P₃P

Lorrie Faith Cranor

October 10, 2013

8-533 / 8-733 / 19-608 / 95-818: Privacy Policy, Law, and Technology Carnegie Mellon University CyLab



Engineering & Public Policy



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

- Idea discussed at November 1995 FTC meeting
- Ad Hoc "Internet Privacy Working Group" convened to discuss the idea in Fall 1996
- W3C began working on P3P in Summer 1997
 - Several working groups chartered with dozens of participants from industry, non-profits, academia, government
 - Numerous public working drafts issued, and feedback resulted in 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
- P3P issued as official W3C Recommendation on April 16, 2002
 - http://www.w3.org/TR/P3P/

P3P1.0 – A first step

- Offers an easy way for web sites to communicate about their privacy policies in a standard machine-readable format
 - Can be deployed using existing web servers
- This will enable the development of tools that:
 - Provide snapshots of sites' policies
 - Compare policies with user preferences
 - Alert and advise the user

P3P is part of the solution

- P3P1.0 helps users understand privacy policies but is not a complete solution
- Seal programs and regulations
 - help ensure that sites comply with their policies
- Anonymity tools
 - reduce the amount of information revealed while browsing
- Encryption tools
 - secure data in transit and storage
- Laws and codes of practice
 - provide a base line level for acceptable policies

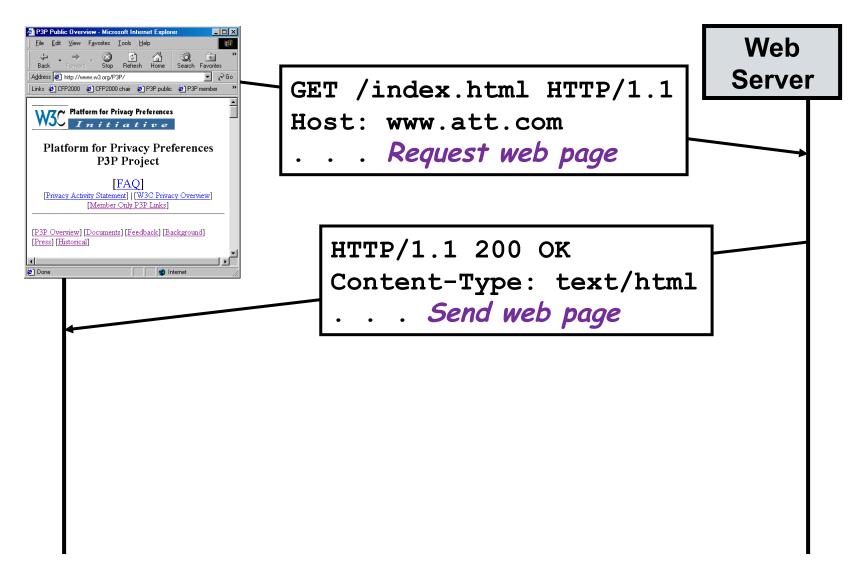
The basics

- P3P provides a standard XML format that web sites use to encode their privacy policies
- Sites also provide XML "policy reference files" to indicate which policy applies to which part of the site
- Sites can optionally provide a "compact policy" by configuring their servers to issue a special P3P header when cookies are set
- No special server software required
- User software to read P3P policies called a "P3P user agent"

