Internet Architecture Board (IAB)
Request for Comments: 6462
Category: Informational
ISSN: 2070-1721

A. Cooper

January 2012

Report from the Internet Privacy Workshop

Abstract

On December 8-9, 2010, the IAB co-hosted an Internet privacy workshop with the World Wide Web Consortium (W3C), the Internet Society (ISOC), and MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL). The workshop revealed some of the fundamental challenges in designing, deploying, and analyzing privacy-protective Internet protocols and systems. Although workshop participants and the community as a whole are still far from understanding how best to systematically address privacy within Internet standards development, workshop participants identified a number of potential next steps. For the IETF, these included the creation of a privacy directorate to review Internet-Drafts, further work on documenting privacy considerations for protocol developers, and a number of exploratory efforts concerning fingerprinting and anonymized routing. Potential action items for the W3C included investigating the formation of a privacy interest group and formulating guidance about fingerprinting, referrer headers, data minimization in APIs, usability, and general considerations for non-browser-based protocols.

Note that this document is a report on the proceedings of the workshop. The views and positions documented in this report are those of the workshop participants and do not necessarily reflect the views of the IAB, W3C, ISOC, or MIT CSAIL.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Architecture Board (IAB) and represents information that the IAB has deemed valuable to provide for permanent record. Documents approved for publication by the IAB are not a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6462.

Table of Contents

1. Introduction 3
2. Workshop Overview 3
2.1. Technical Discussion 4
2.2. SDO Discussion 5
3. Design Challenges 6
3.1. Ease of Fingerprinting 6
3.2. Information Leakage 7
3.3. Differentiating between First and Third Parties 8
3.4. Lack of Transparency and User Awareness 9
4. Deployment and Analysis Challenges 9
4.1. Generative Protocols vs. Contextual Threats 9
4.2. Tension between Privacy Protection and Usability 11
4.3. Interaction between Business, Legal, and Technical Incentives 12
4.3.1. Role of Regulation 12
4.3.2. P3P: A Case Study of the Importance of Incentives ..13
5. Conclusions and Next Steps 14
5.1. IETF Outlook 14
5.2. W3C Outlook 15
5.3. Other Future Work 15
6. Acknowledgements 15
7. Security Considerations 15
8. Informative References 16
Appendix A. Workshop Materials 19
Appendix B. Workshop Participants 19
Appendix C. Accepted Position Papers 21

1. Introduction

On December 8-9, 2010, the IAB co-hosted a workshop with the W3C, ISOC, and MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) about Internet privacy [Workshop]. The workshop was organized to help the Internet community gain some understanding of what it means for Internet-based systems to respect privacy, how such systems have been or could be designed, how the relationship between the web and the broader Internet impacts privacy, and what specific work the IETF and/or the W3C might pursue to address Internet privacy. An overview of topics discussed at the workshop is provided in Section 2.

The workshop discussions revealed the complexity and broad-based nature of privacy on the Internet. Across numerous different applications, a number of fundamental design challenges appear again and again: the increasing ease of user/device/application fingerprinting, unforeseen information leakage, difficulties in distinguishing first parties from third parties, complications arising from system dependencies, and the lack of transparency and user awareness of privacy risks and tradeoffs (see Section 3). Workshop participants also identified a number of barriers to successful deployment and analysis of privacy-minded protocols and systems, including the difficulty of using generic protocols and tools to defend against context-specific threats; the tension between privacy protection and usability; and the difficulty of navigating between business, legal, and individual incentives (see Section 4).

Privacy challenges far outnumber solutions, but the workshop identified a number of concrete preliminary steps that standards organizations can take to help ensure respect for user privacy in the design of future standards and systems. For the IETF, these included the creation of a privacy directorate to review Internet-Drafts, further work on documenting privacy considerations for protocol developers, and initiating a number of exploratory efforts concerning fingerprinting and anonymized routing. Potential action items for the W3C included investigating the formation of a privacy interest group and formulating guidance about fingerprinting, referrer headers, data minimization in APIs, usability, and general considerations for non-browser-based protocols. These next steps and workshop outcomes are discussed in Section 5.

2. Workshop Overview

The workshop explored both current technical challenges to protecting privacy and the ways in which standards organizations can help to address those challenges. Links to workshop materials are listed in Appendix A.

2.1. Technical Discussion

The workshop explored privacy challenges in three different technical domains: at the network level, at the browser level, and with respect to cross-site data exchanges. Example technologies were highlighted in each area to motivate the discussion.

At the network level, participants discussed IP address hiding in mobility protocols, privacy extensions for IPv6 addressing [RFC4941], and onion routing. Discussion about the Tor project [Tor] was particularly insightful. Tor is a circuit-based, low-latency communication service designed to anonymize protocols that run over TCP. End hosts participating in a Tor exchange choose a path through the network and build a circuit in which each "onion router" in the path knows its predecessor and successor, but no other nodes in the circuit. Each onion router in the path unwraps and decrypts received information before relaying it downstream.

