Privacy in Technology

Standards and Practices for Engineers and Security and IT Professionals

JC Cannon, CIPP/US, CIPT

An IAPP Publication
# Contents

Acknowledgments ................................................................. ix
Marla Berry, CIPT

Introduction ................................................................. xi
J. Trevor Hughes, CIPP

**CHAPTER 1**
Understanding the Need for Privacy in the IT Environment ........................................... 1
1.1 Evolving Compliance Requirements ........................................... 1
1.2 IT Risks ................................................................. 3
1.3 Stakeholders’ Expectations for Privacy ........................................... 13
1.4 Mistakes Organizations Make ........................................... 15
1.5 Privacy vs. Security—What’s Alike and What’s Different ........................................... 17
1.6 IT Governance vs. Data Governance ........................................... 18
1.7 The Role of the IT Professional and Other Stakeholders in Preserving Privacy ........................................... 19
1.8 Conclusion ................................................................. 21

**CHAPTER 2**
Core Privacy Concepts ................................................................. 25
2.1 Foundational Elements for Embedding Privacy in IT ........................................... 26
2.2 The Information Lifecycle: An Introduction ........................................... 51
2.3 Common Privacy Principles ........................................... 52
2.4 Conclusion ................................................................. 53
Figure 4-2: Sign-Up Screen with Link to Privacy Statement .................................................. 111
Figure 6-1: Internet Explorer SmartScreen Filter ................................................................. 151
Figure 6-2: Firefox Security Settings ..................................................................................... 152
Figure 6-3: Chrome Privacy Settings .................................................................................... 152
Figure 6-4: Ad Serving Diagram ........................................................................................... 158
Figure 6-5: Lightbeam for Firefox Screenshot Showing Active Third-Party Websites .......... 160
Figure 6-6: Adobe Flash Player Settings Manager: Website Storage Settings ...................... 162
Figure 6-7: Silverlight Configuration Settings: Application Storage .................................... 163
Figure 6-8: Adobe Flash Player Settings Manager: Local Storage Settings ......................... 163
Figure 6-9: Chrome Cookie Management .............................................................................. 169
Figure 6-10: Firefox Cookie Management ............................................................................ 170
Figure 6-11: Internet Explorer Cookie Management ............................................................. 171
Figure 6-12: Tracking Protection in Internet Explorer .......................................................... 172
Figure 6-13: Expired Certificate as Shown in Chrome ......................................................... 174
Figure 7-1: “Add a Device” Feature in Windows
   Displaying a Detected Bluetooth-Enabled Device ......................................................... 194

TABLE LIST
   Table 2-1: Example Data Inventory ................................................................................... 40
Core Privacy Concepts

When Amy joined her company, its privacy program was in its infancy. No company privacy policies governed the work of the IT department. The company’s privacy notice was a generic one that was created by a law firm and covered only their practices as they had existed several years before. They used vendors extensively to manage some of their data for them, but the contracts had no privacy provisions. She was concerned that the company’s lack of privacy maturity could be detrimental to the company’s fiscal future. At that moment, Bill knocked on her door.

“Amy, I’m so glad you joined our team,” stated Bill. “We’ve had a few lingering issues that we held on the back burner that we were hoping you could help us with.”

“Oh, let me hear them,” responded Amy.

“Well, our developers run tests of their new code on our customer data to make sure it’s working okay. One of our developers was debugging an issue on his laptop at a Starbucks and lost it. It wasn’t encrypted, and we haven’t heard about any of the credit card data being used so we believe we are safe—though we feel we should send a breach notice to someone. Next we received an FTC inquiry because they feel we may be in violation of the CAN-SPAM Act because we started sending e-mails to our cosmetics customers for our new spa products even though they opted out of cosmetic e-mails. We thought we were okay. We are also being sued by a Hollywood actress.”

Bill asked the question written on Amy’s face. “Why would a Hollywood actress be suing us? One of our sales guys got her phone number from the spa she goes to in exchange for a discount on our products. When she complained, he replied that her opt-out didn’t apply to us because we didn’t have an agreement with her. She became enraged and hung up. The next day we received a letter from her lawyer. I should probably stop there. You’re not looking too good.”
“You mean there’s more?” Amy asked. Bill nodded while Amy placed her hand to her head. “That’s enough, Bill. Get the team together. I want everyone on my team to clear their calendars next week because they will be taking privacy boot camp. The time for lax privacy practices has ended.”