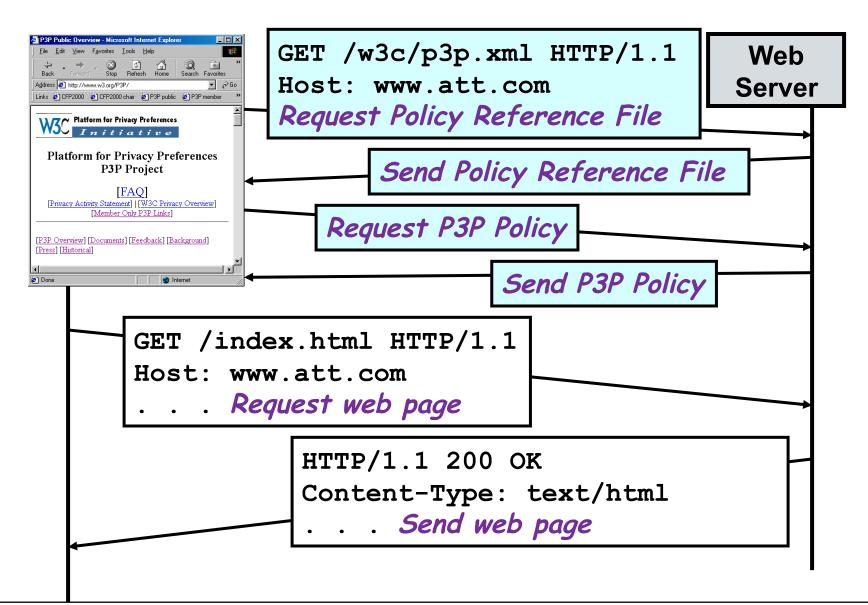
P3P1.0 Spec Defines

- 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 simple HTTP transaction



... with P3P 1.0 added



Transparency

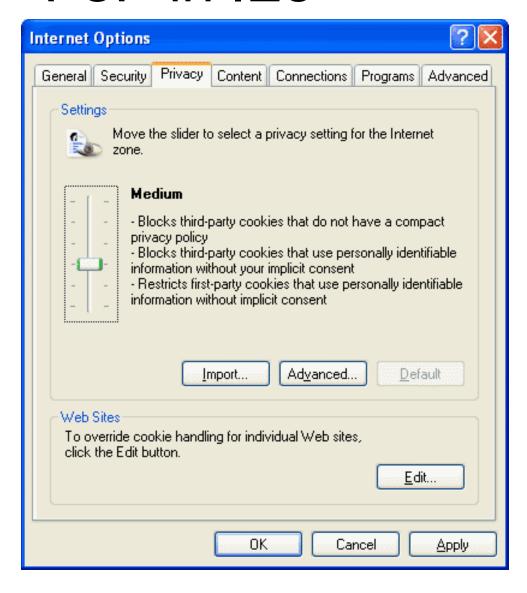
- P3P clients can check a privacy policy each time it changes
- P3P clients can check privacy policies on all objects in a web page, including ads and invisible images

