Best Practices for Online Service Providers

June 2008

ELECTRONIC FRONTIER FOUNDATION
eff.org

Table of Contents

Introduction

Online service providers (OSPs) are vital links between their users and the Internet, offering bandwidth, email, web and other Internet services. Because of their centrality, however, OSPs face legal pressures from all sides: from users, industry, and government. As an intermediary, the OSP finds itself in a position to collect and store detailed information about its users and their online activities that may be of great interest to third parties. The USA PATRIOT Act and other recent legislation have also provided the government with expanded powers to request this information. As a result, OSP owners must deal with requests from law enforcement and lawyers to hand over private user information and logs. Yet, compliance with these demands takes away from an OSP's goal of providing users with reliable, secure network services. In this paper, EFF offers some suggestions, both legal and technical, for best practices that balance the needs of OSPs and their users' privacy and civil liberties.

1. Are you an OSP?

If you think you might be an OSP, you probably are. As defined by the Digital Millennium Copyright Act (DMCA)[1], an OSP is any "entity offering the transmission, routing, or providing connections for digital online communications" or any "provider of online services or network access, or the operator of facilities therefor." The Electronic Communications Privacy Act (ECPA) defines two subcategories of OSPs: "electronic communication services"[2] and "remote computing services."[3] Access to users' information under ECPA is determined in large part by whether your OSP's services fit within one (or both) of these subcategories. As a general rule, email and connectivity services would be electronic communication services, while website hosting would be considered a remote computing service. This means that virtually *any* website or access intermediary, not just established subscriber-based businesses, can be considered an OSP under the law. Indeed, even individuals may be "accidental OSPs" if they set up WiFi access points to share Internet connectivity with friends and neighbors.

2. How OSPs Can Develop Sane Network Policies to Protect Themselves From Legal Liability and Respond to Subpoenas and Court Orders

A key strategy is to minimize the amount of information OSPs collect and store in the first place. Unless they are in a specially regulated industry (finance or health care, for example), no law requires OSPs to collect and store information about their users. In some cases, there are restrictions on what information may be collected.[4]

This means that OSP owners and operators are generally free to develop and implement reasonable data retention policies.[5] Our suggestions for these policies, elaborated below, are not legal advice and are for informational purposes only. If you have any specific questions or concerns about your OSP, please consult an attorney.

---

1 *See* **http://www4.law.cornell.edu/uscode/17/512.html**.

2 "any service which provides to users thereof the ability to send or receive wire or electronic communications . . ." **http://www4.law.cornell.edu/uscode/18/2510.html**.

3 "the provision to the public of computer storage or processing services by means of an electronic communications system." **http://www4.law.cornell.edu/uscode/18/2711.html**.

4 *See e.g.* Children's Online Privacy Protection Act of 1998 ("COPPA"), which governs online collection and use of information from children under the age of 13. 15 U.S.C. §§ 6501 *et seq*.

5 In certain circumstances, such as when there is litigation pending against the OSP, the OSP may be required to retain potential evidence. If you are facing litigation, please consult an attorney regarding data retention.

3. Scope of the EFF Best Practices for Online Service Providers

This EFF Best Practices white paper is focused on our recommended practices for protecting user privacy for United States OSPs, based upon U.S. law. International and multinational OSPs should consult with counsel to determine the legal landscape in the various jurisdictions in which they operate.

OSPs can face many other legal issues beyond user privacy, from DMCA takedown requests to defamation claims to adult materials. While these are outside the scope of this paper, EFF recommends that OSPs review the EFF Bootcamp materials, **http://www.eff.org/bootcamp/**, which provides the basics on a number of key legal issues for Web 2.0 companies. We also recommend reading EFF's Legal Guide for Bloggers, **http://w2.eff.org/bloggers/lg/**, which provides a basic roadmap to the legal issues one may confront as an online publisher.

4. Summary of Recommendations

a. Develop procedures for dealing with legal information requests and providing notice to users.

b. Work with both attorneys and engineers to develop a privacy policy that fits your OSP's practices.

c. Collect the minimum amount of information necessary to provide OSP services.

d. Store information for the minimum time necessary for operations.

e. Effectively obfuscate, aggregate and delete unneeded user information.

f. Maintain written policies addressing data collection and retention.

g. Enable SSL as much as possible throughout your site to secure users' information and communications.

h. Understand threats to the security of sensitive information and communications on your systems, and mitigate them appropriately.

i. Follow best-practice principles for the use of cookies on your site.

j. Insist that the OSPs and other service providers you work with observe these best practices, too.