2.1 Foundational Elements for Embedding Privacy in IT

Without privacy policies in place and proper training for employees, it is extremely difficult to ensure that employees are following proper privacy practices. An organizational privacy policy is the first essential piece of collateral material needed for an effective privacy program. The privacy policy must come before a privacy notice is created as the notice is a reflection of a company’s privacy policy. Likewise, setting up privacy training is not practical until a policy is in place. As important as the privacy policy is, its creation should not be taken lightly.

Within this book, to avoid confusion, the term “privacy notice” will be used to denote the external privacy statement and “privacy policy,” the internal privacy standards.

2.1.1 Organizational Privacy Notice

The privacy notice is an external instrument that informs the outside world about an organization’s privacy practices. It is a reflection of the organization and its commitment to privacy and can affect its brand. It provides transparency for consumers and comfort for privacy advocates; in some jurisdictions, it’s a legal requirement. To avoid legal entanglements and preserve its brand, an organization must ensure that its privacy notice is based on the organization’s internal privacy policies.

The sheer size of a privacy notice can be daunting to the average website visitor. Having a privacy notice that fits on one page and covers the most important aspects of an organization’s privacy practices will go a long way toward setting the website visitor at ease. This will also set the organization apart from most Internet sites. For example (as of press time), Facebook’s privacy notice achieves both of those goals. It also provides “Learn More” links that permit site visitors to get access to more detailed information along with videos and privacy settings. Naming it their Data Use Policy is also more descriptive of the content.

2.1.1.1 Outline of a Privacy Notice

Below is a list of sections that should exist in a privacy notice. They include information that consumers as well as privacy advocates and regulators will want to know about. Organizations with fewer or more complex data-handling practices may require fewer or more sections.

- **What data is collected.** This is one of the most important pieces of the privacy notice, as it lets everyone know what data is collected and, by extension, what data is not collected. A distinction should be made between what data is received
Core Privacy Concepts

by web servers and the data that is processed and used within the organization. This declaration should include data that is observed, inferred and declared directly from users as well as data collected from third parties.

• **How collected data is used.** This section describes how collected data is used across the organization. It should provide a general description of data usage, including how the data may be used by all groups across the organization. This description should also include any third-party usage of the data.

• **How collected data is shared.** This section should cover how data is shared not only outside the organization but with which teams across the organization. Any law enforcement or regulatory requirements to share data should be described here as well. Individuals should be provided with controls to manage the sharing of their data where possible.

• **User control over collected data.** This section should describe how users can control the collection and use of their data. There should be a description of any preference or configuration management system that will help users manage how their data is collected or used. While it may not be possible to discontinue the normal transmission of data logs to web servers during a website visit, users should have some control over how their data is used.

• **Controlling marketing contact.** Users should be able to control how and when an organization contacts them. Service and transactional e-mails are expected when an individual creates a relationship with an organization or purchases a product. However, those same individuals should not have to be subjected to marketing e-mails in order to receive service e-mails. In addition, if an individual has a relationship with one product group within an organization, other product groups within the company should not take the liberty to contact the individual without first obtaining consent. For example, if an individual purchases Duracell batteries from Procter & Gamble, he or she shouldn’t expect to receive marketing e-mails from the Pampers division. Even after consent has been granted, individuals should have the ability to opt out of receiving unwanted e-mails that are not required for their continued relationship with the organization.

• **Use of cookies and other tracking mechanisms.** Cookies are the main mechanisms by which organizations keep track of visitors to their websites. However, companies may use Flash cookies, locally stored objects, HTML5 storage, fingerprinting or some other means to track users. A description should be provided to users about the tracking mechanisms an organization uses as well as some ability to limit their use.

• **Gaining access to data.** The notice should explain how users can access the data an organization holds on them. This access should be provided online where possible, but can be provided in a file format or printout as a last resort. This
access can be limited to just the information collected while the user is in an authenticated state in order to avoid delivering the data to the wrong person.

- **Resolving privacy issues.** This section should describe how users can resolve privacy issues they may have with an organization. This can be handled via an online form, e-mail address, phone number or postal address. Users should also have the option of contacting a third-party organization to address any issues as needed, via either arbitration or a self-regulatory organization.