http://www.att.com/accessatt/



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

P3P in IE6

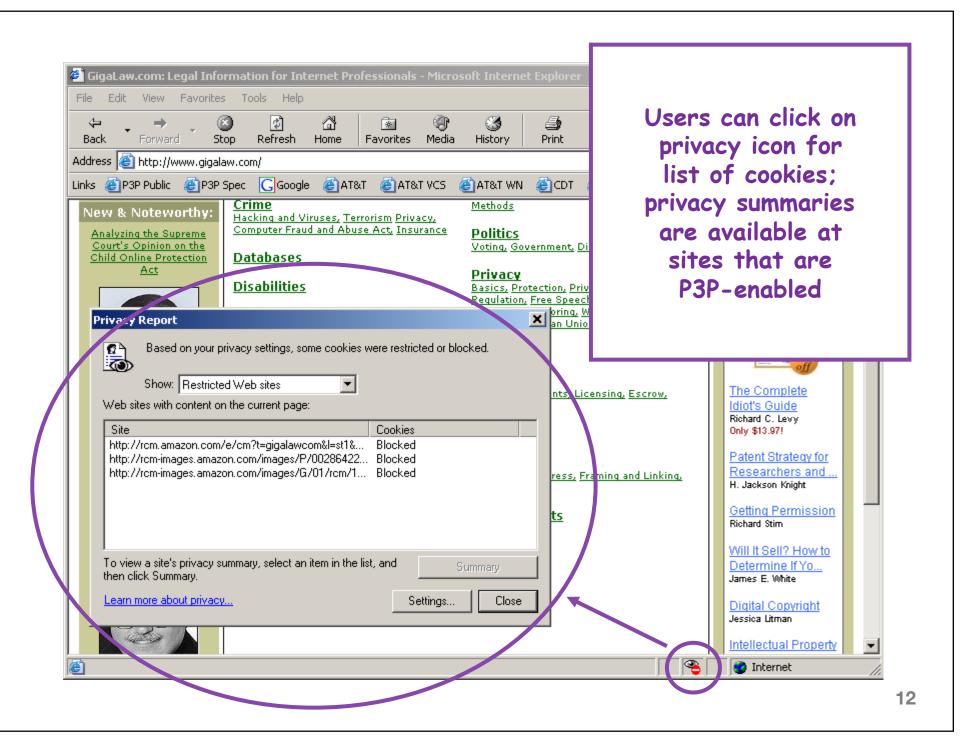


Automatic processing of compact policies only; third-party cookies without compact policies blocked by default



Privacy icon on status bar indicates that a cookie has been blocked - pop-up appears the first time the privacy icon appears



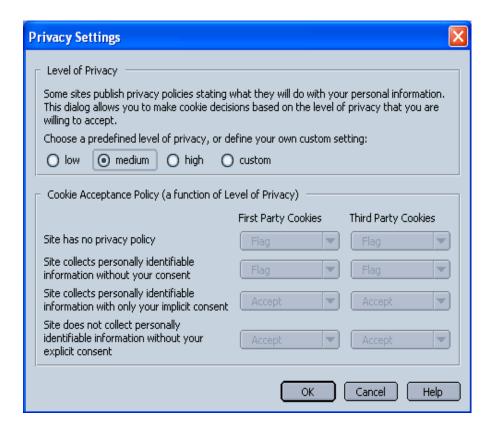




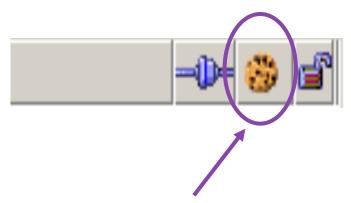
Privacy summary report is generated automatically from full P3P policy



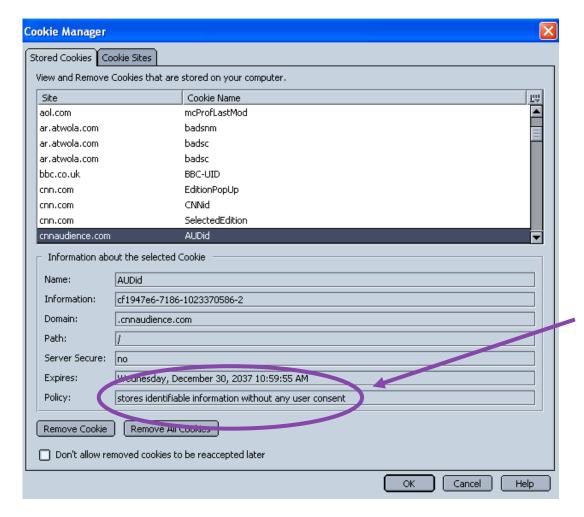
P3P in Netscape 7



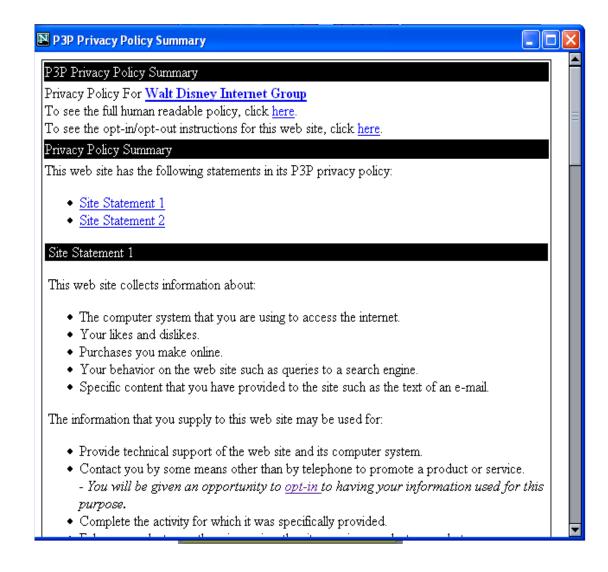
Preview version similar to IE6, focusing, on cookies; cookies without compact policies (both first-party and third-party) are "flagged" rather than blocked by default



