Position Paper for
W3C Workshop: Do Not Track and Beyond
26-27 November 2012

Respectfully submitted to the W3C by:
David Wainberg
Counsel & Senior Director of Technology
Network Advertising Initiative

Please find below some brief remarks submitted for consideration during the W3C's upcoming "Do Not Track and Beyond" workshop. These remarks include background on the Network Advertising Initiative (NAI), a general explanation of the indispensible role that NAI members play in supporting content and services online, and an explanation of why the NAI's self-regulatory model is a success. I hope that this will be useful input to the particular topic of Do Not Track, which primarily affects NAI members, as well as the general discussion of the future of privacy public policy.

About the NAI

The NAI is the leading self-regulatory organization in the U.S. focused exclusively on the third-party online advertising ecosystem. With over 90 member companies, ranging from the largest online businesses in the world, to small, nascent startups, nearly every ad served online in the United States involves the technology of one or more NAI members. NAI members provide crucial services and infrastructure to the diverse range of ad-supported content and services users enjoy every day, from major news websites to niche blogs to independently developed apps.

The NAI was formed in response to early concerns regarding the collection and use of data by third parties for the purpose of making ads relevant to users. The first NAI Code of Conduct was published in 2000. The current NAI Code,[1] published in 2008, reflects the evolution of third party business models, and the accompanying evolution in privacy concerns. The Code includes requirements for notice and choice, as well as other requirements and restrictions around the collection and use of data for online advertising, including limitations on the use of PII and limitations on the use of various types of sensitive data.

Additionally, the NAI implements and enforces policies to further the goals underlying the Code. For example, in 2009, the NAI enacted a policy limiting the use of Flash Cookies, LSOs, and other technologies that do not provide users "an appropriate degree of transparency and control."[2] This is an example of the NAI's unique ability to quickly build and implement industry consensus around policy issues.

---

[1] http://www.networkadvertising.org/sites/default/files/imce/principles.pdf
[2] http://www.networkadvertising.org/faq#n185

The NAI maintains a rigorous annual compliance program to enforce its Code and associated policies.[3] The NAI's team of lawyers and technologists are given considerable access to members' employees and proprietary information. NAI Staff use this access to assess whether its members are in compliance wit the Code, as well as to develop best practices for online advertising practices. This virtuous cycle is one of the key benefits of the self-regulatory model.

The NAI continues to innovate and raise the bar for online privacy. In 2009, the NAI joined other associations in creating the Digital Advertising Alliance, [4] a fast-growing self-regulatory body that now encompasses the entire online advertising ecosystem, and includes members ranging from small advertising networks to major brands. The NAI participates in and contributes to a wide variety of meetings, workshops, working groups, and other proceedings related to all areas of online privacy. And the NAI is currently in the process of further updating its Code of Conduct to keep pace with the rapidly evolving practices and technology of the online advertising ecosystem.

NAI Members Are an Indispensible Part of the Internet Ecosystem

NAI members are key constituents in supporting diverse content and driving innovation. Their contributions benefit publishers and advertisers, online innovation generally, and therefore ultimately benefit users.

Good for publishers, especially small publishers

NAI members are indispensible for online sites and services that rely on advertising for financial support. For many sites and services, third-party interest-based advertising is the only feasible form of economic support. Without the revenue made possible by interest-based advertising, they would be forced to charge users for the content they provide, or to cease operation. While major publishers have the audience and the market power to garner relatively high rates for their advertising inventory even without interest-based advertising, small- and mid-size publishers do not.

Consider this example. Imagine a niche ad-supported publisher of a site catering to aspiring writers of science fiction. Because the site cannot afford its own sales staff or other overhead associated with making direct sales of its advertising inventory, it must rely on a third-party service to generate ad revenue. And because the market for contextual ads directly aimed at science fiction writers is too small, the site cannot depend on contextual advertising alone.

If limited to contextual ads, the value of the site's ad inventory will be a small fraction of what it could be if sold via interest-based advertising. This is because

---

[3] http://www.networkadvertising.org/code-enforcement/enforcement
[4] http://www.aboutads.info/

inferences about visitors' interests that are carried over from unrelated sites make this site's ad inventory significantly more valuable. Some of the site's visitors may be in the market for a new car. Others may be planning travel. Others may be interested in sporting goods. Still others may be shopping for stereo equipment. There is no way these interests could be inferred from the context of the site; these inferences can only be made based on the users' behavior on unrelated sites. With knowledge about the users' interests (using only non-Personally Identifiable Information), a much broader base of advertisers are willing to purchase the science fiction site's inventory, and they are willing to pay a higher rate than the site would get based on context alone. This is why third-party Interest Based Advertising is responsible for much of the free content and services available to users, and why it is a crucial component of the diverse and democratic Internet users enjoy today.

Good for advertisers, especially small advertisers

For advertisers, this means easy access to quality supply at competitive prices. In other words, advertisers, whether large multi-national brands, or small local businesses, depend on interest-based advertising to efficiently and effectively reach potential customers.

Take the example of a small, family-owned restaurant wanting to advertise online. Limited to contextual advertising, how would they find their customers? Limiting their ad buying to contextual would be inefficient and would limit their reach. Using third party services, however, they can find their audience across the web. They might even buy ads on the science fiction site from the example above.

Similarly, consider the example of a small manufacturer of specialized components for industrial systems. They are trying to reach an international market. The availability of relevant contextual ad inventory is extremely limited. Also, they are young and small, and they must use their marketing budget carefully. Only with interest-based advertising can they reach their relatively small audience across the entire Internet, even on websites not directly relevant to industrial systems. This is good for them, and good for the sites on which they advertise.

Good for innovation and competition

Consider also that this demand for third-party services drives innovation and competition, because it creates a market with minimal barriers to entry, where companies can launch quickly and the network effects compound rapidly. This has given rise to an explosion of competitive young companies. In the US alone, the market includes hundreds of companies, employs thousands (maybe hundreds of thousands) of people, and generates billions of dollars while greatly supporting the economic vitality of small and medium businesses.

Good for users

Of course, in the end, Internet users benefit from this Interest-Based Advertising, because it creates a more dynamic, distributed, and democratic Internet, with a wide range of free content and services.

Self-regulation works

Despite these benefits of third-party online advertising, the industry has acknowledged that users' privacy must be considered. As a result, industry has committed, through self-regulatory associations, to maintain and enforce strict codes of conduct regarding the collection and use of data for online advertising purposes.

Self-regulation is an ideal model for this for several reasons that culminate in fast and effective responses to policy issues:

- Expertise. NAI staff draw on their own considerable expertise in law, business, technology, and privacy, and are also able to draw on the vast expertise of the NAI's member companies, to fully understand issues, and the technology underlying the issues, and the potential consequences of proposed policy responses. This can occur in a cooperative and transparent manner. Because member companies give NAI staff exceptional access to personnel and proprietary information, staff and members can have frank and thorough discussions that would otherwise not occur. This leads to a deep understanding of issues and problems, and so leads to more effective solutions.
- Buy-in. The NAI has buy-in from its members and from the industry as a whole. This helps the NAI gain wide adoption of its policies and best practices.
- Enforceability. Through NAI members' public commitments to abide by the NAI Code, they become legally bound to do so, and face liability when they do not. Beyond that, however, the NAI's expertise, in conjunction with its access to inside personnel and information, make it uniquely situated to discover and mitigate issues on an ongoing basis. In fact, this happens proactively; NAI member companies regularly contact NAI staff to discuss the privacy implications of technology and product ideas before they are launched.

Thanks to these advantages, the NAI continues to have ongoing success in raising the bar for online privacy. And the entire ecosystem benefits from this success.

Conclusion

In the interest of brevity, I have not gone deeply into any of the points above. Rather, I hope these general points will be topics of deeper discussion during the upcoming workshop. Given the difficulties and importance of developing sensible, effective policy to regulate practices on the Internet, and given the potential consequences from any "Do Not Track" standard, the group should include the concepts I have introduced into any thinking regarding ways forward with Do Not Track and other policy efforts.

Do Not Track: Children Should be a Top Priority

Society has an obligation to protect our children and online safety for children should be a priority. We need a three-pronged approach to address this issue: policy changes; industry self-regulation; and more parental tools, monitoring and education.

A 2010 *Wall Street Journal* investigation into online privacy, found that popular children's websites install more tracking technologies on personal computers than do the top websites aimed at adults. In fact, the Federal Trade Commission (FTC) says 8 percent of the ID theft complaints in 2010 involved children. Further, Consumer Reports notes that one million children were harassed, threatened, or subjected to other forms of cyberbullying on Facebook in the past year – and that's just one social media site.

Current legislative efforts have failed to act to update COPPA, the industry is failing miserably at self-regulating. As such, it is essential that thought leaders such as the World Wide Web Consortium come to agreement on standards that legislators will support and the industry will implement – all without stifling innovation. No easy task!

Simply, consumers must be allowed to control their personal data without losing all the benefits offered by the Internet. As an example, many Facebook apps require absolute access to all the personal data of a user in order to participate. This is a false choice. Consumers should continue to be able to participate online without being forced to give up their personal privacy. Children should be able to go online, play and learn without leaving a digital track that haunts them for life.

As the media continues to expose the prevalence of the collection and aggregation of personal information, consumers will demand change. Even with action from Congress and more regulation of the online marketing industry, we believe that to keep up with new technologies, the private market must be a large part of the solution.

We need to look at the broader implications and responsibility of Silicon Valley, and not simply the profit margins of data aggregators and marketing agencies. The short-term “get it while we can” mentality has long-term consequences for the profitability of technology and online marketing companies: they will continue to lose trust with consumers unless they step up and support best practices that protect their own revenue source – people. How many more articles like the New York Times’ *You for Sale: Mapping, and Sharing, the Consumer Genome*” will it take before consumers begin to exodus Facebook and Google. How long will it take before McDonald’s HappyMeal.com and CartoonNetwork.com are blacklisted from the family computer?

Parents certainly play an important role in modulating their kids’ online lifestyle, and parents should be educated about best practices and the tools already at their disposal to monitor and moderate what content their kids see.

However, our kids should be allowed access to the wealth of information and entertainment the Internet has to offer...just not at the cost of their safety and privacy..

It is currently a big line between data use and abuse. The debate around Do Not Track standards should serve to define that line. Focusing on children first is not only a societal obligation but the path to a faster resolution of a growing problem that threatens technological innovation and our access to the Web.

Privacy – From Principles to Technology Standards

Information Privacy is a critical success factor for the success of the emerging global digital market place[1]. Consumers will only shop in a market place that they trust. Technology standards will play an important role in progressing support for privacy in the information society accessing this market place[2][3]. Nokia supports Privacy by Design[4]. Privacy by Design is essential to assuring consumers that the underlying technology infrastructure of information society meets their privacy expectations.

Solving the privacy challenge requires multi-party efforts. Regulators and policy makers need to work on privacy principles and identify clear privacy objectives. Advocacy needs to raise awareness and help bring issues to public discussion. Industry and technology representatives need to identify best practices for translating privacy principles into concrete technology solutions and drive privacy objectives into underlying technology infrastructure of information society.

The role of technology standardization is essential for the above objective. Currently, many standards development organizations (SDO) and industry fora are working on technology standards and industry specifications for the information society infrastructure[5]. At the same time, the number of engineers with Privacy Engineering[6] skills is limited. To avoid fragmentation and to help build privacy engineering competency, SDOs and industry fora should:

1. Document the group's privacy commitment and endorse it at the highest level in the organization's management;
2. Identify a permanent group or an individual with responsibility to oversee privacy implications of the organization's work items;
3. Include a section on "privacy considerations" in each of the organization's specifications and make it mandatory for all future specifications;
4. Review and update existing standards to include a "privacy considerations" section;
5. Begin to document the Best-Practices for Privacy Controls and publish them as Privacy Design Patterns, so they can be shared across the industry.

Some SDO and industry fora, such as ISO and IETF have formalized a central steering group or body of experts to address privacy related matters in their organizations. This approach means that privacy considerations are addressed in a timely and coordinated manner. Nokia supports this approach and believes it should be adopted commonly by all SDOs and industry fora.

---

[1] Kroes, N., Cloud Computing and Data protection reform, http://blogs.ec.europa.eu/neelie-kroes/cloud-data-protection, 2012.
[2] Cranor, L., The Role of Technology in Self-Regulatory Privacy Regimes, NTIA submission, http://lorrie.cranor.org/pubs/NTIA.html, 1996.
[3] IPO, Privacy Dividend, http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/Privacy_Dividend.pdf, 2010.
[4] Information & Privacy Commissioner, Ontario, Canada, Introduction to Privacy by Design, http://www.ipc.on.ca/english/Privacy/Introduction-to-PbD/, 2012.
[5] Ernst & Young, Privacy Trends 2012, http://www.ey.com/Publication/ vwLUAssets/Privacy_trends_2012/$FILE/ Privacy-trends-2012_AU1064.pdfpp 15-16, 2012.
[6] Privacy Engineering discipline is emerging as a university major study area. See University of Maryland reference at http://www.csee.umbc.edu/2012/04/privacy-engineering/.

A best-practice for businesses is the transparent articulation of their privacy vision. Similarly, SDOs and industry fora should specify the privacy mission for their standardization work. For example, this can be achieved by the management of these organizations explicitly committing to building privacy into their standards and specifications.

SDOs and industry fora need to assure that specifications published by the group include a common section on ***Privacy Considerations***, analogous to the Security Considerations section required in all IETF specifications[7] and is being proposed also by the Internet Architecture Board[8]. The process used to create the Privacy Considerations is analogous to that used in a Privacy Impact Assessment (PIA)[9]. The Privacy Considerations section should identify applicable privacy principles[10], describe possible control points that will allow for inserting privacy enabling technologies, document the personal data collected/used/stored/ transferred/ otherwise processed, identify vulnerabilities that could present a threat to privacy, specify what the risk of harm would be in these events, and document proposed mechanisms such as privacy controls[11] to mitigate the threats.

SDO and industry fora should review their portfolio of active projects. Each of the projects must be reviewed to assess the privacy impact that they might have. When a specification under development is identified as having a privacy impact, they must commit to adding a Privacy Considerations section prior to publication. As existing standards are considered for update or amendment, they too, should be edited to include a Privacy Considerations section.

Privacy is likely to become a major category of technology innovation[12]. Capturing leading and best practice solutions for privacy problems in the standards and specifications will promote the reuse of these solutions and enable growth of Privacy Engineering competency. Documenting these solutions as Privacy Design Patterns[13] using a standard template such as the POSA Template[14] would facilitate that.

Nokia is making these recommendations because of the importance of technology in solving the "privacy challenge" and the importance of open standards and industry specifications to build the required infrastructure. The proposed five recommendations have the potential to be instrumental in promoting SDOs and industry fora development of privacy friendly standards and specifications that leads to the proliferation of a global digital market place that assures consumer trust.

---

[7] RFC 3552, Guidelines for Writing RFC Text on Security Considerations, http://www.ietf.org/rfc/rfc3552.txt, IETF, 2003.
[8] IETF, Privacy Considerations for Internet Protocols, http://tools.ietf.org/html/draft-iab-privacy-considerations-02, 2012.
[9] See UK IPO, Privacy Impact Assessment Handbook, Version 2.0, http://www.ico.gov.uk/upload/documents/pia_handbook_html_v2/files/PIAhandbookV2.pdf, 2009.
[10] For example, the OECD Privacy Principles, http://oecdprivacy.org/, 1980.
[11] For example, NIST, Security and Privacy Controls for Federal Information Processing Systems and Organizations, http://csrc.nist.gov/publications/drafts/800-53-rev4/sp800-53-rev4-ipd.pdf, 2012.
[12] Golfarb, A. Tucker, C, Privacy and Innovation, http://www.nber.org/chapters/c12453.pdf, 2011.
[13] Hafiz,M.: A collection of privacy design patterns. Proc. 2006 Conference on Pattern Languages of Programs, ACM, NY, pp. 1-13. http://dl.acm.org/citation.cfm?id=1415472.1415481, 2006.
[14] F. Buschmann, R. Meunier, H. Rohnert,P. Sommerlad, and M. Stal. A System of Patterns:Pattern-Oriented Software Architecture. Wiley, 1996.

ESPC
Email Sender & Provider Coalition

October 15, 2012

To the Program Committee:

Re: ESPC Position Paper for W3C Workshop: Do Not Track and Beyond

The Email Sender and Provider Coalition ("ESPC") submits this position paper in connection with the forthcoming W3C workshop entitled "Do Not Track and Beyond," which is scheduled for November 26-27, in Berkeley, California.

The ESPC is a cooperative group of industry leaders working to create solutions to the continued proliferation of spam and the emerging problem of legitimate email deliverability. ESPC's membership provides mail delivery services to over one million clients, from small businesses to large Fortune 500 enterprises. The ESPC's mission is to advocate on behalf of email senders, providers, and other digital marketers operating globally in the online, mobile and social media environments in favor of global laws and self-regulatory efforts that balance consumer protection and business innovation; to educate its membership on current and emerging business and legal developments affecting its membership; and to continue to develop and refine best practices (such as our Information Security Best Practices guidelines for email service providers ("ESPs")) that foster innovation, industry growth, and consumer trust.

The W3C's announcement of the Workshop expressly invites participants from industries that "might respond to a Do Not Track ("DNT") preference or use DNT and related technologies for user transparency and choice beyond online behavioral advertising: including, for example, email marketing, mobile application development and online social networking." As the trade association representing the ESP industry, ESPC is uniquely qualified to represent the views of the email service provider industry as a whole.

The ESPC applauds the work done by the W3C in developing, through a consensus process, standards and specifications for web technologies. Recently, the W3C's Tracking Protection Working Group ("TPWG") has taken on complex issues of public policy; in particular, web tracking. One of the key goals of that effort has been to reach, again, through a consensus process, a common meaning to be attributed to a consumer's Do Not Track request to a user agent. That has proved to be a difficult process, and, as of the date of this writing, it is not complete.

ESPC believes that, ultimately, policymakers should resolve public policy issues. We further believe that the experience of the TPWG has demonstrated the difficulty with applying the W3C's technical expertise and consensus approach to public policy issues. That said, we *do* believe that the W3C provides a unique forum to bring together industry, academics, advocates, and other stakeholders to exchange views, engage in fruitful dialog, and to learn from one another. There is no reason why its activities must be limited to technical standards. Public policy, and, in this case, the important issues surrounding consumer privacy, are good candidates

62 Portland Road, Suite 44, Kennebunk, ME 04043 - P 207.467.3500 - Fax 207.985.2523

for healthy and productive dialog in a W3C setting. If the TPWG process has shown anything, it is that the online ecosystem is complex, that there is not a common set of knowledge shared among all stakeholders, that there are a variety of opinions, and that we can all learn from one another. Ultimately, good privacy is good business, and ESPC is eager to learn from other stakeholders how it can, through self-regulation, continue to improve on its' members commitment to consumer privacy. Likewise, we believe that to the extent there may be misperceptions about how our industry operates, we may be able to provide clarity and, thereby, further a productive dialog.

Substantively, the W3C has proposed the following topics for discussion at the workshop. These include:

- Directions for, and input to, the W3C Tracking Protection Working Group's ongoing work on Do Not Track.
- Preliminary implementation experience and impact evaluations of Do Not Track and related approaches.
- Candidates for future W3C standardization on tracking protection in particular, and on user privacy on the Web in general.
- Trends in online privacy issues and potential techniques to address new concerns.

ESPC is unable to comment in detail on the TPWG's DNT standard or implementation experience because it is not yet final, and, as we understand it, key issues remain to be resolved. That said, we recognize that the general framework of the developing standard addresses first parties, service providers, and third parties. In the context of email delivery, our members believe that the advertisers (our members' clients) whose products are services are featured in emails our members cause to be delivered should be considered first parties. We believe that our members should be treated as service providers under the current version of the Tracking Protection and Scope document. As of the date of these comments, there is not a consensus on the definition of "service provider," but we believe that, at its core, the elements necessary to qualify as a service provider should be that the ESP act as an agent for the advertiser, that it silo the data it collects in this capacity so that it cannot be accessed by other parties, that its use of the data it collects is limited to those uses necessary to perform the services for the advertiser, and to those that the advertiser expressly directs, and to a limited set of other uses, such as auditing, reporting, and fraud prevention, akin to the "permitted uses" being debated by the TPWG. We do not believe that any special designation, as a "service provider" or otherwise, of an ESP acting in this capacity is necessary.

ESPC
Email Sender & Provider Coalition

ESPC Position Paper
W3C Workshop: Do Not Track and Beyond
October 15, 2012
p. 3

On behalf of the ESPC, I thank you for this opportunity to submit this position paper. The ESPC looks forward to participating in the Workshop.

Sincerely,

D. Reed Freeman, Jr.
Outside Counsel

cc: ESPC Board of Directors

62 Portland Road, Suite 44, Kennebunk, ME 04043 - P 207.467.3500 - Fax 207.985.2523

Harlan Yu
Principal, Robinson & Yu LLC

W3C Workshop: Do Not Track and Beyond
November 26-27, 2012

I participated early on in the W3C's tracking protection process. I served on the Program Committee for the Workshop on Web Tracking and User Privacy in April 2011. Since then, I've watched the ongoing Tracking Protection Working Group process with great interest, and growing concern.

The Working Group has made many significant strides, and has forged meaningful consensus, on how the technical mechanism for Do Not Track should work. The mechanism appropriately describes not only the HTTP header field, but also a consistent JavaScript representation of the user's preference, and a way for websites to establish exceptions for tracking from the user. These achievements open the possibility for more transparency in corporate data collection practices, and for consumers to better understand how their information is used and to express preferences about such use. Consensus on the technical protocol is critical to paving the path for policies that can improve consumer privacy.

One guiding goal for the Working Group has been to devise a single, global substantive meaning for the tracking preference, that all sites on the Internet would abide by. If that were feasible, it would carry many obvious advantages. However, after a sustained and robust effort, the group has been unable to define tracking, unlinkable data, first parties, or third parties, and has been unable to agree on how servers might address "rogue" user agents, among a host of other hotly disputed substantive questions—including the scope and goals of compliance.

Given the participants' positions, if the Working Group does publish a Compliance and Scope document, either or both of two outcomes seem likely: (1) key stakeholders in industry will vehemently object to its contents, making clear that they aren't part of a consensus on what DNT should mean, and will decline to implement the standard, and/or (2) privacy advocates will press policymakers to institute regulations that are more stringent and exacting than the W3C document specifies.

In either case, it seems unlikely that the resulting Compliance and Scope document will be able to stand on its own to protect user privacy, absent formal regulatory action. While the Working Group's charter specifies that one success criterion is the "[a]doption of deliverables by user agents and compliance by industry," a standard permissive enough to achieve widespread industry adoption will likely, by the same token, be too permissive to address the underlying concerns that triggered this process in the first place.

What this means in practice is that the non-consensus substance of the draft Compliance and Scope document will serve as a starting point for governmental regulators, who can, when and where they so decide, incorporate elements from the W3C draft into their respective regulatory requirements. This may indeed fracture Do Not Track into various regulatory regimes, and while this may not be an optimal outcome, it may in fact be the final result of this process. The end goal of one single global standard is the right one. But whether the W3C can itself establish a meaningful Do Not Track compliance policy, based on the discussions that have taken place, seems unlikely.

I believe the Working Group needs to be realistic about what it is able to accomplish on its own, and seriously contemplate how formal regulatory action can complement the technical consensus that the Working Group has already achieved.

An Advertisers Paradise: An Adventure in a Dystopian Post-"Do Not Track World"?

Ian Oliver
Nokia Location and Commerce
Espoo, Finland

October 2012

1 Introduction

The W3C proposal on Do Not Track (DNT)[1] represents just one technical item in the quest to build into the existing web infrastructure for users to assert their preferences over how the information being collected is processed, and of course, in what manner it is being collected in the first place.

The current DNT proposal at minimum represents a single control point placed in the transport layer, in this case, the HTTP stack. DNT's failings come from a number of misunderstandings and misrepresentations of this, often stemming from the fact that this is just one control point with a certain scope and that the DNT protocol (because of this) is honour based in that it relies upon the good will of the receiver and not on the abilities of the client software closest to the user.

Secondly the scope of DNT is not sufficiently defined and as we have seen, this has extended to cookie management and lately fingerprinting technologies; such as tracking over browser characteristics such as the user agent string.

This paper does not attempt to address DNT's strengths and weaknesses. Indeed it is the belief of the author that this attempt to develop such protocols has positively focussed the discussion of what tracking really means rather than the underlying technology for implementing tracking prevention. At this stage this is an extremely good thing as it finally starts to bring technology itself and the engineering of privacy to the fore in the whole privacy discussions which is somewhat currently overwhelmed by legal and consumer advocacy positions.

[1] W3C (2012) Tracking Preference Expression (DNT) W3C Working Draft 02 October 2012

As the focus here is to look beyond DNT, rather than concentrating on the technology, we can look at some expectations and consequences of such a technology, or at least some of the unexpected consequences that this might lead to.

The format of this is initially presented in the form of a short 'science-fiction' story[2], which may be viewed as somewhat extreme and something that most likely will not happen. However as we have seen in law making, science, economic theory etc, the consequences of one decision can be quite dramatic and unexpected.

It is the author's belief that the somewhat dystopian future described here probably won't happen, or at least, won't happen in the form described here. The purpose of this paper is to direct some focus on to what kind of developments might happen in the light of unexpected consequences of our collective decisions relating to privacy.

As background to this piece, it is an adaption of a short story written by Stephen Baxter called Glass Earth Inc., which appeared in the short story collection Future Histories edited by Stephen McClelland which collectively address possible futures enabled by developments in communication technologies which are enabled by said technologies.[3]

2 One Day in October 2018

Since the World Privacy Laws of 2013 all browsing was legislated to be anonymous. Even to the point that detailed semantic analysis to reveal the user was no-longer permissible. Indeed this had triggered some of the deepest research and insights into semiotics, semantics and information theory and its application into everyday life as revolutionary as the original World Wide Web. By 2017 two Nobel Prizes had been won directly stemming from this work such was its profound nature: one in physics on the subject of semantics and information theory, the other in economics. From the perspective of advertising, today's web is quite unlike the spam filled, chaotic, intrusive and unstructured advertising mess that so amply characterised the first fifteen or so years of the 2000s.

Initially there was a backlash amongst the advertisers and near war between them, the privacy evangelists and the technology providers. The outcry and resultant, hastily passed laws - initially starting in the EU and Canada and (surprisingly) rapidly spreading to the USA enforced anti-tracking compliance. By mid 2015 most advertisers had given up and the once mighty search and social media giants struggling with a need to find a new business model.

