the direct linking of known user information, such as from an organization's customer relationship management system or authentication service. In the library context, using these types of services would not be consistent with privacy protection since it involves sharing library patron data in bulk with an external organization:

> Recently, data brokers such as Facebook and Google have introduced a new feature on their advertising interfaces: custom audiences. Instead of creating audiences based on user attributes, advertisers can now upload personally identifying information (PII) about specific users; the platform then locates matching accounts and creates an audience consisting of only these users. The advertiser can then use this audience when placing ads, thereby showing their ads only to the specific users whose information they uploaded. For example, a small business may know the names and addresses of its customers; using custom audiences, the business can upload this information to Facebook, and then target these users with advertising directly. The custom audience feature has proven popular with advertisers: it allows them to directly select the users to whom their ad is shown, as opposed to only selecting the attributes of the users.[13]

In this study, a cursory screening is performed

to determine which websites may include tracking agents related to advertising or social networks. These trackers are not easily identified by the source code of the websites. A next phase of enhancements to the parsing scripts is planned that can more accurately identify these trackers.

Notes

1. Tony Scott, "Policy to Require Secure Connections across Federal Websites and Web Services," memorandum, Office of Management and Budget, June 5, 2015, https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2015/m-15-13.pdf.
2. Examples from Register accessed July 26, 2019, register.com.
3. NetApplications, "Browser Market Share," June 2019, https://netmarketshare.com/browser-market-share.aspx.
4. Emily Schechter, "A Secure Web Is Here to Stay," *Google Security Blog*, February 8, 2018, https://security.googleblog.com/2018/02/a-secure-web-is-here-to-stay.html.
5. Scott Helme, "Alexa Top 1 Million Analysis—February 2019," *Scott Helme* (blog), March 11, 2019, https://scotthelme.co.uk/alexa-top-1-million-analysis-february-2019/.
6. J. Hodges, C. Jackson, and A. Barth, "HTTP Strict Transport Security (HSTS)," Internet Engineering Task Force, proposed standard, request for comments, RFC 6797, November 2012, https://tools.ietf.org/html/rfc6797.
7. American Library Association, *Privacy: An Interpretation of the Library Bill of Rights*, (Chicago: American Library Association, 2002, amended 2014 and 2019), http://www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy.
8. See J. Clement, "Google's Revenue Worldwide from 2002 to 2018," Statista, last edited May 22, 2019, https://www.statista.com/statistics/266206/googles-annual-global-revenue/.
9. See Andrew Cormack, "IP Addresses, Privacy and the GDPR," *Jisc Community* (blog), April 4, 2018, https://community.jisc.ac.uk/blogs/regulatory-developments/article/ip-addresses-privacy-and-gdpr.
10. See Google Analytics Help, "Check if a Web Page Uses Analytics," accessed July 16, 2019, https://support.google.com/analytics/answer/1032399?hl=en.
11. See Cormack, "IP Addresses, Privacy and the GDPR."
12. Julia Angwin, "Google Has Quietly Dropped Ban on Personally Identifiable Web Tracking," ProPublica, October 21, 2016, https://www.propublica.org/article/google-has-quietly-dropped-ban-on-personally-identifiable-web-tracking.
13. Giridhari Venkatadri, Athanasios Andreou, Yabing Liu, Alan Mislove, Krishna P. Gummadi, Patrick Loiseau, and Oana Goga, "Privacy Risks with Facebook's PII-Based Targeting: Auditing a Data Broker's Advertising Interface" (paper, IEEE Symposium on Security and Privacy, San Francisco, CA, May 20–24, 2018), https://doi.org/10.1109/SP.2018.00014.