Legal Issues with Requests for User Data or Transactional Information

1. Subpoenas, Warrants, and Court Orders

When law enforcement officers conduct civil or criminal investigations, they must obtain subpoenas, warrants or court orders to retrieve personal information from OSPs.

The government may obtain basic subscriber information[6] with only a subpoena, but generally needs a warrant or a court order for more detailed records. These court orders might request the identity of the user, email message content, visited URLs, search queries, or any other kind of recorded information. While the ECPA requires OSPs to disclose information in response to a legal process, it also prohibits certain disclosures without a proper request. For example, the ECPA prohibits an electronic communications service provider from producing the contents of electronic communications (*i.e.*, the body of an email message or arguments in a URLs query string), even if served with a subpoena, except in limited circumstances.

Thus, the OSP must evaluate the legal process carefully before retrieving the information and furnishing it to law enforcement. Often, this takes a great deal of time and resources, and the OSP should consult an attorney.

An OSP can keep its costs and risks down by setting clear policies about data retention. Under ordinary circumstances, there are no laws that require OSPs to retain personally identifiable information (PII) or activity logs about users, unless this information is subject to other government regulation (such as financial transactions) or the OSP has received a backup preservation request from the government.[7] EFF believes that PII and activity logs about users should be kept only so long as it is operationally necessary. (We explore this issue in more detail in the technical section below.)

OSPs cannot be forced to provide data that does not exist. EFF suggests that OSPs draft an internal policy that states that they collect only limited information and do not retain any logs of user activity on their networks for more than a few weeks. If a court order requests data that is more than a few weeks old, the OSP can simply point to the policy and explain that it cannot furnish the requested data. Likewise, if unnecessary PII is regularly deleted, the OSP cannot supply what it does not retain. This saves the OSP time and money, while also providing the OSP with sufficient data for its own administrative and business purposes.

Several states currently require OSPs to give notice to users prior to turning over PII. Similar laws may soon be enacted in other states. Giving notice may also protect the OSP against lawsuits from users. Put a data retention policy and procedure for responding to subpoenas, court orders and warrants in place now that will protect your users' privacy and your own legal liability.

a) Account Information (PII)

Civil or criminal subpoenas may be issued for identifying information called "subscriber information." This includes personally identifiable information like name, address, phone number and any other personal information that the OSP has collected from the user. Subpoenas for subscriber information are usually aimed at uncovering the identities of people who are posting anonymous comments.

It is important to remember that there are always different sides to a dispute, and by creating a procedure to deal with external legal demands for information without including procedures to consider the customer's rights, OSPs may be unfairly penalizing their own customers. A typical scenario would be someone posting negative comments about a company. The company lawyer sends a subpoena for subscriber information about the poster, perhaps to determine whether it is an employee who can be fired or sued. Sometimes, these demands

6 Such as the user's name, address, records of session times and duration, IP or other network address.

7 *See* **http://www4.law.cornell.edu/uscode/18/2704.html**.

are simply used as a form of harassment, without any sound legal basis or intent to follow through with the legal process. In many cases, once the user's identity has been forcefully revealed, the requesting company takes extra-legal action against the user by firing or taking other forms of retribution against him.

Another common civil subpoena is a DMCA "Subpoena To Identify Infringer," which requires an OSP that hosts allegedly infringing material to disclose "information sufficient to identify the alleged infringer . . . to the extent such information is available to the service provider." Unlike an ordinary subpoena, the DMCA subpoena does not require a lawsuit to be filed first, but it must be accompanied or preceded by a notification of alleged infringement that has specific requirements. However DMCA subpoenas only apply to OSPs that actually host a work, not ISPs that merely provide connectivity, such as in the case of peer-to-peer filesharing. DMCA subpoenas also only apply to claims of copyright infringement.

In other circumstances, individuals may request information about a particular user, complaining that the user has engaged in harassment or other bad acts. In such cases, the OSP may be sympathetic to the alleged victim and be tempted to provide the information directly. However, an OSP has no way to verify the truth of the story and providing this information without legal process could subject the OSP to liability from the user. The safest course is to require a subpoena or other legal process before providing user information to anyone.