[2]NB: The author does not write science-fiction and has no aspirations of doing so as a career - this might be welcomed by the reader(s)

[3]Stephen Baxter (1997) Glass Earth, Inc. In: (eds.) Stephen McClelland. Future Histories. Horizon House and Nokia Corp. 0-9530648-0-8

This new anonymity for users proved to something of a new freedom for users but left much of the commercial side of the internet stagnating. Out of this emerged a compromise: a centralised advertising proxy run as a government-independent private enterprise who would (sic.) *guarantee anonymity* from the producers and advertisers at the expense of each individual being required *not just to view* a certain amount of advertisements but to *interact* with them to ensure that the advertisement had actually been read. A person's quota would become the new currency of the internet and be based upon your social network, your willingness to promote products and ultimately to purchases.

Because of the necessity for personal anonymity, the specific details of the mechanisms of how this worked were somewhat confidential. That didn't seem to overly concern users, nor the privacy advocates, nor the advertisers - everyone got their share - *for privacy's sake...*

* * *

John sat down at his office computer that morning, coffee in hand. London was never easy in the mornings, but a 45 minute trip on the tube gained him 45 advertisement credits - a bargain despite the veritable bombardment of special offers, new products, old products, brands and names on the senses.

With 200 advertisement credits left for the day, including the deduction for the London-McDonalds Bonus. The arch across the Thames was hideous, but 100 credits deducted from the normal daily tally was worth it. Some even said that next year's proposed Coca-Cola's branding of Tower Bridge might even bring another 100 credits deduction - who knows how many credits they would get for Pepsi's proposed sponsorship of Big Ben?!

Two hundred credits usually meant an hour or so of viewing and interacting with advertisements. Hell, he might just have to make that purchase of a computing device: a great offer this month tempted him with its bundled 2000 advertisement credits. Not that John needed another device, but almost two weeks without having to go through this daily routine...

Funny how one remembers the days when people complained about targeting advertising on the screen saver or when they were playing a game...oh those halcyon days of 2012...

It was an inevitable part of the deal..an anonymous internet for forced advertisement consumption via some centralised proxy, or whatever they were - part search engine, part advertiser, part social network.

That always bothered him to a point - most didn't really care - but they always seemed to know what advertisements to show him...a little too good given that the rest of the internet was anonymous; then again they were the only provider of advertisements now. Maybe that's why here was here, despite his job seemingly

almost futile now: privacy has been solved hadn't it?

The computer played the advertisements and almost subconsciously he clicked each strategically to demonstrate that he had sufficiently read the contents - it took him a while to acquire the skill to do that well enough to fool the system but once gained it freed him to perform some degree of multitasking. People were offering courses in 'strategic advertisement interaction' now; illegal in some US states apparently - a new black market forming?

A new breed of advertisements were coming too - multiple, cross-referenced adverts that demanded your understanding too.

He used a pen and paper, his little eccentricity...he toyed with various ideas, maybe even a novel of some kind, a great adventure possibly to relieve his frustration and boredom? He scribbled the line: *"In a hole in the ground there lived..."*. John smiled and continued writing, penning the title:

Globally Targeted Advertisement Tracking Preference Expression (DNT)
W3C Working Draft - October 2018

3 Discussion

Despite the whimsical nature of the story and the very real lack of technical explanation, the point here is to focus discussion on what we want to achieve through prevention-of-tracking technologies and what the unexpected consequences of those decisions might be based upon the very real needs to business, infrastructure and the end users: What would an internet without tracking look like?

If we succeed in limiting tracking do we destroy an industry(ies) or do we open ourselves for new business opportunities and in what form will these opportunities take? Would advertisers and anti-tracking product makers team up and provide for a self-fulfilling prophecy? Do anti-virus software vendors write viruses?

There is some current criticism on the speed of technical development of DNT - somewhat unfairly due to the scope of the problem and only now the technical aspects of privacy are being addressed. We still do not have a good theory of privacy or even common terminological framework that unifies the engineers, scientists, mathematicians, lawyers and consumer advocates - let alone the end-user - yet. Even when technical solutions appear are the three major factions[4] in this area even agreeing on the speed of development? What might be slow, painful work to one group is too fast for another.

[4] I split this area into engineers/scientists (technical), legal and consumer advocacy

If we lose tracking with its identification issues, what does identity, linking and anonymity really mean in this environment? Would the pseudo-anonymity be acceptable? Would individual people be subject to a single identity or single pseudo-identity, or even, would multiple persona become the norm for everyone.

Compliance and enforcement would potentially lead to some kind of software certification for privacy. What effect might this have? Obviously all software would have to be certified before deployment in much the same way as medical, avionics and other safety-critical system are. Do we even have the necessary software engineering skills to fulfill this?

To summarise, while there are criticisms of DNT and related technologies in this area from various viewpoints, DNT **has** succeeded in opening up a more general discussion on the nature of privacy. Whatever DNT and related, forthcoming and future technologies lead to (as well as the development of legal and consumer advocacy), we must not be complacent regarding their effects and thus exceptionally careful not to create a dystopian disaster by being careless or too eager.

TRUSTe Position Paper October 21, 2012

Taking a Balanced Approach to Privacy

Microsoft's recent steps around setting the DNT default to "on" [DNT:1], and Apache's response to ignore the signal, has sparked discussions around what it means to provide choice, especially since online tracking and the use of technology to collect data play a critical role in the online eco-system enabling a rich online experience providing personalization, improvements to data security and usability, and efficient advertising and monetization models.

By virtue of the definition of the word - choice is the act of demonstrating a preference that a particular activity be engaged or avoided. In privacy parlance choice is about an individual indicating a preference about when a particular data collection or use practice is exercised against information about an individual. Choice can either be inferred or expressed. In its program requirements, TRUSTe defines these terms in the context of whether a company can engage in a practice based upon actions taken (or not taken) by an individual. It is up to the individual to exercise a preference over whether a company can or cannot engage in a particular data collection activity.

TRUSTe has taken a similar approach when developing its behavioral advertising solutions [Trusted Ads and Trusted Mobile Ads] that enable users to exercise their preference around having data collected and used for ad targeting purposes. Fundamental to these solutions is providing users clear and conspicuous notice (transparency) regarding why data is being collected thus allowing users to exercise their preference (choice) in an effective and meaningful way so the preference being communicated is what the user intended. The underlying technology for Trusted Ads utilizes a cookie-based approach.

DNT extends beyond the cookie-based approach utilizing the browser (user agent) headers to communicate a preference. Here a set preference is global and applies to all websites and parties operating on those sites. Some view this as an all or nothing proposition thus creating complexities around standardizing behaviors around when a DNT:1 signal is received. As noted above, the user needs to exercise their preference meaning the default is unset. There are instances where a user agent may set the default on the user's behalf. Even though setting the default does not reflect TRUSTe's viewpoint that the preference must be set by the user - ignoring a DNT:1 signal from a user agent with a preset default does not solve the problem either because in this instance users who indicated a preference not to be tracked will not be honored.

Resolving this requires a balanced approach that provides transparency and the user an opportunity to indicate their preference. One approach is to examine how DNT and cookie-based approaches can work in tandem. This can be done through utilizing out-of-band exceptions requests. A user visits a site, site or a third party integrated into the site recognizes the DNT:1 signal, provides the user notice explaining the DNT setting is 1, and provide the user the option to grant a site-wide or web-wide exception. The solution also gives the user the option to exercise preference at a more granular level using the cookie-based approach – opting-out of those parties they do not wish to be tracked. This can be done based upon the party's function (i.e. analytics or social plugin) or by domain (i.e. example.com) allowing the user to allow for functionality they find most beneficial or trustworthy

companies. One of the challenges with this solution, and one area the W3C can explore in developing standardization, is how to address choice collision and develop standards around the logic needed to ascertain what the user intended.

Highlighting trustworthy companies enables users to make an informed choice about whether they are tracked and by whom. A key component of this is Accountability. Companies agreeing to some level of oversight by an independent third party, such as TRUSTe's Trusted Data Collection certification program, demonstrates those companies are accountable not only to themselves but to users. The ability to demonstrate Accountability needs be built into any solution that enables users to exercise choice over the collection and use of their data. The W3C, as it has done in the DNT Preference Expression Specification, should ensure privacy considerations outlined in any of its standards support a mechanism for trustworthy companies to demonstrate accountability.

Many questions remain around how DNT will be implemented; including how this will work in the mobile space. The W3C Tracking Protection Group is working through these very complex questions, which at some point need to extend to mobile. TRUSTe believes a balanced solutions-based approach that includes an accountability component is needed to address these and other privacy related questions.

TRUSTe is actively participating in this discussion and others being had by regulators and various industry groups. TRUSTe looks to continue to play an active role in these discussions promoting solutions that support a balanced framework, and enables businesses to innovate and users to express preferences based on a privacy framework built upon transparency, choice, and accountability.

Future of the Cookie

An IAB Initiative and Working Group

As someone who's been involved with web-based software development since before Netscape went public, I can confidently say that the use of cookies has gotten out of hand.

Originally designed for simple temporary data storage, the cookie now forms a fundamental infrastructure of the Web, being used for user profiling, segmentation and optimization, targeting and retargeting, mapping user IDs between platforms, buying and selling of data, end-user privacy controls, frequency capping of ads, Web analytics, online advertising attribution and verification, session management, shopping cart management …. the list goes on.

The problem, speaking from a Product Manager perspective, is that the use cases and requirements for a persistent and anonymous online tracking mechanism have long surpassed the capabilities of the cookie. The square peg has been hammered into the round hole for too many years, evidenced by numerous issues experienced by online publishers, consumers and the online ad industry as a whole.

For online publishers, the proliferation of 3rd party pixels has increasingly slowed page loads, increased discrepancy counts, degraded their user experience, and justified concerns of data leakage. It's also led to a broken compensation model – publishers risk revenue loss if they don't support 3rd party pixels, revenue loss for the users that block or delete cookies, and a tilted playing field favoring large consumer Web site brands who can track users for longer. And publishers are certain to have additional burden as various initiatives such as Do Not Track, browser opt-in defaults, and regulatory measures gain traction.

For online consumers, the proliferation of 3rd party pixels has degraded their online experience and increased anxiety over online privacy. With data collection so widely fragmented over countless domains, devices, browsers, apps, etc. it's impossible for them to apply privacy controls consistently, let alone make their online choices persist over time.

For the ad industry as a whole, the reliance on cookies (and 3rd party pixels) combined with the magnitude of cookie deletion (churn) has resulted in a battle between a rapidly degrading economic model, and the costly, continuous and high-volume deployment of cookies. Even though cookies are unreliable as a user tracking mechanism, especially across devices, the industry continues to deploy them at an escalating pace, causing excessive network traffic and related costs, "internet bloat", regulatory threats, and anxiety among consumers and publishers.

But there is a future for the cookie and an opportunity to turn the industry's most negative issue into a positive (or at least a neutral) one that will result in a win/win/win for publishers, consumers, and advertisers. Along with Philip Smolin (Turn), Susan Pierce (Google), Amy Kuznicki (Verizon Wireless) and Brian Murphy (AOL), I'm proud to be co-chairing an ambitious IAB Advertising Technology Council initiative called The Future of the Cookie along with the IAB's Mobile Marketing Center of Excellence.

We've recruiting leaders from the top companies in the digital advertising industry to join us in a mission to discuss and propose responsible solutions to the problems that exist today—and then execute a plan that leads us into the future of online user tracking, transparency, and control.

The working group is currently considering 6 distinct solutions, and would like to share our work with the W3C, as well as learn more about your efforts. Current working group members include representatives from the following companies:

- Rubicon Project
- Google

- Turn
- Rocket Fuel
- nPario
- Electronic Arts
- Adobe
- Media Mind
- Digital First Media
- Operative
- NBC
- Facebook
- DataXu
- eXelate
- Blue Kai
- Microsoft
- Yieldex
- Mocean Mobile
- Verizon Wireless
- Greystripe
- AOL
- AT&T AdWorks
- Pandora
- InMobi
- NY Times
- 24/7 Media

JORDAN MITCHELL | Vice President, Product

Rubicon Project

••• M +1 425 443 1819

••• Skype: jordan.mitchell.rubicon

••• Twitter: @Kickstand

KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

P +1-202-637-9800
F +1-202-637-0968
E info@cdt.org

POSITION PAPER: TRACKING PREFERENCE BEYOND THE UA

22 October 2012

W3C Workshop on "Do Not Track and Beyond", 26-27 November 2012
Joseph Lorenzo Hall (joe@cdt.org)

While aimed squarely at defining the technical and compliance requirements around a protocol for user agents (UA) that speak HTTP, the Do Not Track (DNT) header from w3c's Tracking Protection Working Group (TPWG)[1] will need to be set by users interacting with a user interface (UI) and it will need to work in the complex environment of applications built using HTML5. This raises a number of privacy-relevant questions across w3c Working Groups that the TPWG — or the new Privacy Interest Group (PING)[2] — may be able to advance, while not straying from the scope of charters that specify flexible UI design.

I. One DNT Preference, or One for Every UA?

Platform-level implementations for DNT-like tracking preference expression (TPE) settings are already in production. Apple's iOS 6 includes a "Limit Ad Tracking" (LAT) setting that, when set, signals to applications that only a limited set of restricted uses are permitted for a unique ubiquitous identifier on that platform, the "Identifier for Advertisers" (IFA).[3] As LAT is a more general type of tracking preference expression, it is not set at the level of individual user agents (browsers, apps, etc.) but at the OS level. It is not hard to imagine a more granular analog where every application on the iOS platform has a tracking preference setting. This is similar to what iOS does for location sharing preferences (*See:* Fig. 1), and iOS has extended this to other types of data sharing preferences including Contacts, Calendars, Reminders, Photos, etc.

Are there considerations at the platform-level within the purview of TPWG when TPE becomes pervasive? I have two broad classes of concerns with intra-platform TPE. First, they may work, like IFA, as a cross-party "super-cookie"; e.g., IFA is the same value for every application on a device until the device settings are reset. This would be equivalent to, for example, desktop browsers returning a unique identifier to each site visited upon request (until browser uninstall). This drastically increases the potential for server-side sharing of browsing-related activities per device without any notice to the user.

---

[1] *See:* http://www.w3.org/2011/tracking-protection/

[2] *See:* http://www.w3.org/Privacy/

[3] G.S. Hans and Joseph Lorenzo Hall, "Apple iOS 6 and Privacy", Center for Democracy & Technology Blog, 1 October 2012, available at: https://www.cdt.org/blogs/0110apple-ios-6-and-privacy-0

Second, with intra-platform TPE, an IFA-like value is allowed for a narrow range of permitted uses but it is unclear how those restrictions can be ensured or, equivalently, how users can convince themselves parties only engage in permitted uses. Should compliance be mechanistic? That is, the OS might refuse to provide the IFA value unless the requesting party attests that it is engaged in a specific permitted use? Or should compliance be controlled by policy? That is, the platform controller (Apple, in the case of iOS) might identify calls to the IFA query function during app certification and require developers to attest that the uses they engage in are from the set of permitted uses.

II. Tracking Preference Expression in HTML5

In a similar vein, we at CDT have been thinking increasingly about privacy issues in HTML5 and to what extent tracking preference expression and other privacy concerns may suffer if HTML5 WGs operate without a common privacy model or some level of coordination. HTML5 has a dizzying array of APIs,[4] many of which implicate core privacy interests of users. For example, clearly HTML5 elements such as the Calendar API and Contacts API will mediate potentially privacy-sensitive user interactions as they deal with sensitive data (time-location data and social network data/personal contact details). However, more generic and more powerful HTML5 elements exist, such as WebRTC — for peer-to-peer in-browser audiovisual interaction — and Web Intents — which allows seamless remote computation on local resources. These can create what appear to be local user interactions but that are fulfilled by remote or "cloud" infrastructure. How does a user specify that they do not want "local" (or user-contributed) resources to be used or tracked for certain activities? The familiar debate of "collect vs. use" raised in TPWG discussions becomes somewhat more serious when very rich user content and activities applied to that content is at issue (e.g., tracking the exact steps a user employs to edit a photograph using Web Intents functionality that provides a remote image editing application).

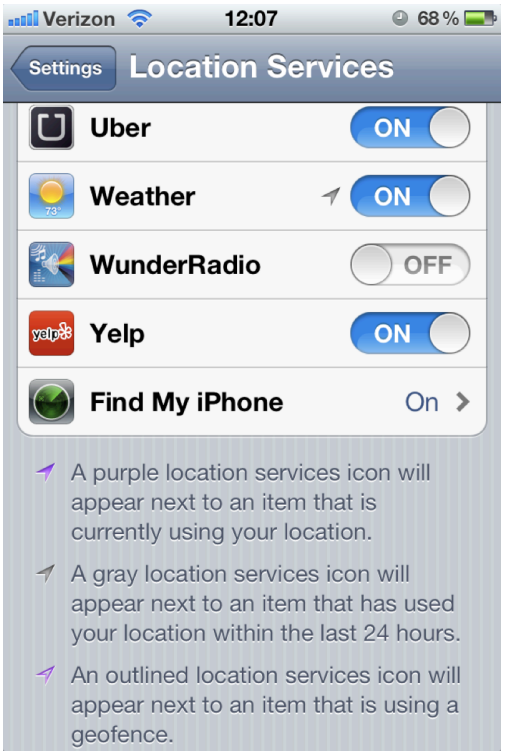


Fig. 1: iOS' "Location Services" preference pane showing granular control and notice iconography.

The need here is for a coherent and sensible model for how privacy protection might work in web applications built with HTML5 APIs. Unfortunately, the independent development of HTML5 APIs means that understanding the privacy implications across the suite of APIs is exceedingly difficult. We hope to discuss these more compounded privacy tensions in rich web applications at November's workshop.

[4] Erik Wilde, "HTML5 Landscape Overview", dretblog, 18 October 2012, available at: http://dret.typepad.com/dretblog/html5-api-overview.html

EFF POSITION PAPER: Unlinkability/auditability
W3C Workshop: Do Not Track and Beyond
26-27 November 2012

Defining "unlinkability" is a major issue for the Tracking Protection Working Group in crafting the TPE standard. On the W3C mailing list, Shane Wiley suggested a practical approach to handle some of the problems that have arisen in the TPWG discussions: roughly, that we develop a high-level definition as normative text, supplemented by non-normative text examples of aggressive approaches that clearly satisfy "unlinkability" (k-anonymity, URL filtering, super campaign structures, client-side storage, etc.).

EFF agrees that this may be a fruitful approach, but we believe it is critical to include non-normative text of approaches that would fail "unlinkability." The remainder of this position paper sets forth some of our current thinking.

Discussion

The current editor's draft defines "unlinkability" in two different ways. (The text set forth below corrects errors in that draft and thus is not quite verbatim.)

3.6.1 Option 1: Unlinkable Data

A party renders a dataset **unlinkable** when it
1. takes commercially reasonable steps to de-identify data such that there is confidence that it contains information which could not be linked to a specific user, user agent, or device in a production environment
2. publicly commits to retain and use the data in unlinkable fashion, and not to attempt to re-identify the data
3. contractually prohibits any third party that it transmits the unlinkable data to from attempting to re-identify the data. Parties should provide transparency to their delinking process (to the extent that it will not provide confidential details into security practices) so external experts and auditors can assess if the steps are reasonable given the particular data set.

3.6.2 Option 2: Unlinkable Data

A dataset is **unlinkable** when there is a high probability that it contains only information that could not be linked to a particular user, user agents, or device by a skilled analyst. A party renders a dataset unlinkable when either:
1. it publicly publishes information that is sufficiently detailed for a skilled analyst to evaluate the implementation, or
2. ensures that the dataset is at least 1024-unlinkable.

Discussion

We focus on the initial definition of "unlinkable," not the additional safeguards aimed

against re-identification attacks.

Option 1 is a variant of the Federal Trade Commission's (FTC) definition of "not reasonably linkable":

Data is not "reasonably linkable" to the extent that a company:
(1) takes reasonable measures to ensure that the data is de-identified;
(2) publicly commits not to try to re-identify the data; and
(3) contractually prohibits downstream recipients from trying to re-identify the data.

This definition turns on the meaning of "de-identified," which the FTC defined as: "the company must achieve a reasonable level of justified confidence that the data cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer, computer, or other device."

The FTC noted that this principle applies even when the data has not yet been linked to a particular consumer, computer or device, so long as it may reasonably become so linked. It also noted that if a company maintains and uses both identifiable data and data that has been de-identified, it should silo the two datasets separately. See generally Federal Trade Commission, Protecting Consumer Privacy in an Era of Rapid Change 20-22 (2012).

Option 1 is also similar to language used by the Digital Advertising Alliance.

A. De-Identification Process

Data has been De-Identified when an entity has taken reasonable steps to ensure that the data cannot reasonably be re-associated or connected to an individual or connected to or be associated with a particular computer or device.

An entity should take reasonable steps to protect the non-identifiable nature of data if it is distributed to non-Affiliates and obtain satisfactory written assurance that such entities will not attempt to reconstruct the data in a way such that an individual may be re-identified and will use or disclose the de-identified data only for uses as specified by the entity.

An entity should also take reasonable steps to ensure that any non-Affiliate that receives de-identified data will itself ensure that any further non-Affiliate entities to which such data is disclosed agree to restrictions and conditions set forth in this subsection V.A.

As with the FTC approach, the key language for present purposes is the actual definition of de-identification.

Significant questions have been raised about Option 1 on the mailing list. For instance, Shane Wiley suggested that "performing a one-way secret hash (salted hash) on identifiers (Cookie IDs, IP Addresses) and storing the resulting dataset in a logically/physically separate location from production data with strict access controls, policies, and employee education" would meet the definition set forth in Option 1. His

stated goal was to "find the middle-ground between complete destruction of data and an unlinkable state that still allows for longitudinal consistency for analytical purposes BUT CANNOT be linked back to a production system such that the data could be used to modify a single user's experience."

Prof. Felten, however, argued: "hashing IP addresses (with or without salting) does not render them unlinkable. After hashing, it's easy to recovery the original IP address. The story is similar for other types of unique identifiers--there are ways to get to unlinkability, but hashing by itself won't be enough."

Option 2, which derives from the EFF/Stanford/Mozilla proposal, differs mainly from Option 1 by requiring a "high probability" rather than "reasonableness," defined either objectively (1024-unlinkable) or by expert analysis.

This approach is similar to the treatment of health data under the HIPAA de-identification rule. HIPAA establishes a high standard that patient information is "de-identified" only when "there is no reasonable basis to believe that the information can be used to identify an individual." 45 C.F.R. § 164.514 (a). The current HIPAA rule then provides two "safe harbors": data is de-identified if one follows a prescribed approach of removing identifiers, or if certified as de-identified by a statistician.

The first safe harbor involves the removal of eighteen specified patient identifiers, including but not limited to, patient name, location (other than state or 3-digit ZIP codes with populations greater than 20,000), email address, telephone number, Social Security Number, and the like. 45 C.F.R. § 164.514(b)(2)(i). Significantly, the eighteenth identifier that must be removed is "any other unique identifying number, characteristic, or code." Prof. Sweeney's research suggests that under this safe harbor, about 0.04% of the population can be re-identified.
http://dataprivacylab.org/projects/identifiability/pharma1.pdf

The second safe harbor requires a formal determination by a qualified statistician who, applying statistical and scientific principles and methods for rendering information not individually identifiable, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information. The statistician must reach a conclusion that the risks of re-identification are "very small" in order for the patient information to be properly "de-identified." Unfortunately, "very small" is not defined. An obvious suggestion here is that "very small" be taken to mean identifiability of 0.04%, analogous to Option 2's 1024-unlinkable rubric.

Auditability

A further weakness of the "statistician" safe harbor under HIPAA is that there is no oversight of the statistician's "very small" risk determination. Even for good actors, opacity of the de-identification techniques and assumptions about background data can make it hard to know whether a risk that was "very small" will become higher as data and

re-identification techniques improve.

Option 1 and Option 2 each attempt to address this problem. Under Option 1, parties should provide transparency to their delinking process (to the extent that it will not provide confidential details into security practices) so external experts and auditors can assess if the steps are reasonable given the particular data set.

Under Option 2, the alternative to 1024-unlinkability is that the entity claiming de-identification publicly publish information that is sufficiently detailed for a skilled analyst to evaluate the implementation.

We believe that some form of these requirements is critical to assuring that data remains unlinkable under the standard.

Conclusion

It may be possible to define "unlinkability" along the lines suggested by Shane Wiley: a high-level definition in normative text, with non-normative text setting forth examples of acceptable or even aggressive implementations. In order to set a floor, however, we believe that the non-normative text must also include examples of technical methods that are not acceptable. In addition, the standard must contain a meaningful process for auditing/verification.

CHAPELL & Associates

October 21, 20012

Nick Doty, W3C
Jan Schallaböck, ICPP
Event Co-chairs
32 Vassar Street, 32-G519
Cambridge, Massachusetts 02139

Re: November 26 & 27 Workshop

Dear Nick & Jan –

Thank you for the opportunity to share my position on the W3C's proposed expansion further into areas of privacy, public policy and marketing standards.

I believe that continued innovation across the Internet is best served by having the W3C continue to do what it does best: facilitate the creation of technical standards. Moving into areas of policy and digital marketing are bad for privacy, bad for innovation and bad for the Consortium.

Expansion of scope moves the W3C well outside of its core competency
Prior to the creation of the Tracking Protection WG, the W3C's most noteworthy entrée into public policy was the P3P standards - the limitations of which are widely documented.

The W3C currently lacks the expertise to address issues of policy or marketing best practice in a meaningful and productive way. The core participants W3C working groups tend to be those with technical and/or operational acumen. Moving into areas of digital marketing would require a significant expansion of both membership and W3C staff expertise. I strongly encourage the Consortium to consider whether such a shift would take too much focus from its important work on technical standards.

An expansion weighs against smaller marketplace participants

The natural answer to the issue of core competency raised above might be to simply bring in additional participants into the W3C process. And perhaps, that's a key driver here – as I would imagine that the work on a Do Not Track (DNT) standard has increased the visibility as well as the membership rolls of the W3C. Moving into new disciplines almost certainly would bring additional revenue opportunities to the Consortium.

However, a nearly $8,000 per year membership fee is simply out of the question for most smaller companies with limited budgets. Concerns over the W3C process being dominated by large organizations are well documented. Moving into areas of policy and digital marketing are only going to exacerbate such concerns.

Similarly, while the Tracking Protection WG has participation across the United States and European Union, its worth noting that a large swath of the rest of the world has had very little input into the DNT standards.

Policy issues are extremely difficult to harmonize on a pan-world basis

The various cultural mores, legal jurisdictions and industry practices vary widely across the globe. While this may also be true when it comes to technical standards, it is more so in policy and other areas.

Email marketing provides an interesting illustration of some of these challenges. Generally speaking, the EU is opt-in; the U.S. opt-out and a good portion of the rest of the world has not taken a position. Harmonizing the opt-in, opt-out tension that exists between the U.S. and EU is as challenging as determining whether American Football is better than fútbol played in Europe – and about as productive. Each region is going to have its own distinct point of view, and the W3C should think long and hard before imposing the views of a limited set of participants in a handful of jurisdictions upon the rest of the world.