- **Date of privacy notice.** The date the privacy notice was released should be prominently displayed as part of the privacy notice. If the privacy notice was recently updated, that should be stated.

- **Changes to privacy notice.** Users should be made aware of how often a privacy notice may be changed and how they will be informed of changes. Users should also be provided with a means to see previous versions of a privacy notice. Organizations should track which privacy notice was in effect when any data was collected or processed.

2.1.1.2 Creative Formatting

Formatting can make a privacy notice easy to read and navigate. Microsoft’s updated privacy notice (Figure 2-1) uses large colored tiles and a layered design that delivers on both capabilities and makes their notice engaging.²

Figure 2-1: Microsoft.com Privacy Statement

Used with permission from Microsoft.
2.1.1.3 Consolidation

Consolidating multiple privacy notices into a single one can make it easier to find the desired privacy information. Google consolidated more than sixty disparate privacy notices into a single notice (see Figure 2-2), making it easier to find the privacy information of interest without having to go to multiple web pages.

Figure 2-2: Google Privacy Policy

Google and the Google logo are registered trademarks of Google Inc.; used with permission.

2.1.1.4 Multilayered Notices

Multilayered privacy notices provide an abbreviated form of an organization’s privacy notice while providing links to more detailed information when required by the reader. The U.S. Postal Service was one of the early organizations to implement a multilayered privacy notice. An approach to the creation of multilayered notices was introduced back in 2007 in a paper published by The Center for Information Policy Leadership at Hunton & Williams LLP. Since then, many companies have deployed variations of the idea. Layered notices offer brevity at one level, making an overview of an organization’s privacy practices accessible to site visitors who are looking for a simple notice, while a detailed notice is just a click away for those who want more in-depth information. Those wishing to implement a layered notice should be certain to give readers enough information at the overview level to make an informed decision about the site and the organization without having to look at details.
2.1.1.5 Privacy Icons

Privacy icons provide visual cues to a company’s or application’s privacy characteristics. The icons can be shown on a website’s homepage or in the footer of each page for easy viewing by site visitors. Applications can show the icons at startup, during install or within a privacy interstitial, application policy or settings dialog. The icons can be useful as part of a just-in-time privacy consent notice because the data being collected is often very specific. They are less practical when used for a company’s entire privacy notice, as a company can have multiple data collection, usage, sharing and retention practices. For this reason privacy icons never took off. The Center for Internet and Society at Stanford Law School held a privacy icon hackathon at Mozilla’s headquarters. During the event, participants found it difficult to decide which set of icons would be most appropriate based on the privacy notices selected from the Internet.

Disconnect created a set of privacy icons (shown in Figure 2-3) that have been in use since June 2011. The Association for Competitive Technology created a set of privacy icons for developers. However, adoption has been slow. One of the concerns with using such a system is that if the icons are too broad, they give the impression that all data is collected, shared with everyone and kept forever. Alternatively, if the icons are too specific and do not accurately cover all privacy practices, the company can leave itself susceptible to litigation.

Figure 2-3: Disconnect Privacy Icons

![Disconnect Privacy Icons](image)

Used with permission from Disconnect, Inc.
2.1.1.6 Privacy Nutrition Label

A privacy nutrition label is similar to the nutrition labels seen on products in a store. Instead of listing the nutritional value of a product, the labels provide the abbreviated form of a company’s privacy practices. The privacy nutritional label is more informative than the privacy icons, though it is only practical as part of the company’s privacy notice or a privacy notice for a newly installed application. However, like the privacy icons, it is difficult for one label to provide a clear, unambiguous view of a large firm’s privacy practices when there could be multiple collection, usage, sharing and retention practices across the company. A sophisticated privacy nutrition label was defined by Carnegie Mellon University’s CyLab Usable Privacy and Security (CUPS) Laboratory (see Figure 2-4). The CUPS Laboratory also sponsors a yearly event, Symposium on Usable Privacy and Security (SOUPS), which takes place in different cities across the United States.

