W3C, P3P & DNT

Lorrie Faith Cranor
October 2, 2014

8-533 / 8-733 / 19-608 / 95-818:
Privacy Policy, Law, and Technology
Today’s agenda

• Quiz

• What’s on the midterm?

• Lots of TLAs
  – W3C
  – P3P
  – DNT
By the end of class you will be able to:

• Understand what W3C is and what it does, and how to read a W3C specification

• Understand the history of P3P

• Understand the major components of P3P

• Understand the history and current status of DNT
W3C

- International member organization
- Founded in 1994 by Web inventor Tim Berners-Lee
- Mission: Lead the web to its full potential
- Most work revolves around standardization of web technologies
  - Structured process for developing standards
  - Working drafts -> Last call ->
    Candidate Recommendation ->
    Proposed Recommendation -> Recommendation
Original Idea behind P3P

• A framework for automated privacy discussions
  – Web sites disclose their privacy practices in standard machine-readable formats
  – Web browsers automatically retrieve P3P privacy policies and compare them to users’ privacy preferences
  – Sites and browsers can then negotiate about privacy terms
P3P history

• November 1995 - Idea discussed at FTC meeting

• Fall 1996 - Ad Hoc “Internet Privacy Working Group” convened

• Summer 1997 - W3C began working on P3P
  – Several working groups chartered with dozens of participants from industry, non-profits, academia, government
  – Numerous public working drafts issued, many changes
  – Early ideas about negotiation and agreement ultimately removed
  – Automatic data transfer added and then removed
  – Patent issue stalled progress, but ultimately became non-issue

• April 16, 2002 - P3P issued as official W3C Recommendation http://www.w3.org/TR/P3P/

• 2012 – Microsoft complains that companies are circumventing P3P
P3P1.0 Spec

• A standard vocabulary for describing set of uses, recipients, data categories, and other privacy disclosures

• A standard schema for data a Web site may wish to collect (base data schema)

• An XML format for expressing a privacy policy in a machine readable way

• A means of associating privacy policies with Web pages or sites

• A protocol for transporting P3P policies over HTTP
  – A format for expressing optional P3P compact policy headers
A simple HTTP transaction

GET /index.html HTTP/1.1
Host: www.att.com

. . . Request web page

HTTP/1.1 200 OK
Content-Type: text/html

. . . Send web page
... with P3P 1.0 added

```
GET /w3c/p3p.xml HTTP/1.1
Host: www.att.com
Request Policy Reference File
```

```
GET /index.html HTTP/1.1
Host: www.att.com
Request web page
```

```
HTTP/1.1 200 OK
Content-Type: text/html
Send web page
```
Transparency

- P3P clients can check a privacy policy each time it changes.
- P3P clients can check privacy policies on all objects in a web page, including ads and invisible images.

http://adforce.imgis.com/?adlink=2|68523|1|146|ADFORCE

http://www.att.com/accessatt/
P3P in IE6

Automatic processing of compact policies only; third-party cookies without compact policies blocked by default.

Privacy icon on status bar indicates that a cookie has been blocked - pop-up appears the first time the privacy icon appears.
Users can click on privacy icon for list of cookies; privacy summaries are available at sites that are P3P-enabled.
Privacy summary report is generated automatically from full P3P policy.
P3P in Netscape 7

Preview version similar to IE6, focusing, on cookies; cookies without compact policies (both first-party and third-party) are “flagged” rather than blocked by default.

Indicates flagged cookie
Users can view English translation of (part of) compact policy in Cookie Manager.
A policy summary can be generated automatically from full P3P policy.
What’s in a P3P policy?