CHAPELL & Associates

While ill advised, if the W3C decides to head down this path, I would strongly encourage the following PRIOR to the creation of any additional working group:

- First, the W3C should have a very clear list of the specific issues that the group is trying to address and the goals to be met.
- Second, the W3C should wait until it is in position to characterize the work of the Tracking Protection WG as an unmitigated success before attempting to extend its bailiwick into other areas. Moving into areas where the body of work is limited to P3P (unclear privacy protections, not widely implemented) and DNT (an incomplete grade at this point) seems premature if not entirely inappropriate.

Thanks for the opportunity to share my thoughts. I look forward to the opportunity to speak in November and share additional color.

Sincerely,

Alan Chapell
Chapell & Associates

customer commons

October 22, 2012 By Customer Commons, customercommons.org

Abstract: Proposal User Agent Button Development

The R-Button User-Agent is an indispensable tool -- your personal agent -- representing you as you assert your preferences and negotiate terms of service for privacy, tracking, and data sharing on the web, phone, tablet and other devices.

Customer Commons proposes to build a tool to create Do Not Track Plus, or DNT+, that will allow the user to assert more than just a preference not to be tracked. DNT+ will act as a user agent to express preferences for how long data is retained, for what purposes and under what terms or policies. DNT+ will allow users to see how their preferences match up, at first and over time, as entities change their data policies.

The Project:

The R-button, or "relationship button" is meant to allow a person to set their preferences for privacy, tracking, and personal data sharing. Among these are choices for "Do Not Track" (DNT), which several entities including Internet Explorer, Firefox and Chrome are all working toward, as well as the W3C specification. But choices could encompass much more.

Each time a person interacts with a phone app or web site, or shares data, they face the tracking of their activities and often a desire to set some level of privacy. But, doing this with every entity we work with on a case-by-case basis would be tedious and time consuming.

The R-button proposes to give a person a User-Interface method to see when their pre-chosen preferences are being respected, and make changes or find other ways to do what they want without losing control of their data, privacy or being tracked. The R-Button would connect what a person has chosen in the past at a personal cloud or other preference tool, or let a person set simple preferences for privacy, tracking and preferences at the R-button UI if they didn't have anything else. The R-button would then share those preferences to those entities at the other end of wherever a user goes on the web, in their phones or anywhere a person can be tracked or share data.

Interface:

We propose to start with a hypothesis for design and meaning here:

⊂ ⊃ - shown in the state of "no relationship"

The symbol on the left, ⊂, represents the individual user, the first party to the relationship. The symbol on the right, ⊃, represents the second party to the relationship, typically an organization or service. The two symbols resemble magnets, suggesting attraction to each other. Their parallel lines and equal size also suggest equality, correspondence and interoperability.

The left and right pieces could close when a relationship has been made between the person's preferences and the entity the person is interacting with online, on their phone or elsewhere.

Clicking on the symbols would allow a person to set preferences for that moment, see what is set already by default, and see what the entities terms are to the user.

We expect to test and build the UI and functionality that works best for a broad range of users. We expect that what user's need may be different that what we propose. We think building for user's is the best method for getting this right.

No Company Will Own This

The R-button specifications, UI and code, and usability research will be put out under an open source license and as an open standard for any entity to use. The R-button will live at Customer Commons (customercommons.org), a non-profit, which will shepherd the project as well as maintain it and future releases as part of its non-profit mission to support individuals and their ability to be treated fairly wherever they interact.

The R-button was originally conceived at Harvard's ProjectVRM by Doc Searls, author of the Intention Economy.

Development Scope

The **R-button Research and Development Project** will follow an agile usability and development model to professionally build and test the concept of the R-*button* (including the name "*R-button*") as a standardized user interface affordance independent of any vendor, device, or application.

The R-button will depend and inter-operate with many other efforts including policy tools like Open Notice, Privacy Icons and the Information Sharing Group's Nutrition labels, the W3C user agent architecture under discussion now, trust frameworks in use and under development, and identity agents such as Mozilla's Persona agent. We will work with all entities to develop a symbol for the user, and interactions that will translate and tie together at the UI level all the efforts that are needed for user's to assert preferences and see in simple fashion what their data and identities are doing.

Developed and tested affordances we want to include are *choice*, *privacy*, *security*, *resilience*, *interoperability*, *auditability* and *reduction or elimination of cost*. Among a person's *choices* will be:

- *Specifying identity providers* (including the individual himself or herself)
- *Establishing and maintaining relationships with relying parties (those guys on the other side of your surfing, apps or purchases)*
- E*stablishing and maintaining connections with personal data stores and lockers*, personal clouds and
- E*stablishing and maintaining personal terms, preferences and policies toward relationships with other parties* (including "do not track" me and "here are the permitted uses of my data")
- *Accessing audit trails of agreements and actions that take place within relationships*

between the individual (⊂) every other party (⊃).

The Emergence of a New Paradigm

Fortunately a new paradigm is emerging, one based on the concept of individual independence and agency and peer-to-peer relationships among equals. One example is the *Information Sharing Agreement*, which requires parties of equal agency, power, and ability to form non-coercive agreements. Another is the *personal cloud*, an autonomous collection of data and capacities for individuals that corresponds — and can relate directly to — the clouds maintained by organizations (or other individuals). Personal clouds can interact with servers, but outside the subservient role required by the cookie model. (In other words, it will work with, rather than against, the client-server paradigm.) Another is the *trust framework*, which is comprised of autonomous individuals making voluntary assertions about each other, and does not require a single dominating party.

The need for the *R-button* come from this paradigm shift. It will become not only the literal symbol of this new paradigm — the goals for user-driven control of identity and personal data — but the actionable trigger point for simple, safe, secure identity and relationship management across all types of digital devices. Furthermore, adoption will be driven both by users and by vendors, because both want the convenience, trust, and relationship value it represents.

We've spent the last four years at Harvard's ProjectVRM studying and defining the problem. Now, with Customer Commons, we want to create a tool that will give people controls and a way to share their preferences everywhere they go. And because it's open source, it will constantly evolve with the help of a global community of developers.

The W3C Event

Currently Customer Commons is in the research phase of this project and we would like to share user research conducted to date at the W3C event. We are in the process of obtaining data and writing it up for public release. This research could be shared at the W3C DNT event, as well as the User Agent / R-Button plans for developing a user interface to knit together the various architecture and policy initiatives.

Who We Are

The mission of Customer Commons, a California-based non-profit, is to restore the balance of power, respect and trust between individuals and organizations that serve them. We stand with the individual and therefore do not take contributions from commercial entities. We are run and funded by individuals.

Building on the work of ProjectVRM at Harvard's Berkman Center, Customer Commons provides individuals with the knowledge and tools they need to shape their online relationships. We believe this is increasingly important in a world where online and offline activities are rapidly merging. Big data, behavioral tracking, mobile geo location and other emerging data collection practices are altering the fabric of our shared social and economic relationships.

Customer Commons holds a vision of the individual as an independent actor who retains autonomous control over his or her personal data, desires and intentions. In this vision, each of us will act as the optimal point of integration and origination for data about us. Individuals must be able to share their data and intentions selectively and voluntarily. Individuals must also be able to know exactly what information is being held about them by those who gather it, by whatever means. To achieve this, people must be able to assert their own terms of engagement, in ways that are both practical and easy to understand for all sides.

Customer Commons believes that informed and empowered individuals, free to make unbiased choices based on their own intentions, are more powerful and engaged participants in the marketplace. Free customers are more valuable to themselves, to vendors and to the market economy, than captive ones. Data that is volunteered intentionally is a far more useful form of economic signaling than either coerced data from captive customers, or from invisible surveillance. As the Internet increasingly comes to dominate global communications and commerce, informed and autonomous individuals will become critical to a healthy, civil, democratic society.

What We Do

As a public-facing organization focused on the emerging issues at the intersection of the empowered customer and the public good, Customer Commons seeks to both educate and inspire change for customers worldwide. We do this through research, educational initiatives and promotion of customer side tools.

Customer Commons seeks to meet individuals where they naturally are, while offering a way forward that demonstrates practical alternatives. Our work is grounded in research exploring public attitudes toward customer relationships with vendors and organizations, online and offline. We also research solutions under development, seeking to understand, present and advocate viable approaches. This information supports the development of new tools and methods while also informing the larger Internet policy conversation and our vision of an empowered individual in the context of a healthy and dynamic market-based system.

All of Customer Commons' programs reflect the values of truly transparent and flexible methods for individuals to maintain their unique voices in relationship to other entities in the marketplace. Therefore, we support technological approaches that are primarily personal, not social, designed from the outset to help people clearly express their own intent. Our work will help individuals engage in equal and open relationships that do not lock them in to a single set of solutions. With truly engaged customers, businesses, governments and institutions of all kinds will have countless more willing hands, heads and hearts to solve the problems of our world, while creating prosperity for all.

Nitin Badjatia serves as Director, Market Strategy at Oracle Corporation. Prior to his current role, Nitin was Director, CX Strategy – Knowledge Solutions at RightNow Technologies (acquired by Oracle in 2011). Prior to RightNow, Nitin was Director, Business Strategy at Knova Software, and was a part of the financial services practices at Siebel Systems and Oracle (his first stint at the software giant).

Jennifer Cobb is Principal at Spruce Advisers, a strategy and communications consultancy dedicated to the intersection of technology and social good. Jennifer has served as a senior staff member at both venture-backed start-ups and in the non-profit sphere. She currently works with organizations working to further the positive impact of technology.

Iain Henderson is Product Director at Allfiled a builder of personal data services for individuals. Iain is a

long term CRM practitioner and customer data specialist who has long advocated the need to build personal data services that work for individuals in order to complement the tools available to organizations. You can find Iain on Twitter at @iainh1.

Leyla Hill is VP Business Affairs and General Manager at Hearts of Space, the long-running public radio music program and online music service. Her diverse background includes data processing administration (in the pre-PC heyday of mainframes) and business management, editing, and publishing administration. She currently oversees all business and legal affairs and customer service for Hearts of Space.

Mary Hodder is an entrepreneur, founder, user researcher, user advocate and early adopter. She founded Dabble.com in 2005 and she also founded a mobile app called "wellness mobile" to self track and share one's own wellness. Hodder has worked with large and small organizations as an information architect and interaction designer, creating algorithms, and conducting user research in the form of usability studies, needs assessments and heuristics.

Doc Searls is co-author of The Cluetrain Manifesto, a business classic, and The Intention Economy: When Customers Take Charge. He is also Senior Editor of Linux Journal, a fellow at the Center for Information Technology & Society at the University of California, Santa Barbara, and an alumnus fellow at Harvard's Berkman Center for Internet & Society, where he runs ProjectVRM, which coordinates development of tools for customer independence.

Steven Tulsky is Principal of The Benemetrics Consulting Group, advising nonprofit organizations and emerging businesses in the areas of financial strategy, planning, and management. He enjoys lending his quantitative expertise to organizations that understand their clients and their services better than they understand their numbers. Prior to developing his consultancy, Mr. Tulsky held roles in industry including Chief Financial Officer, Director of Finance, Treasurer, and Assistant Treasurer of large, medium, and small public and private firms.

Joyce Searls is a serial entrepreneur with a varied background in the fashion, restaurant and real estate development fields. In addition to Customer Commons, she has multiple management and board roles, including Linux Journal and Project VRM. She also collaborates with her husband, Doc in their consultancy, The Searls Group.

Do Not Beg:
Moving Beyond DNT through Privacy by Design

Mike Perry
The Tor Project, Inc
mikeperry@torproject.org

Abstract

The Do Not Track header (henceforth DNT:1) seeks to provide privacy protections against third party tracking through user request and regulation. It is our position that while DNT:1 is potentially useful as a purely informational tool for browser vendors and service providers, enforcement of the header suffers from a number of issues including covert circumvention, enforcement jurisdiction, manipulation, regulatory capture, and abuse. Moreover, every privacy property that DNT:1 aims to provide through regulatory enforcement can be better provided through technical changes to browser and network behavior during private browsing modes. We therefore suggest that the W3C standards body focus on standardizing these technical measures, rather than attempting to broker negotiations over regulatory policy and law.

1 Introduction

In this position paper, we describe the current and potential issues with DNT:1 and associated regulation, and also describe our prototype browser implementation[9] that aims to provide the same third party tracking resistance properties as DNT:1, but without relying on costly regulation and auditing.

We also believe that third party privacy can become a competitive feature for browser vendors, Internet service providers, and privacy preserving overlay networks.

2 Overview of DNT:1

The Do Not Track header seeks to provide users with a uniform mechanism to opt-out of third party tracking. Third party content elements are supposed to honor the header by declining to set cookies and record user activity on their servers. The draft standard[6] states that first party sites do not need to alter behavior with respect to the header. It also states a number of exceptions where third parties may still choose to retain and analyze data.

2.1 Benefits of DNT:1

The primary benefit of the Do Not Track header is that it provides a strong signal to browser vendors and websites with respect to their user's interest in privacy. Within a few months of the header's appearance, 7% of desktop and 18% of mobile Firefox users dug through the Firefox privacy settings to enable it.[4]

However, despite the value of sizing the market segment for frictionless privacy enhancing web technologies, it is very likely that DNT:1 will become a total disaster once the transition to regulatory enforcement draws near.

2.2 Shortcomings and Dangers of DNT:1

The primary shortcoming of DNT:1 is that it in no way alters the behavior of numerous browser technologies that enable and facilitate third party tracking, and instead relies entirely on ad-hoc auditing and potentially even regulatory enforcement.

Should strict auditing and direct regulatory requirements be enforced in some jurisdictions, it is very likely that at least some portion of the advertising industry would relocate to more favorable jurisdictions. In fact, they would be incentivized to do so, since it would allow them to offer advertising services at more favorable rates than their competitors who do not.

Similarly, it introduces serious risks of regulatory capture, especially in jurisdictions where advertising is able to wield considerable political influence over the selection of elected officials.

To answer these concerns, some DNT:1 advocates claim they favor "Carrot and Stick" incentive schemes that do not involve direct regulation, but instead will rely on web crawls to determine suspicious third party activity[7]. Their claim is that violators can be added to an always-on adblocker filter, and good actors could even be given immunity from data breach notification requirements and related privacy regulations.

However, without changes to the underlying browser technologies, there are simply too many ways for advertisers to covertly encode identifying information in third party elements. Even seemingly innocuous changes such as minor Javascript and CSS alterations across multiple elements can be used to encode covert identifiers that are stored in the browser cache for use as third party tracking cookies. This doesn't even begin to scratch the surface of covert third party identifier storage and supercookie vectors, let alone IP address and fingerprinting-based vectors, all of which we will discuss in more detail in later sections.

Further, behavioral targeting can be made very subtle, and difficult to distinguish from random chance. For example, Target has begun taking great pains to obscure behavioral targeting in its catalogs, to avoid alienating customers. Their targeted advertisements are still present, but they are merely blended with off-target messaging to provide a false sense of security and privacy[2]. It is extremely likely that such techniques will be employed by bad actors in the third party advertising world as well.

Further still, because of the various exemptions allowed in the DNT:1 standard, it is hard for users to know when the header is being honored, and if their activity is still being recorded, exchanged, and sold.

2.3 Hidden Costs of DNT:1

We believe that DNT:1 has seen such favorable adoption by browser vendors because of the ease of deployment for them. Adding a single HTTP header is substantially simpler than devoting research and development resources to addressing the network adversary in private browsing modes.

However, DNT:1 merely shifts the costs of privacy development and enforcement off of the browser vendors and onto every other party involved in the Internet economy, as well as onto new parties who were previously not involved (such as auditors, vigilantes, and governmental regulators).

Further, DNT:1 demands that standards organizations such as the W3C shift gears away from producing and reviewing technical standards to instead broker policy deals between regulators, legislators, and industry.

We believe that standards bodies and regulatory agencies shouldn't be wasting resources asking themselves how, when, and why advertisers don't obey DNT:1. Instead, they should be asking themselves why browser vendors whose revenue streams are often directly related to advertising markets continue to deploy technologies that facilitate and encourage covert third party tracking with no technical alternatives, even when their users enable their so-called "private browsing modes".

3 Do Not Track through Privacy By Design

Remarkably, the very same third party tracking resistance properties suggested by the DNT:1 draft standard are possible through a combination of browser and network behaviors.

All of these properties flow from a very simple core idea: two different first party domains should not be able to link or correlate activity by the same user, except with that user's explicit consent.

Initially, consent can be interpreted as link-click navigation. However, as federated login technologies such as web-send[1] and Persona[8] (formerly BrowserID) evolve, link-click based tracking vectors (such as the Referrer header) can be reduced or eliminated.

To understand the scope of the changes to the browser and network service providers to provide third party tracking resistance, we need to break down the problem into roughly four main areas of linkability and privacy: identifier sources, fingerprinting sources, disk activity, and IP address utilization.

3.1 Browser Behavior: Identifier Sources

Obviously, the primary vector through which third party tracking operates is the third party cookie.

Mozilla has a wonderful example of a first party isolation improvement written by Dan Witte and buried on their wiki[10]. It describes a new dual-keyed origin for cookies, so that cookies would only be transmitted if they matched both the top level origin and the third party origin involved in their creation. Thus, third party features could still function, if the user authenticated to that third party within the context of their first party url bar domain (perhaps using Mozilla's Persona, for example).

With respect to cache identifiers, the earliest relevant example of isolation work is SafeCache[5]. SafeCache eliminates the ability for 3rd party content elements to use the cache to store identifiers across first party domains. It does this by limiting the scope of the cache to the origin in the url bar origin. This has the effect that commonly sourced content elements are fetched and cached repeatedly, but this is the desired property. Each of these prevalent content elements can be crafted to include unique identifiers for each user, in order to track users who attempt to avoid tracking by clearing cookies.

Other identifier storage mechanisms that require such isolation include HTTP Auth, window.name, DOM Storage, IndexedDB, SPDY, HTTP-Keepalive, and cross-domain automated redirects. In Tor Browser[9], we either disable or isolate these technologies.

Properly isolating browser identifiers to the first party domain also has other advantages as well. With a clear distinction between 3rd party and first party cookies, the privacy settings window could have a user-intuitive way of representing the user's relationship with different origins, perhaps by using only the favicon of that top level origin to represent all of the browser state accumulated by that origin. The user could delete the entire set of browser state (cookies, cache, storage, cryptographic tokens, and even history) associated with a site simply by removing its favicon from their privacy info panel.

3.2 Browser Behavior: Fingerprinting Sources

After identifier isolation, the next source for covert tracking is through browser fingerprinting. Advertising networks can probe various browser properties known to differ widely in the userbase, thus constructing an identifier-free mechanism of tracking users.

Unfortunately, just about every browser property and functionality is a potential fingerprinting target. In order to properly address the network adversary on a technical level, we need a metric to measure linkability of the various browser properties that extend beyond any stored origin-related state.

The Panopticlick project by the EFF provides us with this metric[3]. The researchers conducted a survey of volunteers who were asked to visit an experiment page that harvested many of the above components. They then computed the Shannon Entropy of the resulting distribution of each of several key attributes to determine how many bits of identifying information each attribute provided.

While not perfect[1], this metric allows us to prioritize effort at components that have the most potential for linkability.

[1] In particular, we believe it is impossible to eliminate inter-browser fingerprinting vectors. Instead, fingerprinting metrics and defenses should focus on distinguishing features amongst a population with the same user agent. The Panopticlick test is not currently set up to do this.

This metric also indicates that it is beneficial for to standardize on implementations of fingerprinting resistance where possible. More implementations using the same defenses means more users with similar fingerprints, which means less entropy in the metric. It is for this reason (among others) that the Tor Project seeks to share its Firefox-based browser implementation[9] with any interested parties.

The fingerprinting defenses deployed by the Tor Browser include reporting the desktop resolution as the content window size, reporting a fixed set of of system colors, disabling plugins by default, limiting the number of fonts a document is allowed to load, and disallowing read access to the HTML5 canvas without permission.

The DNT:1 header itself is also fingerprinting vector for bad actors if we allow our users to set it, and the related scandal between Microsoft and Apache will likely cause us to entirely remove the DNT:1 option from Tor Browser's privacy preferences as a result.

3.3 Browser Behavior: Disk Activity

In addition to protecting against the network adversary, we believe that private browsing modes should not force the user to go without disk access. The two defenses are orthogonal, and private browsing mode users should still be allowed to store history, bookmarks, and even cookies and cache if they so choose.

Interestingly, a unified toplevel privacy UI could provide easy access to quickly clear all of these disk records on a per-site basis, using the same UI window for both tracking privacy and network storage.

3.4 Network Behavior: IP address utilization

Currently, there are many ways users can obtain a fresh IP address in an ad-hoc fashion. Users can use open wireless networks or tether to their phones. In fact, it is common practice for ISPs in many parts of the world to rotate user IP addresses daily, to discourage servers and to impede the spread of malware. This is especially true of cellular IP networks.

Obviously, only technically savvy users are likely to take full advantage of these properties correctly. However, there is no reason why an IP address allocation approach can't be generalized and standardized. One could imagine any privacy proxy (perhaps even one provided by your primary ISP) that intelligently isolates your first party page loads, along with all of their associated third party content, to a given IP address. By standardizing such a mechanism, privacy preserving networks can compete on network properties such as privacy or performance, rather than some combination of network and user agent.

The mechanism Tor has chosen to convey this information to the overlay network is the SOCKS username and password fields. Our plan is for the Tor Browser to inform the Tor client which network requests correspond to a given first party URL bar domain. The Tor client will then ensure that all first party loads use a different path through the Tor overlay network.

In fact, the Tor Project has concluded that it is in the best interests of the organization to share user agent development and standardization with other privacy preserving networks, both to reduce our development efforts, and to lead to a wider browser fingerprint population for our userbase. The German privacy company JonDos, GmbH has already joined this effort.

4 Conclusions

We discussed the Do Not Track header, the privacy properties it seeks to provide, and its shortcomings. We believe it is possible to provide these very same privacy properties through privacy by design.

While the DNT:1 header appears to be a simple change on the browser side, it has numerous hidden costs in terms of regulators, auditors, and server-side changes, in addition to serious regulatory challenges. We believe that it will actually be less costly in total to make the equivalent changes to the browser, and these changes will have the advantage of supporting markets for privacy proxies and related privacy enhancing technologies.

References

1. Tyler Close, Rajiv Makhijani, Mark Seaborn, Kenton Varda, Johan Apelqvist, Claes Nilsson, and Mike Hanson. Web Introducer. http://web-send.org/introducer/.
2. Charles Duhigg. How Companies Learn Your Secrets. https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=9.
3. Peter Eckersley. How unique is your web browser? In *Proceedings of the 10th international conference on Privacy enhancing technologies*, PETS'10, pages 1–18, Berlin, Heidelberg, 2010. Springer-Verlag.
4. Alex Fowler. Mozilla Led Effort for DNT Finds Broad Support. https://blog.mozilla.org/privacy/2012/02/23/mozilla-led-effort-for-dnt-finds-broad-support/.
5. Collin Jackson and Dan Boneh. Protecting browser state from web privacy attacks. In *In Proceedings of the International World Wide Web Conference*, pages 737–744, 2006.
6. J. Mayer, A. Narayanan, and S. Stamm. Do Not Track: A Universal Third-Party Web Tracking Opt Out. https://tools.ietf.org/html/draft-mayer-do-not-track-00.
7. Jonathan R. Mayer and John C. Mitchell. Third-Party Web Tracking: Policy and Technology. https://www.stanford.edu/~jmayer/papers/trackingsurvey12.pdf.
8. Mozilla Developer Network. Persona. https://developer.mozilla.org/en-US/docs/persona.
9. Mike Perry. The Design and Implementation of the Tor Browser. https://www.torproject.org/projects/torbrowser/design/.
10. Dan Witte. https://wiki.mozilla.org/Thirdparty.

Priv3: A Third Party Cookie Policy

Mohan Dhawan
Rutgers University
mdhawan@cs.rutgers.edu

Christian Kreibich
ICSI & UC San Diego
christian@icir.org

Nicholas Weaver
ICSI & UC San Diego
nweaver@icir.org

ABSTRACT

In today's World Wide Web, there exists significant economic pressure to track user activity, a development users may find objectionable. Safari's third-party cookie policy works well to block tracking from advertisers and other pure third-party content, but is insufficient to block the multi-function tracking present in *Third-party* widgets such as "Like" buttons and other "social plugins" offered by the likes of Facebook, Google, and Twitter. These elements provide desired user functionality, but also expose the user to the possibility of cookie/referrer-based *tracking* by those third-party sites. Naïve approaches that completely disable third-party interactions prevent such tracking, but at the same time break desired tasks, such as "liking" a page, or engaging in a discussion forum. In order to enable a middle ground, we present Priv3, a web browser extension which uses conditional suppression of third-party cookies and selective reloading of elements on a web page to provide a generic mechanism to protect user privacy from these trackers without compromising usability, creating an "allow with user intent" third-party cookie policy. We have made Priv3 available as Firefox extension that has been downloaded 97,000 times to date, featuring an average user base of 17,000 users daily.

1 Introduction

We believe that any "Do Not Track" mechanism which relies solely on voluntary compliance will not work. Thus absent the force of law, we need to develop technical mechanisms which prevent tracking while still enabling functionality. The Safari third party cookie policy, which we describe as "allow due to previous interaction", works to block tracking from advertisers, but it can't block the tracking performed by various social widgets as when a user logs into one of the social sites, Safari now allows the site's cookies as valid third-party cookies.

By design, these widgets also allow their providers—often companies that specialize in advertising—to track user activity, leading one to wonder as to the extent to which these widgets were designed to track and profile users.

Classic examples of such dual-purpose trackers include Facebook's "Like" button and the "Comment" box. These elements both provide information to users (the number of "likes" and the comments on the article) and enable mechanisms for the users to interact with these elements. However, they also conveniently notify Facebook that a particular, logged-in Facebook user is currently viewing the page, a circumstance valuable for crafting targeted advertisements. Not infrequently, users remain ignorant of the fact that this *third-party privacy leakage* is taking place.

Unlike pure third-party tracking as employed e.g. by advertising networks, simply blocking third-party cookies disrupts desired functionality: users find themselves unable to "Like" pages or enter comments. Such blanket prohibitions may even break some of the functionality provided by the web sites themselves. The central bit of information that allows third parties to associate the rendering of a widget with a particular user is the HTTP cookie that the third party plants in the user's browser upon first contact. Preventing the cookie information from reaching the third party prevents (or at the very least substantially complicates) the association with the user. Accordingly, recent solutions to the problem [3, 15] disable third-party elements completely, replacing or removing them until a user performs an explicit action. Such user intervention to enable third-party cookies is annoying and hampers usability.