For Tor to provide anonymity guarantees, Tor nodes need to be able to strip out information elements that can be used to re-identify users over time. For example, web technologies such as cookies, large portions of JavaScript, and almost all browser plug-ins (including Flash) need to be disabled in order to maintain Tor's privacy properties during web use, significantly hampering usability.

At the browser level, the discussion focused first on experiences with "private browsing" modes. Private browsing puts a browser into a temporary session where no information about the user's browsing session is stored locally after the session ends. The goal is to protect the user's browsing behavior from others who may make use of the same browser on the same machine. Private browsing is not designed to protect the user from being tracked by malware (e.g., keyloggers), remote servers, employers, or governments, but there is some evidence that users fail to understand the distinction between protection from snooping among users who share a device and these other forms of tracking. The specific protections offered by private browsing modes also vary from browser to browser, creating privacy loopholes in some cases.

The browser discussion also addressed proposals for "Do Not Track" (DNT) technologies to be built into browsers to provide users with a simple way to opt out of web tracking. At the time of the workshop, various different technical proposals had been designed to offer users the ability to indicate their preference to opt out or to block communication to certain web sites altogether. The discussions at the workshop illustrated a lack of agreement about what type of

tracking is acceptable, which technical mechanisms would be best suited for different scenarios, and how the mechanisms would interact with other aspects of privacy protection (such as notices to users).

The cross-site data-sharing discussion focused on current uses of Open Authorization (OAuth) (with Facebook Connect, for example). While improvements have been made in obtaining user consent to sharing data between sites, challenges remain with regard to data minimization, ease of use, hidden sharing of data, and centralization of identity information.

2.2. SDO Discussion

Participants discussed past experiences in approaching privacy within the IETF and the W3C. Individual protocol efforts within the IETF have sought to address certain privacy threats over the years. Protocol designers have taken steps to reduce the potential for identifiability associated with protocol usage, such as in the IPv6 privacy extensions case [RFC4941]. Protocols architected to rely on intermediaries have sought to minimize the user data exposed in transit, most notably in SIP [RFC3323]. Protocol architectures used in interpersonal exchange have sought to give users granular control over their information, including presence [RFC2778] and geolocation information [RFC3693]. Efforts to square privacy with usability are ongoing; the ALTO working group [ALTO], for example, is working out how to balance the needs of users and network operators to share data with each other about content preferences and network topologies against legitimate concerns about revealing too much of either kind of information.

The IETF also has experience to draw on in building a culture of security awareness. Beginning with [RFC1543], RFCs were required to contain a Security Considerations section. But that simple mandate did not immediately translate into the extensive security consciousness that permeates the IETF today. Over many years and with much effort invested, a more systematic approach to security has evolved that makes use of a variety of tools and resources: the security area itself, guidelines to RFC authors about security considerations [RFC3552], the security directorate, security advisors assigned to individual working groups, security tutorials at IETF meetings, and so on.

The W3C likewise has a number of past efforts to draw on. One of the earliest large-scale standards efforts aimed at improving web privacy was the Platform for Privacy Preferences [P3P]. The idea behind P3P was to have web sites provide machine-readable privacy policies that browsers could vet and possibly override according to the user's preference. The P3P policy expression language was robust enough to

allow sites to make complex assertions about how they intended to make use of data related to users, but market developments have created a number of challenges with deployed policies.

More recent work at the W3C centered around the appropriateness of various privacy features to be included in the Geolocation API [Geolocation], which gives web sites a way to access the user's precise location. The API requires that implementations obtain user consent before accessing location information and allow users to revoke that consent, but decisions about retention, secondary use, and data minimization are left up to individual web sites and applications. The geolocation effort and the P3P experience both raise questions about how to navigate usability, regulation, business incentives, and other aspects that normally lie outside the scope of standards development organization (SDO) work.

3. Design Challenges

Workshop discussions surfaced a number of key issues that can make designing privacy-sensitive protocols and systems difficult: the increasing ease of user/device/application fingerprinting, unforeseen information leakage, difficulties in distinguishing first parties from third parties, complications arising from system dependencies, and the lack of transparency and user awareness of privacy risks and tradeoffs.

3.1. Ease of Fingerprinting

Internet applications and protocols now share so many unique identifiers and other bits of information as part of their ordinary operation that it is becoming increasingly easy for remote nodes to create unique device or application fingerprints and re-identify the same devices or applications over time [Panopticlick]. Hardware identifiers, IP addresses, transport protocol parameters, cookies, other forms of web storage, and a vast array of browser-based information may be routinely shared as users browse the web. The ease of fingerprinting presents a significant challenge for any application that seeks to guarantee anonymity or unlinkability (such as [Tor], which uses onion routing to strip out data that identifies communications endpoints).