- Name and contact information for site
- The kind of access provided
- Mechanisms for resolving privacy disputes
- The kinds of data collected
- How collected data is used, and whether individuals can opt-in or opt-out of any of these uses
- Whether/when data may be shared and whether there is opt-in or opt-out
- Data retention policy
Assertions in a P3P policy

• General assertions
  – Location of human-readable policies and opt-out mechanisms – `discuri`, `opturi` attributes of `<POLICY>`
  – Indication that policy is for testing only – `<TEST>` (optional)
  – Web site contact information – `<ENTITY>`
  – Access information – `<ACCESS>`
  – Information about dispute resolution – `<DISPUTES>` (optional)

• Data-Specific Assertions
  – Consequence of providing data – `<CONSEQUENCE>` (optional)
  – Indication that no identifiable data is collected – `<NON-IDENTIFIABLE>` (optional)
  – How data will be used – `<PURPOSE>`
  – With whom data may be shared – `<RECIPIENT>`
  – Whether opt-in and/or opt-out is available – required attribute of `<PURPOSE>` and `<RECIPIENT>`
  – Data retention policy – `<RETENTION>`
  – What kind of data is collected – `<DATA>`
P3P/XML encoding

```xml
<POLICIES xmlns="http://www.w3.org/2002/01/P3Pv1">
  <POLICY discuri="http://p3pbook.com/privacy.html" name="policy">
    <ENTITY>
      <DATA-GROUP>
        <DATA ref="#business.contact-info.online.email">privacy@p3pbook.com</DATA>
        <DATA ref="#business.contact-info.online.uri">http://p3pbook.com/</DATA>
        <DATA ref="#business.name">Web Privacy With P3P</DATA>
      </DATA-GROUP>
      <ACCESS><nonident/></ACCESS>
      <STATEMENT>
        <CONSEQUENCE>We keep standard web server logs.</CONSEQUENCE>
        <PURPOSE><admin/><current/><develop/></PURPOSE>
        <RECIPIENT><ours/></RECIPIENT>
        <RETENTION><indefinitely/></RETENTION>
        <DATA-GROUP>
          <DATA ref="#dynamic.clickstream"/>
          <DATA ref="#dynamic.http"/>
        </DATA-GROUP>
      </STATEMENT>
    </ENTITY>
  </POLICY>
</POLICIES>
```
Why web sites adopt P3P

• Demonstrate corporate leadership on privacy issues
  – Show customers they respect their privacy
  – Demonstrate to regulators that industry is taking voluntary steps to address consumer privacy concerns

• Distinguish brand as privacy friendly

• Prevent IE6 from blocking their cookies

• Anticipation that consumers will soon come to expect P3P on all web sites

• Individuals who run sites value personal privacy
P3P early adopters

• News and information sites – CNET, About.com, BusinessWeek

• Search engines – Yahoo, Lycos

• Ad networks – DoubleClick, Avenue A

• Telecom companies – AT&T

• Financial institutions – Fidelity

• Computer hardware and software vendors – IBM, Dell, Microsoft, McAfee

• Retail stores – Fortunoff, Ritz Camera

• Government agencies – FTC, Dept. of Commerce, Ontario Information and Privacy Commissioner

• Non-profits - CDT
Web site adoption of P3P

- AT&T study surveyed 5,856 websites in 2003, found 538 P3P policies
  - Adoption highest among popular websites (~30% of top 100 sites)
  - Web site adoption increasing slowly, but steadily
  - Low adoption for government sites – but changed with new regulations

- Large number of P3P policies contain technical errors
  - Most errors due to old version of P3P spec or minor technical issues
  - 7% have severe errors such as missing required components

Legal issues

- P3P specification does not address legal standing of P3P policies or include enforcement mechanisms.
- P3P specification requires P3P policies to be consistent with natural-language privacy policies:
  - P3P policies and natural-language policies are not required to contain same level of detail.
  - Typically natural-language policies contain more detailed explanations.
- In some jurisdictions, regulators and courts may treat P3P policies equivalently to natural language privacy policies.
- The same attorneys and policy makers involved in drafting natural-language policy should help create P3P policy.
<table>
<thead>
<tr>
<th>Privacy policy</th>
<th>P3P policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designed to be read by a human</td>
<td>Designed to be read by a computer</td>
</tr>
<tr>
<td>Can contain fuzzy language with “wiggle room”</td>
<td>Mostly multiple choice – sites must place themselves in one “bucket” or another</td>
</tr>
<tr>
<td>Can include as much or as little information as a site wants</td>
<td>Must include disclosures in every required area</td>
</tr>
<tr>
<td>Easy to provide detailed explanations</td>
<td>Limited ability to provide detailed explanations</td>
</tr>
<tr>
<td>Sometimes difficult for users to determine boundaries of what it applies to and when it might change</td>
<td>Precisely scoped</td>
</tr>
<tr>
<td>Web site controls presentation</td>
<td>User agent controls presentation</td>
</tr>
</tbody>
</table>
P3P Interface design challenges