We present Priv3, a browser extension which uses a generic mechanism of conditional suppression of third-party cookies and automatic reloading of selected web page components to protect user privacy from both social and web trackers, without compromising usability, which we describe as "allow with user intent". Priv3 takes advantage of these social features also supporting anonymous (non-logged-in) users, by initially loading all the elements associated with Google, Facebook, Twitter, and LinkedIn without cookies. Thus even without cookies, these widgets will still display the number of "likes," comments, and other features. When a user then chooses to interact with a third-party widget visible on the page, such as entering a comment using Facebook's "Comment" box, Priv3 detects user intention to interact with the page and automatically reloads with cookies only the selected components on the web page belonging to the third party. The entire process remains unobtrusive and does not interfere with the user's overall web browsing.

We demonstrate that conditional suppression of third-party cookies and automatic reloading of selected elements on the web page provides a generic defense against both social and web trackers, and that we can achieve the desired functionality without hampering end-user experience. We have built and released Priv3 as Firefox browser extension on AMO [1]. To date, users have downloaded Priv3 97,000 times and we have an average user base of about 17,000 users daily.

The rest of the paper proceeds as follows. Section 2 discusses web tracking and the existing defenses against it. In Section 3, we

present details about the design of Priv3. In Section 4, we discuss our experience with Priv3and finally conclude in Section 5.

2 Overview

2.1 Background

Most web sites include scripts and widgets from third-party sources for the purpose of providing useful services, such as personalization or social interaction. These elements generally run within <iframe> instances, allowing them to set and access cookies within the third party source, with information about the page containing the <iframe> passed within the URL of the <iframe> itself.

For example, a site which wishes to include the "Like" button in a web page either directly creates an <iframe> pointing to a Like button URL on Facebook's site or includes a script and a HTML which the script replaces with the desired <iframe>. In either case, the URL within the <iframe> includes the URL of the desired page, and since the <iframe>'s domain is facebook.com, any user cookies are also transmitted. Thus Facebook sees that the identified user visited the specific page. Such third-party elements need to render within their own <iframe> instances in order to properly transmit the social site's session cookies.

Note that while these elements work to track users by their session cookies, they can still provide useful information to users in the absence of session cookies, i.e., when the user has not logged on. The "Like" button, for example, still renders meaningfully, displaying the total number of "Likes." The Facebook comment plugin still displays the comments on the page.

Unlike the tracking employed by advertising networks and analytics tools, these trackers are *multi-functional*: the designer of the widget-including site would like the visitors to gain some direct benefit. Thus a policy which simply blocks third-party cookies, or the widgets altogether, would prevent users from adding comments, "liking" pages, or performing other relevant activities.

2.2 Current defenses

The problem of privacy violation through web tracking is well documented in prior work [6, 8, 10, 15]. Several solutions have been proposed by browser vendors and researchers alike to ensure user privacy while browsing. Based on their implementations on the client, we categorize these approaches as follows:

Safari's Third Party Cookie Policy. Most web browsers allow preference settings for clients to suppress acceptance of third-party cookies. Safari uses a unique default policy for these third-party cookies, which may be described as "allow due to previous interaction". Only if the user has previously interacted with the site are the third-party cookies allowed. This protects the user from tracking by pure advertising or analytics sites, as these sites are unable to set or read third party cookies. Although effective at blocking advertising-related tracking, this can't protect the user from tracking by social widgets.

Do Not Track. Modern browsers also implement the Do-Not-Track proposal [9]. When this header is present in the HTTP requests, it signals a user's intent to opt out of third-party web tracking. While Do-Not-Track provides users with a simple mechanism to safeguard their privacy, it is not binding on the third-parties to honor the directive and requires strict regulation to ensure compliance.

Extensions. Due to the limited customization available in the built-in functionality provided by the web browsers, developers often use browsers' APIs to build privacy-enhancing extensions. Most of these extensions require user input to enable a flexible cookie management policy, i.e., either a complete or universal cookie blocking or even selective cookie suppression for specific domains. A major problem with the above approach is its inflexibility: once the browser has loaded the page, it will not relax its cookie policies dynamically.

To overcome this issue, another line of extensions, like Ghostery [3] and ShareMeNot [14], explicitly request the user to reload the web page with selected third-party cookies enabled.

- **Ghostery** is a browser extension which provides complete blocking of a large number of third-party trackers and advertisers. Recent versions of the extension also attempt at blocking social tracking by removing all social widgets by default. The user must explicitly click on additional Ghostery introduced buttons to reload the page with widgets enabled. Although effective, this not only affects site layout and introduces a significant page change when an element is enabled, it also prevents the anonymous viewing of comments and similar features.

- **ShareMeNot** is a privacy extension specifically designed for protecting users from tracking by social networks. ShareMeNot replaces the social network buttons with ShareMeNot's internal buttons which, when clicked, pass the click to the social network. This approach changes the page layout and cannot process any elements that the extension is not specifically configured to support. ShareMeNot does not support the Facebook Comments widget, for example, so the extension simply removes those elements from the page without providing a mechanism to see the comments anonymously and browse the content.

HTML and JavaScript. Both <iframe>s and the same-origin policy are useful but they by themselves are not sufficient to stop web or social trackers. Recent advances in web technology allow content publishers, i.e., hosting web sites, to limit transmission of information. HTML5 proposes a new "noreferrer" attribute value [4] which directs the browser to remove Referer headers from the specified HTTP request, while the recently introduced Content Security Policy (CSP) [11] by Mozilla can be used to suppress web tracking by specifying a list of pre-approved domains available for communication.

3 Design and Implementation

Priv3 is a browser extension designed for protecting user privacy against web trackers. As mentioned in Section 2.2, prior work provides limited usability, and such extensions often require user intervention or introduce rendering artifacts. Thus, a key idea we incorporated into Priv3 was to observe the user's intent and enable automatic reloading of cookies. We also desired that our design be generic: while we target specific sites (Google, Facebook, Twitter, and LinkedIn), our tool does not need to understand the implementation of specific elements. Finally, we desired that our tool be both transparent and, to as large a degree possible, unnoticeable. Thus we introduce no changes within the page design, and any content refreshes only represent the transition from an logged-out to a logged-in state on any social widgets. We describe this third-party cookie policy as "Allow with user intent".

User intent. We define user intent as an explicit action on behalf of the user to interact with visible elements on the web page and allowing release of any personally identifiable information associated with it. For example, we would like that unless the user explicitly clicks the "Like" button to record his preference, no third party would receive the cookies, but that once a user expresses an

intent to Like something, the button should work normally. We infer user intent by keystroke and mouseclicks. If the user directs a keystroke or a mouseclick to a social element, we believe it is clear that the user wishes to interact with the element, making it safe to fully enable.

When a user visits a web page on a browser enhanced with Priv3, all the targeted multi-purpose trackers are loaded without cookies, causing them to behave as if the user is not logged in. During a browsing session if the user expresses intent to interact with a particular third-party widget, Priv3 reloads all the DOM elements on the web page belonging to the third-party site with access to the domain cookies.

This selective cookie suppression and reloading of parts of the web page ensures that only trusted components on the web page have access to the user's personal information. This mechanism is seamless and unobtrusive thereby user-friendly.

Implementation. We implemented Priv3 for the Firefox web browser and it is available on Mozilla's addon gallery.[1] We now discuss a few of the salient issues in the implementation.

(1) Capture intent. Priv3 intercepts user mouse clicks and key strokes to identify the target of the event, usually a visible component, like a hyperlink. But in several cases the hyperlink might itself be embedded within an `<iframe>` from another domain. Priv3 uses Firefox's APIs to precisely identify the exact event target and later uses this information to selectively reload third-party components.

(2) Access control. By default, when a web page loads, Priv3 removes the cookie headers in the HTTP requests to third-party domains. This ensures that if the user does not wish to interact with the third-party then it does not learn the identity of the user. But, simply scrubbing out the HTTP cookie header does not prevent the remote server from not learning the identity of the user.

A third-party JavaScript script code can invoke `document.cookie` to gain access to the domain cookies, which could later be transferred to the third-party by circumventing the same-origin policy. To disable such accesses, Priv3 uses script execution event handlers [12, 13] which are fired just before and after the browser executes the corresponding JavaScript code. These handlers control access to the browser's cookie store, i.e., before the third-party script is to be executed, the handler disables access to the cookie store to all JavaScript and later restores the permission after the script has finished execution. Thus, all third-party JavaScript code executing on the web page has no access to either the third-party cookies or the domain cookies until the user chooses to interact with it.

(3) Selective reload. Once the user's intent to interact with a third-party has been established, Priv3 reloads all components from the intended third-party domain. If the target of the user intent was within a third-party `<iframe>`, Priv3 reloads the `<iframe>` with the domain cookies enabled. In certain cases, a click on a hyperlink opens up a new popup or window pointing to a third-party domain. Priv3 identifies such user intent to navigate to a third-party site and ensures that the ensuing HTTP request to the third party has access to the domain cookies. It also reloads JavaScript code from a third-party domain with access to domain cookies.

4 Discussion

Priv3 is a relatively small Firefox extension, requiring less than 700 lines of code. This small size complements the simple nature of the selective-reloading mechanism. Although we currently only block the four major social trackers, it would be trivial to extend to other domains. In particular, the extension does not contain code to recognize individual social elements, only the domains which host social elements. We initially released Priv3 a year ago, with largely positive feedback from our users.

It also remains transparent to the user, as we do not replace the implementation of buttons or social elements. The social networks themselves ensure that the logged-in and logged-out states share similar dimensions, thus the page layout does not deviate from the site's intent. When the user authorizes the social element, only the `<iframe>` instances themselves reload, preventing any disruptive visual effect. We find the clearest indication of the effect manifests on Facebook comment elements: when the user clicks the mouse to input text, the *only* noticeable change on the page consists of the user avatar image changing from the default logo to the user's profile photograph.

We believe this transparency matches user expectations. We believe that users generally do not expect the "Like" button to report to Facebook that a user views a page, but do expect that a *click* on the "Like" button does. Our goal was to make the behavior match this expectation: only when the user "likes" something should Facebook learn about it.

Cookie access control. Priv3 uses access control over cookies to perform its functionality. Its access control utilizes a blacklist of domains to prevent transfer of third-party cookies. Thus, the blacklist must be updated every time the third party introduces resources fetched from a new affiliate domain. This proves essential for preventing the third party from subverting Priv3 by storing tracking meta-data in the cookies from the new domain. Fortunately this is a very coarse-grained blacklist, as it does not need to understand new social elements. For example, to block access for Google+, Priv3 blacklists the following domains: `google` `gstatic`, `google-analytics` and `youtube`.

Priv3 does not enable fine-grained cookie management for end users or modify any existing user-defined cookie policies (such as "allow-for-session"), or introduce any new rules. It simply prevents JavaScript fetched from third-party sources from accessing the hosting web site's domain cookies, until the user expresses his intent.

Aggregators. A few third-party widgets, including AddThis [2] or ShareThis [5], provide access to a number of social networks. These aggregators are themselves web trackers which in turn facilitate social tracking. Priv3 supports user intent based cookie blocking mechanism for such aggregators as well. If a user visits a page with the aggregator widget, Priv3 blocks the aggregator's cookies unless the user specifically visits or logs into the aggregator's website. Once the user interacts with a social network of his choice as displayed on the aggregator widget, cookies corresponding to the social network are enabled across the web page. Thus, Priv3 proves effective against both social and web trackers.

Evasion. It would be possible for a social network to evade our current implementation. Because we still enable loading, albeit without cookies, passive tracking techniques [7] and IP-based tracking may still identify users' history. Similarly, our technique does not currently block HTML5 or flash local storage. However, actually exploiting these weaknesses would be politically dangerous for the social networks, as they would be actively subverting user privacy expectations, and the technique could be extended to ensure that the `<iframe>` instances do not have access to local storage, with the browser string replaced with a known common string.

[1] Priv3 can be downloaded from `https://addons.mozilla.org/en-US/firefox/addon/priv3/`.

5 Conclusion

We have presented Priv3, a browser extension which uses conditional suppression of third-party cookies and automatic reloading of selected web page components to provide a generic defense against both social and web trackers. Social widgets can only track a user by cookie when the user actually interacts with the element, instead of simply viewing the web page containing the element.

We show that this "allow on user intent" cookie policy enables Do Not Track functionality using purely technical means which applies not only to advertisers (which are blocked by Safari's "allow on previous interaction" policy) but also social widgets which serve to both track users and provide functionality. This policy protects users from unwanted tracking without requiring either voluntary compliance or the force of law.

Priv3 accomplishes the desired functionality without hampering end-user experience, as the social elements all remain in the page with no cosmetic change, apart from the transition from the unlogged-in to the logged-in state when the user interacts with an element.

We have built and released Priv3 as Firefox browser extension on Mozilla's addon gallery. It has been downloaded 97,000 times to date, and has an average user base of about 17,000 users daily.

6 References

[1] Addons mozilla. https://addons.mozilla.org.

[2] Addthis. http://www.addthis.com/.

[3] Ghostery. http://www.ghostery.com/.

[4] HTML5 noreferrer. http://www.whatwg.org/specs/web-apps/current-work/multipage/links.html#link-type-noreferrer.

[5] Sharethis. http://sharethis.com/.

[6] Berkeley Law, UC Berkeley. Web Privacy Census. http://www.law.berkeley.edu/privacycensus.htm.

[7] EFF. Panopticlick. https://panopticlick.eff.org/.

[8] Dongseok Jang, Ranjit Jhala, Sorin Lerner, and Hovav Shacham. An empirical study of privacy-violating information flows in javascript web applications. In *Proceedings of the 17th ACM conference on Computer and communications security*, CCS '10, pages 270–283, New York, NY, USA, 2010. ACM.

[9] Jonathan Mayer. Do Not Track. http://tools.ietf.org/id/draft-mayer-do-not-track-00.txt.

[10] Jonathan R. Mayer and John C. Mitchell. Third-party web tracking: Policy and technology. In *IEEE Symposium on Security and Privacy*, pages 413–427, 2012.

[11] Mozilla. Content Security Policy. https://dvcs.w3.org/hg/content-security-policy/raw-file/tip/csp-specification.dev.html.

[12] Mozilla. onafterscriptexecute. https://developer.mozilla.org/en/DOM/element.onafterscriptexecute.

[13] Mozilla. onbeforescriptexecute. https://developer.mozilla.org/en/DOM/element.onbeforescriptexecute.

[14] Franzi Roesner. Sharemenot. https://addons.mozilla.org/en-US/firefox/addon/sharemenot/.

[15] Franziska Roesner, Tadayoshi Kohno, and David Wetherall. Detecting and defending against third-party tracking on the web. In *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation*, NSDI'12, pages 12–12, Berkeley, CA, USA, 2012. USENIX Association.

Opening up the Online Notice Infrastructure

An 'Open Notice' Call For Collaboration

Abstract

Many people who use the web express concern about both the increasing amounts of their personal data being collected online and the terms to which they agree. The existing global infrastructure for managing data and terms relies heavily on the use of 'notices', and these come from the data collector, rather than from the individual. Notices are the legal regulatory mechanisms that refer to privacy policies and terms of service. This system of notices is unfit for use in the contemporary digital context. The notices are often unreadable, unclear, and not useful in practice. They provide limited guidance for people and few benefits to the companies and organisations who issue them. Open Notice is an effort that calls for an open, global, and public infrastructure for legally required, digitally necessary consent-based notices. The Open Notice effort seeks to foster collaboration among the various projects working on these problems around the world. This paper outlines a framework for greater multi-stakeholder collaboration between these projects, regulators, and governments, with the aim of creating an open and global notice infrastructure.

1. Introduction

The current infrastructure for managing the collection of personal data online is strongly influenced by pre-existing privacy laws. At the core of all modern Data Protection regimes around the world is the notion of "informed consent." Before collecting and processing personal data, or to form an agreement to terms of service or a license, a data controller must provide a person with a "notice" and seek "consent." Online, this notice requirement is currently negotiated in the context of terms-of-service and opt-in's with privacy policies. Do Not Track is a welcome mechanism by which individuals can signal that they do not want to be tracked; but in the long term, the online notices that data collectors provide to individuals need to be made more usable.

The current notice infrastructure is "closed." By this we mean that notices are not *automatically findable* and *systematically usable*. There is not currently a common way for an individual to locate an online notice at the point when it becomes relevant to him or her. And, if located, the use of these notices is not

systematic; there are no standard icons, policy layers, terms, or structure. Some patient individuals may attempt to read these notices, but by and large it is not yet possible to assess, evaluate, nor track notices in a systematic way.

The result is that people are not aware of what policies and notices they are subjected to at any one point in time online. There is no way to see how policies may interact with each other, or to see how many policies are active for any one single online interaction. This is because, notices are systematically closed, developed ad hoc by and for lawyers, as to have the impact that people agree to notices but fail to understand them. Closed notices are very expensive, they hinder and stop the free flow of information and the ability for people to control their own information in a performative way. THe effect is that the economic performance of existing notices is very low. It takes people hours to use notices and days to interact with organisations about the use of them. Opening notices would dramatically improve the performance of notices for not only privacy, but security, health and safety information, in multiple ways. As a result, web users increasingly find their expectations for privacy and use of service are in conflict with the legal realities of the notices they have "agreed" to.

An "open" notice infrastructure would be one in which these notices *are* automatically findable and systematically usable. The average web user would, with minimal cost, be able to access, assess, track and compare the notices they have agreed to. There are a number of ways this could be achieved, and the solution is likely to involve a mixture of top-down, bottom-up, and crowd-sourced technical and policy efforts. The important point is that under an "open" system, notices would no longer demand an unrealistic level of patience and diligence on the part of the user.

Opening up notice may not, in itself, solve all the problems of privacy and fair service use. However, it is the most relevant starting point and a fundamental building block for further solutions across all jurisdictions. The positive opportunities inherent in the emerging market for personal data are predicated on the informed consent of individuals. Until they can actually *use* online notices, people will not be in a position to make informed consensual choices.

2. A framework for opening up notice

There are a number of projects working on solving different aspects of this problem. They each would benefit from a more open notice infrastructure. There are a number of steps in the process of opening up notice, including [see

appendix A for flowchart of Open Notice efforts and components, (contributed by collaboration between Solon Barocas' and Par Lannero's)];

The components are:

- Locating and recording the URL of a notice/policy.
- Capturing and cleaning the notice, and putting it into a centralised repository.
- Tagging notices.
- Providing a method for remotely tracking changes to notices.
- Establishing standardised terms/language.
- Developing schema that breaks a notice into different facets.
- Coding notices according to a schema – either by the organisation who created the notice, or hand coded by lawyers and law students, crowd-sourced, and / or natural language processing techniques.
- Converting the notice into a machine readable format.
- Subsequently, they can be:
    - a) parsed into simpler formats, such as icons, simple language, or nutrition labels.
    - b) compared with each other, and rated according to a variety of metrics.
    - c) compared to and checked for compliance with a user's own privacy preferences (P3P style).

There are different projects working on different stages of this process. Our intention is to facilitate basic notice standards, privacy, and terms to help all projects in this area interoperate, and even build upon each other's work. Ideally, an Open Notice coalition would continue to develop and contact all the groups and projects working on this problem. There may be additional stages in the process which we have neglected, which other projects might be working on. The overall aim is for a broad, multi-stakeholder process which encourages any relevant projects to get involved, collaborate, innovate, and find funding.

What could be achieved through collaboration?

- Common standard, location and structure for notices.
- Common formatting, layering of notices to polices and icons.
- Common legal terms, vocabulary, and ontology.
- Common way to link/augment existing policies so they are open.
- Common data formats, and places to share code development.

- Mapping the global legal environment, and combining this with common ways to refer (or even better, link) to immediate remediation as well as the regulation/legislation.
- Developing common metrics for evaluating notices across multiple factors, from legal compliance within a given regime, to attributes like the clarity and usability of notice.
- Standard ways to inform users of changes to notices.
- Standard ability to enable people to independently track, manage and control consent.

3. The potential benefits

There are many potential benefits of greater collaboration between these projects. For those working on building this infrastructure, there will be greater interoperability between projects. By using common data formats and places to share code, we can hopefully minimise duplication of existing work and be more effective. By agreeing on common terms and ontologies, projects working on different aspects of the ecosystem can normalise legal terminology and ensure that the whole process flows more smoothly for each other and for end users of notice. For instance, an individual should be able to apply the rating system (or legal metrics) provided by one project to the notices that are captured and tracked by another project. Or alternatively, use one service to track changes to the notices they have agreed to, and another service to translate them into simplified versions. Common terms and formats could create this kind of interoperability.

The benefits to society at large would be significant. Common standards and open infrastructure for notice would pave the way for incredible personalisation, new services, and alternative forms of notice which are bilateral and very performative. New commercial and social tools and services might also be built on top of this open infrastructure. For instance, emerging "Vendor Relationship Management" tools could be used in conjunction with this infrastructure to go beyond tracking and evaluating notices towards two-way negotiation between individuals and companies.

4. Conclusion

Opening up the notice infrastructure is not in and of itself the solution to all current and potential privacy problems, but usable transparency over notices and consents is an important and necessary first step. Privacy laws and principles around the world emphasise openness as informed consent is a core prerequisites for legitimate information processing. The existing notice

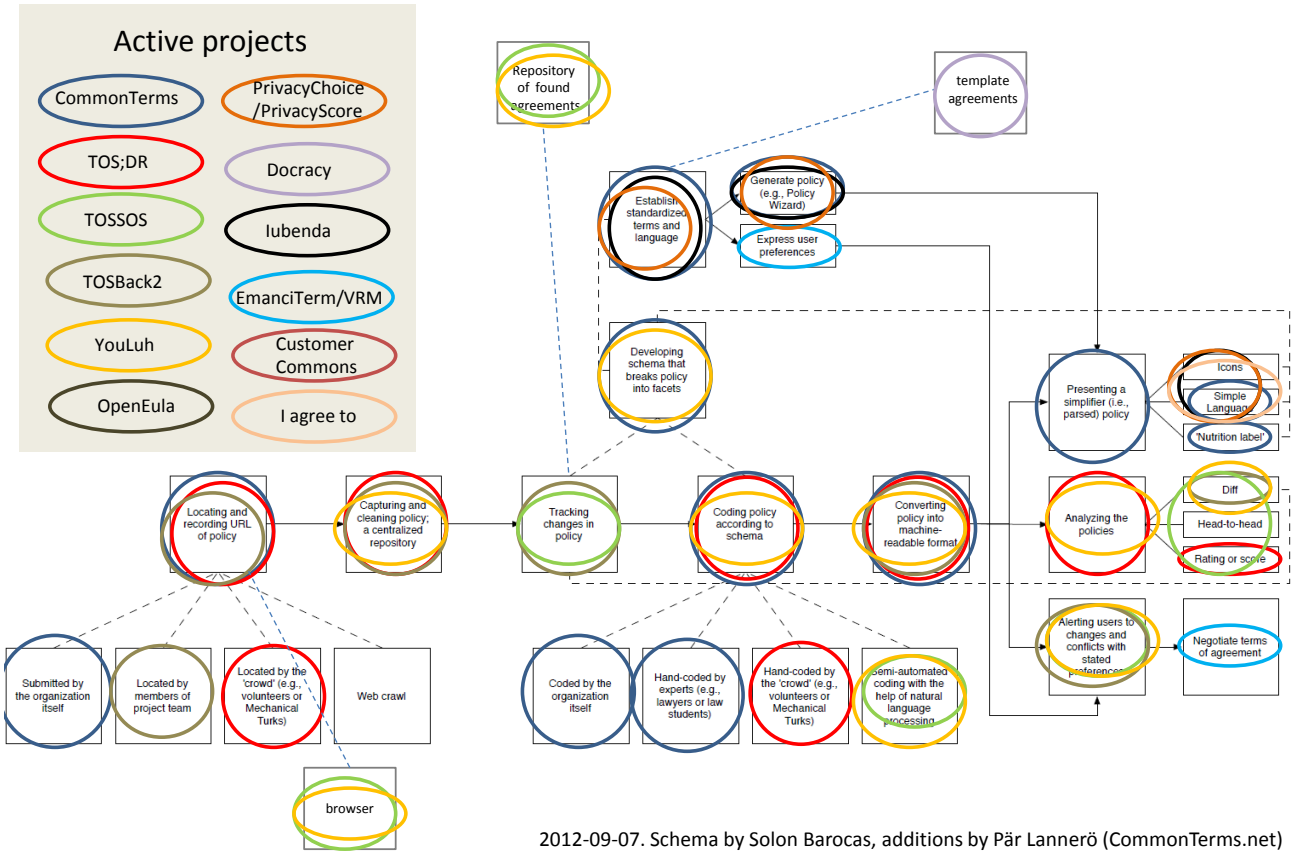
infrastructure does not live up to its stated purpose, rendering its legality to administer terms of service and the like questionable. With further regulation on the horizon, the existing framework will cease to meet the demands of modern information infrastructure.

Many projects today are working on updating, adapting, and bridging this infrastructure. By working together on Open Notice, collaboration with these projects will create an open infrastructure, helping each other and people who interact with digital policies across the web.

Appendix A

Ecosystem of Privacy Policy/TOS management

The road to solving the Biggest Lie problem



2012-09-07. Schema by Solon Barocas, additions by Pär Lannerö (CommonTerms.net)

Appendix B: List of Founding Participants and Projects

Mark Lizar & Reuben Binns
Open Notice Project Coordinators

Pär Lannerö
CommonTerms project
“CommonTerms is a non-profit project with the same focus as Open Notice: to change the current broken status of online Terms & Conditions (Terms of Service, Privacy Policies and to some extent EULAs and similar). We have drafted a *preview format* for online Terms & Conditions. We have tried to find all other similar projects, in order to get some cooperation going. We try to contribute to putting the "Biggest Lie" problem in focus, by presenting at conferences, talking with journalists, policy makers, web developers, being active in social media and launching the confession booth www.biggestlie.com. We did a detailed analysis of 22 TOS documents selected from popular online services to see what terms are commonly appearing, and we are planning to use the resulting database as a start for a growing database of common terms, which will be accessible by API and as open as possible for anybody to read, add to and reference.”

Sebastian Lemery
TOSSOS Project
TOSSOS was built to archive Terms/Privacy Policies, to allow users the ability to compare products side by side and see legalese summarized into plain English. There is currently a early release Chrome Plugin and a simple, general purpose API.

Veronica Picciafuoco
Docracy.com
“Docracy's general goal is to be the home of free legal documents. We call it "github for legal documents" as we encourage users to branch and improve documents (anybody can publish a new document). As you can imagine, terms of service and privacy policies are some of the most requested documents. In an effort to foster the standardization of privacy policy for mobile apps, we launched this project: https://www.docracy.com/mobileprivacy/ - the call to action is: improve the draft, create new versions for different types of apps, etc.... getting to a crowdsourced standard privacy policy language that mobile developers can rely upon... we welcome any contribution / partnership with projects that share our goals. Also, we are a source of publicly available

documents and we host some famous TOS/PP that we keep updated, so users can see differences between versions."