In many cases, the information that can be used to fingerprint a device was not originally shared for that purpose; identifiers and other information are provided to support some other functionality (like IP addresses being shared in order to route packets), and may incidentally be used to fingerprint. This complicates the task of preventing fingerprinting, because each application or protocol likely needs its own identifiers and information to function.

Furthermore, some services are increasingly coming to rely on fingerprinting in order to detect fraud or provide customized content, for example. Finding privacy-friendly substitutes for fingerprinting will only become more difficult as these services become more entrenched (see Section 4.3).

The space of fingerprinting mitigations requires further exploration. For example, workshop participants discussed the use of JavaScript queries to obtain a browser's (often highly unique) font list, and the tradeoffs associated with browsers instead (or additionally) supporting some small subset of fonts in order to reduce browser identifiability. As with many other privacy features, such a restriction presents a tradeoff between privacy and usability, and in the case of fingerprinting writ large, it may be difficult to find consensus about which mitigations appropriately balance both values. As a first step, the IETF may consider documenting the fingerprinting implications for widely used IETF protocols (TCP, HTTP, SIP, etc.).

3.2. Information Leakage

Internet protocols and services tend to leak information in ways that were not foreseen at design time, as explored during the IETF 77 technical plenary [IETF77] and in recent research [PrivLoss] [PrivDiffus]. For example, the HTTP referrer header [RFC2616] (misspelled in the original specification as "Referer") provides a way for a web site to obtain the URI of the resource that referred the user to the site. Referrer headers provide valuable insights to web sites about where their users come from, but they can also leak sensitive information (search terms or user IDs, for example), because URI strings on the web often contain this information. The infrastructure of an individual web site is often designed solely with a view to making the site itself function properly, and embedding search terms or other user-specific information in URIs may serve that goal, but when those URIs leak out to other sites via a referrer header, it creates the potential for third parties to use and abuse the data contained therein.

The use of URIs for authentication of identity or capabilities can be susceptible to the same kinds of problems. Relying on a "possession model" where any user in possession of an authentication or capability URI can gain access to a resource is only suitable in situations with some means of control over URI distribution, and can lead to wide leakage when used on the open web.

3.3. Differentiating between First and Third Parties

Distinguishing between "first-party" interactions and "third-party" interactions is important for understanding the implications of data collection, sharing, and use that take place during the normal course of web use. Unfortunately, the traditional meanings of these concepts do not always clearly match up with user expectations or evolving web technologies. Traditionally, the term "first party" has been used to refer to the domain of a web site to which a user agent directs an explicit request on behalf of a user. The term "third party" has been used to refer to the domain of a web resource that a user agent requests as a result of a first-party request, with the third-party resource hosted at a different domain from the first-party domain.

This distinction between first-party and third-party domains is in part a result of long-standing user agent practices for handling HTTP cookies. Typically, HTTP cookies are returned only to the origin server that set them [RFC6265]. Cookies set from first-party domains may not be read by third-party domains and vice versa. In some cases, cookies set from first-party domains that contain subdomains are accessible by all subdomains of the first-party domain. The distinction between first-party domains and third-party domains is reflected in browser-based cookie controls: major web browsers all offer distinct first-party cookie settings and third-party cookie settings.

However, a user's perception or expectation of the difference between a "first party" and a "third party" may not fall neatly within these distinctions. Users may expect that content hosted on a first-party subdomain, but provided or used by a third party, would be treated as third-party content, but browsers often treat it as first-party content. Conversely, when third-party content appears from a source with which the user has an established relationship -- such as the Facebook "Like" button or other social widgets -- users may consider their interaction with that content to be a desirable first-party interaction, even though the content is hosted on a third-party domain.

Handling these expectations programmatically is difficult, since the same identifier structures (domains, subdomains) can correlate to different user expectations in different contexts. On the other hand, prompting users to express a preference about what kinds of data collection and use should be allowable by each party encountered on the web is not practical. Web and browser developers are actively seeking novel ways to address this challenge, but there are few clear-cut solutions.

3.4. Lack of Transparency and User Awareness

There is no question that users lack a full understanding of how their information is being used and what the tradeoffs are between having their data collected and accessing services at little or no cost. Much of the tracking that takes place on the web is passive and invisible to users. Most companies disclose their data usage practices in written privacy policies, but these policies are rarely read, difficult to understand, and often fail to disclose salient details (such as data retention lifetimes). Even when web tracking is associated with some visual indication -- a highly targeted Gmail ad or the Facebook "Like" button, for example -- users often do not realize that it is occurring.