Remember, Internet users have a right to anonymous free speech under the First Amendment, which includes the right to have a court rule on whether or not a legal claim is sufficient to require disclosure of the speakers identity. An OSP receiving one of these subpoenas should notify the user as quickly as possible before responding to it.[8] The user should be provided with a reasonable amount of time, such as 30 days, to respond before the OSP produces the requested information. This will give the user an opportunity to object to disclosure of his or her identity (technically, by filing a "motion to quash the subpoena"). If the user indicates his or her intent to file a motion, the OSP should wait until the court rules on the motion before producing.

b) Transactional Information and Activity Logs

The ECPA also provides protections against the disclosure of the "contents of a communication" and "a record or other information pertaining to a subscriber to or customer of" a covered OSP. Activity logs that show the behavior of a particular user would be, at minimum, records pertaining to that customer.

In some cases, transactional information provided by the customer would also be the more protected category of the "contents of a communication." For example, a court has held that the search terms a user transmits to an OSP to be processed into search results are the contents of a communication.[9] The contents of a communication generally may not be obtained with a civil subpoena to the OSP.

Special rules apply to providers of online video services. The Video Privacy Protection Act (VPPA)[10] provides additional rules protecting "information which identifies a person as having requested or obtained specific video materials or services"—*i.e.*, transactional records that show which video a particular customer watched. The VPPA also requires the destruction of such

8 On occasion, court orders to provide user information to the government may be accompanied with a request not to notify the user. In such circumstances, OSPs should consult with an attorney.

9 *In re United States for an Order Authorizing the Use of a Pen Register & Trap*, 396 F.Supp.2d 45, 49 (D. Mass. 2005).

10 **http://www4.law.cornell.edu/uscode/18/2710.html**.

records "as soon as practicable, but no later than one year from the date the information is no longer necessary."

OSP should require legal process before providing transactional information or activity logs to third parties, and consult with legal counsel before compliance.

2. Privacy Policies

As a matter of law and as a good information practice, OSPs should publish privacy policies describing what they do with the data they collect from consumers: what purposes they will use it for, how long they will keep it, whom (if anyone) they will share it with, and under what terms.[11] There is no model privacy policy that universally works for all or even most OSPs. Rather, each OSP is different, and should work closely with an attorney to craft a policy specific to its practices. The privacy policy also requires the close attention of the OSP's technical staff. Systems engineers are often in the best position to know exactly what the OSP collects and stores, and a privacy policy written by marketing and legal staff without an in-depth understanding of the technical details may ultimately be inaccurate.

As a general matter, though, a privacy policy should accurately and completely disclose the privacy practices of the site. Avoid using weasel words like "we will disclose your personal information where permitted by law," as they provide little information or assurance to the users. In drafting a privacy policy, OSPs should be mindful of the Federal Trade Commission's Fair Information Practice Principles, **http://www.ftc.gov/reports/privacy3/fairinfo.shtm**, though note that a policy can comply with the FTC's principles and still fail to strongly protect privacy. Once you've settled on a privacy policy, stick to its terms. An OSP that doesn't follow its privacy policy may violate state and federal consumer protection laws.

If you change your privacy policy, you should give notice to the users and make the prior versions available (with the dates those policies were effective). If the changes would be difficult to identify, OSPs can provide a tracked changes version or a written summary of the changes documents. If the changes significantly effect users privacy, the OSP should obtain customer consent to the new policy, and treat those customers who did not consent under the policy in place at the time of collection.

To the extent that you work with other OSPs, such as companies that serve advertisements on your site, you should require them to follow these best practices. After all, an OSPs' privacy practices are only as good as the sum of all the practices on its website. If you do share data with a third party that does not meet the same standards as your own privacy policy, you will not be able to give your users strong assurances.

3. International Considerations

While this White Paper is focused on U.S. law, we recognize that even relatively small Internet OSPs operate globally. If you are dealing with users in regions with strong data protection laws, such as Europe, you should investigate how privacy regulation there may interact with your internal practices. International law is outside the scope of this document, but it should be noted that the EU and United States have negotiated a "safe harbor" set of data principles that mesh well with standard best practices (including privacy notices, opt out procedures, and third

11 At least one state (California) requires certain commercial OSPs to publish privacy policies conspicuously on their sites describing their information practices. California Online Privacy Protection Act of 2003 (OPPA), Cal. Bus. & Prof. Code §§ 22575-22579.

party transfer rules) and will save you from liability under EU privacy law. Information on the safe harbor and self-certification is available at **http://www.export.gov/safeharbor/**.[12]

Technical Issues

Up until now, we have discussed EFF's recommendations for best practices to help OSPs minimize the cost of legal overhead. There is also a technical side to this issue. By being consumer-conscious about logging PII and activity logs, network administrators can proactively save company resources, ensure compliance with published privacy policies, and protect the privacy of their users at the same time. Upon receipt of a court order, OSPs are compelled by law to comb through their logs to extract the requested data using their own resources.[13] Thus, the cost of handling court orders scales proportionally with the retention of user traffic logs.