Hugo Roy
ToS;DR Project
"ToS;DR (Terms of Service; Didn't Read) is creating a transparent peer-review process to rate terms of service to produce open data and free software." Hugo suggests "standardised information about services so that we can share information about a specific service, building a strong archiving system to track changes (see tosback2), tools to scrape these archives to help humans read the important stuff and leave out the unimportant stuff, and a way to build knowledge around these data for users (one example: a rating system)".

Matt Snyder
Youluh Project
"Youluh will be a service for end users. Users will have a client app by which they can upload any EULA, ToS, or other electronic agreement they encounter, and receive back a brief report card containing statistics about the clauses in the agreement. The service takes into account that many clauses are borrowed and adapted by agreement authors, so one clause may appear largely unchanged in many agreements. The report card tells the user at a glance about clauses he/she has already accepted as part of prior agreements, clauses accepted by friends, clauses with commentary by other users and/or experts, and clauses that are new or contain new wrinkles. The report card is a decision-support mechanism, meant to assist users who feel they do not have time to read every agreement they encounter, but want to be alerted to anything seriously bad before agreeing. Think of it as a virus scanner for electronic agreements. Youluh is being built by a commercial enterprise, Double Crossroads LLC, but there is no specific profit goal for Youluh, and it may eventually become non-profit."

Gregg Bernstein
iagreeto.org (MFA thesis project)
UX Designer, Researcher, and Educator
"I translate complex information into something usable. For my thesis, I worked with an attorney to translate a ubiquitous EULA (iTunes) into something human-readable and visually more intuitive at iagreeto.org. I'm interested in plain language and simplified interface design, and this is the lens through which I view ToS, EULA, and Privacy issues. I also consult for CommonTerms."

Ashkan Soltani
Know Privacy
Approach: A comparison of users' expectations of privacy online and the data collection practices of website operators.

Goal: To identify specific practices that may be harmful or deceptive and attract the attention of government regulators.

Result: Recommendations for policymakers to protect consumers and for website operators to avoid stricter regulation.

Mary Hodder & Renee Lloyd
Customer Commons
CREATING A WORLD OF LIBERATED, POWERFUL AND RESPECTED CUSTOMERS.
Customer Commons is a non-profit for customers who are tired of just complaining about the "powers that be" and want to contribute to making tools for the rest of us

Joe Andrieu
Standard Information Sharing Label
“I'm a developer and entrepreneur. I've been working with Iain Henderson and others at the Information Sharing Work Group (ISWG) to develop a framework for helping individuals take control over the information they share online. The ISWG recently launched the Standard Information Sharing Label, aka Standard label, as a clear, consistent way people to understand what happens with their information when they share it, at the point they make the decision to share. I'd love to coordinate promotional efforts, collaborate on standards, etc. In particular, the Standard Label has a field for third party ratings, which allows a user-agent, i.e., a browser plugin, to display reputations from user-configured sources, independent of who authored the Label.”

Brian Erdelyi
www.clearware.org

Clearware.org was founded to help make sense of these software agreements and is guided by the fundamental principle that all computer users have a right to understand the terms or conditions of software end-user license agreements (or online terms of services) before deciding whether to purchase, install or use the associated software or service. Clearware.org proposes a set of simple, colour coded and easily recognizable symbols that depict terms impacting the user's experience, privacy and system security. These symbols are presented in a human-readable format similar to care labels on clothing, nutrition

facts on food or warnings on hazardous materials. A system readable-format to programatically notify users and a crowd sourced model for creating and storing the labels were also proposed.

Behavioral Targeting
Legal Developments in Europe and the Netherlands

Frederik Zuiderveen Borgesius

Ph.D researcher, focusing on behavioral targeting and privacy law
Institute for Information Law, University of Amsterdam
F.J.ZuiderveenBorgesius [at] uva.nl

Position Paper for the W3C Do Not Track Workshop, November 2012

Introduction

This paper discusses legal developments in Europe and the Netherlands. Recent decisions show that European data protection law, or privacy law, applies to behavioral targeting in most cases. Dutch law explicitly presumes that data protection law applies to behavioral targeting. This means that companies have to comply with data protection law's fair information principles. For example, companies must refrain from secret or excessive data collection. Perhaps the principles could provide inspiration for future W3C projects. Could technology design foster fair information processing?

I would like to speak at the workshop on such issues, and look forward to discussing them with the workshop participants.

Legal developments in Europe

In Europe, the right to privacy and the right data protection are fundamental rights.[1] Like most privacy laws in the world, European data protection law is triggered when a company processes "personal data". Many behavioral targeting companies process pseudonymous profiles (individual but nameless profiles). Do these companies process "personal data"? Yes, say European data protection authorities. This is compatible with case law of the highest court of the European Union.

[1] Article 7 and 8 of the Charter of Fundamental Rights of the European Union, and article 8 of the European Convention of Human Rights.

The European Data Protection Directive defines personal data as: “any information relating to an identified or identifiable natural person ('data subject').” A person is identifiable when he or she can be directly or indirectly identified. To determine whether a person is identifiable, it’s not decisive whether it’s the company holding the data, or another party that can identify a person.[2]

The Court of Justice of the European Union, the highest authority on the interpretation of European Law, has not ruled on behavioral targeting yet. But there is relevant case law. The discussion about behavioral targeting is similar to the debate about IP addresses. In November 2011, the Court ruled that the IP addresses in that case are personal data.[3] The Court thus reaffirms that information without a name can constitute personal data.[4]

European national Data Protection Authorities, cooperating in the Article 29 Working Party, say that data that can distinguish a person within a group are personal data.[5] The Working party adds that pseudonymous profiles, for example tied to a cookie, are personal data because they “enable data subjects to be 'singled out', even if their real names are not known”.[6] Although not legally binding, the Working Party’s opinions are influential, since it usually takes decisions by consensus.

Many, although not all,[7] commentators agree that data protection law applies to behavioral targeting.[8] The Privacy Commissioner of Canada[9] and the American Federal Trade Commission reach similar conclusions.[10] The proposal for a new European Data Protection Regulation also applies to pseudonymous profiles and “online identifiers” in most cases.[11] Taking all this into account, it seems safe to assume that data protection law generally applies to behavioral targeting.

Legal developments in the Netherlands

In June 2012, the new Dutch Telecommunications Act entered into effect.[12] The Dutch Act essentially copies the ‘cookie clause’ of the European e-Privacy Directive,[13] and only allows the use of tracking technologies after prior informed consent of the user. (A translation of the

[2] Article 2(a) and recital 26 of the Data Protection Directive 95/46/EC.
[3] CJEU, 24 November 2011, Case C70/10 (Scarlet/Sabam), par. 51.
[4] See for example CJEU, 9 November 2010, Joined cases C-92/09 and C-93/09 (Volker und Markus Schecke and Eifert), par 52; CJEU, 24 November 2011, Joined cases C-486 and C-469-10 (Asociación Nacional de Establecimientos Financieros de Crédit), par. 42.
[5] Article 29 Working Party, Opinion 4/2007 on the concept of personal data (WP 136). 20 June 2007, p. 12-20.
[6] Article 29 Working Party, Opinion 2/2010 on online behavioral advertising (WP 171). 22 June 2010, p. 9.
[7] See e.g.: G-J. Zwenne, Over IP-adressen en persoonsgegevens, en het verschil tussen individualiseren en identificeren (About IP addresses and personal data, and the difference between individuating and identifying), Tijdschrift voor Internetrecht, February 2011, p. 4-9.
[8] See e.g.: P. Traung, ‘EU Law on Spyware, Web Bugs, Cookies, etc., Revisited: Article 5 of the Directive on Privacy and Electronic Communications’, Business Law Review 2010-31, p. 216–228.
[9] Office of the Privacy Commissioner of Canada, Privacy and Online Behavioural Advertising (Guidelines), December 2011, www.priv.gc.ca/information/guide/2011/gl_ba_1112_e.pdf, p. 2.
[10] The FTC says that privacy rules should apply when a company can reasonably link information to a consumer or a device (FTC, Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers (March 2012), www.ftc.gov/os/2012/03/120326privacyreport.pdf, p. 22).
[11] The Regulation’s definition of personal data includes “online identifiers” in the list of examples that may be used to identify a person (article 4(1)). But see also recital 24 (Proposal for a regulation on the protection of individuals with regard to the processing of personal data and on the free movement of such data, COM(2012) 11 final).
[12] The official Dutch text can be found at: http://wetten.overheid.nl/BWBR0009950.
[13] Directive 2002/58/EC, as amended by Directive 2009/136/EC.

provision is in the appendix). Companies may not infer consent from inactivity of the user. Silence is not consent. The Dutch provision is technology-neutral: it applies to cookies and other tracking technologies such as device fingerprinting.[14] Without their consent, Dutch internet users may not be tracked. This also applies to users that haven't set a Do Not Track preference in their browser.

But the Dutch Act goes further. It also contains a legal presumption regarding tracking technologies for behavioral targeting. The use of such technologies is presumed to entail the processing of personal data. The legal presumption shifts the burden of proof. It's up to behavioral targeting companies to prove that they don't process personal data. The provision basically codifies the view of the European data protection authorities. The Dutch legislator added the legal presumption to emphasize that default browser settings could never be interpreted as consent for tracking cookies or similar technologies.[15]

The Dutch Telecommunications Authority OPTA oversees compliance with the provision. OPTA says that the provision also applies to foreign website publishers and behavioral targeting companies. OPTA can issue fines of up to 450.000 euro.[16]

If a company processes "personal data", the Dutch Data Protection Authority also enters the picture. Because of the legal presumption, the Data Protection Authority doesn't have to prove that a company employing tracking technologies processes personal data. The Data Protection Authority can't impose fines, but it can impose large preventive penalties if a company doesn't comply with its administrative orders.[17]

The legal presumption enters into effect on 1 January 2013.[18] The Dutch Senate said that this delay could enable the online marketing industry to come up with a user-friendly system to obtain consent, for instance by developing a meaningful Do Not Track standard.[19]

Fair Information Processing

If a company processes personal data, it has to comply with all the data protection principles. Most importantly, data processing has to be transparent. Secret data collection is not allowed.[20] But there's more. For example, the data minimization principle prohibits the collection or storage of excessive amounts of data.[21] The security principle requires companies to ensure a reasonable level of security of data they process.[22] The law grants people whose data are being processed several rights. For instance, everyone has the right of

[14] Eerste Kamer, vergaderjaar 2011–2012, 32 549, G, 17 February 2012, p. 4-6 (answers of Minister of Economic Affairs, Agriculture and Innovation to the Senate).
[15] Explanatory memorandum to the amendment by Van Bemmel and Van Dam to the Bill to amend the Telecommunications Act (Dutch Parliament 2010-2011, 32549, nr. 39.
[16] OPTA, 'Veelgestelde vragen over de nieuwe cookieregels. Update' (Frequently asked questions about the new cookie rules. Update'), 2 August 2012 www.opta.nl/nl/actueel/alle-publicaties/publicatie/?id=3636, p. 3, p. 9.
[17] The website of the Data Protection Authority is at www.dutchdpa.nl.
[18] Article VII, 1(c) of the Besluit implementatie herziene telecommunicatierichtlijnen (decision implementation telecommunications directives), https://zoek.officielebekendmakingen.nl/stb-2012-236.html.
[19] Handelingen Eerste Kamer van de Staten Generaal, Vergaderjaar 2011-2012, Vergaderingnummer 28, Telecommunicatiewet en Wegenverkeerswet 1994, 32549, https://zoek.officielebekendmakingen.nl/h-ek-20112012-28-9.pdf.
[20] Article 10 and 11 of the Data protection Directive.
[21] Article 6(c) and 6(e) of the Data Protection Directive.
[22] Article 16 and 17 of the Data protection Directive.

access to data that have been collected concerning him or her, and the right to have data rectified. People can always withdraw their consent.[23]
At the core of the European data protection regime are the fair information principles. These forty-year old principles are well established.[24] The principles are contained in international instruments such as the OECD Data Processing Guidelines,[25] and the Data Protection Convention (ratified by 44 countries).[26] Although the national implementation varies, the principles express a worldwide consensus on how to ensure fair information processing.

Conclusion

In sum, European data protection law most probably applies to behavioral targeting. Dutch law is more explicit and presumes this is the case. Therefore companies must comply with data protection law's fair information principles.

Perhaps the fair information principles could provide inspiration for future W3C projects. Could the W3C help to put the principles in practice? For instance, maybe technology could help to make data processing transparent. Or technology might enable people's right to access data concerning them. As the Mission of the W3C puts it: "technology design can foster trust and confidence."[27]

* * *

[23] Article 12, 24 and 15 of the Data protection Directive; article 8 of the EU Charter of Fundamental Rights.
[24] See for an early example: US Department of Health Education and Welfare (HEW), Records, computers and the rights of citizens: report of the Secretary's Advisors Committee on Automated Personal Data Systems, Washington: US Government Printing Office 1973.
[25] See www.oecd.org/internet/interneteconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.
[26] See www.coe.int/dataprotection.
[27] W3C, Web of trust, http://www.w3.org/Consortium/mission#principles.

APPENDIX

Article 11.7a of the Dutch Telecommunications Act (unofficial translation by the author)

1. Without prejudice to the Data Protection Act, anyone who wishes to access information that has been stored in a user's terminal equipment, or wishes to store information in a user's terminal equipment via an electronic communications network, must:

a. provide the user with clear and complete information, in accordance with the Data Protection Act, at least about the purposes for which he wishes to access the information concerned and/or for which he wishes to store information, and

b. have obtained the user's consent for this activity.

Any activity as referred to in the preamble, with a view to collecting, analyzing or combining information about the user's or subscriber's use of various services of the information society, for commercial, charitable or idealistic purposes, is presumed to be the processing of personal data, as defined in article 1(b) of the Data Protection Act.

2. The requirements of paragraph 1 a and b shall also apply in the event that (other than by means of an electronic communications network) anyone causes information to be stored, or information stored in the terminal equipment to be accessed, by means of an electronic communications network.

3. The provisions of the first and second paragraph shall not apply if they relate to technical storage of, or access to, information, with the sole purpose of:

a. carrying out the communication over an electronic communications network, or

b. providing a service of the information society requested by the user, and the storage of, or access to, information is strictly necessary.

4. Regarding the requirements set out in paragraph 1(a) and (b), further rules can be given by governmental decree in agreement with Our Minister of Justice and Security. The Data Protection Authority shall be consulted about the draft of such a governmental decree.

* * *

Jim Brock
@privacychoice
101 Cooper Street
Santa Cruz CA 95060
831-239-0095

W3C Workshop: Do Not Track and Beyond
Position Paper for Participation

PrivacyChoice Background

Founded in 2009
Mission: make privacy easier

For pros:
Trackerlist database and API (10+ deployments)
Free site and app tracker scanning
Free mobile privacy policies

For folks:
Privacyfix - personal privacy control-panel
Privacyscore - site privacy ratings
TrackerBlock - browser add-ons and tracking protection lists

Research and commentary:
blog.privacychoice.org

Interest in the Workshop

Get educated about Do Not Track technical and policy standards
(with particular interest in mobile implementations).
Contribute research and learning from our users and ad-industry research.
Connect with open projects that can use our dataset.
Identify potential new features that leverage W3C standards.

It's The Users, Stupid! Towards User-Centered Privacy Standards by Considering Default Settings

Serge Egelman
Computer Science Division
Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA 94720
egelman@cs.berkeley.edu

Position

"If you choose not to decide, you still have made a choice."

-Rush

In order to create truly effective privacy standards, default settings need to be included in the specifications. The corollary to this is that if default settings are not specified, implementers will interpret the standards in different ways, and other stakeholders will cite this divergence as evidence that the standard is not working. We saw this occur with an early version of the W3C's Do Not Track (DNT) standard: Microsoft proposed enabling DNT by default for users of Internet Explorer 10 [1]. In response to Microsoft's move, the advertising industry wrote a letter to Microsoft claiming that their decision will harm consumers [2], leaving open the possibility that advertisers may simply refuse to recognize DNT headers when transmitted by IE 10 users. Mozilla has taken a similar position, claiming that the decision to enable or disable DNT "must be the user's choice" [3].

While stakeholders on both sides continue to argue which default setting is most supportive of user choice, none of these stakeholders are bothering to ask the user directly. There are essentially two states for a DNT setting: enabled or disabled. Whichever one remains the default, a choice is being made on behalf of users. In order to avoid presumptuous arguments about what users actually want, it is imperative that users' preferences be adequately represented.

McDonald and Cranor conducted early work on user acceptance of behavioral advertising [4,5]. They found that when outright asked if users support being tracked across websites so that they can receive tailored advertisements, only 20% support the notion.

At the same time, this is not a decisive argument for enabling DNT by default: users have different reactions to DNT based on how they are asked. In the aforementioned studies, users were simply asked whether they would be comfortable being tracked. In a subsequent study that I performed, when users were provided with monetary incentives to keep behavioral advertising enabled, the vast majority of them supported tracking [6]. Thus, we need to pay careful attention to methodology when gauging user preferences.

Moving forward, all stakeholders need to come together to agree on both the benefits and concerns surrounding DNT. These pros and cons need to be presented to users when studies are performed so that their results are ecologically valid. Until we do this, it is impossible to say which setting supports the most users by default.

References

1. Worstall, Tim. "Microsoft Sticks with Do Not Track Default: And Boy Are The Advertisers Angry." Forbes.com, October 3, 2012. http://www.forbes.com/sites/timworstall/2012/10/03/microsoft-sticks-with-do-not-track-default-and-boy-are-the-advertisers-angry/
2. Association of National Advertisers. "ANA Board Opposes Microsoft's Decision to Implement `Do-Not-Track' Default Function for Internet Explorer 10 Browser." October 1, 2012. http://www.forbes.com/sites/timworstall/2012/10/03/microsoft-sticks-with-do-not-track-default-and-boy-are-the-advertisers-angry/
3. Clarke, Gavin. "Top Admen Beg Microsoft to Switch Off 'Do Not Track' in IE 10." The Register, October 3, 2012. http://www.theregister.co.uk/2012/10/03/ie_10_dnt_default_advertisers_letter/
4. McDonald, Aleecia and Cranor, Lorrie. "Beliefs and Behaviors: Internet Users' Understanding of Behavioral Advertising." In Proceedings of the 38th Research Conference on Communication, Information and Internet Policy (Telecommunications Policy Research Conference), October 2, 2010.
5. McDonald, Aleecia and Cranor, Lorrie. "Americans' Attitudes About Internet Behavioral Advertising Practices." In Proceedings of the 2010 Workshop on Privacy in an Electronic Society (WPES), October 4, 2010. http://www.aleecia.com/authors-drafts/wpes-behav-AV.pdf
6. Egelman, Serge and Felt, Adrienne, and Wagner, David. "Choice Architecture and Smartphone Privacy: There's A Price for That." In Proceedings of the Workshop on the Economics of Information Security (WEIS), June, 2012.

October 22, 2012

To the W3C "Do not track and beyond" Program Committee:

Constant Contact® helps small organizations create and grow customer relationships in today's socially connected world. Through its unique combination of online marketing tools and free personalized coaching, Constant Contact helps more than half a million small businesses, associations, and nonprofits find, connect to, and engage with their next great customer, client, or member. Launched in 1998, Constant Contact has long championed the needs of small organizations, providing them with an easy and affordable way to create and build successful, lasting customer relationships.

Constant Contact understands the importance of data privacy and has continually participated in organizations and working groups targeted at developing industry-wide standards. Constant Contact is a full member of the Messaging Malware and Mobile Anti Abuse Working Group (M^3AAWG), and is a founding member of the Email Sender and Provider Coalition (ESPC).

Constant Contact supports the efforts of the W3C's Tracking Protection Working Group (TPWG) to define consensus standards that provide consumers choice for third party online behavioral advertising. While much progress has been made on the technical and compliance standards, they have not progressed enough for us to make specific recommendations. However, we do have select comments to share with the working group and extended W3C community.

We are concerned about the ability of small organizations to comply with the TPWG's proposed technical and compliance policies.

To date, most of the debate and consensus decisions have been between internet service providers, web client software providers, representatives from the online advertising community and privacy advocates. These discussions were mostly about how large companies with knowhow and resources can (or cannot) comply with the complexities of these new standards. Thus far, there has been little discussion about the effect these standards will have on small business and nonprofit organizations. Small organizations are generally headcount and capacity constrained and may be uniquely challenged ensuring ongoing compliance with these new tracking policy standards. We would advocate for a standard that establishes a level playing field for both smaller and larger organizations and does not put small organizations at a disadvantage from a compliance perspective.

As an agenda topic for the upcoming November meeting, we would like to see more discussion about compliance with the TPWG's draft documents as it relates to small organizations.

Constant Contact looks forward to actively participating in the forthcoming W3C workshop entitled "Do Not Track and Beyond," which is scheduled for November 26-27, in Berkeley, California.

Sam Silberman
Director, Standards
Constant Contact

Introduction

The emergence of digital technologies including the Internet, smartphones, tablets and other digital devices has increased both the complexity of the core definition of this construct, the ways in which privacy may be viewed, and exacerbated the consequences of failing to adequately protect privacy at both individual and systematic levels. Privacy as a complex construct has been approached by researchers from various disciplines and from a number of different perspectives[1]. As Kilger and Jovanova (2012) point out

> Some researchers view the general concept of privacy from a political perspective as an essential right to protect of individuals that is supposed to be protected by the state (Rosen, 2000). Others view privacy as a commodity subject to the forces of costs and benefits that may be able to be described by a particular calculus (Li et al, 2010). Privacy has also been framed as the ability to control the dissemination and use of personal information (Margulis, 1977).

Understanding more about how people feel about different dimensions of privacy and how these relate to the use of digital technology may assist in providing some guidance in developing digital privacy standards and their implementation among new and existing digital technologies.

Data Source

The source for the analyses is a recent 2011 Experian Simmons' National Consumer Study (current 2012 data is also available and would be used for the actual discussion but I could grab the 2011 data quickly). This study is a national probability sample of the United States adult population that is conducted continuously throughout the year. The sample design contains disproportionate sampling for Hispanics, higher income households and certain geographic areas and is design weighted to account for this as well as stratum non-response and then projected to current population estimates. The study contains approximately 25,000 respondents and over 60,000 variables – including measures of Internet, smartphone and tablet usage, extensive demographics and attitudes and behaviors directly related to privacy issues.

There are currently 19 privacy measures available in the National Consumer Study and these measures cover a number of topics including the respondent's assessment of the risk and impact of online information on their lives, the amount of control they feel they have over their personal information, their tendency to use the Internet less because of privacy concerns, their interactions with companies concerning the collection and use of their personal information and more. A full list of the privacy measures can be found in Appendix A.

Examples of U.S. Adult Attitudes Towards Privacy Issues

Examining perceptions and behaviors regarding privacy one can examine a number of things such as the size of the segment of the U.S. population that feels a particular privacy issue is of concern. Also, one might be able to compare the subjective or objective evaluations of the actual, estimated risk that issues

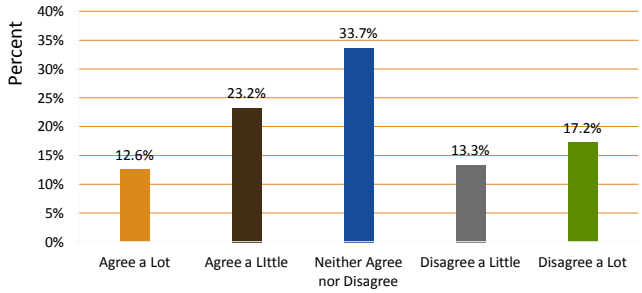
[1] For the reader interested in a comprehensive review of the privacy literature see Smith et al (2011).

in each of these areas poses with the perception of these risks by the U.S. population. Balancing actual with perceived threats may provide practitioners with data that may assist in prioritizing privacy standards as wells as providing insight into where better educational efforts may be most effective.

How much control people have over personal online information is one measure about how people feel about online privacy. As can be seen in Figure 1, only about 1 in 8 people agree a lot that they have control over the online information about them. A little over 30% of adults disagreed to some extent that they have a fair amount of control and about 1 out of 3 adults did not agree or disagree with the statement.

Basic Distributions

Figure 1
I Feel Like I Have A Fair Amount of Control Over the Personal Information About Me That Can Be Found Online

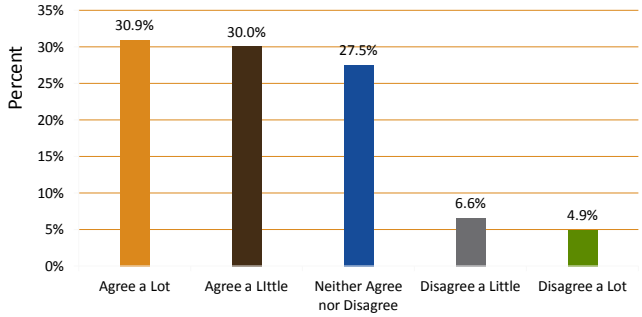


Source: Experian Simmons Fall 2011 12-Month National Consumer Study

Do people feel that personal online information is just as risky as providing it offline? According to the data, over 60% of respondents said that they agreed to some extent with this statement as shown in Figure 2. Only 11.5% of U.S. adults disagreed to some extent with this statement. Thus a majority of American adults are inclined to believe that the risk of providing personal information online is equivalent to that of providing it offline.

Basic Distributions

Figure 2
Providing Personal Information Offline is Just as Risky as Providing it Online

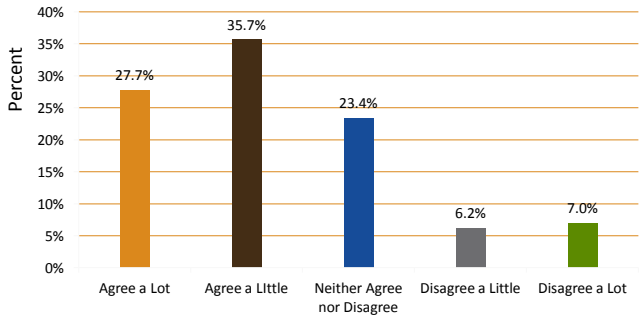


Source: Experian Simmons Fall 2011 12-Month National Consumer Study

There is a lot of debate about the actual safety of online information and much of that debate often involves complex technical evaluations of information security measures, corporate infosec policies, legal protections and other relevant factors. It is unlikely, given the uncertainty and lack of consensus with which privacy and information security experts view the safety of online that the average American has a very accurate estimate of how secure and private information they have provided online actually is. How do people feel about the level of understanding they have about the risks of providing personal information online?