Efforts abound to attempt to present information about data collection and usage in a more digestible way. P3P was one early effort, but because it sought to support the expression of the vast expanse of potential policies that companies may have, it developed more complexity than the average user (or user interface) could sustain. More recent efforts have focused on using a limited set of icons to represent policies or provide an indication that tracking is taking place.

4. Deployment and Analysis Challenges

Workshop participants identified a number of barriers to both deployment of privacy-protecting technologies and the analysis of the privacy properties of technological systems. These included the difficulty of using generic protocols and tools to defend against context-specific threats; the tension between privacy protection and usability; and the difficulty of navigating between business, legal, and individual incentives.

4.1. Generative Protocols vs. Contextual Threats

Privacy is not a binary state. Rather than operating either entirely in private or entirely in public, individuals experience privacy contextually, resulting in differing requirements for privacy protection, depending on the circumstance and the individual. On the Internet, the contextual nature of privacy means that threats against it can vary, depending on the deployment scenario, the usage scenario, the capabilities of different attackers, and the level of concern that different kinds of attackers generate among different users.

Addressing the full waterfront of privacy threats within generic protocols and tools is largely intractable. As a result, existing privacy features developed at the network and application layers have taken more targeted approaches. For example, privacy extensions for stateless address autoconfiguration in IPv6 [RFC4941] support addresses constructed dynamically rather than generating addresses based on interface Media Access Control (MAC) addresses, which for most users are persistent and unchangeable unique identifiers that could be used for long-term tracking. While IPv6 privacy extensions provide important protection against tracking and re-identification by remote endpoints, they do not prevent -- and were not meant to prevent -- all parties from being able to associate an IP address with a particular user. ISPs and governments still have means to make such associations, and remote endpoints have many other mechanisms at their disposal to attempt to identify users persistently, albeit without using IPv6 addresses.

This kind of experience with developing privacy tools shows that designing privacy features into systems and protocols requires a clear understanding of the scope of the threats they are designed to address. This scope is currently being debated in discussion about developing "Do Not Track" (DNT) mechanisms for the web and other online contexts. A number of different approaches have been proposed, including browser functionality to retain opt-out cookies, an HTTP header that expresses the user's preference not to be tracked, and a browser-based block list mechanism that prevents the browser from communicating with tracking sites (for an overview, see [OptOuts]). Regardless of the approach, these mechanisms function based on some understanding of which "tracking" users should be able to control, which in turn is based on some notion of the threats presented by different kinds of tracking conducted by different kinds of entities on the web. Should DNT mechanisms apply to sites with which the user already has an established relationship? Or sites that use only aggregate, non-individualized data? Does tracking for fraud prevention or customization present different threats than tracking for advertising or marketing purposes? The answers to these questions will dictate DNT design choices.

The space of privacy threats on the Internet may appear particularly broad from a protocol design perspective, because many of the protocols in widest use are designed generically to support a variety of applications and functionality. HTTP, for example, is used for a wider variety of purposes than its original designers likely anticipated; it is unsurprising that some of these purposes include obtaining and using data about web users in ways that may be privacy-infringing. It is unreasonable to ask protocol designers to mitigate the potential privacy risks of every possible deployment that may result from a particular protocol design; the key questions are about

how the responsibility for protecting against privacy intrusion should be split between protocols, APIs, applications, and services, and which kinds of privacy features can best be implemented in each place.

4.2. Tension between Privacy Protection and Usability

The workshop discussions highlighted the tension between providing privacy protections and maintaining usability. Tor [Tor] provides some salient examples of this tradeoff. Tor seeks to provide protection against network surveillance, but by lengthening the routing path, it may significantly increase round-trip time. Tor obscures endpoint IP addresses; thus, it also interferes with IP-based geolocation. Web browsing using Tor is particularly challenging, as most browser plug-ins, much of JavaScript, and a number of other browser-based features need to be blocked or overridden in order to meet Tor's anonymity requirements. With Tor, privacy clearly comes at a price.

Even less aggressive privacy features may come with usability tradeoffs. One example is the blocking of HTTP referrer headers for privacy protection reasons. Some sites provide a customized experience to users based on the referring page, which means that disabling referrer headers, as some browsers allow users to do, may sacrifice user experience features on certain sites. Part of the challenge is the level of nuance involved in making decisions about privacy -- how can users be made to understand the privacy tradeoffs of blocking HTTP referrer headers, for example, when the effects of doing so will vary from site to site, or when there is limited UI space to communicate the tradeoffs? Even seemingly simple privacy controls like private browsing are not well understood.

The feature set that implementors choose to make available is often reflective of the tension between usability and privacy. For example, SIP [RFC3261] supports Secure/Multipurpose Internet Mail Extensions (S/MIME) to secure SIP request bodies, but given its user experience impact, few implementations include S/MIME support. Although usability challenges are generally thought of as user-level issues that are out of scope for the IETF, to the extent that they trickle down into implementation decisions, they are highly relevant.