A general best practice to mitigate this problem is to log only enough information to maintain and upkeep the OSP's intended services—no more, no less. Logs should be stored for a minimal amount of time. The "correct" strategy for a particular OSP will depend on the services they provide to their users. We outline some possible strategies below.

1. Identifying and Minimizing Personally Identifiable Data

OSPs that have registered users will usually have account pages that contain the PII provided by the user. These PII locations are generally easy to identify, and the OSP can minimize the PII by limiting the questions asked in the sign up process to information necessary for their business. To protect the privacy of user account information, we recommend following the best practices for legal compliance in the section above.

However, user account pages are not the only location for PII and other sensitive information on an OSP's systems. OSPs should also pinpoint, on every server, all logs where PII is being recorded. It's important to remember that IP addresses and MAC addresses are crucial sources of identity-revealing information, and they are often requested in court orders. The most common locations for PII include:

- DHCP logs (IP address-to-MAC address assignments, session times)
- RADIUS logs (user name, IP address assignment, callback telephone number, session time, etc.)
- Web and FTP server logs (client IP address, files accessed, request time, query string, etc.)
- Web application servers and content management systems (which may store IP, file accesses, account logins, etc, separately from the web server itself)
- Email server logs (sender/recipient addresses, message date and time, relay hostnames, etc.)
- Firewall and IDS logs (IP addresses, packet payloads, date and time of connections, protocol used, etc.)
- User contact information databases (mailing address, phone number, billing information, etc.)

12 See also **http://ec.europa.eu/justice_home/fsj/privacy/index_en.htm**,

13 In some cases, OSPs can seek reimbursement for the costs of compliance. *See, e.g.*, **http://www4.law.cornell.edu/uscode/18/2706.html**. However, reimbursement may not capture all the costs associated with legal compliance.

PII can crop up in surprising places. For example, web server logs often record the "HTTP referer" information supplied by a web browser's request for a web page. The referer is the URL of the webpage from which a link was followed to the subsequent webpage. The URL of the referring page may contain PII or other sensitive information, usually as part of a query string. For example, if the referring page's URL was

```
www.example.com/customer?name=jane+doe&address=123+main+street
```

the referer information would contain PII (Jane Doe's name and address). Therefore Example.com would be transmitting PII to third party websites, and those websites would likely record that information.

The originating web server could help protect its users' privacy by configuring the site to ensure that no sensitive information is contained in a URL, or to strip the query string from the URL. Never include a username and password in a query string (referer strings aside, there are lots of other reasons why this is dangerous). Likewise, as discussed in more detail below, the subsequent web server can protect privacy by either not recording, aggregating or obfuscating the query string from the referer.

For each piece of PII being recorded, it is imperative that network administrators justify why they are keeping the information and consider a realistic time limit for retaining the information. These decisions should be recorded in an internal data retention policy. We outline three possible methods for PII-elimination below: these are obfuscation, aggregation and deletion.

a) Data Obfuscation

The easiest, but least protective, strategy is to periodically scrub the logs to obfuscate all explicit or deducible PII. Since virtually all OSPs maintain multiple logs and user information databases, providers should ensure that user identity cannot be gleaned when matching two or more processed logs. Setting a reasonable time duration before PII obfuscation allows OSPs to administer and troubleshoot their networks in real-time. The amount of time PII-exposed logs are stored will depend on the service requirements, but, of course, PII should never be kept any longer than necessary.

For example, in some cases, an OSP may need to associate logged information with a particular user, such as distinguishing between visitors and hits for purposes of determining how many people visit the OSP's web site. Data obfuscation is useful for this task, since the site does not need to know the particular IP number. Rather, the OSP only needs to define a temporary unique identifier to track visitors' sessions, and can then erase the IP while retaining the identifier.[14] Section 3(b) describes one reasonable way to create such an identifier using short-duration cookies.