Basic Distributions

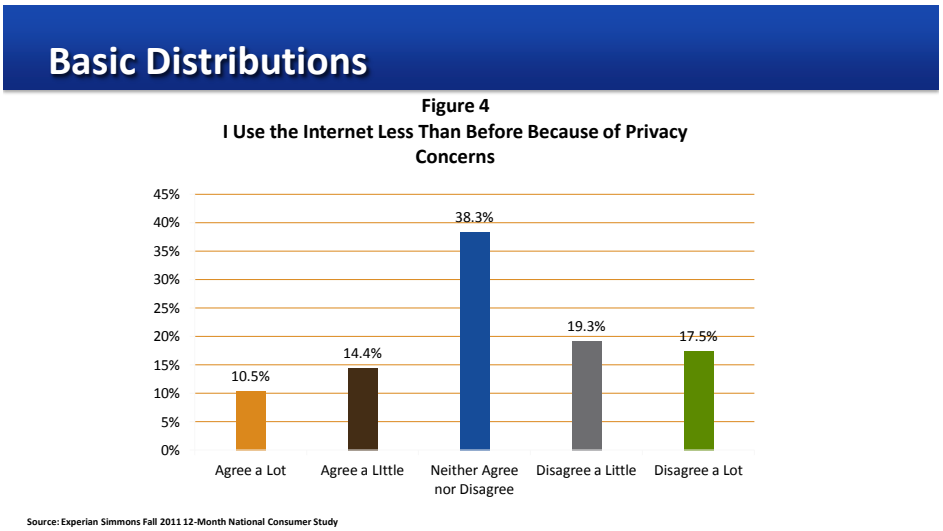
Figure 3
I Feel that I Understand the Risks of Providing Personal Information Online



Source: Experian Simmons Fall 2011 12-Month National Consumer Study

Over 63% of adults in the U.S. agree at least to some extent that they understood the risks of providing personal information online. This is perhaps somewhat surprising given some of the media attention to recent data breaches and it is likely that this perception varies depending upon the length of time individuals have been using the Internet.

Given that there is the perception of at least some risk involved with being online, has this perception of risk discouraged individuals from using the Internet as much as perhaps they did previously? In Figure 4 below, we see that about 1 in 4 U.S. adults state that they use the Internet less than before because of privacy issues. This underscores the need to build in privacy protections that individuals can count on and have some confidence in. As threats to online privacy continue to become more sophisticated and accumulate more serious consequences, this portion of the U.S. population that has backed off to some degree from the Internet may continue to grow and along with it losses in opportunities for more economic benefits from the net may occur.

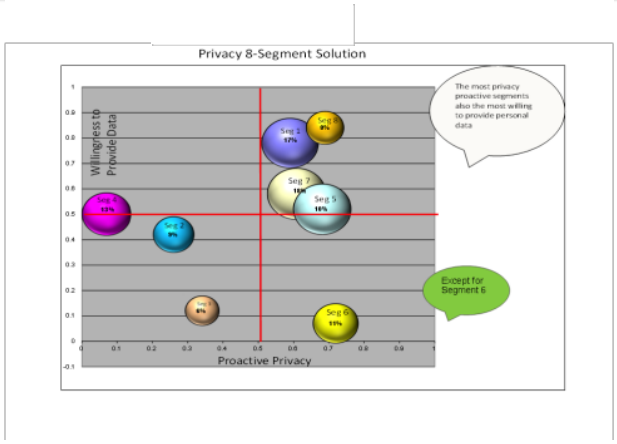


Finally, there is the persistent idea of the paradox of privacy (Smith et al ,2011). That is, a number of research studies have suggested that what people say about protecting privacy and what they do sometimes are quite different. Gaining an understanding of this paradox is important in that it may help us better understand the willingness of people to provide information online in relation to ecommerce and other important objectives even when there are threats to privacy.

In the last figure below, we performed a segmentation of respondents from some early privacy data that had come in. The axis represents the extent to which people were willing to proactively protect their privacy. The y axis represents the willingness to provide personal data in exchange for something of value. Segment 6 is the archetype that one thinks of when thinking of how people structure their attitudes and behaviors about privacy – that is, they state that they take proactive action to protect their privacy and they are not willing to provide personal information in exchange for something of value.

However, what is interesting is the upper right hand quadrant – this is where people who are at least somewhat proactive about their privacy are also willing to provide personal information in exchange for something of value. This lines up with the privacy as commodity perspective previously cited as well as providing some additional empirical evidence for the privacy paradox.

Privacy Segmentation



In summary, there is a lot to learn about how people perceive online privacy risks as well as whether there are intervening factors such as the willingness to exchange personal information for objects of value. It is hoped that by providing a better understanding of how individuals perceive constructs such as online privacy and the risks of providing personal information online, whether actively or passively, this information maybe useful in assessing web privacy standards priorities as well as helping to communicate to people how these standards operate.

References

Kilger, M. and D. Jovanova. “Predictors of Personal Privacy Attitudes and the Consequences for Survey Researchers.” Paper presented at the 2011 American Association of Opinion Researchers, April, 2011, Phoenix.

Li, H., Sarathy, R. and Xu, H., 2010. “Understanding situational online information disclosure as a privacy calculus.” Journal of Computer Information Systems, volume 51 (1), pp. 62-71.

Margulis, S., 1977. “Conceptions of privacy: Current status and next steps.” Journal of Social Issues, volume 33(3), pp. 5-21.

Rosen, J., 2000. The Unwanted Gaze: The Destruction of Privacy in America. New York: Random House.

Smith, H., Dinev, T. and H. Xu, 2011. “Information privacy research: An interdisciplinary view.” MIS Quarterly,volume 35(4), pp. 989-1015.

Appendix A

All of the following measures utilize a five point Likert scale for response category:

Agree a lot / Agree a little / Neither Agree nor Disagree / Disagree a little / Disagree a lot

1. I sometimes use a search engine to find out what information about me might be online
2. I feel like I have a fair amount of control over the personal informationabout me tha can be found online
3. I often look up a company or organization online before I give them information about myself
4. I often read the privacy statements that companies have on their websites or in print
5. Most of the personal information about me that is online is relatively harmless
6. I don't mind companies using information about me to better understand products and services that I might want
7. I feel that I understand the risks of providing personal information online
8. I am willing to provide some personal information to a company in order to get something that I want
9. I would feel more comfortable providing personal information that display a trusted seal of approval
10. I don't mind companies sharing my product preferences as long as it's anonymous
11. I like knowing how companies are using information about me
12. I use the Internet less than before because of privacy concerns
13. Providing personal information offline is just as risky as providing it online
14. If there were a program to inform companies about my privacy preferences I would participate in that program
15. I want more personal control over information that companies might have about me
16. I know many people who have had something negative happen to them because of personal information available online
17. I trust the federal government to make the best decisions about how to protect my privacy
18. Once a piece of personal information becomes available online, there is nothing I can do about it
19. I have experienced a situation where online information about me has had negative consequences

October 28, 2012

To: Nick Doty, W3C
Jan Schallaböck, ICPP

Re: W3C Workshop Call for Papers – Do Not Track & Beyond

The Online Trust Alliance submits the following for consideration for participating at theW3C workshop being held on November 26-27, 2012

Background

The Online Trust Alliance's (OTA) mission is to enhance online trust, while promoting innovation and the vitality of the internet. OTA is a non-profit organization representing the broad ecosystem with supporters and sponsors based in the US, Canada, EU, Singapore, Latin America and Australia. As a "voice of reason" OTA's goal include:

- Developing a pragmatic and holistic view of the issues and tradeoffs incorporating the long-term aggregate view of impact to consumers, commerce and society.
- Help educate businesses, policy makers and stakeholders while developing and advancing best practices and tools to enhance the protection of users' security, privacy and identity.
- Supports collaborative public-private partnerships, benchmark reporting, meaningful self-regulation and data stewardship.

OTA has working on the concept of do not track since proposed by the Center for Democracy & Technology. Subsequently OTA has submitted papers and / or testimony to the FTC, Commerce Department, FCC, White House and ICANN. OTA has participated in briefings and listening sessions contributing to the White House Privacy Bill of Rights and Commerce Department Privacy white paper.

Review the progress self-regulatory efforts, OTA believes the status quo is no longer acceptable, and both sides must move toward the center. OTA believes we must appreciate the long-term impact of the absence of user notice and control of the collection, use, sharing and retention of their online data. At the same time there is a significant void in the appreciation of the value users receive from the data exchange which funds the content and services they consume.

Historically the interactive marketing and advertising industries have faced similar challenges. For example in the absence of regulation, the industry deployed technical counter measures which today are on-by default and embraced by nearly every browser and ISP. These include popup blockers and disabling of links and images from unknown senders in email to anti-phishing filters. While admittedly they impact legitimate advertising and marketing, can result in false positives and were disruptive to their operations and practices, industry has evolved and prospered.

Interest in the Workshop - Specific to the workshop, OTA proposes a review of the potential techniques and scenarios which may evolve from the deployment of DNT. We suggest a review of the browser adoption curve and review of potential scenarios that may result over time, from both the support and lack of support for DNT. A partial list could include default adoption of third party cookies blocking, dynamic block lists and integrated ad-blocking mechanisms. In addition we propose an examination of scenarios sites may deploy in response to DNT and the pros and cons from the user, publishers and advertisers perspective. These may include redirecting users to limited content sites, driving site visitors to subscription models, serving more ads on a page, to posting popup notices their site does not support DNT.

I look forward to continued participation with W3C and others to advance privacy enhancing technologies and practices while promoting innovation and the value users receive from advertising supported services,

Craig Spiezle
Executive Director & President
Online Trist Alliance
+1 425-455-7400

W3C “Do Not Track and Beyond” Workshop
Position paper from Aleecia M. McDonald

Dear Nick et. al.,

My interest in this workshop stems from co-chairing the Tracking Protection Working Group, which works on Do Not Track. I also research privacy at the Stanford Center for Information & Society as a Resident Fellow.

I do not wish to speak.

Thanks for your consideration.

Sincerely,

Aleecia M. McDonald

"Do not track" and beyond – Frank Wagner, Deutsche Telekom

"Do not track" is a developed Internet standard that enables users to indicate their own tracking[1] preferences to the websites they visit and the apps they use. Based on the European legal framework, users' opportunities for indicating their preferences are simplified.

Provided that the selection options are designed appropriately in the user agent, users can be "forced" to make a clear, intentional decision for or against tracking by prompting them to make a decision the first time the user agent is called.

This "decision" is made in the configuration of the user agent. The user only has to learn once where this setting is located and what it does. When opt-out scenarios are used, in contrast, users have to find out where the opt-out function is located, how to activate it, and how comprehensive it is anew for each individual website. Moreover, the fact that users are often forwarded to other websites to activate the opt-out function does not exactly give the impression that their interests are being followed.

In contrast to opt-out scenarios, where the website operators are clearly interested in having as few opt-outs as possible, the situation for opt-in scenarios is quite different. In this case, the website operators must assume major interest in opting in by the users. The corresponding mechanisms are positioned prominently; options for users to cancel a granted opt-in much less so.

In this context, "do not track" gives users much simpler options for indicating their tracking preferences. With "do not track", users no longer have to hunt through websites and apps to find the desired settings. As such, it would be consistent to make generic information about visited websites and used apps available to users at a centralized point. In this case, users would only have to learn once where to find this generic information for the respective website or app, similar to the do-not-track concept.

A minimal amount of generic information would be suitable for this purpose, but additional information can also be provided optionally. Options for extending this provided information for specific websites or apps should be provided.

Experiences from P3P should be utilized to identify the relevant information.

This implementation method would not only make it possible to address scenarios in conventional Internet portals; it would also be conceivable to use generic information to support decision-making in the business domain. This could involve cloud computing scenarios, however, in which personal data is processed. For SaaS (software as a service) offerings in public clouds, specific standardized criteria could be used to support decision-making. In this context, relevant factors for potential customers include where the data is saved, where it can be accessed, which sub-providers are involved in the production chain, which contractual foundation was selected between the contracting parties and which security levels are available. The possibility of positioning such information in a standardized place would be a major step toward improving the comparability of cloud services.

[1] Using "tracking" as a generally valid term in this paper requires that the term be defined and specified clearly and made transparent for Internet users. If it is not, it will not be possible to use a do-not-track function adequately.

We have to assume that the ability to compare similar offers would not only be relevant in the cloud computing domain, but could be transported to many other areas as well.

Therefore, now that "do not track" has created a standard for configuring user preferences, one objective of a future standard should be to improve transparency. The experiences from P3P should be taken into account accordingly.

Comcast supports the development of Internet standards that promote transparency about the collection and use of user activity data and that permit web publishers and others in the Internet ecosystem to offer consumers choice about how their data is used.

Comcast has contributed technical resources, attention and analysis to the development of Do Not Track tools from very early in the process of developing these tools. For example, Comcast was the first company to work with TrustE to implement an online icon on a trial basis, to see whether icons would be a viable means for consumers to elect not to participate in certain uses of their online activity data. In the case of this trial, we were asked by TRUSTe to offer users an opt-out from behavioral advertising, even though there was no behavioral advertising on ComcastÆs participating website at the time of the trial. This trial took place before the advertising industry offered its own icon program through the Digital Advertising Alliance.

Comcast has been a participant in the W3CÆs Tracking Protection Working Group, participating in the development of a set of widely implementable technical standards for offering consumers choice about the use of their data. As a large distributor of online content, we believe we have the scale and technical expertise to represent web publishers and content distributors who have a key stake in the outcome of this process and to evaluate the practicality of implementing proposed standards.

We participate in the W3CÆs Tracking Protection Working Group because we believe that this standards work can be very valuable and can lead to greater awareness, choice, and control for users and greater innovation and opportunity for online publishers, distributors, and others in the Internet ecosystem. However, the W3C process may not be the best forum for addressing the many complex technical, policy, and philosophical perspectives all at once. The result of trying to accommodate all of these very diverse perspectives may unfortunately be to distort the process and make it difficult to create technical standards that make sense and will be widely adopted.

We would support privacy work by the W3C that strives for progress on pieces of the overall privacy environment, leading to meaningful incremental changes that can be evaluated and managed as pieces of a larger privacy environment are put together. By addressing several small components of user privacy separately, working groups are more likely to develop useful technical standards and tools that together can create a privacy toolbox for all members of the Internet ecosystem to use constructively.

Some areas worthy of consideration in the future include: methods for communicating data collection and retention practices separately from ôdo not trackö options, consideration of standards for offering users a range of choices about privacy and data use (rather than making ôdo not trackö a proxy for all user choices), offering those user preferences outside of an immediate transaction or installation, and considering how these preferences can be offered consistently by many different kinds of systems and interfaces. Comcast would welcome the opportunity to untangle some of these issues from Do Not Track and participate in the development of open and consistent voluntary standards that could be widely implemented and effective.

Do Not Track and Beyond Workshop Position Paper
John M. Simpson
Privacy Project Director, Consumer Watchdog
Oct. 31, 2012

Consumer Watchdog's Interest In Do Not Track

Consumer Watchdog is known for its success in media advocacy, where our nonprofit, nonpartisan public interest group focuses attention on vital issues of public interest, catalyzing opinion leaders and policymakers. For instance our Privacy Project has helped frame the privacy debate in the media, in Washington, DC, and in the technology industry's hub of California.

Consumer Watchdog was an early advocate of Do Not Track and with good reason. A poll conducted for us in the summer of 2010 by Grove Insight found 80 percent of Americans supported a Do Not Track option managed by the Federal Trade Commission. Eighty-four percent favored preventing online companies from tracking without the user's explicit written consent. Ninety percent supported more laws to protect your personal information. (http://insidegoogle.com/wp-content/uploads/2010/07/MemInternetPrivacy-0727101.pdf)

Consumer Watchdog endorsed U.S. Rep. Jackie Speier's "Do Not Track Me On Line Act", HR 654, and sponsored California Sen. Alan Lowenthal's Do Not Track bill, SB 761. When it became clear that the World Wide Web Consortium's Tracking Protection Working Group offered a genuine possibility of developing a meaningful standard for Do Not Track, specifying how the user's preference would be communicated and what a site's compliance obligations would be, I eagerly joined as an invited expert. Since joining the Working Group a year ago, I believe I have been an active and constructive participant in what I continue to hope will develop a meaningful DNT specification that will benefit both consumers and business.

Goals and Scope of Workshop

When the First Public Working Drafts of the Tracking Preference Expression specification and the Compliance and Scope specification were released, I was impressed with the extent to which the documents relied upon user expectations to develop the language. As the documents have gone through various iterations over the last year, I fear they have too frequently moved away from that guiding principle. For example, what the standard would now consider allowable data sharing amongst affiliates, exceeds what most consumers would expect. Why would you ever expect See's Candy and GEICO insurance to be the same party through corporate affiliation?

Combined with what I perceive to be an ever expanding list of "permitted uses" I think the Working Group is in danger of producing a standard that has no relationship whatsoever to the plain English meaning of Do Not Track. Here is an analogy: Suppose I discover that my neighbor has decided to use a video camera to monitor my bedroom. I find this out because he sends me copies of the video he has taken. I am outraged and ask him to stop tracking my

activities in my bedroom. He agrees, says he has received my do not track message and stops sending me the videos. Nothing else changes.

Exaggerated, perhaps, but I'm trying to make the point that most people believe Do Not Track means exactly that -- do not collect information about my Web browsing activities. Or, in my example analogy, don't record the videos. The gap between what ordinary users will understand DNT to mean and what the standard seems likely to require appears large enough to undermine consumers' trust in the Web. I believe this is a very real problem that must be addressed. At the very least there would need to be considerable education about what the W3C standard means, why it differs from user expectation and how it enhances user privacy. Who will do that? As the W3C DNT standard is emerging, I am hard pressed to understand how DNT gives a user any more protection than blocking third-party cookies and clearing cookies after each browser session.

Another issue that concerns me is the extent to which the Do Not Track issue has been conflated with Online Behavioral Advertising. It doesn't really matter that much to me that I receive ads that are thought to be of interest to me. Indeed, if I knew I was in the market for a particular item, I might well willingly indicate that I was shopping for it so I would see relevant ads. What is troubling are the digital dossiers that are collected and indefinitely maintained about what sites I've visited. Is there a way to determine broad categories into which I might fall, use those for ad targeting, but forgo the digital dossier replete with all the URLs of every site I've ever visited?

Finally, another topic to explore are the privacy concerns that are unique to the mobile market. Admittedly, I have much to learn about this space and its business practices, but I fear the mobile ecosystem is rapidly becoming the Wild West of the Internet.

####

W3C Workshop: Do Not Track and Beyond
26-27 November 2012

I would like to attend the W3C Workshop on "Do Not Track and Beyond" on November 26-27, 2012. We are interested in following the discussions and hearing what future issues might be addressed by the W3C or other stakeholders. Below, I have included information about the company and also about the Amazon Web Services cloud computing business.

Thank you.

Brian Huseman
Director, Public Policy
bhuseman@amazon.com

About Amazon.com

Amazon.com, Inc., a Fortune 500 company based in Seattle, opened on the World Wide Web in July 1995 and today offers Earth's Biggest Selection. Amazon.com, Inc. seeks to be Earth's most customer-centric company, where customers can find and discover anything they might want to buy online, and endeavors to offer its customers the lowest possible prices. Amazon.com and other sellers offer millions of unique new, refurbished and used items in categories such as Books; Movies, Music & Games; Digital Downloads; Electronics & Computers; Home & Garden; Toys, Kids & Baby; Grocery; Apparel, Shoes & Jewelry; Health & Beauty; Sports & Outdoors; and Tools, Auto & Industrial.

About Amazon Web Services

Launched in 2006, Amazon Web Services, Inc. began exposing key infrastructure services to businesses in the form of web services -- now widely known as cloud computing. The ultimate benefit of cloud computing, and AWS, is the ability to leverage a new business model and turn capital infrastructure expenses into variable costs. Businesses no longer need to plan and procure servers and other IT resources weeks or months in advance. Using AWS, businesses can take advantage of Amazon's expertise and economies of scale to access resources when their business needs them, delivering results faster and at a lower cost. Today, Amazon Web Services is a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of enterprise, government and startup customers businesses in 190 countries around the world. AWS comprises of over 28 different services, including Amazon Elastic Compute Cloud (Amazon EC2), Amazon Simple Storage Service (Amazon S3) and Amazon Relational Database Service (Amazon RDS). AWS services are available to customers from data center locations in the U.S., Brazil, Europe, Japan, Singapore and Australia.

Machine Interpretable Expression of Compliance

Dave Raggett <dsr at w3 dot org>, W3C,
Rigo Wenning <rigo at w3 dot org>, W3C

This work contains ideas that were conducted as part of the PrimeLife project with funding from the European Union's 7th Framework Programme. The work reported is experimental and the examples shown are ficticious, and taken from a working demonstator. In a slightly different form, this work was already submitted to an earlier W3C Workshop on Privacy and data usage control

Introduction

W3C began as early as 1996 to think about technologic remedies to the privacy issues created by the Web. It started with PICS for privacy. The W3C Platform for Privacy Preferences (P3P) 1.0 was published as a W3C Recommendation in July 2002 [1]. It defines a machine interpretable format for websites to express their privacy practices. A revised format (P3P 1.1) was published as a W3C Note in November 2006, but failed to reach Recommendation status [2]. Since 2004, W3C has been involved with Privacy Research Projects in the European Commission's 7th Framework programme.[3][4] And now W3C is working on defining a Tracking Preference Expression Specification TPE[5] and a Tracking Compliance and Scope Specification (TCS)[6].

During the work on the TPE and the TCS the question came up what it means to be compliant and how to express compliance regimes. Having tokens representing complex concepts in long human readable documents is not scalable. The Privacy Policies on Web sites and their pages of legal language did not help to improve the feeling of the users of the Web.

But how can a Service actually describe why a certain data item is needed, why it is not harmful to send it, what they have done to take the privacy worries of end users into account, convince them to use the service. Privacy is more and more in the center of marketing strategy as not addressing the topic will keep users away from services. Cloud computing, where nobody really knows where the data actually is, where the complexity of the system is leading to lack of understanding that in turn nurtures doubts about misuse of collected data. The air of transparency is the best remedy to counter fear, uncertainty and doubt about the unknown. This was the use case for P3P. This use case is higher on the agenda than ever.

To counter fear of the unknown, in summary, P3P described the business name and address responsible for the website, the dispute resolution procedures, the means (if any) for users to access personal data collected by the website, the kinds of data collected, the purposes it will be used for, the data retention policy, and the recipients of the data. It lacked a way to give a short notice to the user on how the different facts expressed in a machine readable form are relating to each other. Software was supposed to explain this and that failed. Thus the need for a new approach.

P3P supports a notice and consent model of privacy, where websites describe their privacy policies and users can review the policy and decide whether to walk away or to proceed to interact with the site, and by so doing indicate their consent to that policy.

Rather than expecting users to review the privacy policy for each website that they visit, a P3P enabled web browser performs an automatic comparison of the user's recorded preferences with the website's policy, and only alerts the user if there is a mismatch.

With the work on DNT and its exception mechanism, this has taken a new turn. The Tracking Preference Expression Specification will contain an API to ask the user for his permission to personalize content and thus collect personal data. In Europe, the service needs some kind of consent to continue storing information client side. But how would one convince the user to accept the data collection, how do we reassure the user? Again, 22 pages of legalese haven't done the trick in the past and they won't do the trick here. The P3P statement vocabulary contains a lot of the semantics actually needed to have an internationalized interface to tell people what is collected and what the service intends to do with the collected data. This can be taken as a basis and extended to cater for the new needs. This paper tries to go first steps in that direction.

The PrimeLife Dashboard

With increasing public awareness of the amount of information being collected by websites, it seems timely to consider new approaches covering more than just cookies, whilst enabling a practical treatment of the user interface for expressing privacy preferences.

To investigate this, a Firefox extension was developed to look at the issues involved. This had to support:

a. auto-generation of a human readable version of the policy
b. automatic comparison of the user preferences with the policy
c. automatic generation of a human readable report on any mismatches
d. user interface for viewing and changing user preferences

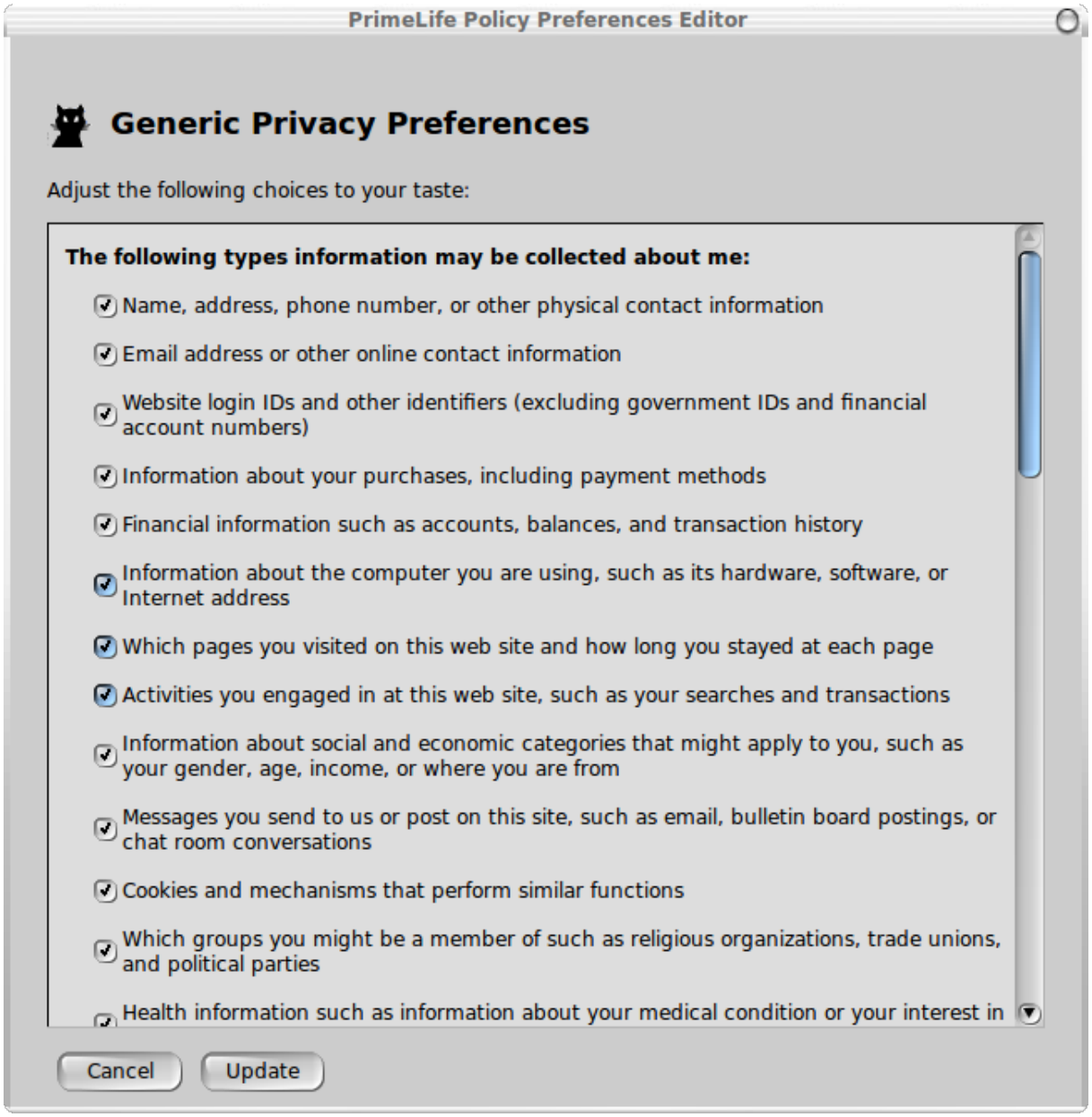
The scope was taken as the data that websites can collect from HTTP request headers during a session. This includes the IP address, cookies, the user agent header, information on user preferences for language and data formats, the requested URL, the date and time of day, and more.