Although workshop participants reached few firm conclusions about how to tackle usability issues arising from privacy features, the group agreed that it may be beneficial for the W3C to do some more thinking in this area, possibly toward the end of including usability considerations in individual specifications. The challenge with such an effort will be to provide useful guidance without being overly prescriptive about how implementations should be designed.

4.3. Interaction between Business, Legal, and Technical Incentives

4.3.1. Role of Regulation

The Internet has sustained commercial content for decades. Many services are offered at little or no cost in exchange for being able to sell advertising or collect user data (or both). As the commercial value of the web in particular has exploded in recent years, the paradigm for regulating privacy has also begun to change, albeit more slowly.

At the dawn of the commercial Internet, few web sites had written privacy policies that explained what they did with user data. Under regulatory pressure, sites began to document their data collection and usage practices in publicly posted policies. These policies quickly became lengthy legal documents that commercial sites could use to limit their liability, often by disclosing every possible practice that the site might engage in, rather than informing users about the salient practices of relevance to them.

Because so many businesses are fueled by user data, any move to give users greater control over their data -- whether by better informing them about its use or providing tools and settings -- often requires the force of regulatory influence to succeed. In recent years, regulatory authorities have put pressure on companies to improve their privacy disclosures by making them simpler, more concise, more prominent, and more accessible (see the 2010 Federal Trade Commission privacy report [FTC]). Certain companies and industry sectors have responded by developing privacy icons, using short notices in addition to privacy policies, and making the language they use to describe privacy practices more accessible and easier to understand.

Regulators play an important role in shaping incentive structures. Companies often seek a balance between acting to limit their liability and pushing the envelope with respect to uses of consumer data. If regulators take a strong stand against certain practices -- as, for example, European legislators have against cookies being set without user consent [Directive] -- legitimate businesses will feel compelled to comply. But where there is regulatory uncertainty, business responses may differ according to different market strategies. The variety of potential responses to the emerging discussion about mechanisms to control web tracking demonstrates this variation: some businesses will embrace support for enhanced user control, others may restrict their offerings or charge fees if they are unable to track users, and still others may elect to circumvent any new mechanisms put in place. The absence of regulatory pressure tends to make the line between "good" and "bad" actors less evident.

4.3.2. P3P: A Case Study of the Importance of Incentives

That absence of regulatory pressure revealed itself in the case of P3P. The first version of P3P was standardized in the early 2000s, when legalistic privacy policies were the norm and users had only elementary controls over the data collected about them on the web. P3P challenged that paradigm by providing a way for web sites to express machine-readable privacy policies for browsers to vet and possibly override according to the user's preference. The P3P policy expression language was designed to allow sites to make complex assertions about how they intended to make use of data related to users.

The designers of Internet Explorer 6 made a crucial decision to only allow sites to use third-party cookies if they had installed adequate P3P policies. To avoid having their cookies blocked, most commercial sites adopted some P3P policy, although many sites merely cut and pasted from the example policies provided by the W3C. Today, large numbers of sites are misrepresenting their privacy practices in their P3P policies, but little has been done in response [Policies], and browser support for P3P outside of IE is limited.

While theories abound to explain the current status of P3P implementations, there is no doubt that the relationship between regulatory and commercial incentives played a significant role. The P3P policy expression language provided support for companies to be able to express in granular detail how they handle user data, but the companies had little reason to do so, preferring to protect themselves from the liability associated with revealing potentially unsavory practices. In theory, the threat of regulatory backlash could have served as an incentive to publish accurate P3P policies, but at the time of P3P's release, there was little regulatory interest in moving beyond long, legalistic privacy policies. Even today, regulators are reluctant to bring enforcement actions against companies with misleading policies, perhaps because their own incentive structure compels them to focus on other, more prominent matters.

The P3P experience is instructive in general for attempts at crafting privacy features that require the active participation of both ends of a communication. Actors that are meant to articulate their own privacy preferences, whether they be companies or individuals, require incentives to do so, as do those that are meant to process and react to such preferences. For example, the IETF's GEOPRIV architecture allows for expression of user preferences about location information [RFC4119]. While users may have more incentive to disclose their privacy preferences than companies did in the P3P case, successful use of the GEOPRIV model will require endpoints that

consume location information to abide by those preferences, and in certain contexts -- commercial or employment-related, for example -- they may be unwilling, or regulatory pressure may be required to spur a change in practice.

It is clearly not the prerogative of Internet protocol developers to seek to change existing incentive structures. But acknowledging what motivates businesses, individuals, and regulators is crucial to determining whether new privacy technologies will succeed or fail.