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

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

Byers, S., Cranor, L. F., and Kormann, D. 2003. Automated analysis of P3P-enabled Web sites. ICEC '03, vol. 50. ACM Press, New York, NY, 326-338. DOI=http://doi.acm.org/10.1145/948005.948048

Web site data practices 2003

- Most sites collect PII, but few collect sensitive information
- Most sites use data for marketing and pseudonymous profiling
 - Telemarketing and identified profiling is less common
 - 72% of sites offer choices about marketing
- 49% of sites share data with parties other than agents using data for purpose it was provided, but 46% of these offer choice
 - We suspect percentage offering choice is actually higher but sites using old version of P3P spec can't disclose this
- 92% sites that collect identified data provide some access provisions
- 34% of sites offer privacy-related dispute resolution options involving an independent organization (such as a privacy seal)
- 63% of sites do not have data retention policy for all data

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

Privacy policy	P3P policy
Designed to be read by a human	Designed to be read by a computer
Can contain fuzzy language with "wiggle room"	Mostly multiple choice – sites must place themselves in one "bucket" or another
Can include as much or as little information as a site wants	Must include disclosures in every required area
Easy to provide detailed explanations	Limited ability to provide detailed explanations
Sometimes difficult for users to determine boundaries of what it applies to and when it might change	Precisely scoped
Web site controls presentation	User agent controls presentation

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

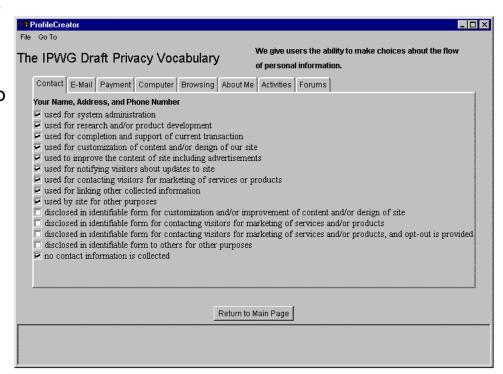
- Privacy policies are complex
- User privacy preferences are often complex and nuanced
- Users tend to have little experience articulating their privacy preferences
- Users are generally unfamiliar with much of the terminology used by privacy experts

Iterative design approach

- Four P3P user agent prototypes developed over 4-year period while P3P specification was under development
 - 1997 W3C prototype
 - 1999 Privacy Minder
 - 2000 AT&T/Microsoft browser helper object
 - 2001 AT&T usability testing prototype
- AT&T Privacy Bird beta released publicly Feb.
 2002
 - August 2002 user study
 - Beta 1.2 released Feb. 2003

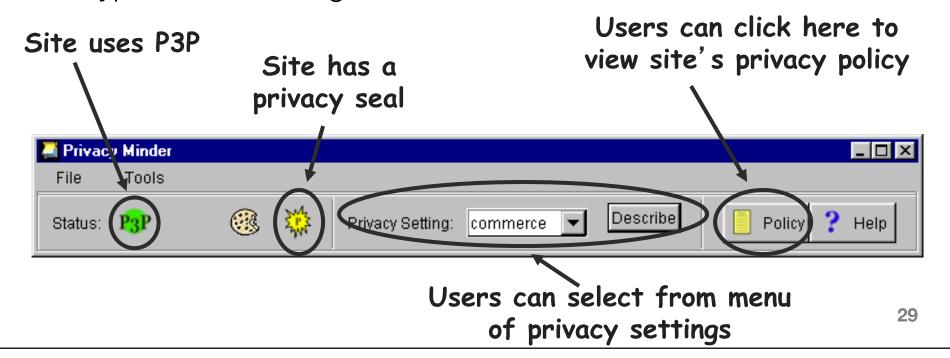
W3C prototype

- 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



Privacy Minder

- Proxy-based P3P user agent based on early P3P draft
- All configuration done through APPEL files
- Privacy Minder came with several APPEL files representing typical user settings