Other key solutions to wipe PII from logs include:

- Obscuring the third, fourth and sixth octets of all MAC addresses in the same way as above. This will obfuscate both the exact manufacturer ID (first three octets) and the specific device ID (last three octets) being used.

14 This example assumes the OSP is using logfile analysis to determine website visitors. Some OSPs use third parties to conduct traffic analysis, which often involves inserting a "web bug" on the OSPs page. The web bug allows the third party to collect information directly form the OSPs users. Such practices mean that the privacy protections for the user are the lower of the OSPs practice and the third party's practice. See discussion of "Working with Third Parties" above.

- Obscuring the last four digits of phone numbers, or replacing it with '0000', but keeping the area code and exchange (but only if you need the area code and exchange for providing your service).
- Obscuring or deleting all usernames in e-mail addresses. (Note that this technique doesn't work for people who register and use their own names as domain names.)
- Obscuring or deleting all query strings in HTTP referer records and certain URLs (http://www.google.com/search?q=electronic+frontier+foundation). Most of the privacy risk from URL parameters comes from very rare but highly revealing examples, so it is usually safe to create records of the most frequent queries before erasure.
- Minimizing the records of HTTP headers to a basic minimum, omitting or erasing unusual or unique data that may inadvertently serve as PII.
- Replacing very specific User Agent records with more generic information

Some tactics that should not be used include:

- Encrypting PII with either symmetric or asymmetric keys, if decryption keys are also stored. Any subpoena or court-order can still force OSPs to turn over the decryption keys along with the encrypted data. (Some archiving techniques use temporary keys to encrypt logs; rather than delete the data itself, the temporary keys are destroyed, making older data unreadable.)
- Hashing PII with a non-random, well-known one-way hash. Using trial-by- error, one could match hashed candidate IP addresses with the encrypted IP address to reveal the original data.

When implemented in a timely fashion, obfuscation gives OSPs the flexibility to glean general usage patterns without retaining PII; implemented poorly, OSPs will continue to be subject to the legal consequences of information requests.

b) Data Aggregation

A better strategy is to use aggregation techniques to compile general usage statistics followed immediately by log deletion. This allows OSPs to fully discard all logs, including PII-obfuscated logs, after a specified duration of time, but still keep tabs on network access patterns. OSPs can save a substantial amount of resources using this technique, since aggregation requires minimal hard disk space. It also ensures that no specific PII will be retained on OSP servers in the long term.

Consider an OSP which hosts an Internet search engine and wants to track popular search queries. The OSP may be very interested in logging the referer information, since the query string can indicate what search terms were used by the new visitor to find the OSP's site. Obfuscation of the query string would not work because it would mask the data the OSP wants to track. Obfuscation of only the IP address (while exposing the query string) could still lead to potential IP address matches and PII leaks. Using aggregation techniques, the OSP can simply extract the query strings from the log file, tally the number of times each query was made, and then delete the file completely.[15] One OSP reported to us they automatically aggregate their web server logs every night, then immediately delete the previous day's logs. This method

15 Sometimes people will include their own PII in a search string. OSPs can minimize the potential to collect this PII by only listing common search terms, i.e. those which occur above a threshold number of time in a data period.

decouples users' identities from their search queries while allowing the OSP to keep track of popular search topics.

c) Data Deletion

Obfuscation and aggregation are only effective when used in tandem with log deletion. A strict policy which dictates when the OSP should fully purge logs from hard drives is a mandatory step in minimizing the potential challenges of legal compliance. Decisions on log retention time intervals will vary drastically. Free, open WiFi providers may delete connection logs immediately after log-off, while pay-per-use WiFi providers must keep logs for weeks until billing and collection have been completed. OSPs should note that different types of log files may have different data retention intervals.

Even after logs have been deleted from disk, the PII may still reside on the disk until that memory segment is reused and written over. Even then, advanced forensic searches of server hard drives could still reveal past data stored on them. These processes may cause OSPs significant disruptions. If possible, you should use secure deletion utilities to fully scrub the hard drives containing deleted logs (for instance, on Unix-based servers, use "shred" instead of "rm")[16]. This will help ensure the removal of all sensitive PII.