5. Conclusions and Next Steps

5.1. IETF Outlook

The workshop demonstrated that the understanding of how to address privacy within the Internet standards community is nascent. The IETF faces particular challenges, because IETF protocols generally do not mandate implementation styles or pre-conceive particular deployment contexts, making the space of potential privacy threats attributable to any single protocol difficult to foresee at protocol design time.

Workshop participants nonetheless outlined a number of potential next steps. Work has already begun to attempt to provide guidance to protocol designers about the privacy impact of their specifications [PrivCons]. In refining this guidance, many of the questions raised at the workshop will need to be confronted, including those about how to properly model privacy threats against generic protocols, how to anticipate privacy risks that have been exposed in the previous design efforts, and how to document risks that are more difficult to foresee and mitigate. Workshop participants acknowledged that developing such guidance is likely necessary if document authors are expected to incorporate "Privacy Considerations" sections in their documents, but even with guidance, this is likely to be an uphill battle for many authors for some time to come.

As preliminary steps, those with privacy expertise may seek to apply the current guidance to existing IETF protocols. The security area directors have also created a privacy directorate where privacy reviews of documents coming before the IESG are being conducted.

Participants also expressed an interest in further pursuing a number of the technical topics discussed at the workshop, including lessons learned from the experience of Tor and the fingerprinting implications of HTTP, TCP, SIP, and other IETF protocols. These and other efforts may be explored within the Internet Research Task Force (IRTF) in addition to, or in lieu of, the IETF.

5.2. W3C Outlook

The W3C is likewise in a position of seeking a more comprehensive approach to privacy within the SDO. Because the work of the W3C operates within a more defined scope than that of the IETF -- namely, the web -- the questions before the W3C tend to lie more in the space of distinguishing between what can appropriately be accomplished within W3C specifications and what should be left to individual implementations, a theme that repeated itself again and again at the workshop.

To further develop its approach to privacy, the W3C will investigate an interest group to discuss privacy topics. Some potential topics that emerged from the workshop include the fingerprinting impact of W3C protocols, data minimization in APIs, dealing with referrer header privacy leakage, developing privacy considerations for non-browser-based protocols, and developing usability considerations as part of specification design.

5.3. Other Future Work

The workshop covered a number of topics that may deserve further exploration in the IETF, the W3C, and the privacy community at large. These include development of privacy terminology; articulation of privacy threat models; analysis and experimentation with "Do Not Track" mechanisms for the web; work on cross-site data sharing, correlation, and linkability in web and non-web contexts; and investigation of policy expression languages.

6. Acknowledgements

Thanks to Bernard Aboba, Nick Doty, and Hannes Tschofenig for their early reviews.

7. Security Considerations

Workshop participants discussed security aspects related to privacy, acknowledging that while much of the standards community may have once viewed most relevant privacy concerns as being encompassed by security considerations, there is a growing realization of privacy threats that lie outside the security realm. These include concerns related to data minimization, identifiability, and secondary use. Earlier security work provided minimal provision for privacy protection (e.g., the definition of "privacy" in [RFC2828] and some guidance about private information in [RFC3552]).

8. Informative References

[ALTO] IETF, "Application-Layer Traffic Optimization (alto)", 2011, <http://datatracker.ietf.org/wg/alto/charter/>.

[Directive] European Parliament and Council of the European Union, "Directive 2009/136/EC of the European Parliament and of the Council", November 2009, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:01:EN:HTML>.

[FTC] Federal Trade Commission Staff, "A Preliminary FTC Staff Report on Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers", December 2010, <http://www.ftc.gov/opa/2010/12/privacyreport.shtm>.

[Geolocation] Popescu, A., Ed., "Geolocation API Specification", September 2010, <http://www.w3.org/TR/2010/CR-geolocation-API-20100907/>.

[IETF77] Krishnamurthy, B., "Privacy Leakage on the Internet", March 2010, <http://www.ietf.org/proceedings/77/slides/plenaryt-5.pdf>.

[OptOuts] Cooper, A. and H. Tschofenig, "Overview of Universal Opt-Out Mechanisms for Web Tracking", Work in Progress, March 2011.

[P3P] Wenning, R., Ed., and M. Schunter, Ed., "The Platform for Privacy Preferences 1.1 (P3P1.1) Specification", November 2006, <http://www.w3.org/TR/P3P11/>.

[Panopticlick] Electronic Frontier Foundation, "Panopticlick", 2011, <http://panopticlick.eff.org/>.

[Policies] Leon, P., Cranor, L., McDonald, A., and R. McGuire, "Token Attempt: The Misrepresentation of Website Privacy Policies through the Misuse of P3P Compact Policy Tokens", September 2010, <http://www.cylab.cmu.edu/research/techreports/2010/tr_cylab10014.html>.

[PrivCons] Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., and J. Morris, "Privacy Considerations for Internet Protocols", Work in Progress, October 2011.