Figure 2-4: Bell Group Privacy Nutrition Label

Used with permission from CyLab Usable Privacy and Security Laboratory, Carnegie Mellon University.
2.1.1.7 Combination Privacy Statement

Some companies, like TRUSTe, have combined several simplification techniques to create a privacy notice where icons and short descriptions make it easy to find and grasp the company’s privacy content (see Figure 2-5).¹⁰

Figure 2-5: TRUSTe Privacy Statement

Image courtesy of TRUSTe.
2.1.1.8 Blended Mobile Statements
PrivacyChoice created Policymaker, which is a combination of icons and a nutrition label.\textsuperscript{11} TRUSTe has also created a program to provide mobile applications with a privacy statement that contains icons and a layered notice (Figure 2-6).\textsuperscript{12}

Figure 2-6: TRUSTe Privacy Policy Optimized for Mobile Devices

Image courtesy of TRUSTe.
2.1.1.9 Privacy Notice Services

One way to simplify the task of creating, hosting and supporting a privacy notice is to outsource the effort to another company. There are several companies that have a legal staff to guide companies through the process of a privacy evaluation (different from a privacy impact assessment), in which they determine an organization’s data collection, usage, sharing, retention and disposal practices in order to create a privacy notice with the appropriate content.

Websites such as www.freeprivacypolicy.com provide an online tool for the creation of a privacy policy. Companies such as TRUSTe provide privacy notice creation, localization and hosting services, which can be expensive for small companies to manage on their own. In addition, a privacy notice service can act as a privacy response center for an organization and respond to privacy inquiries from its website visitors.

2.1.2 Organizational Privacy Policies

The privacy policy is the guiding set of privacy principles used by teams within an organization to help them understand their privacy obligations as they develop software and services, create marketing campaigns, work with vendors and engage the general public. A privacy policy also serves as a guide for all organizational activities and drives the commitments made within the privacy notice. Even though an organization might be a sprawling global enterprise with dozens of business groups, it should have a single overall policy that sets the tone for privacy across the entire organization. There can still be separate privacy policies that cover specific groups, products or programs, but they should align with the overall organizational policy. The importance of the privacy policy dictates that its creation should involve members from all major business groups across the organization.

Because of the eventual impact of the privacy policy, those creating it should be careful not to make the policy overly prescriptive and/or restrictive as this could limit the business potential of some groups. The organizational privacy policy should provide general guidelines while leaving business groups the flexibility to provide more prescriptive rules that support their business goals. An organization’s privacy policy should at a minimum cover the following topics:

- **Types of data classification.** To identify data and the rules that may apply to it, a classification system should be defined. A classification system can be a simple set of sensitivity levels or a complex taxonomy, understanding that complexity can be a barrier to adoption and accuracy.

- **Data collection principles.** The policy should describe when and how data should be collected as well as list obligations for data collection. The description should indicate the notification, control, protection required, minimization requirements and sharing limits for collected data.

- **Protection of data.** The privacy policy should indicate how data is to be protected during collection, storage and handling. The type of protection
required, be it encryption or access control, will typically vary based on the classification of the data collected and regulatory requirements.

- **Data retention period.** All data should have an associated retention period. This period will vary based on regional regulatory requirements. The retention period may be indefinite, but there should always be a justification for the period selected. The policy should also describe what happens to the data once the retention period expires.

- **Treatment of sensitive data.** The definition of sensitive data can be elusive, as everyone has a different idea of what is sensitive, and that can change with context. Most jurisdictions agree that specific medical and financial data is sensitive. However, religion, political views, ethnicity, sexual preference and entertainment choices can be considered sensitive to certain people and locales. An organization should decide which categories of data are sensitive, how they should be handled and the type of consent, if any, that should accompany the use of sensitive data.

- **Sharing of data across groups.** A policy should describe the management of data as it is shared across groups within an organization. It is important that the policy take into account the privacy commitments made at the time by the group originally collecting the data. This includes complying with the retention commitment based on when the data was originally collected. The organization's sharing policy should eventually be covered in the privacy notice.

- **Sharing of data with partners and vendors.** Privacy policies should cover how and when data can be shared with external parties like partners and vendors. Contracts can help to ensure that third parties process the data they receive in a way that does not conflict with the organization's privacy policies. The organization's retention policies should be reflected in the sharing policy. The privacy notice should describe any third-party sharing that might occur as well as a mechanism to opt out of this sharing where appropriate.