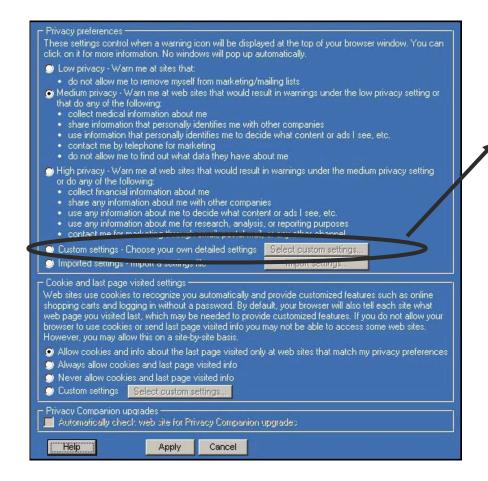
AT&T/Microsoft browser helper object

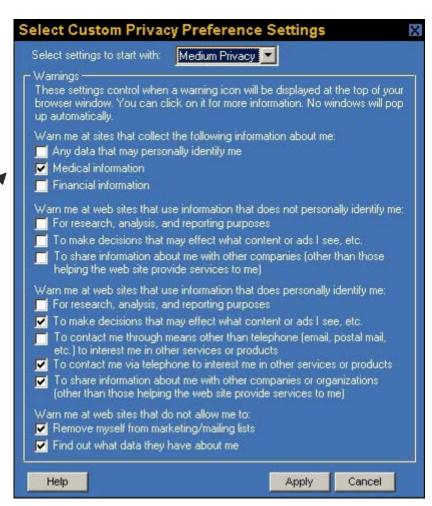
- Based on nearly-finished P3P spec
- Implemented as IE5 browser helper object, added privacy button to browser toolbar
- Preference configuration designed to fit on one screen
- Instead of offering every combination of preferences, we used survey data to focus on 12 areas of concern
- Included glossary of privacy jargon on preference screen, but users ignored it
- Asked users to indicate acceptable practices, too difficult
- Stored preferences as APPEL files

Set My Preferences	Privacy Check Results	View Privacy Rules (APPEL)	Site's Privacy Policy (in XML)		Demo User Agent Help		
Web sites can:	Nosaits	Trailes (ATT EE)	OK for visited site	Visited site can share this info	Definitions of italicized terms Internal uses includes such things as completing transactions, troubleshooting		
1. Collect only the data necessary to process my specific request			V		customer problems, and		
2. PLUS: Collect data for internal uses only (choose one):					customizing website content to customers' interests, but		
Only data that does not reveal my identity			0		not marketing or profiling		
 Data that does identify me except for medical and/or financial information 			•		Profiling means collecting data about your interests and habits to predict other things you might want or do		
• Any data about me, including medical and/or financial information			0				
3. PLUS: Collect data for other purposes:					Marketing purposes means		
 Collect data that does not identify me for profiling 			V		contacting you to try to interest you in other products Sharing info means selling or giving data to organizations or people external to the organization		
Collect data that does identify me for <i>profiling</i>			V				
Collect data that does identify me for <i>marketing purposes</i>			V				
Features you may require from Web site:		I Require		represented on the website			
Ability to remove myself from marketing/mailing lists			V				
Ability to find out what data they have about me							
informs you when the site	does not prevent a web site f 's policy violates your privacy and/or supply information.	rom collecting data it only perferences so you can decide					

AT&T usability testing prototype

- Another browser helper object implementation
- Simplified language to eliminate need for glossary
- Preferences asked for unacceptable rather than acceptable practices
- Users presented with high, medium, low, and custom settings
- Custom settings offered 13 choices
- Users found preference setting navigation confusing