[PrivDiffus] Krishnamurthy, B. and C. Wills, "Privacy Diffusion on the Web: A Longitudinal Perspective", Proceedings of the World Wide Web Conference, pages 541-550, Madrid, Spain, April 2009, <http://www.cs.wpi.edu/~cew/papers/www09.pdf>.

[PrivLoss] Krishnamurthy, B., Malandrino, D., and C. Wills, "Measuring Privacy Loss and the Impact of Privacy Protection in Web Browsing", Proceedings of the Symposium on Usable Privacy and Security, pages 52-63, Pittsburgh, PA USA, ACM International Conference Proceedings Series, July 2007, <http://www.cs.wpi.edu/~cew/papers/soups07.pdf>.

[RFC1543] Postel, J., "Instructions to RFC Authors", RFC 1543, October 1993.

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[RFC2778] Day, M., Rosenberg, J., and H. Sugano, "A Model for Presence and Instant Messaging", RFC 2778, February 2000.

[RFC2828] Shirey, R., "Internet Security Glossary", RFC 2828, May 2000.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[RFC3323] Peterson, J., "A Privacy Mechanism for the Session Initiation Protocol (SIP)", RFC 3323, November 2002.

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.

[RFC3693] Cuellar, J., Morris, J., Mulligan, D., Peterson, J., and J. Polk, "Geopriv Requirements", RFC 3693, February 2004.

[RFC4119] Peterson, J., "A Presence-based GEOPRIV Location Object Format", RFC 4119, December 2005.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

[RFC6265] Barth, A., "HTTP State Management Mechanism", RFC 6265, April 2011.

[Tor] The Tor Project, Inc., "Tor", 2011, <https://www.torproject.org/>.

[Workshop] IAB, W3C, ISOC, MIT CSAIL, "Internet Privacy Workshop 2010", 2011, <http://www.iab.org/activities/workshops/internet-privacy-workshop-2010/>.

Appendix A. Workshop Materials

Main page: http://www.iab.org/activities/workshops/internet-privacy-workshop-2010/

Slides: http://www.iab.org/activities/workshops/internet-privacy-workshop-2010/slides/

Minutes: http://www.iab.org/activities/workshops/internet-privacy-workshop-2010/minutes/

Position papers: http://www.iab.org/activities/workshops/internet-privacy-workshop-2010/papers/

Appendix B. Workshop Participants

- Fu-Ming Shih, MIT
- Ian Jacobi, MIT
- Steve Woodrow, MIT
- Nick Mathewson, The Tor Project
- Peter Eckersley, Electronic Frontier Foundation
- John Klensin, IAB
- Oliver Hanka, Technical University Munich
- Alan Mislove, Northeastern University
- Ashkan Soltani, FTC
- Sam Hartman, Painless Security
- Kevin Trilli, TRUSTe
- Dorothy Gellert, InterDigital
- Aaron Falk, Raytheon - BBN Technologies
- Sean Turner, IECA
- Wei-Yeh Lee, NAVTEQ
- Chad McClung, The Boeing Company
- Jan Seedorf, NEC
- Dave Crocker, Brandenburg InternetWorking
- Lorrie Cranor, Carnegie Mellon University
- Noah Mendelsohn, W3C TAG Chair
- Stefan Winter, RESTENA
- Craig Wittenberg, Microsoft
- Bernard Aboba, IAB/Microsoft
- Heather West, Google
- Blaine Cook, British Telecom
- Kasey Chappelle, Vodafone Group
- Russ Housley, IETF Chair/Vigil Security, LLC
- Daniel Appelquist, Vodafone R&D
- Olaf Kolkman, IAB Chair
- Jon Peterson, IAB/NeuStar, Inc.
- Balachander Krishnamurthy, AT&T Labs--Research
- Marc Linsner, Cisco Systems

- Jorge Cuellar, Siemens AG
- Arvind Narayanan, Stanford University
- Eric Rescorla, Skype
- Cullen Jennings, Cisco
- Christine Runnegar, Internet Society
- Alissa Cooper, Center for Democracy & Technology
- Jim Fenton, Cisco
- Oshani Seneviratne, MIT
- Lalana Kagal, MIT
- Fred Carter, Information & Privacy Commissioner of Ontario, Canada
- Frederick Hirsch, Nokia
- Benjamin Heitmann, DERI, NUI Galway, Ireland
- John Linn, RSA, The Security Division of EMC
- Paul Trevithick, Azigo
- Ari Schwartz, National Institute of Standards and Technology
- David Evans, University of Cambridge
- Nick Doty, UC Berkeley, School of Information
- Sharon Paradesi, MIT
- Jonathan Mayer, Stanford University
- David Maher, Intertrust
- Brett McDowell, PayPal
- Leucio Antonio Cutillo, Eurecom
- Susan Landau, Radcliffe Institute for Advanced Study, Harvard University
- Christopher Soghoian, FTC In-house Technologist, Center for Applied Cybersecurity Research, Indiana University
- Trent Adams, Internet Society
- Thomas Roessler, W3C
- Karen O'Donoghue, ISOC
- Hannes Tschofenig, IAB/Nokia Siemens Networks
- Lucy Elizabeth Lynch, Internet Society
- Karen Sollins, MIT
- Tim Berners-Lee, W3C