The best way to protect against the risk of log artifacts on disk is to never create any user logs in the first place. This is the ideal and safest solution even though it is often impractical. By reconfiguring the logging preferences in server applications, one can easily change the log level to record nothing about network events. But for most OSPs, these logs are necessary for network troubleshooting and security precautions. This is also virtually impossible for large, for-profit providers that need to maintain billing and subscriber contact information. Thus, the best tactic for an OSP is to come up with a safe and sane network policy in which logs are retained for the shortest possible time.[17]

2. Securing Personally Identifying Data

Users often transmit PII through OSPs' web sites, which OSPs should take measures to protect against unintended third-party access if unforeseen events occur (i.e., hacking, data breaches). Protecting user data doesn't just mean protecting it at the OSP: it also means endeavoring to protect it from others who might intercept it as it travels between the user and the OSP. This is more than a theoretical threat: users on Wi-Fi networks, shared networks at work or while traveling, can all be at risk from relatively simple third-party surveillance.

a) Security on the Wire

One indispensable requirement for the privacy and security of sensitive information is encrypting it when it travels over the network. HTTP, and many other important protocols, can be encrypted with SSL (Secure Sockets Layer).[18] The "https" one sees in a browser means HTTP

16 *shred* is a software utility that is part of the GNU Core Utilities. While *rm* deletes a file, shred overwrites the file.

17 Under the Electronic Communications Privacy Act, the "contents of a communication" become significantly less well protected after 180 days. While we recommend keeping records for far shorter periods (like a few weeks), we strongly recommend that you take care not to unnecessarily hold onto the contents of a communication for more than 180 days.

18 Technically speaking, the most recent versions of SSL are called TLS (Transport Layer Security), though the two are effectively synonyms in common usage.

over SSL. There are many considerations related to the right uses of encryption by OSPs; we will just work through the question of when web pages should be encrypted.

For each page on a website, OSPs must answer several questions: is this page going to be available over HTTPS? If so, is it going the page going to be encrypted *by default,* or will the user have to actively navigate to an HTTPS page? If the default is going to be HTTPS, will the page be *unavailable* over unencrypted HTTP?

In terms of CPU resources, it is more computationally expensive for an OSP to serve pages over HTTPS than regular HTTP pages. (Were this not the case, we would recommend always using HTTPS). As things stand, the more sensitive the information is, the more "yes" answers there should be to the questions in the previous paragraph.

If there is absolutely nothing sensitive about the information on a site (perhaps it is the website of a pen manufacturer, simply listing the pens they make), then there is no reason to incur the expense of HTTPS.

If a site contains a great deal of non-sensitive information, but has some use-cases which are sensitive, then it should offer users the ability to switch to an HTTPS version of the site. An encyclopedia – or a search engine – should be run this way, since people might use the site to research information about medical conditions, or controversial political opinions, or sexuality, and might wish to do so without their government, or systems administrators, or other computers on their WiFi network, being able to see this.

An OSP should always use SSL by default to protect users' login names and passwords, payment information, and other sensitive personal data. Once a user has logged in, he need *not* be redirected back to HTTP pages unless additional steps are taken to protect his authenticated session.[19]

Depending on the nature of the account, it may be acceptable to allow an HTTP session instead in the (unlikely) event that a browser does not support HTTPS. An OSP might reasonably permit unencrypted logins to a bulletin board. Sites that collect financially consequential information, or other highly sensitive information, should never allow unencrypted connections.

b) Security at the OSP

OSPs also need to mitigate the risk of data leaks and unauthorized access to information they keep on their servers. OSPs risk significant harms to their business if they fail to adequately mitigate the threats posed by hacker intrusion or by inappropriate use of that information by employees.

It is extremely difficult, if not impossible, for an OSP to make IT infrastructure hack-proof. The problem should be regarded as an exercise in risk mitigation. OSPs must hire competent computer security staff or consultants. If OSPs store significant amounts of sensitive data, they must take proportionate steps to protect that information. These should include (1) hiring competent penetration testers to learn how difficult it is to break into their systems, and to close off the easiest paths for hackers, and (2) storing older logs and similarly sensitive data archives on computer systems that are not connected to the Internet, either directly or indirectly.

19 Simple website authentication models use a cookie to identify the logged-in user. If they are redirected to an HTTP page after logging in, someone on their WiFi connection can trivially observe the cookie and hijack their session.

OSPs must also mitigate the threat of employee misuse of customers' data. OSPs must have clear policies stating if, when and how they will examine their clients' information. These policies should be outlined in the privacy policy, and they must be documented clearly within the organization. If an OSP stores significant amounts of sensitive data, it needs to have employees and software in place to monitor internal compliance with these policies.