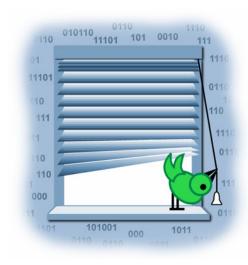


AT&T Privacy Bird

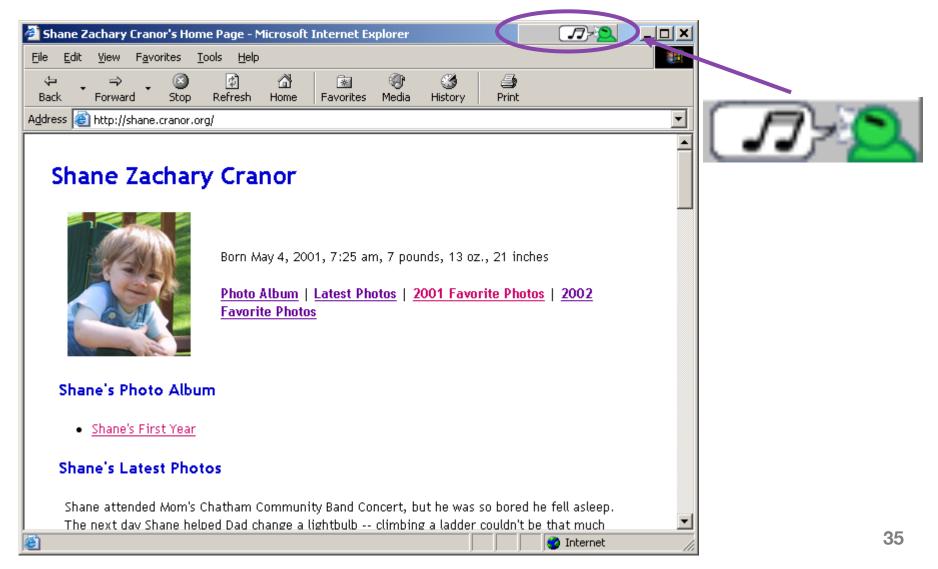
- Free download of beta from http://privacybird.com/
- "Browser helper object" for IE 5.01/5.5/6.0
- Reads P3P policies at all P3P-enabled sites automatically



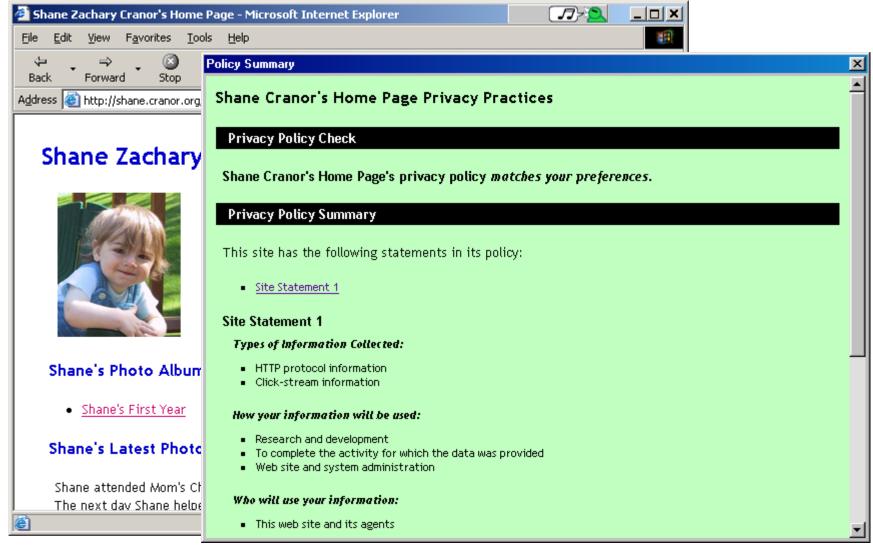
- Clicking on bird icon gives more information
- Current version is information only no cookie blocking