To simplify the user interface for preferences, a subset of P3P was chosen. This has the following object model:

- The URI for the site's full (human readable) policy
- The URI for instructions that users can follow to request or decline to have their data used for a particular purpose (optional)
- The name of the business responsible for the website
- The set of categories of collected data as defined by P3P 1.1
- The set of purposes collected data can be used for as defined by P3P 1.1
- The set of recipient types as defined by P3P 1.1
- The data retention policy type as defined by P3P 1.1

Note this uses P3P's data categories rather than the taxonomy of data items. This was found to be a much better fit to the needs for describing the kinds of data collected from HTTP requests.

The simple object model allows the preferences user interface to be provided as a set of grouped checkboxes, as shown below:



Accessing the policy and generating a human readable version

To reach a website, the user can type in a URL, follow a bookmark, or follow a link from another site, e.g. on the results page from query on a search engine like Google. The browser extension intercepts the Firefox location change event and cancels the HTTP request before it is sent. The extension then sends an HTTP HEAD request to the website's root. The response is examined to find a refererence to the site's generic privacy policy. This is represented as an HTTP Link header (analogous to the HTML link element), e.g.

```
Link: <http://localhost/w3c/policy.json>;
   rel="http://primelife.eu/generic-privacy-policy"
```

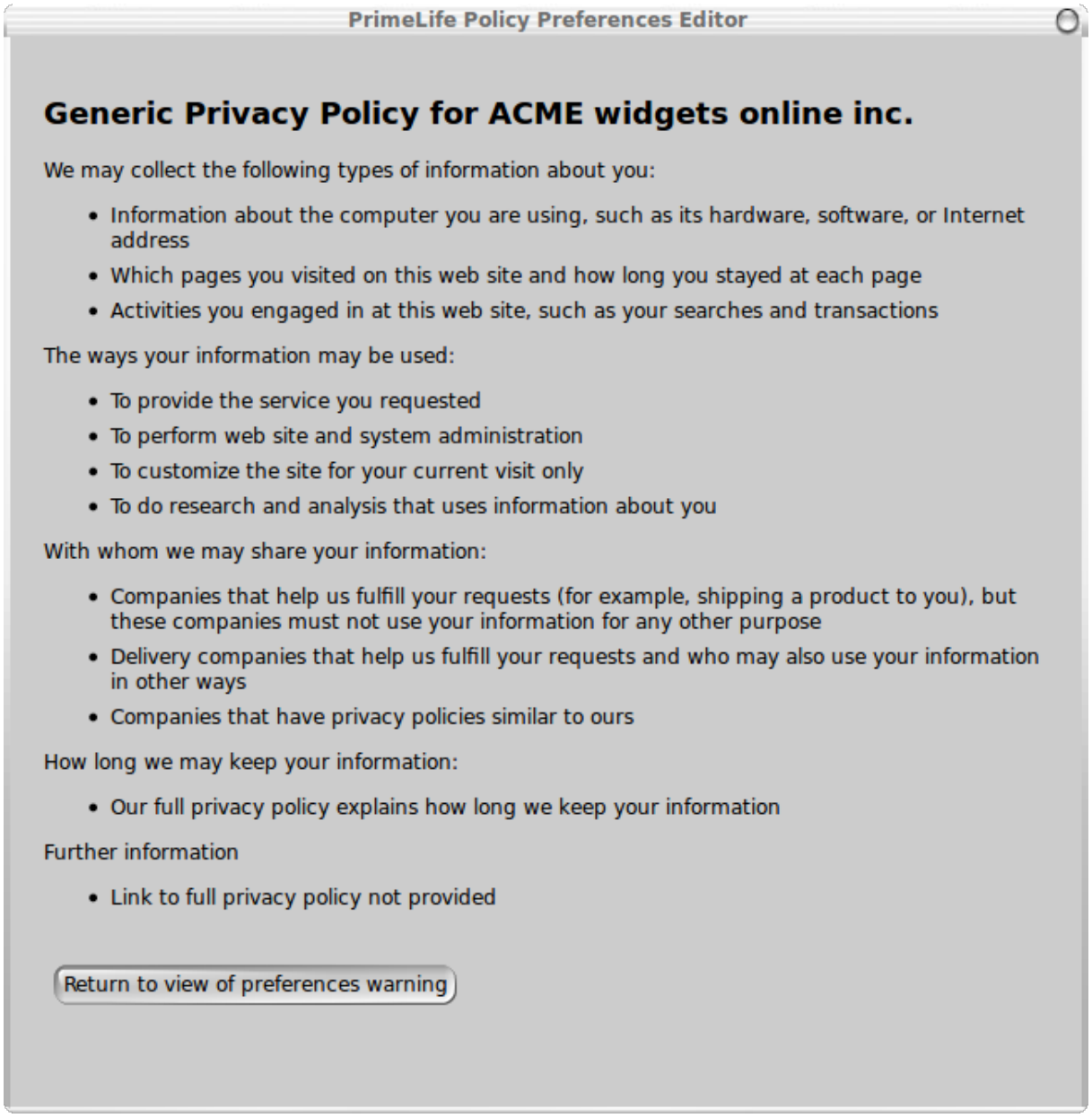
This header is easy to add to pages generated via PHP. The URI for the policy is then dereferenced to obtain the policy itself. Note P3P 1.0 defined a P3P HTTP header rather than using the generic Link header. This is something that could be considered if and when this work is brought into the standards track.

The object model for policies is decoupled from the on-the-wire transfer format, but from a practical point of view it was easiest to implement the transfer format with JSON [3]. Here is an example policy in JSON:

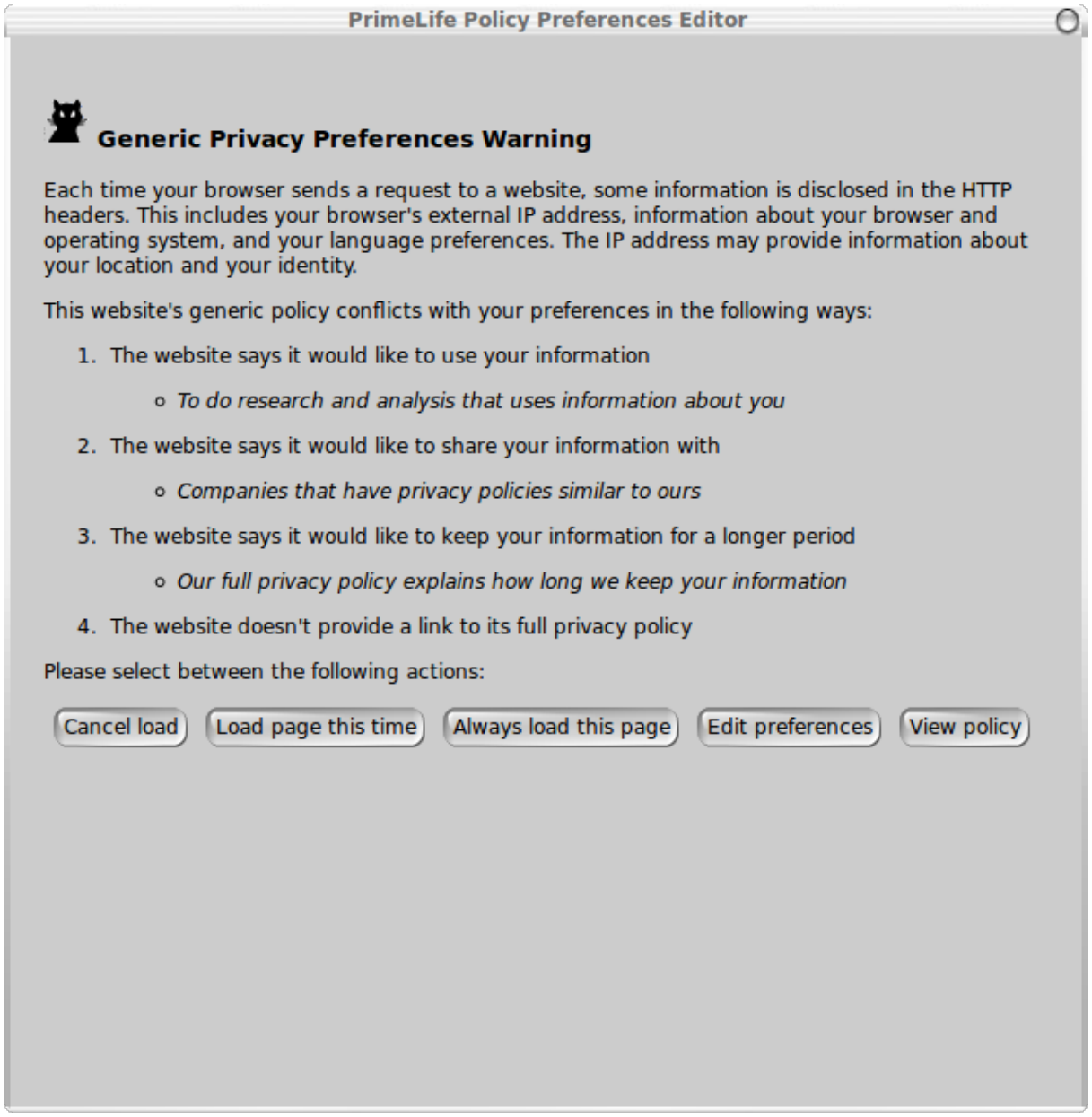
```
{
    "fullURI": null,
    "optURI": null,
    "name": "ACME widgets online inc.",
    "purposes": ["current", "admin", "tailoring", "individual-analysis" ],
    "recipients": [ "ours", "delivery", "same" ],
    "retention": "business-practices",
    "categories": [ "computer", "navigation", "interactive" ]
}
```

Generating a human readable version of the privacy policy

The P3P 1.1 specification includes suggested text for each element in the taxonomy. This was copied into JavaScript and used to generate a human readable version of the policy. Here is an example:

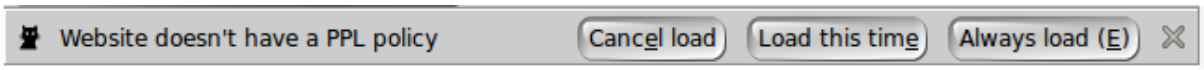


The same text was also used for constructing a dialog summarising the mismatch between the user's preferences and the website's policy, for example:

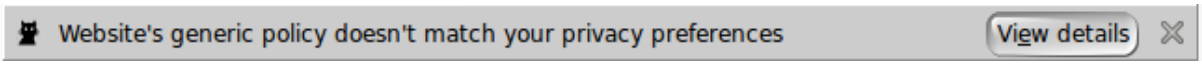


If the site's policy matched the user's preferences, or the user decided to override the mismatch, the browser extension then proceeds to relaunch the HTTP request for the original URL.

The Firefox notification bar is shown when a site is found to lack a privacy policy.



The Firefox notification bar is shown when a mismatch is found.



Clicking "View details" brings up the warning dialog shown earlier.

A local SQLite database was used to capture the user's preferences, and to cache the policy for sites as a performance optimization.

Anonymising Proxies

The act of making an HTTP HEAD request on a website's root discloses the browser's external IP address. This can be avoided by routing the request through an HTTP proxy. This could be configured via a user preference.

Summary and suggestions for further work

This paper has described a fresh take on P3P that goes beyond the limitations of compact policies, whilst still enabling a simple user interface for setting preferences. The object model lends itself to the use of JSON as a policy transfer format. The restricted semantics for a machine readable policy covering data collected in HTTP requests, is supplemented by a link to the site's full human readable policy. The proposal starts to think about how to integrate the ontology created by P3P statement vocabulary into the HTML5 and javascript interactions to allow for higher transparency and suggests further work in that area.

A further consideration is the privacy policy for other kinds of personal information collected by websites, for example, credentials coupled to a user's public or partial identity. Can the P3P taxonomies be extended to support these?

P3P and the approach described in this paper are couched in legal terms relevant to the obligations extended by websites to their users. Websites also have the challenge of operationalizing privacy policies when it comes to controlling access and usages of personal data in the website's backend. This suggests the need for transforming privacy policies into data handling policies. The PrimeLife project is looking at extending the XACML access control language to cover data handling policies, see H5.3.2 [4].

Widespread support for machine readable privacy policies is likely to involve a legislative mandate with measures in place to ensure that sites conform to the policies they disclose. However, this would only apply to the countries with the corresponding laws. A way is needed to allow the browser to verify the jurisdiction a given website is subject to. This could take the form of digital certificates issued by national agencies.

A separate issue is many people aren't sufficiently motivated to set privacy preferences. One reason is the desire to just get to the website in question without having to bother with reviewing the policy. Another is a lack of knowledge sufficient for an informed decision. This points the way to the use of independent third parties for help with setting privacy preferences, and for monitoring the data handling practices of websites. Some progress has been made with the latter in terms of a browser extension (Privacy Dashboard) that tracks what information is collected by the websites you visit, together with a means to set your preferences on a site by site basis [5].

Further reading

[1] http://www.w3.org/TR/2002/REC-P3P-20020416/
[2] http://www.w3.org/TR/2006/NOTE-P3P11-20061113/
[3] http://www.json.org/
[4] http://www.primelife.eu/results/documents/activity-5-policies
[5] http://www.primelife.eu/results/opensource/76-dashboard

Standardization for Privacy Management

Position paper for the W3C "Do Not Track and Beyond" Workshop, November 2012

Mark Frigon, Arnaud Le Hors – IBM Corporation
October 22nd, 2012

Today, policymakers, businesses and society grapple with privacy issues at a time when advances in technology help us personalize service and solve complex problems through analysis of user data. Privacy discussions began with the emergence of the Internet. As the Internet has grown so has the gathering and use of varying amounts of information including at times sensitive information about individuals. Good data stewardship by businesses and governments can help address some of the privacy concerns , but is this enough?

Too often industry practices and/or technology solutions tend to be ones that are the least disruptive for the particular industry from which the proposal to protect consumer privacy emanates:

- Browser-based initiatives such as "Do Not Track" seek to preserve the browser capabilities, such as javascript and cookies, and thus recommend pushing enforcement of decisions around privacy from the browser-side and on to the server-side.
- The advertising network-based proposals, such as AdChoices, appear to ignore a more consumer-friendly universal browser-based "opt-in" approach, in favor of one that would require users to explicitly "opt-out".

While some of these technologies have gained considerable traction, most approach the problem with differing, and at times, conflicting implementations:

- Advertising-based initiatives rely on a browser cookie to store any consumer's "opt-outs". However, the EU Directive suggests that no such cookies may be placed on a consumer's browser without their explicit consent, rendering any consumer opt-out preferences moot.

- Browser-based “Do Not Track” approach allows consumers to request that websites do not track them. However, such approach relies on websites to honor the browser request and in no way prevents websites from actually storing sensitive data.

Counterproductively, all the differing approaches have only created more confusion for consumers as to how and what personal data is being collected and how to sufficiently manage one’s privacy. In addition, companies face uncertainty around which approaches will ultimately be adopted or codified. Accordingly website operators face a choice between the expense of implementing several approaches or, conversely, choosing complacency.

IBM’s Customer Experience Digital Data Acquisition proposal to the W3C offers a framework for a more standardized and flexible approach to consumer privacy management and enforcement. The proposed “Registrant” object in the Digital Data Acquisition proposal offers a framework from which to standardize privacy efforts. The Registration object offers a unique identifier field (“digitaldata.registrant.Registrantid”) which website operators can use to store requisite visitor identification. Additionally, other forms for sensitive data could be stored in the digitaldata.registrant.attributes array. This array provides an extensible container to store and access additional attributes that may tie to the user: user preferences, permission, membership levels, etc.

Additionally a common problem exists when a website operator might not be fully aware of all the cookies being placed on its own website due to 3rd party javascript “includes” for advertising, social media plug-ins, analytics, and other capabilities. Standardization around a common data model and common consumer object offers website operators the possibility of having more control around which information its vendors collect from its own customer-base. For example, an advertising network might “look” for a consumer’s social network IDs to provide more targeted advertising without the knowledge of the website operator. Such technical possibilities make it very difficult for a website operator to manage its own privacy policy.

In order to best balance the interest of all parties involved (consumers, website operators, browser manufactures, advertisers, policymakers, etc) we believe a standard backed by the

W3C to be the best way to gain adoption of a common privacy standard. Because of this we have submitted the Customer Experience Digital Data Acquisition specification to W3C, and are eager to work with workshop participants to explore the following questions:

1. Can browsers facilitate user management of the registration data while allowing websites to request access to potentially sensitive information?

2. What controls can website operators be given to enforce 3rd party "includes" collect and store data consistent with the website operator's privacy policy?

3. How might a standard help manage user data elements that may be defined as personally identifiable under different conditions?

4. What distinctions, if any, should be made between 1st party and 3rd party data access? Should the consumer or the website operator determine access based on a privacy policy?

We are committed to working through these questions, discussing feedback, and finding possible resolutions and solutions to provide consumers with more transparency, website providers with more control, and policymakers with an example of a self-regulatory approach. In addition to discussion during this workshop, we encourage all interested parties to join in the review and further development of the Customer Experience Digital Data Acquisition specification within a new Community Group. Our goal is to eventually have the specification "graduate" from the Community Group and move onto the Recommendation track as the basis for a new Working Group.

References

Customer Experience Digital Data Acquisition submission
http://w3.org/Submission/2012/04/

Tracking Protection Working Group
http://www.w3.org/2011/tracking-protection/

AdChoices
http://www.youradchoices.com/

The Paradox of Privacy Empowerment: The Unintended Consequences of "Do Not Track"

Position paper for W3C Workshop: Do Not Track and Beyond
Berkeley, California, November 26-27, 2012

Berin Szoka[1]

The debate over "Do Not Track" offers an excellent microcosm for understanding the larger privacy policy discourse. Arguments for giving users a tool to express their privacy preferences exert enormous rhetorical appeal. Those arguing for versions of DNT that are more restrictive of the collection and use of information about user behavior essentially insist that "We're merely giving users a choice!" Who could possibly be against letting users choose for themselves? Why should anyone else get to choose *for us*—especially companies that seem to be profiting from the ignorance or helplessness of users?

Tools like "Do Not Track" (and "privacy-friendly" interfaces more generally) are usually justified as simply offering users a means of expressing their true preferences. But such choice architectures[2] are anything but neutral: even with the best of intentions and in the name of facilitating user choice, choice architects will produce outcomes that users would not have chosen if they could make fully rational decisions in a frictionless world without transactions costs. This is the essential paradox of user empowerment.

"Privacy advocates" regularly cite opinion polls showing that users demand greater privacy protection—and thus conclude that privacy-friendly choice architectures simply facilitate the true preferences of users. But listening to what consumers *say* they want tells us much less about their preferences than seeing what preferences they *reveal* in the process of making real-world decisions about trade-offs among values. As much as users value privacy, they do not value privacy in isolation or inherently, but relative to other values—including other forms of privacy.

To avoid the paradox of user empowerment to the greatest extent possible, choice architects must understand how their proposed choice architecture will shape real-world outcomes, and the impact that will have on these many competing values. Let us consider the unintended consequences of three contested aspects of DNT:

---

[1] This position paper draws testimony I gave to the Senate Commerce Committee in June 2012, http://techfreedom.org/node/185

[2] On term "choice architecture" and its inherent non-neutrality, *see generally* Richard H. Thaler University of Chicago, Cass R. Sunstein & John P. Balz, Choice Architecture, April 2010, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1583509.

1. **Default setting** - How, and by whom, may a browser be set to send DNT:1?
2. **Definition of tracking** - What is it DNT:1 tells servers not to do?
3. **Architecture of negotiation** - How do sites get users who send DNT:1 headers to opt-in to tracking—and to remain opted-in?

Each is a complicated issue. But all three may be understood, to a degree, in terms of the traditional opt-in and opt-out paradigms. DNT:1 is nothing more than a signal sent by the user's browser expressing a preference not to be "tracked," however defined—after which website publishers, advertisers and other data collectors must somehow negotiate with the user to get him or her to "opt back in" (a term actually used in the TPE[3]) to "tracking" (by granting a site or network a "user-granted exception"). If browsers and other user agents may turn on DNT:1 by default, then the adoption rate of DNT will quickly exceed publishers' "maximum acceptable loss threshold." Below that point it makes little practical sense for publishers and advertisers to bother building an architecture of negotiation, because it is more cost-effective to let DNT:1 users free-ride off those allow tracking (either by setting DNT:0 or by not having it set at all).

Put more simply, if browsers are allowed to turn DNT:1 on by default, most users will live in a world where "tracking" is opt-in. This will be a choice made *for*, not *by*, users. But either way, all of the problems of more general "Opt-In Dystopias" described by Nicklas Lundblad and Betsy Masiello would apply once DNT:1 is turned on. They distill their concerns into four categories:

> **Dual cost structure**: Opt-in is necessarily a partially informed decision because users lack experience with the service and value it provides until after opting-in. Potential costs of the opt-in decision loom larger than potential benefits, whereas potential benefits of the opt-out decision loom larger than potential costs.
>
> **Excessive scope**: Under an opt-in regime, the provider has an incentive to exaggerate the scope of what he asks for, while under the opt-out regime the provider has an incentive to allow for feature-by-feature opt-out.
>
> **Desensitisation**: If everyone requires opt-in to use services, users will be desensitised to the choice, resulting in automatic opt-in.
>
> **Balkanisation**: The increase in switching costs presented by opt-in decisions is likely to lead to proliferation of walled gardens.[4]

The problem is that DNT, like any choice architecture, affects not only "demand" (empowering users to choose) but also the "supply" (the choices available to users). The difficulty of obtaining opt-ins (user-granted exceptions) will serve as a barrier to entry, protecting larger, established incumbents against competition from new entrants. This will be true on some level for individual sites: absent dual-cost structure problem, one might think that any site a user

---

[3] http://www.w3.org/TR/tracking-dnt/#exceptions-principles

[4] N Lundblad and B Masiello, "Opt-in Dystopias", (2010) 7:1 SCRIPTed 155, http://www.law.ed.ac.uk/ahrc/script-ed/vol7-1/lundblad.asp

visits will easily be able to get an opt-in. But obtaining such opt-ins is costly, both for user and for sites, which must implement a mechanism for obtaining user-granted exceptions. Some sites will simply decide not to risk alienating users, and forego potential additional revenue, while other better established sites or sites less subject to competition, will gain a competitive advantage.

But the greater problem lies with web-wide exceptions, opt-ins to data collection by an ad network or other data collector across the web. To be sure, these are essential to making DNT work without breaking business models that depend on third-party ad networks, but they will also necessarily favor certain established players in the data and advertising ecosystem over other, generally smaller players. One might dismiss these competitive effects as the necessary consequence of restructuring an industry that is loathed by many (despite the benefits it confers),[5] but this consolidation would likely be accompanied by a qualitative change in the *kind* of information collected. Once a network obtains a web-wide exception, why *not* collect more data across the web? Why not associate it in a richer profile? As Masiello and Lundblad explain:

> service providers may attempt to maximise data collection in every instance that they are forced to use an opt-in framework; once a user consents to data collection, why not collect as much as possible? And the increased transaction costs associated with opt-in will lead service providers to minimise the number of times they request opt-in consent. In combination these two behaviours are likely to lead to an excessive scope for opt-in agreements. In turn, users will face more complex decisions as they decide whether or not to participate.[6]

Indeed, why not require users to log-in and provide more information about their real identity? Of course, requiring users to go through an account-creation process would likely turn off many users—if only because it took longer than simply clicking on a dialog box that asked about enabling personalized content. But consumers have become quite accustomed to using Single Sign On systems to log into websites with their Facebook, Twitter, Google or Microsoft Live accounts (and so on). It is not difficult to see such networks becoming federated content networks—the new walled gardens so feared by Tim Wu, Jonathan Zittrain and many others. Leaving a website inside one network and going to the other would require granting another web-wide exception to another network. This isn't necessarily bad but if it ultimately means that *more* information is collected about Internet users, DNT will leave many of its advocates sorely disappointed—and it is certainly not a result any user would have chosen.

This perverse potential (but likely) result simply one example of a larger problem: human rationality is bounded; we are simply not capable of weighing the full implications of choices as complicated as those over privacy. This does not mean that user empowerment is not a

---

[5] See generally, Comments of Berin Szoka, *Privacy Trade-Offs: How Further Regulation Could Diminish Consumer Choice, Raise Prices, Quash Digital Innovation & Curtail Free Speech*, Dec. 7. 2009 http://ftc.gov/os/comments/privacyroundtable/544506-00035.pdf

[6] *Opt-in Dystopias*.

worthy goal; it is (and it is generally preferable to more top-down alternatives such as regulatory prescriptions on the use of data). But it *does* mean we should not pretend that choice architects are not, in fact, making important choices for users in the process of designing choice mechanisms like Do Not Track.

The problems described above will become more acute the more broadly "tracking" is defined, the more users turn on DNT:1, and the more cumbersome negotiation is. Two particular contested issues within the TPWG will significantly aggravate the opt-in dystopias problem:

1. **Default Settings** - Although the TPWG has always rested on the consensus that DNT headers must be set by users not user agents like browsers,[7] Microsoft breached that consensus earlier this year when it announced earlier this year that it would choose *for* users by setting DNT:1 on by default in its new IE10 browser. European regulators have essentially endorsed this position, calling for users to "told about any default setting; and prompted to keep or to change it"—even if that setting is DNT:1, and therefore not compliant with the DNT spec—and insisting that servers must not disregard DNT headers, even when sent by browsers that turn on DNT:1 by default.[8] It remains unclear how this issue will be resolved.
2. **Configuration** - The TPWG co-chairs recently rejected a proposal to clarify that, to "reflect the user's preference," user agents must "require equal effort to configure [DNT]"[9]—prompting the first formal objection filed in the TPWG.[10] Thus, unless this decision is ultimately reversed by the W3C, a user agent need not set DNT:1 by default if doing so proved problematic; it need only design a user interface that will achieve the same result.

Ultimately these concerns are likely to be dismissed by insistence that sites and services will simply negotiate around DNT to reach the same outcome they would have reached anyway. But in the real world (as opposed to a frictionless perfect market), transactions costs often swamp the gains created by transactions such as the negotiation between site and user. The online advertising ecosystem currently works because it generated tiny amounts of value from enormous volumes of transactions. Even the small transactions costs of forcing today's implicit quid pro quo to become explicit could produce dramatically different outcomes. Nor is it clear that negotiation or payments would generate as much revenue as advertising—meaning that rising transactions costs would be borne by publishers, and passed on to users in the form of reduced quality, quantity or innovation, or higher prices (if they can actually charge prices).