3. Offering User Choice on Privacy Preferences

As discussed above, many OSPs keep logs of the interactions that each of their users have with the service, and allow users to provide PII as part of a registration process. There are legitimate and useful reasons why some OSPs may collect this information, but whenever possible, OSPs should offer users who wish to have a greater degree of privacy an easy way to minimize or eliminate tracking.[20]

a) "Opt Out" Data Collection

In many cases, allowing users a way to minimize data collection will take the form of "opting out" of tracking or allowing registered users to limit the amount of information provided during registration (and to limit the public disclosure and sharing of such information).

When offering web users the ability to opt out of logging mechanisms, it is essential to make the opt out persistent. Users should not have to repeatedly opt out of being tracked by your service.

If a website is going to offer a tracking opt-out feature, we would recommend offering a separate URL that obfuscates or eliminates tracking. For example, you could let users visit your site at "https://private.yourservice.com" rather than "http://www.yourservice.com," and turn off logs on the privacy sub-domain.

Some sites and third-party ad server use preference cookies. While an opt-out preference cookie may be a useful addition to a cookie-less method, it is not a substitute. Because the most privacy sensitive users are also the most likely to block cookies or to preserve them for only short periods of time, it may end up be the least effective for the most privacy sensitive users.

For sites that have public facing user pages (like a social networking service), we recommend providing the user with flexibility in deciding what information is disclosed to other users and/or the public. To maximize user control, OSPs may want to provide variable setting for each data item, letting the user determine whether, for example, their home address is provided to various categories like the public, friends, family, third parties, friends of friends, etc.

b) A Word About Cookies

Many OSPs use cookies to deliver and optimize their services for users. Because cookies can contain or confer access to sensitive consumer information, we recommend that OSPs use them in the most privacy-protective way possible.

---

20 In general, it is very difficult for users to avoid having their history of interactions with a service logged, since there are so many sources of information that an OSP may use to re-correlate their identity. We have illustrated this situation in detail for the case where the OSP is a search engine; *see* http://www.eff.org/wp/six-tips-protect-your-search-privacy.

Conceptually, we would divide the uses of cookies into three categories:

1. *Preference cookies* are usually set as permanent cookies with a site; they might indicate information such as a user's preferred language, whether they want adult material filtered for them by the OSP, etc.

2. *Tracking cookies* contain a unique piece of information; no two tracking cookies from a site will be the same. Some sites set them as permanent cookies, and some as session cookies.

3. *Authenticated cookies* are tracking cookies that indicate that a user is logged into a particular account on the service.

Best practices for dealing with these kinds of cookies are as follows:

As much as possible, a site should always function if cookies happen to be blocked entirely by the user, so the site should never *require* a preference cookie to operate. Rather than (for instance) demanding that the user select a language before showing him any content, a website should pick a language from hints in the client's HTTP headers, and fall back to a reasonable guess (English, or whatever the plurality of users for the site prefers) and let the user change it by clicking on a flag icon to receive a cookie or URL parameter if he wishes. Preference cookies should be easily understandable so that browser plugins can customize them for users if desired. Preference cookies should not also be tracking cookies.

The use of tracking cookies should be minimized unless they are authenticated cookies. Some websites make non-authenticated tracking cookies necessary for the operation of their user interface. This is a very poor design practice, since it causes unnecessary inconvenience for users who prefer to block cookies by default.

In some cases it may be reasonable to use a tracking cookie in order to keep aggregate statistics about the usage of a site. For instance, an OSP might need to record the number of unique visitors to each page. For purposes of this sort, tracking cookies should be set with a very short duration (such as 30 minutes or an hour). If you need to retain access logs for a site that uses tracking cookies, replace all of the identifying information in the logs – such as IP addresses – with the random cookie ID.

Authenticated cookies are necessary if users are going to log into your site. If the security of accounts on your site matters at all, you must consult detailed sources on "cross-site request forgery" and other forms of "cross-site scripting" attacks in order to design your authenticated cookie scheme correctly. If you are not communicating securely using SSL, consider the risks of long-term authentication cookies being intercepted and re-used.

As mentioned above, your privacy practices are only as good as those of everyone else who works with you to provide services to users.