- **Creation of departmental privacy policies.** Groups within an organization will often create a privacy policy that covers their specific processing of collected data. These policies must not conflict with the organizational policy. The organizational privacy team should be involved in the sign off of departmental privacy policies. As the organization's privacy policy evolves, departmental policies must be updated to reflect any changes that might impact those policies.

- **Performance of privacy reviews.** A privacy policy should outline when and how privacy reviews should occur. This description should cover how privacy reviews apply to both new and completed projects to ensure that they maintain the level of privacy protection mandated by privacy policies as they evolve.
• **Participation in a privacy response center.** All organizations should have a privacy response center in place that responds to external privacy incidents. The privacy policy should detail how groups within the organization should be involved in the privacy response center and indicate the obligations of each group in response to a privacy incident. The policy should also describe how the organization should be integrated with the response programs of external organizations such as the Better Business Bureau.

• **Responding to privacy inquiries.** Privacy inquiries can be received by an organization that are not caused by a privacy incident. Regulators, consumer advocates or journalists may have general questions about an organization’s privacy practices, notifications or plans. The privacy policy should describe how privacy inquiries will be handled and who should be involved in the response to privacy inquiries.

• **Responding to data requests.** Data requests can come from users who own the data, law enforcement agencies or third parties wishing to have access to data. An organization’s privacy policy should indicate the conditions under which data requests will be honored and outline the process for verifying the rightful owner of the data. The policy also needs to describe the management of takedown requests and should cover whether the organization should handle the request or pass the request on to the person who posted the content to be taken down.

2.1.2.1 Data Classification

Data classification is key to effective and efficient data inventories and policy-based data management. While the idea of data classification may sound simple, it can be very complex. A data record may contain a simple user ID, date and a web address. What would be the classification of each field and the entire record? What if the user ID is an index in a table with personally identifiable information (PII)? What if the web address points to a user’s Facebook account? Now imagine that a new record contains a person’s birth date, gender and zip code. How would each field be classified? How would the record be classified? Research done by Latanya Sweeney shows that 87 percent of the U.S. population have reported characteristics that likely make them unique based only on a five-digit zip code, their gender and date of birth.14 Furthermore, 53 percent of the U.S. population are likely to be uniquely identified using only a person’s city, gender, date of birth. How does the record classification change when a record is split up or combined with other data? How do the retention, access and usage policies change? These are all questions that must be worked out before a data classification system can be implemented.

The approach to classifying data should not be rushed or taken lightly. The chosen classification system will have a long-term effect on the processing of data by the organization. There will be a desire to classify the data at a granular level unnecessarily. The data classification system should not be any more granular than the rules developed
to regulate it. For example, a database table may include phone number, postal address and e-mail address. These three different types of data will probably be treated the same in regard to privacy. Therefore, they should be given the same classification, such as PII.

There are several ways to classify data. Research should be done to determine the best approach for the organization. It should not be so complex that it is difficult to understand or discourages adoption. There should also be clear rules about how each category should be treated. Classification can be as simple as high, medium and low; can be based on data sensitivity; or can be as sophisticated as a system based on multilevel security (MLS) used by various militaries. MLS makes it easy to develop a strong role-based or attribute-based access control system. These provide a great means for programmatically protecting data based on policy, though they require a lot of maintenance, which can lead to out-of-date and inconsistent classifications.

One error that organizations make is having different classifications for the same data or different rules for processing the same classification of data. These disparities can occur when teams create classifications in isolation, data is collected because of an acquisition or local laws or contractual agreements influence classification rules. Until the disparities can be rectified, organizations should be careful to ensure that the treatment of the data matches the commitments made at the time that the data was collected. This may mean an extra classification value to allow for the application of different rules to certain pieces of data in the same category.

2.1.2.2 Data Retention

All data should have a retention period associated with it based on the data’s classification and regulatory requirements. When data hasn’t been classified or where no retention period has been applied to the classification, an agreed-upon maximum period should be applied. Organizations should resist simply stating “as long as needed for business purposes.” We all know how frustrating it can be when repair people give us vague time frames for when they will show up, possibly disrupting our plans for the day.

Sometimes a set of data will have a retention period of several years due to regulatory requirements; however, not all of the data in a class may be subject to the same requirements. As a result, data in a class may need to be stored for different periods of time.