• P3P 1.0 specification focuses on interoperability, says little about user interface
  – P3P 1.1 spec will provide explanations of P3P vocabulary elements suitable for display to end users

• P3P user agents typically need user interfaces for:
  – informing users about web site privacy policies
  – configuring the agent to take actions on the basis of a user’s privacy preferences
Informing users about privacy is difficult

- Privacy policies are complex
  - Over 36K combinations of P3P “multiple choice” elements
- Users are generally unfamiliar with much of the terminology used by privacy experts
- Users generally do not understand the implications of data practices
- Users are not interested in all of the detail of most privacy policies
- Which details and the level of detail each user is interested in varies
Specifying privacy preferences is difficult

• Privacy policies are complex

• User privacy preferences are often complex and nuanced

• Users tend to have little experience articulating their privacy preferences

• Users are generally unfamiliar with much of the terminology used by privacy experts
Iterative design approach

- Four P3P user agent prototypes developed over 4-year period while P3P specification was under development
- AT&T Privacy Bird beta released Feb. 2002
  - August 2002 user study
  - Beta 1.2 released Feb. 2003
W3C prototype

- Based on pre-W3C draft of P3P vocabulary with 3 fields, 7x9x2 = 126 combinations of elements
- Preference interface eliminated the impractical combos, combined 2 dimensions → 7x14 = 98 combinations
- Matrix represented by tabbed interface
- Feedback: too complicated, too many choices
AT&T Privacy Bird

- Free download of beta from http://privacybird.com/
- “Browser helper object” for IE 5.01/5.5/6.0
- Reads P3P policies at all P3P-enabled sites automatically
- Puts bird icon at top of browser window that changes to indicate whether site matches user’s privacy preferences
- Clicking on bird icon gives more information
- Current version is information only – no cookie blocking
Chirping bird is privacy indicator

Shane Zachary Cranor

Born May 4, 2001, 7:25 am, 7 pounds, 13 oz., 21 inches

Photo Album  | Latest Photos  | 2001 Favorite Photos  | 2002 Favorite Photos

Shane’s Photo Album

- Shane’s First Year

Shane’s Latest Photos

Shane attended Mom’s Chatham Community Band Concert, but he was so bored he fell asleep. The next day Shane helped Dad change a lightbulb -- climbing a ladder couldn’t be that much
Click on the bird for more info

Policy Summary

Shane Cranor’s Home Page Privacy Practices

Privacy Policy Check

Shane Cranor’s Home Page’s privacy policy matches your preferences.

Privacy Policy Summary

This site has the following statements in its policy:

- Site Statement 1

Site Statement 1

Types of Information Collected:

- HTTP protocol information
- Click-stream information

How your information will be used:

- Research and development
- To complete the activity for which the data was provided
- Web site and system administration

Who will use your information:

- This web site and its agents
Privacy policy summary - mismatch

1-800-Flowers.com, Inc. Privacy Practices

Privacy Policy Check

1-800-Flowers.com, Inc.'s privacy policy does not match your preferences:

- Unless you opt-out, site may share financial information or information about your purchases with other companies (other than those helping the site provide services to you).
- Unless you opt-out, site may share information that personally identifies you with other companies (other than those helping the site provide services to you).

Privacy Policy Summary

This site has the following statements in its policy:

- Site Statement 1 - All users and customers

Site Statement 1 - All users and customers
Types of information collected:
Expand/collapse added in beta