Conclusion

OSPs need to understand their legal risks and obligations when codifying their logging practices. They must adopt a reasonable internal data retention policy and follow this policy consistently. Being strict about deleting all PII on servers will protect OSPs from many hidden costs. OSPs need to understand security risks threatening sensitive information, both on the network and within their systems. By taking proactive technical steps, and knowing their legal rights and obligations, OSPs can simultaneously maximize the privacy of users and protect themselves from the damaging effects of the DMCA, the ECPA and other data disclosure laws.

NOTE: This White Paper is Not Legal Advice

This Best Practices white paper provides EFF's recommendations for best practices for online service providers to protect their user's privacy. It is not legal advice. To determine how to implement good data practices for your OSP you should consult with an attorney. EFF is a small, grassroots legal advocacy nonprofit supported by member contributions. We provide pro bono (free) legal assistance in cases where we believe we can help shape the law. Unfortunately, we have a relatively small number of very hard-working attorneys, so we do not have the resources to defend everyone who asks, no matter how deserving. If we cannot assist you, we will make every effort to put you in touch with attorneys who can. If you need assistance in finding counsel, you can contact us at information@eff.org.

Glossary of Terms and Acronyms

Cookie. On the web, a cookie is a bit of text sent by a web server to a browser, which can be retrieved later. See **http://en.wikipedia.org/wiki/HTTP_cookie**.

COPPA. Children's Online Privacy Protection Act of 1998, 15 U.S.C. §§ 6501 et. seq.

CPU resources. Central Processing Unit resources. A CPU is the chip at the heart of a computer. As used in this document, CPU resources refers to the expense associated with having ones servers conduct computations.

DHCP. Dynamic Host Configuration Protocol, see **http://en.wikipedia.org/wiki/DHCP**.

DMCA. Digital Millennium Copyright Act, 17 U.S.C. § 512.

ECPA. Electronic Communications Privacy Act, 18 U.S.C. § 2501 et seq.

ECS. Electronic communications service is broadly defined as "any service which provides to users thereof the ability to send or receive wire or electronic communications." 18 U.S.C. § 2510(15).

EFF. The Electronic Frontier Foundation is a non-profit, member-supported civil liberties organization working to protect rights in the digital world. EFF actively encourages and challenges industry, government and the courts to support free expression, privacy, and openness in the information society. Founded in 1990, EFF is based in San Francisco.. See **http://www.eff.org**.

FTP. File Transfer Protocol. See **http://en.wikipedia.org/wiki/FTP**

HTTP and HTTPS. Hyper-Text Transfer Protocol and HTTP over a Secure Socket Layer. See **http://en.wikipedia.org/wiki/HTTP**

ICS. Interactive computer service. This legal term originated in Section 230 (47 U.S.C. § 230(f)(2)), and has been interpreted broadly to include websites and other online services.

IDS. Intrusion Detection System, a software system that detects unwanted manipulations to systems.

IP Address. Under the IPv4 protocol, the Internet Protocol address is a set of four numbers that provides the location of a computer. Also called an IP number.

MAC Address. Media Access Control address, a quasi-unique identifier attached to most network adapters (NICs), such as WiFi cards or ethernet cards.

OSP. Online service provider. This is an omnibus term, encompassing legal terms of art such as electronic communications service providers, remote computing services, and interactive computer services.

PII. Personally identifiable information. Information that can identify a particular person, such as names, phone numbers, or addresses. IP numbers can be PII, since there are often records that associate a particular IP at a particular time with a particular user.

Query String. A query string is a portion of a URL that contains data (called arguments) to be passed to web server applications, such as Common Gateway Interface (CGI) programs. It is often the text following a question mark (?) in a URL. See **http://en.wikipedia.org/wiki/Query_string**.

RADIUS. Remote Authentication Dial In User Service, see **http://en.wikipedia.org/wiki/RADIUS**

RCS. Remote computing service. The term is defined as "provision to the public of computer storage or processing services by means of an electronic communications system." 18 U.S.C. § 2711(2).

SSL. Secure Socket Layer, a cryptographic protocol that provide secure communications on the Internet, see **http://en.wikipedia.org/wiki/Secure_Sockets_Layer**. See also Transport Layer Security (TLS).

URL. Universal Resource Locator, also known as Uniform Resource Locator or Uniform Resource Identifier. Commonly means a web page address. See http://en.wikipedia.org/wiki/URL.

VPPA. The Video Privacy Protection Act of 1998, 17 U.S.C. § 2510.