---

[7] "The goal of this protocol is to allow a user to express their personal preference regarding tracking to each server and web application that they communicate with... Key to that notion of expression is that it MUST reflect the user's preference, not the choice of some vendor, institution, or network-imposed mechanism outside the user's control." TPE § 3.

[8] Neelie Kroes, An update on Do Not Track The Centre for European Policy Studies (CEPS)/Brussels, 11 October 2012, http://europa.eu/rapid/press-release_SPEECH-12-716_en.htm

[9] http://lists.w3.org/Archives/Public/public-tracking/2012Sep/0197.html

[10] http://lists.w3.org/Archives/Public/public-tracking/2012Oct/0104.html

Building on Ronald Coase's seminal work on the importance of transactions costs, Harold Demsetz offered the basic insight that continues to guide the law and economics of setting defaults (which economists generally refer to as "property rights"): in a frictionless world, if the initial assignment of rights is inefficient, negotiation will inevitably and costlessly solve the problem; but in the real world, that initial assignment may prove sticky, thus we should not assign rights in ways that are inefficient.[11] Once again, choice mechanisms are not neutral. If, the day before Microsoft announced their decision to set DNT:1 by default, it was true that "majority default DNT is not the world this standard will exist in. DNT is going to be a 10% solution,"[12] and DNT:1 creates the negative unintended consequences described above (among others), why should choice architects not set the initial assignment to the setting that is more likely to be efficient: DNT:1 *off* by default and not privileged when users configure their browser? An argument could be made to the contrary if it could be shown that "tracking" (as defined by the DNT spec) actually lead to real harm, but as yet, no such argument has been substantiated, and the question of harm has repeatedly been sidestepped within the TPWG.

It is understandable, if ironic, that privacy advocates should desire outcomes that could actually reduce privacy and make consumers worse off—because the chain of causation is attenuated and unclear compared to the noble intentions behind restrictive defaults. Nobody wins Nobel Prizes in Economics for explaining things that are completely obvious, and even once they do, it can take decades (or more) for their insights to permeate areas of discourse outside of economics—such as Internet standard-setting.

It is much more understandable what some market players have to gain by joining forces with well-intentioned but short-sighted privacy advocates: competitive advantage. This is simply another example of the well documented alliance of "bootleggers and baptists."[13] Microsoft, in particular, stands to lose little by disrupting the online advertising market, in which it has struggled to compete. It is by no means clear whether a world of high DNT adoption rates would benefit, in relative terms, Microsoft more than Google (or, for that matter, Facebook), but it might well help Microsoft, since it would generally favor large incumbents with directs relationships with users, such as through the browser and OS. And Microsoft would hardly be the first company to wager that it held a losing hand, and that its odds would be better with a fresh deck of cards.

What lies ahead for choice architects "beyond DNT?" The perpetually difficult task of weighing costs and benefits, and attempting to foresee the unpredictable, in shaping users' choices.

---

[11] Harold Demsetz, *Toward a Theory of Property Rights*, 57:2 Am. Econ. Rev 347 (1967). http://www.econ.ucsb.edu/~tedb/Courses/Ec100C/Readings/Demsetz_Property_Rights.pdf

[12] *See* Lauren Gelman, "Re: tracking-ISSUE-150: DNT conflicts from multiple user agents [Tracking Definitions and Compliance]", public-tracking@w3.org mailing list, May 30, 2012, http://lists.w3.org/Archives/Public/public-tracking/2012May/0341.html.

[13] Bruce Yandle, "Bootleggers and Baptists-The Education of a Regulatory Economist," Regulation 7, no. 3 (1983): 12. http://www.cato.org/pubs/regulation/regv7n3/v7n3-3.pdf

Position Paper for W3C DNT Workshop

Rebecca Balebako, Pedro G. Leon, and Blase Ur
Carnegie Mellon University CUPS Lab

The current multi-stakeholder process to define the Do Not Track standard has considered inputs from many parties, including browser vendors, network advertisers, privacy researchers, and the government. However, we feel that average users, the millions of people who will be directly affected by a Do Not Track standard, have been underrepresented. These are the users whose information is collected, and also those users who benefit from having more customized and relevant advertising.

Over the past year, our research group has published a series of studies analyzing online behavioral advertising from the perspective of end-users. These studies have added the voice of these users to the debate about DNT and online behavioral advertising. At the W3C Workshop on Do Not Track, we hope to continue to give voice to users as part of the debate on the future of online privacy.

The W3C Tracking Protection Working Group is currently working on the definition of the DNT standard, aimed to improve user privacy and user control.[1] Of course, DNT is not the first attempt at letting users express online privacy choices. In a study published at CHI 2012, our group evaluated the usability of nine popular privacy tools provided by the advertising industry, third-party developers, and browser vendors. Our study found substantial usability flaws in all of the tools we tested, ranging from inappropriate default settings to inappropriate feedback to confusing configuration options [Leon et al. 2012a].

While analyzing the usability of existing privacy tools is essential in understanding the current situation for end-users, a detailed look at what users understand and think about online behavioral advertising is also essential to moving consumer privacy forward. We have provided such an analysis through in-depth interviews with 48 average users, as described in a paper we titled "Smart, Useful, Scary, Creepy" [Ur et al. 2012]. The duality of this title reflects the conflicting nature of end-users' opinions of OBA. While the idea of targeting advertising based on past behaviors on the web seemed smart and useful to many study participants, the tracking that provided this data without consumers' knowledge struck participants as scary and creepy.

In particular, the results of these interviews provide directions for improving current transparency and control mechanisms to support non-experts' online privacy preferences. This task is challenging for a number of reasons. The online behavioral advertising ecosystem is complex, and users often have difficulty

[1] http://www.w3.org/2011/tracking-protection/

understanding the different players in this ecosystem. In particular, how their information flows is confusing, even for privacy experts. Users would need to reason about the privacy practices of the many companies in the advertising ecosystem to make decisions about online tracking.

Among the lessons from these interviews is that participants had difficulty making decisions about different companies involved in online behavioral advertising. On one hand, this result lends support to current Do Not Track implementations in which users make blanket decisions about tracking, in contrast to many existing privacy tools that requires users to make filtering decisions on a per-company basis.

However , the idea that Do Not Track enables users either to send or not to send a blanket Do Not Track signal conflicted with the context-sensitive nature of our participants' opinions about online behavioral advertising. In particular, we learned that context of browsing matters. Both utility and privacy fuel users' attitudes towards allowing or not allowing advertising to be targeted based on their browsing. As a result, an all or nothing DNT approach does not seem to support user preferences fully. Furthermore, the solution is not as simple as specifying globally in which browsing circumstances DNT would apply. We observed that some browsing scenarios that were innocuous to certain users were perceived as privacy violations by other users, and vice versa. These results suggest that users would benefit from a Do Not Track variant enabling them to communicate their preferences dynamically.

Through these interviews, we also found that users strongly fear the collection of personal information for the purpose of tailored advertising. Of course, the form that users' collected information takes in databases is often a mystery to users. While most network advertisers claim that no personal information is collected, research has found instances in which popular websites leak personal information to third parties [Krishnamurthy et al. 2011]. Furthermore, the distinction between first-party and third-party advertisers blurs for companies like Google, Yahoo, Adobe, and Microsoft, who in certain cases are a first party, while in other cases are a third party. This line is further blurred by "social plugins," such as Facebook "Like" buttons, Google "+1" buttons, and Twitter widgets. We believe that there is tremendous opportunity for added transparency about the detailed types of data collected about users.

While we advocate increased transparency concerning the data that has been collected about individual users, we have also investigated transparency about the practice of online behavioral advertising in general. In particular, we investigated the icons and taglines (e.g., "AdChoices") that are displayed across the Internet to determine what message average users glean from these privacy disclosures [Leon et al. 2012b]. We found that the icons and taglines overall did not communicate clearly about behavioral advertising. Our study participants were more likely to think these icons would let them purchase their own advertisements than to understand that they could make choices about their privacy related to targeted

advertising. In contrast, we found that including verbs and other action words could communicate more successfully about online behavioral advertising and eliminate misconceptions.

In order to develop effective privacy options for users, it is also necessary to understand the current privacy situation. Over the past year, we have developed and tested a novel method for measuring the effectiveness of privacy tools that claim to limit online behavioral advertising [Balebako et al. 2012]. This method enables systematic measurements of OBA that is based on past browsing. We automated the collection of advertisements in a way that controlled for the time of visit, IP address, machine setup, and Flash LSOs. In a case study of this method, we found that third-party browser plugins and cookie-based tools we tested were effective in reducing behavioral advertising. However, in our case study, Do Not Track headers were ineffective in reducing behavioral advertising. Our method is particularly relevant in analyzing the efficacy of DNT as its standardization progresses.

Our current work has extended our measurement study across a wider range of advertising agencies to capture more completely the effectiveness of existing privacy tools. We will continue to measure the effectiveness of OBA tools through a longitudinal study, examining the effectiveness of DNT, cookie blocking, and browser plugins over the next two years. Our experience measuring tools has provided insight into the actual result of browsing information being used to target ads, supplementing users' perspectives with a snapshot of current targeting practices.

Moving forward, we believe that the Do Not Track standard should consider providing users with meaningful information that they can use to make informed decisions about tracking. For instance, we believe that sharing, retention and secondary uses of information, as defined by the Privacy Rulesets project,[2] are important aspects to consider. In addition, we believe that further investigating non-expert users' mental models could shed light about other important elements of the Do Not Track standard. Does it make sense to include basic or advanced options in DNT to accommodate the needs of users for whom a simple on/off switch doesn't provide support? What kind of feedback is needed for users to understand the implications of their filtering decisions? What level of access to their collected data would be appropriate for users? When should users be required to make a decision?

We would like to supplement the Do Not Track conversation with our belief that providing users control over data that has been collected about them can better align privacy options with users' wishes. Currently, a number of network advertising companies provide a dashboard that allows user to both learn what information is being collected and request the removal of certain data. We believe that the DNT standard or subsequent privacy tools should provide the same level of

[2] http://dev.w3.org/2009/dap/privacy-rulesets/

access and control without requiring users to identify themselves. This approach could help protect users' privacy and allow ad networks access to richer information. In the same way cookies are used to identify unique users, they can be used to provide access and control capabilities.

Taken as a whole, our work over the past year is particularly valuable to the current debate about online privacy by providing a voice to average users. Many users have little or no knowledge about the mechanisms that enable online tracking, and all humans are subject to a number of cognitive limitations and biases. For instance, to use behavioral economics terms, humans have bounded rationality, employ hyperbolic discounting, and display both overconfidence and a tendency to stick with default options.

Our research has demonstrated that current privacy tools and mechanisms for providing users privacy notice and choice present many opportunities for improvement. Our deep investigations of users' attitudes and abilities to use existing privacy tool have provided insight into designing privacy mechanisms that better align with users' expectations and mental models. While the needs of network advertisers, privacy advocates, and other stakeholders are fundamental to the design of the Do Not Track standard, we believe that the voice of users is also essential. We hope to provide that voice at the workshop.

References

[Balebako et al. 2012] Rebecca Balebako, Pedro Leon, Richard Shay, Blase Ur, Lorrie Faith Cranor, "Measuring the Effectiveness of Privacy Tools for Limiting Behavioral Advertising," In Web 2.0 Security and Privacy Workshop (W2SP), San Francisco, California, May 2012.

[Krishnamurthy et al. 2011] Balachander Krishnamurthy, Konstantin Naryshkin, and Craig E. Wills. Privacy leakage vs. protection measures: The growing disconnect. In Web 2.0 Security and Privacy Workshop (W2SP), Oakland, California, May 2011.

[Leon et al. 2012a] Pedro G. Leon, Blase Ur, Rebecca Balebako, Lorrie Faith Cranor, Richard Shay, Yang Wang, "Why Johnny Can't Opt Out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising," In Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), Austin, Texas, May 2012.

[Leon et al. 2012b] Pedro G. Leon, Justin Cranshaw, Lorrie Faith Cranor, Jim Graves, Manoj Hastak, Blase Ur, Guzi Xu, "What Do OBA Privacy Disclosures Communicate to Users," In Workshop on Privacy in the Electronic Society (WPES) 2012.

[Ur et al. 2012] Blase Ur, Pedro G. Leon, Lorrie Faith Cranor, Richard Shay, Yang Wang, "Smart, Useful, Scary, Creepy: Perceptions of Online Behavioral Advertising,"

In Proc. of the Symposium On Usable Privacy and Security (SOUPS). Washington, D.C., July 2012.

Adobe Systems Incorporated Position Paper for the
***W3C Workshop: Do Not Track and Beyond* 26-27 November 2012**

Submitted for participation in the W3C Workshop on Do Not Track and Beyond

By MeMe Jacobs Rasmussen,
VP, Chief Privacy Officer,
Adobe Systems Incorporated

Adobe has long held that standards, treated in an open fashion, play a crucial role in the growth of society in the information age and in fostering a vibrant, competitive, information technology marketplace. Adobe has created and stewarded some of the most widely used formal and de facto standards in technology today: Postscript, PDF, TIFF, XMP, DNG, etc. Adobe believes that standardization, at an appropriate time in a technology's life cycle, brings key benefits: users can avoid vendor lock-in and benefit from choice, and third parties can attempt to build on and extend existing standards. Standards not only are compatible with innovation; when used appropriately, they promote it.

Adobe works with numerous standards bodies, but none more closely than W3C. Current and former employees of Adobe have contributed time and expertise to numerous W3C groups, including the current Tracking Protection Working Group. Adobe is also a corporate supporter of W3C's valuable work.

Adobe has a keen interest in privacy. Adobe is a steward of others' data, as a provider of various online services, from content management to Web hosting to analytics to content delivery on mobile platforms. Adobe also produces tools and platforms used by others to produce content and host Websites. Across these diverse businesses, Adobe seeks to follow, and encourages companies that use our online services to follow, the Fair Information Practice Principles as a way to meet the expectations of consumers and our business customers.

Because of Adobe's many interests around privacy and our longstanding interest in and support of the W3C, Adobe respectfully requests to participate in the upcoming W3C Workshop: *Do Not Track and Beyond.*

W3C Workshop on Do Not Track and Beyond-Position Paper
Chris Jay Hoofnagle & Jennifer M. Urban

We recently commissioned a national survey focusing upon information privacy that examined several online privacy issues. Among them was the Do Not Track proposal. Our full paper is available online,[1] and we excerpt the main findings here.

Use versus collection

In our pretest of survey questions, we asked respondents what DNT meant, but almost two-thirds of the respondents simply did not know. Indeed, the vast majority of American consumers have never heard of DNT.

As a result of the pretest, we changed approaches, and instead of asking what they expected DNT would do, we asked consumers what they preferred it do. We also explicitly asked the full sample whether DNT was something they had heard of, or not.

We asked American consumers, "Policymakers are considering creating a “do not track” option for the internet. Have you heard of proposals for a “do not track” system, or not?" Thirteen percent had heard of it, and fully 87 percent had not. (N=1203.)

We then asked what Americans would prefer Do Not Track to do. Three options were presented to respondents in random order. One option would prevent collection of information; the other two would prevent two different actions on the part of the website.

A majority—60 percent—said they wanted DNT to stop websites from collecting information about the user. This was the closest option to the FTC’s proposal and the proposals of privacy groups, although simplified. The next largest group, 20 percent, wanted DNT to block advertisements. We presented that option because we suspected that many users expected DNT to simply block ads. Only 14 percent chose

[1] Chris Jay Hoofnagle, Jennifer Urban and Su Li, *Privacy and Modern Advertising: Most US Internet Users Want “Do Not Track” to Stop Collection of Data About their Online Activities*, Amsterdam Privacy Conference 2012, Oct. 8, 2012, available at http://ssrn.com/abstract=2152135.

the option that most closely matches the industry's suggested use-restriction proposal.

Tracking on medical websites

Medical information is recognized as particularly sensitive. It is one of the few data types explicitly protected under federal law. As early as 2000, the advertising industry itself recognized the sensitive nature of medical information and promised not to use it for advertising. In July 2000, the major network advertising companies[2] articulated the Network Advertising Initiative (NAI) "Self-Regulatory Principles for Online Preference Marketing by Network Advertisers." Under this framework, the NAI promised to not use "sensitive" personally identifiable data for "online preference marketing." The group explained, "Network advertisers shall neither use personally identifiable information about sensitive medical or financial data, sexual behavior or sexual orientation, nor social security numbers, for OPM [online preference marketing]."[3]

Almost ten years later, the advertising industry has retreated from its 2000 position, and the 2000 principles can no longer be found on the NAI's website.

When we turned to tracking on medical websites in this survey, we found that large numbers of consumers do not know what the rules are. We asked respondents whether it was true or false that advertisers are not allowed to track users using the internet to learn about medical conditions. While 36 percent correctly answered that this statement was false, 63 percent either did not know or responded incorrectly. Specifically, 22 percent thought permission was necessary, and 41 percent said they did not know the answer.

Attitudes toward online advertising

We first asked, "In general, how often do you find online advertising, such as the advertising that appears on search results webpages and banner advertisements, useful?" Thirty percent do find utility in

[2] 24/7 Media, AdForce, AdKnowledge, Avenue A, Burst! Media, DoubleClick, Engage, L90, MatchLogic.
[3] http://www.ftc.gov/os/2000/07/NAI%207-10%20Final.pdf.

advertising, with 10 percent reporting that they find search and banner advertisements useful often, and 20 percent sometimes find this information useful. Thirty-six percent answered "hardly ever," and 33 percent answered "never."

We followed the utility question by asking whether users actually clicked on advertisements. We asked, "How often do you click on advertisements when using the internet?" Fifty percent claimed that they never click on ads, and 35 percent responded hardly ever. It may be that this is incorrect—that consumers do not remember or selectively remember their interaction with advertising.[4] We have no way to test this directly; however, there is some indirect evidence supporting our respondents' claims. It is true, for example, that accidental and fraudulent clicks are a pervasive problem,[5] and that as of this writing, the most popular add-on for the Firefox web browser is AdBlock Plus, with over 14 million users.

Conclusion

In light of consumer attitudes and marketplace realities, Do Not Track is a modest intervention. Yet the advertising industry has argued for systemically weakening what "Do Not Track" means, and has retreated from earlier, stronger promises to limit tracking.

We found that most consumers want Do Not Track to mean exactly that: do not collect information that allows companies to track them across the Internet. This may seem obvious, but even the definition articulated by the FTC may fall short of these consumer expectations. Further, advertising industry groups presently are lobbying for a different interpretation that would allow pervasive tracking and use of information derived from online experiences, even if the consumer opts out.

---

[4] We note that we did not have a way to verify whether this claim is borne out by actual behavior, but it is clear that our respondents subjectively found online ads of limited value.

[5] Ryan Kim, Report: *40 Percent of Mobile Clicks are Fraud or Accidents*, GIGAOM, Aug. 31, 2012, *available at* http://gigaom.com/2012/08/31/report-40-percent-of-mobile-clicks-are-fraud-or-accidents/

This disconnect appears pervasive and strong. In addition to the fact that a strong majority of respondents prefer that Do Not Track allow them to opt out of collection, there is a lack of understanding about what trackers can do. We found that only about 1 in 5 internet users understands that advertisers can track them on medical sites. Here too, despite broad consensus that medical information is especially sensitive and despite widespread consumer ignorance of the rules governing the collection and use of behavioral tracking on medical websites, advertising lobbying groups have stuck to a "notice and no choice" approach.

Consumers and advertisers seem to be at an impasse on privacy. This impasse is the product of consumers' anxiety about tracking, and advertisers' concern that any imposition upon data collection will undermine an existing and growing business model. Subjectively at least, nearly 70% of consumers say that they find little if any value in online ads. Half claim to never click on ads at all. Yet advertisers' position on tracking is that consumers should be tracked even if they opt out of tracking, suggesting that consumers' subjective opinions about tracking do not matter.

Lost in the present debate is the fact that DNT essentially responds to a specific business model, one in which third parties attempt to build advertising value by tracking individuals in all aspects of their lives. This model seems to require continually ratcheting up data collection and ratcheting down privacy protections in an attempt to show value to ad buyers.

Targeting consumers based upon specific information about them appears to be increasing across a variety of internet and mobile marketing models, with an apparent goal of linking online and offline purchase behavior. In previous work, we have explored mobile payments models that promise to connect more payments ecosystem players with detailed "Level 3" purchase data (lists of the specific things consumers buy) for individual consumers shopping at bricks-and-mortar stores and mobile app models that use app users' address books to target offers.[6] And as this paper was being prepared, for

[6] Chris Jay Hoofnagle, Jennifer M. Urban, and Su Li, *Mobile Payments: Consumer Benefits & New Privacy Concerns* (Apr. 24, 2012); and Jennifer M.

example, newspapers reported that Facebook was beginning to buy data on Facebook users' specific purchases in CVS drugstores in order to show whether targeted ads served to individual profiles actually resulted in increased sales of the advertised products.[7]

If present trends continue, we will soon find ourselves in a world where ultra-large tracking platforms will have data about almost all online and offline consumer transactional behavior. Consumers will find themselves subject to these platforms' power to collect and use that data, and with little recourse or say about that collection and use.

We think that there are ways around the impasse between advertising models and consumers' apparent expectations of privacy. There are alternative approaches to the "track everyone, everywhere" model. Academics including Steven Bellovin,[8] Eric Goldman,[9] and Helen Nissenbaum[10] have proposed alternative models that would allow highly targeted ads without creating dossiers of internet behavior held by third parties.

Urban, Chris Jay Hoofnagle, and Su Li, *Mobile Phones and Privacy* (July 12, 2012), *available at* http://ssrn.com/abstract=2103405.

[7] Rebecca Greenfield, *Facebook Now Knows What You're Buying at Drugstores*, THE ATLANTIC WIRE (Sept. 24, 2012), *available at* http://www.theatlanticwire.com/technology/2012/09/facebook-tracking-you-drug-store-now-too/57183/.

[8] Elli Androulaki and Steven M. Bellovin, *A secure and privacy-preserving targeted ad-system*, *in* Proceedings of the 1st Workshop on Real-Life Cryptographic Protocols and Standardization, Jan. 2010.

[9] Eric Goldman, *A Coasean Analysis of Marketing*, 2006 WIS. L. REV. 1151 (2006).

[10] Vincent Toubiana, Arvind Narayanan, Dan Boneh, Helen Nissenbaum, Solon Barocas, *Adnostic: Privacy Preserving Targeted Advertising*, NDSS 2010, *available at* http://crypto.stanford.edu/adnostic/.

SENDING OUT A PING FOR WEB PRIVACY: LAYING OUT A PLAN
Christine Runnegar and Tara Whalen, Co-Chairs, W3C Privacy Interest Group

This is not so much a position paper, but rather a call to action.

The Privacy Interest Group (PING) was chartered by the W3C in 2012 "to improve the support of privacy in Web standards by monitoring ongoing privacy issues that affect the Web, investigating potential areas for new privacy work, and providing guidelines and advice for addressing privacy in standards development"[i].

Already, PING has made some progress towards these goals[ii], but more concerted work is needed across the W3C community.

Guidelines and advice

PING's mandate includes the development of privacy guidance for Web standards developers.

One of the first questions the PING chairs were asked was:

> "What is the intended audience"?

The primary audience is Web specification authors, however, there is also interest within PING to develop guidance (e.g. best practices) for implementers and deployers, particularly as additional privacy concerns arise at the point where implementation and deployment choices are made. Such guidance needs to be relevant, understandable and useful for the target audience.

Spotting potential privacy risks and vulnerabilities is an art, and possibly difficult for specification authors with no background in privacy. Nonetheless, there are (at least anecdotally) recurring privacy issues with Web standards that could stand to be articulated. There are also a number of different proposed design approaches to improve privacy of a specification (e.g., the permission-based model used in the Geolocation API, data minimisation, etc.) that could be evaluated (e.g., pros and cons).

Privacy reviews

There is consensus that W3C specifications, especially at an early stage in their development, would benefit from privacy reviews. PING has been identified as the logical coordinator for such reviews, and already four specifications have been proposed as candidates: Navigation timing; Web Intents; Content Security Policy 1.0; and the Web Cryptography API (at a future date to be advised). Charters for new groups also specifically require privacy reviews (e.g., Systems Applications Working Group Charter[iii]).

While there is considerable and varied privacy expertise in PING, it is not yet clear whether PING members will have sufficient expertise in the relevant specification subject matters to provide meaningful privacy reviews. There is also the question of resources and the weight that would be given to such reviews. Are PING members willing and able to commit to providing timely privacy reviews? How would those reviews be taken into account? How will the inevitable conflicts between privacy and functionality, usability, security, reliability, etc. be resolved?

Ad hoc informal privacy reviews provided by privacy experts have been useful in identifying possible issues (e.g., the potential for fingerprinting). However, to ensure greater consistency in approach across Web standards, it would be useful to develop methodology for privacy vulnerability and risk assessment, together with a set of preferred design criteria. Such guidance would also help other W3C working groups conduct their own privacy reviews.

This submission is provided by Christine Runnegar and Tara Whalen in their capacity as PING co-chairs.
22 October 2012

Identifying potential areas of new work

Discussions in PING have, to date, identified two key areas of potential new work:

- Fingerprinting
    - [for PING] – explaining what is fingerprinting, the challenges and best ways to mitigate fingerprinting while still allowing for greater integration and functionality
    - [for others] – develop a standard anonymous fingerprint
    - [for others] – develop ways to expose fingerprinting (e.g., make it easier for the browser to detect fingerprinting)
- Privacy indicators for the browser

Questions for the workshop:

- What are known privacy vulnerabilities and risks associated with Web standards?
- What should we do about them?
- What privacy design principles make sense for the Web?
- How do we ensure privacy concerns are raised at an early stage?
- How do privacy standards work and privacy regulation interact (or should interact)?
- When and how should privacy reviews be conducted? How will the reviews be recorded?
- How will conflicts between privacy and functionality, usability, security, reliability, etc. be resolved?

[i] http://www.w3.org/2011/07/privacy-ig-charter
[ii] Please see the informal chairs' summaries at:
http://lists.w3.org/Archives/Public/public-privacy/2012AprJun/0065.html
http://lists.w3.org/Archives/Public/public-privacy/2012AprJun/0083.html
http://lists.w3.org/Archives/Public/public-privacy/2012JulSep/0004.html
http://lists.w3.org/Archives/Public/public-privacy/2012JulSep/0019.html
http://lists.w3.org/Archives/Public/public-privacy/2012JulSep/0053.html
http://lists.w3.org/Archives/Public/public-privacy/2012OctDec/0001.html
[iii] http://www.w3.org/2012/09/sysapps-wg-charter.html