1.2

<table>
<thead>
<tr>
<th>Policy Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Federal Trade Commission</td>
</tr>
<tr>
<td>Privacy Policy Check</td>
</tr>
<tr>
<td>Federal Trade Commission's privacy policy matches your preference</td>
</tr>
<tr>
<td>Privacy Policy Summary</td>
</tr>
<tr>
<td>+ Policy Statement 1 - Basic Information</td>
</tr>
<tr>
<td>Data collected from all Web users: access logs, and search strings (if entered).</td>
</tr>
<tr>
<td>Types of Information that may be collected:</td>
</tr>
<tr>
<td>- search terms</td>
</tr>
<tr>
<td>- click-stream information</td>
</tr>
<tr>
<td>How your information may be used:</td>
</tr>
<tr>
<td>- To complete the activity for which the data was provided</td>
</tr>
<tr>
<td>- To do web site and system administration</td>
</tr>
<tr>
<td>Who may use your information:</td>
</tr>
<tr>
<td>- This web site and the companies that help the site provide services to you</td>
</tr>
<tr>
<td>+ Policy Statement 2 - Data Collection</td>
</tr>
<tr>
<td>Access to your information</td>
</tr>
<tr>
<td>This site allows you to access your data about you from its records</td>
</tr>
<tr>
<td>+ How to reach this site</td>
</tr>
<tr>
<td>+ How to resolve privacy-related issues</td>
</tr>
<tr>
<td>More Information</td>
</tr>
</tbody>
</table>

---

[Note: The image contains a screenshot of a web page with expandable/collapsible sections for policy statements and data collection.]
Bird checks policies for embedded content
Privacy Bird icons
Privacy Preference Settings

These settings control when a warning icon will be displayed at the top of your browser window. You can click on the warning icon for more information.

Select Privacy Level:  ○ Low  ○ Medium  ○ High  ○ Custom  ○ Imported

**HEALTH OR MEDICAL INFORMATION**

Warn me at web sites that use my health or medical information:

- [ ] For analysis, marketing, or to make decisions that may affect what content or ads I see, etc.
- [ ] To share with other companies (other than those helping the web site provide services to me)

**FINANCIAL OR PURCHASE INFORMATION**

Warn me at web sites that use my financial information or information about my purchases:

- [ ] For analysis, marketing, or to make decisions that may affect what content or ads I see, etc.
- [ ] To share with other companies (other than those helping the web site provide services to me)

**PERSONALLY IDENTIFIABLE INFORMATION** (name, address, phone number, email address, etc.)

Warn me at web sites that may contact me to interest me in other services or products:

- [ ] Via telephone
- [ ] Via other means [email, postal mail, etc.]
- [ ] And do not allow me to remove myself from marketing/mailing lists

Warn me at web sites that use information that personally identifies me:

- [ ] To determine my habits, interests, or other characteristics
- [ ] To share with other companies (other than those helping the website provide services to me)

- [ ] Warn me at web sites that do not allow me to find out what data they have about me

**NON-PERSONALLY IDENTIFIABLE INFORMATION** (demographics, interests, web sites visited, etc.)

Warn me at web sites that use my non-personally identifiable information:

- [ ] To determine my habits, interests, or other characteristics
- [ ] To share with other companies (other than those helping the website provide services to me)
But how do you find sites with good policies?
PHILLIP'S
1-800-FLORALS
1-800-356-7257


PICKS OF THE WEEK

FTD® Star Gazer™ Bouquet #3061X $109.95
Multicolor Roses Bowl #0683T $99.95
Pastel Basket Planter #1112T $49.95
Privacy Finder

• Prototype developed at AT&T Labs, improved and deployed by CUPS
• Uses Google or Yahoo! API to retrieve search results
• Checks each result for P3P policy
• Evaluates P3P policy against user’s preferences
• Reorders search results
• Composes search result page with privacy annotations next to each P3P-enabled result
• Users can retrieve “Privacy Report” similar to Privacy Bird policy summary
This site may collect the following types of information about you:

- search terms
- HTTP protocol information
- click-stream information
- use of HTTP cookies
  - Information about your tastes or interests
  - Cookies and mechanisms that perform similar functions
  - Which pages you visited on this web site and how long you stayed at each page
  - Website login IDs and other identifiers (excluding government IDs and financial account numbers)
  - Information about the computer you are using, such as its hardware, software, or Internet address
  - Email address or other online contact information
  - Name, address, phone number, or other contact information
- third party’s name
- home contact information (optional)
- server stores the transaction history
- user’s name (optional)

The ways your information may be used:

- To aid in historical preservation as governed by a law or policy described in this privacy policy
- To contact you through means other than telephone (for example, email or postal mail) to market services or products -- unless you opt-out
- To make decisions that directly affect you using information about you, for example to recommend products or services based on your previous purchases -- unless you opt-out
- To customize the site for your current visit only
- To do research and analysis in which your information may be linked to an ID code but not to your personal identity
- To contact you by telephone to market services or products -- unless you opt-out
- For research and development, but without connecting any information to you
- To perform web site and system administration
- To provide the service you requested

With whom this site may share your information:

- Other companies whose privacy policies are unknown to this site -- unless you opt-out
- Companies that have privacy policies similar to this site’s -- unless you opt-out
- Delivery companies that help this site fulfill your requests and who may also use your information in other ways

Access to your information
P3P Adoption Studies

• Compiled two lists of search terms:
  – Typical: 20,000 terms randomly sampled from one week of AOL user search queries
  – Ecommerce: 940 terms screen scraped from Froogle front page

• Submitted search terms to Google, Yahoo!, and AOL search engines and collected top 20 results for each term

• Checked each result for P3P policy and evaluated policies against 5 “rulesets” and P3P validator

• Saved 1,232,955 annotated search results in database

• Separately checked for P3P policies on 30,000 domains most clicked on by AOL search engine users

Results: P3P deployment

- 10% of results from typical search terms have P3P
- 21% of results from ecommerce search terms have P3P
- More popular sites are more likely to have P3P
Results: Frequency of P3P-enabled hits

- 83% of searches had at least one P3P-enabled site in top 20 results
- 68% of searches had at least one P3P-enabled site in top 10 results
- For top 20 search results returned by AOL search engine for typical search terms:
  - 29% return at least 1 P3P-enabled hit that matches medium privacy preferences
  - 34% return at least 1 P3P-enabled hit in that does not share data
  - 31% return at least 1 P3P-enabled hit that does not market without opt-in
  - Thus, ~ 1/3 of the time AOL users will find site with “good” privacy policy in first 2 pages of results
Does Privacy Finder influence purchases?

• Yes!

• J. Tsai, S. Egelman, L. Cranor, and A. Acquisti. The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study.
P3P deployment overview

- Create a privacy policy
- Analyze the use of cookies and third-party content on your site
- Determine whether you want to have one P3P policy for your entire site or different P3P policies for different parts of your site
- Create a P3P policy (or policies) for your site
- Create a policy reference file for your site
- Configure your server for P3P
- Test your site to make sure it is properly P3P enabled
Generating a P3P policy

• Edit by hand
  – Cut and paste from an example

• Use a P3P policy generator
  – Recommended: IBM P3P policy editor
    http://www.alphaworks.ibm.com/tech/p3peditor

• Generate compact policy and policy reference file the same way (by hand or with policy editor)

• Get a book
  – Web Privacy with P3P
    by Lorrie Faith Cranor
    http://p3pbook.com/
IBM P3P Policy Editor

Sites can list the types of data they collect and view the corresponding P3P policy.
Compact policies

- HTTP header with short summary of full P3P policy for cookies (not for URLs)
- Not required
- Must be used in addition to full policy
- Must commit to following policy for lifetime of cookies
- May over simplify site’s policy
- IE6 relies heavily on compact policies for cookie filtering – especially an issue for third-party cookies
Server configuration

- Only needed for compact policies and/or sites that use P3P HTTP header
- Need to configure server to insert extra headers
Reading the P3P specification

• http://www.w3.org/TR/P3P11/
DNT history

• 2007 – Public interest groups propose Do Not Track (like Do Not Call) to FTC

• 2009 – Google ad-on to make opt-out cookies permanent, Mozilla ad-on implements DNT header

• 2010 – FTC Chairman discusses DNT with Senate committee

• 2011 – W3C launches working group, browsers implement DNT

• 2012 – Advertising industry pledges to support DNT, Microsoft enables DNT by default in IE10

• 2013 – Working group votes to continue working, ad industry quits

• 2014 – W3C issues LC working draft

• 2015 – W3C issues CR draft, EFF issues their own DNT policy
Headlines

• Do Not Track proposal is DOA (July 16, 2013)
  http://money.cnn.com/2013/07/16/technology/do-not-track/