Once a retention period has been reached, the organization needs to determine what to do with the data. It can be deleted, de-identified or aggregated.

- Data deletion requires the complete removal of data once the retention period has passed. This can be a difficult task depending on the access controls on the data. If data is allowed to be shared across teams, copied to personal computers or external devices or shared with third parties, it will be nearly impossible to be certain that all the data has been deleted.

- De-identification is the process of removing all items from the data record that could lead to the identification of an individual. Care should be taken that identifiers or other data within the record cannot be combined with other data to permit identification of an individual.
• Data aggregation is the process of combing data from multiple records so that the combined data reflects the attributes of a group versus an individual. Care should be taken that there are enough records in the collection to avoid the identification of an individual.

Once again, organizations should harmonize their policies across teams to ensure that retention policies are applied consistently to the same data. Custom agreements will always override organizational policies; however, they should not conflict with commitments made in privacy notices to users.

2.1.2.3 Data Deletion
Deleting unneeded or expired data is one of the best ways to remove the risk that can come from having too much data. The less data an organization holds, the lower its privacy risk. Data can be deleted by deleting files, formatting the storage medium or removing specific entries from a database or other data store. When to delete data can be based on the data’s retention period, which can be determined by the data record’s timestamp or file creation date when the data was collected. These dates should persist as the data moves throughout the company. Data deletion can also be triggered by the termination of a contract, acquisition by another organization, completion of a transaction, a regulatory requirement or a deletion request by the data owner. Upon deletion care must be taken to ensure that copies of the data do not exist in offline storage or in backup media. Backup procedures should include regular rotation of backup media such that expired data is removed or the backup media are destroyed.

2.1.2.4 Data Inventory

“Bill, how is privacy boot camp going?” asked Amy.

“So far so good. We’re learning a lot. As a matter of fact, the more we learn, the more worried I get,” responded Bill.

“I feel if we are going to get a better handle on our privacy we need to better understand what data we have. Our three-level privacy categorization for data is great for understanding how to protect it, but it doesn’t help much with understanding our regulatory requirements for the data or our commitments to data that we collect from users. When you get back to boot camp, let me know what you can find out about data classification schemes,” said Amy.

“I like your idea. Validating compliance has been tough for teams since they weren’t sure when to apply policies to data. I’ll let you know what I find out.”

Understanding the risk to data requires that organizations know what data exists under their control. Answering questions like “Do you have any data of type X?,” “Have you complied with your data retention policies?” and “Who has access to data of type X?” is
difficult unless there is a clear idea of where all the data of type X is stored. Better yet, if the data is classified, finding it and applying policies can be much simpler and even automated.

Problems can arise when data is stored off of the organization’s network. When an organization’s data is placed in removable storage, personal devices or transferred to third parties, it is difficult to validate that it is being managed in accordance with organizational policy. In addition, those scenarios increase the risk of data breach, which can be a major liability for an organization. Guidelines for effecting a meaningful inventory include:

- Have rules governing where data can be placed.
- Minimize the use of offline storage. An exception can be made for backups.
- Data placed on thumb drives and personal devices should be encrypted and short lived.
- Contracts should be in place to govern the use of organizational data by third parties.
- Classify data to simplify the discovery of data and the application of appropriate policies.
- Create data flows and list all data stores in a data inventory.

Scripts can be written to help simplify the data inventory process. To make this job easier, field and column names for data should be made consistent across data stores. Adding metadata to the column and record definitions will make it easier to find a certain class of data. For example, within a database, the table and column attributes could contain a classification designation. XML attributes can be used to indicate the data classification for an element. For example:

```xml
<DataItem Classification="SensitivePII">
  <Name Classification="PII">John Smith</Name>
  <Religion Classification="Sensitive">Catholic</Religion>
</DataItem>
```

Once the data classification has been completed, performing a data inventory can be easier and more complete. Scripts can be created that can scan databases and other data stores to not only find data but also to classify it into an inventory table. Once a data inventory is complete it can be placed in a database where it can more easily be accessed via web pages, database tools and scripts. Compliance tools can be run against data stores as well to verify compliance with retention, de-identification and access policies. A data inventory can also simplify audits and record retrievals to respond to data inquiries. Table 2-1 provides an example of what a data inventory might look like.