Appendix C. Accepted Position Papers

1. "Addressing the privacy management crisis in online social networks" by Krishna Gummadi, Balachander Krishnamurthy, and Alan Mislove

2. "Thoughts on Adding "Privacy Considerations" to Internet Drafts" by Alissa Cooper and John Morris

3. "Toward Objective Global Privacy Standards" by Ari Schwartz

4. "SocialKeys: Transparent Cryptography via Key Distribution over Social Networks" by Arvind Narayanan

5. "Web Crawlers and Privacy: The Need to Reboot Robots.txt" by Arvind Narayanan and Pete Warden

6. "I Know What You Will Do Next Summer" by Balachander Krishnamurthy

7. "An architecture for privacy-enabled user profile portability on the Web of Data" by Benjamin Heitmann and Conor Hayes

8. "Addressing Identity on the Web" by Blaine Cook

9. "Protection-by-Design: Enhancing ecosystem capabilities to protect personal information" by Jonathan Fox and Brett McDowell

10. "Privacy-preserving identities for a safer, more trusted internet" by Christian Paquin

11. "Why Private Browsing Modes Do Not Deliver Real Privacy" by Christopher Soghoian

12. "Incentives for Privacy" by Cullen Jennings

13. "Joint Privacy Workshop: Position Comments by D. Crocker" by Dave Crocker

14. "Using properties of physical phenomena and information flow control to manage privacy" by David Evans and David M. Eyers

15. "Privacy Approaches for Internet Video Advertising" by Dave Maher

16. "Privacy on the Internet" by Dorothy Gellert

17. "Can We Have a Usable Internet Without User Trackability?" by Eric Rescorla

18. "Privacy by Design: The 7 Foundational Principles -- Implementation and Mapping of Fair Information Practices" by Fred Carter and Ann Cavoukian

19. "Internet Privacy Workshop Position Paper: Privacy and Device APIs" by Frederick Hirsch

20. "Position Paper for Internet Privacy Workshop" by Heather West

21. "I 'like' you, but I hate your apps" by Ian Glazer

22. "Privicons: A approach to communicating privacy preferences between Users" by E. Forrest and J. Schallabock

23. "Privacy Preservation Techniques to establish Trustworthiness for Distributed, Inter-Provider Monitoring" by J. Seedorf, S. Niccolini, A. Sarma, B. Trammell, and G. Bianchi

24. "Trusted Intermediaries as Privacy Agents" by Jim Fenton

25. "Protocols are for sharing" by John Kemp

26. "On Technology and Internet Privacy" by John Linn

27. "Do Not Track: Universal Web Tracking Opt-out" by Jonathan Mayer and Arvind Narayanan

28. "Location Privacy Protection Through Obfuscation" by Jorge Cuellar

29. "Everything we thought we knew about privacy is wrong" by Kasey Chappelle and Dan Appelquist

30. "TRUSTe Position Paper" by Kevin Trilli

31. "Position Paper: Incentives for Adoption of Machine-Readable Privacy Notices" by Lorrie Cranor

32. "Facilitate, don't mandate" by Ari Rabkin, Nick Doty, and Deirdre K. Mulligan

33. "Location Privacy in Next Generation Internet Architectures" by Oliver Hanka

34. "HTTPa: Accountable HTTP" by Oshani Seneviratne and Lalana Kagal

35. "Personal Data Service" by Paul Trevithick

36. "Several Pressing Problems in Hypertext Privacy" by Peter Eckersley

37. "Adding Privacy in Existing Security Systems" by Sam Hartman

38. "Mobility and Privacy" by S. Brim, M. Linsner, B. McLaughlin, and K. Wierenga

39. "Saveface: Save George's faces in Social Networks where Contexts Collapse" by Fuming Shih and Sharon Paradesi

40. "eduroam -- a world-wide network access roaming consortium on the edge of preserving privacy vs. identifying users" by Stefan Winter

41. "Effective Device API Privacy: Protecting Everyone (Not Just the User)" by Susan Landau

42. "Safebook: Privacy Preserving Online Social Network" by L. Antonio Cutillo, R. Molva, and M. Onen

Author's Address

Alissa Cooper
CDT
1634 I Street NW, Suite 1100
Washington, DC 20006
USA

EMail: acooper@cdt.org
URI: http://www.cdt.org/