• The Internet’s best hope for a Do Not Track standard is falling apart. Here’s why.
  (October 11, 2013)
  http://www.washingtonpost.com/blogs/the-switch/wp/2013/10/11/the-internets-best-hope-for-a-
  do-not-track-standard-is-falling-apart-heres-why/

• How bickering and greed neutered the 'Do Not Track' privacy initiative (May 22, 2014)

• ADVERTISING ALLIANCE TO WEB STANDARDS GROUP: DROP "DO NOT TRACK" (June 23, 2014)

• Do-Not-Track Will Benefit Our Whole Industry (August 29, 2014)
  http://www.mediapost.com/publications/article/233197/do-not-track-will-benefit-our-whole-
  industry.html

• Why We Oppose Do Not Track and How to Fix It: Rules Need to Apply to All Data Collectors -- Including Facebook and Google (July 25, 2014)
  http://adage.com/article/guest-columnists/oppose-track-fix/294319/
What type of protocol?

• List of trackers to block?
• One-way signal from browser to website?
• Two-way communication
  – Browser signals to website
  – Website signals back
Conflicting signals

• What if users have opted out with opt-out cookie or other mechanism but not DNT?
• What if users have opt-in but send DNT=1?
Exceptions

• How can users make an exception for some sites? For some trackers? For some site/tracker combinations?

• How do we prevent sites from tricking users into making an exception or making an exception w/out user consent?
Deliberate choice by user

“Key to that notion of expression is that the signal sent must reflect the user's preference, not the choice of some vendor, institution, site, or network-imposed mechanism outside the user's control; this applies equally to both the general preference and exceptions. The basic principle is that a tracking preference expression is only transmitted when it reflects a deliberate choice by the user. In the absence of user choice, there is no tracking preference expressed.”

http://www.w3.org/TR/2014/WD-tracking-dnt-20140424/
CR working draft specifies

- DNT request header field as an HTTP mechanism for expressing the user's preference regarding tracking
- HTML DOM property to make that expression readable by scripts
- APIs that allow scripts to register site-specific exceptions granted by the user
- Mechanisms for sites to communicate whether and how they honor a received preference
  - “Tk” response header field
  - Well-known resources that provide a machine-readable tracking status
- http://www.w3.org/TR/tracking-dnt/
Definition of tracking

*Tracking* is the collection of data regarding a particular user's activity across multiple distinct contexts and the retention, use, or sharing of data derived from that activity outside the context in which it occurred. A *context* is a set of resources that are controlled by the same party or jointly controlled by a set of parties.
DNT meaning

• 1
  – This user prefers not to be tracked on the target site.

• 0
  – This user prefers to allow tracking on the target site.
No defaults allowed

- A tracking preference expression is only transmitted when it reflects a deliberate choice by the user.

- In the absence of user choice, there is no tracking preference expressed.

- A user agent must offer users a minimum of two alternative choices: *unset* or *DNT:1*. A user agent may offer a third alternative choice: *DNT:0*. 
Tracking status value

- ! — under construction
- ? — dynamic
- G — gateway to multiple parties
- N — not tracking
- T — tracking
- C — tracking with consent
- P — tracking only if consented
- D — disregarding DNT
- U — updated
Tracking compliance

- [http://www.w3.org/TR/tracking-compliance/](http://www.w3.org/TR/tracking-compliance/)

- First party compliance with DNT:1
  - May collect, retain, and use data, including for customizing content, services, and ads

- Third party compliance with DNT:1
  - May collect data with explicit user consent, data is deidentified, or permitted uses:
    - Frequency capping
    - Financial logging
    - Security
    - Debugging
Congress weighs in


- Senators Markey and Franken, and Congressman Barton complain that DNT has different rules for 1st party and 3rd party
EFF privacy-friendly Do Not Track (DNT) Policy

• EFF Privacy Badger blocks tracking, but unblocks for companies that comply with their DNT policy

• Does not make distinction between first and third party

• https://www.eff.org/dnt-policy

Privacy Badger 1.0
New Browser Plugin Blocks Spying Ads and Invisible Trackers.