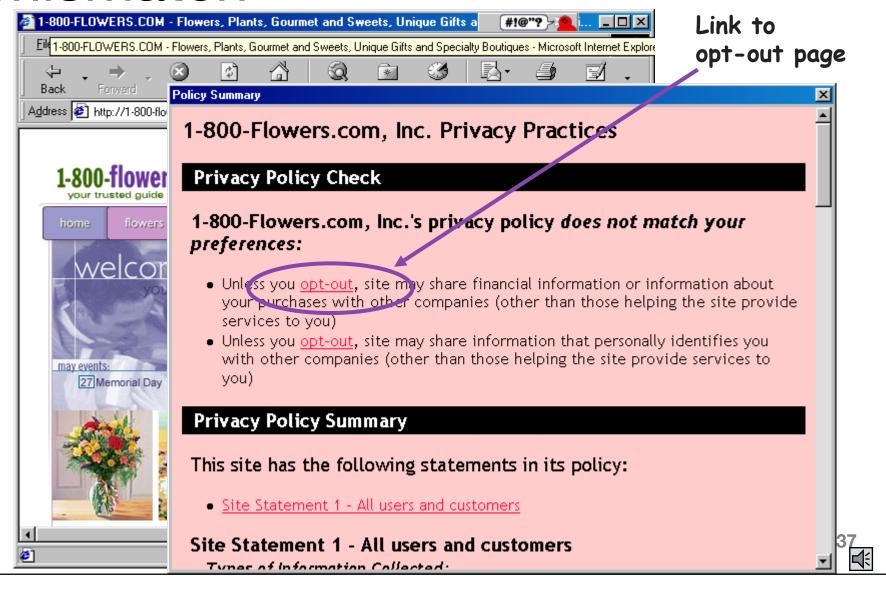
Chirping bird is privacy indicator



Click on the bird for more info



Privacy policy summary - mismatch



Expand/collapse added in beta 1.2

Policy Summary

Federal Trade Commission

Privacy Policy Check

Federal Trade Commission's priva

Privacy Policy Summary

- + Policy Statement 1 Basic Infor
- + Policy Statement 2 Data Colle
- Access to your information
 This site allows you to access your coabout you from its records
- + How to reach this site
- + How to resolve privacy-relate(

More Information

Policy Summary

+ Federal Trade Commission Privacy Practices

Privacy Policy Check

Federal Trade Commission's privacy policy matches your preference

Privacy Policy Summary

- Policy Statement 1 - Basic Information

Data collected from all Web users: access logs, and search strings (if entered).

Types of Information that may be collected:

- search terms
- click-stream information

How your information may be used:

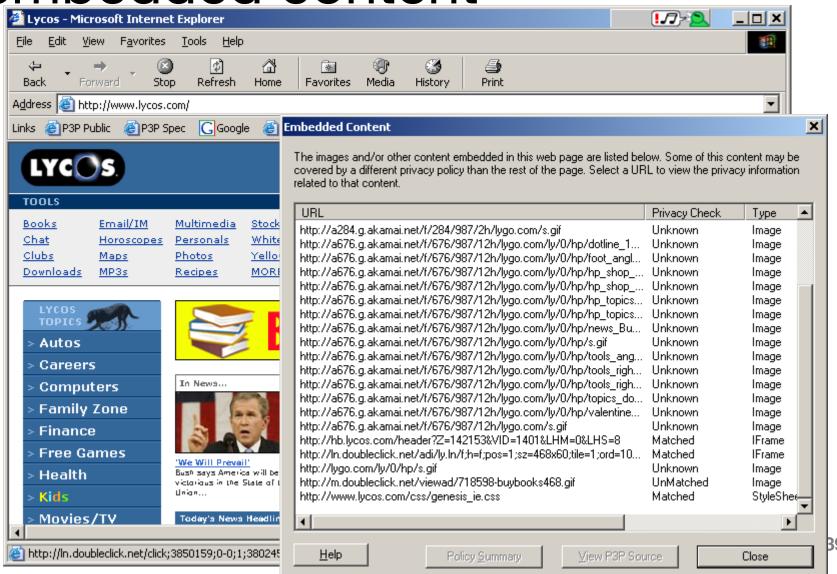
- · To complete the activity for which the data was provided
- To do web site and system administration

Who may use your information:

This web site and the companies that help the site provide services to you.

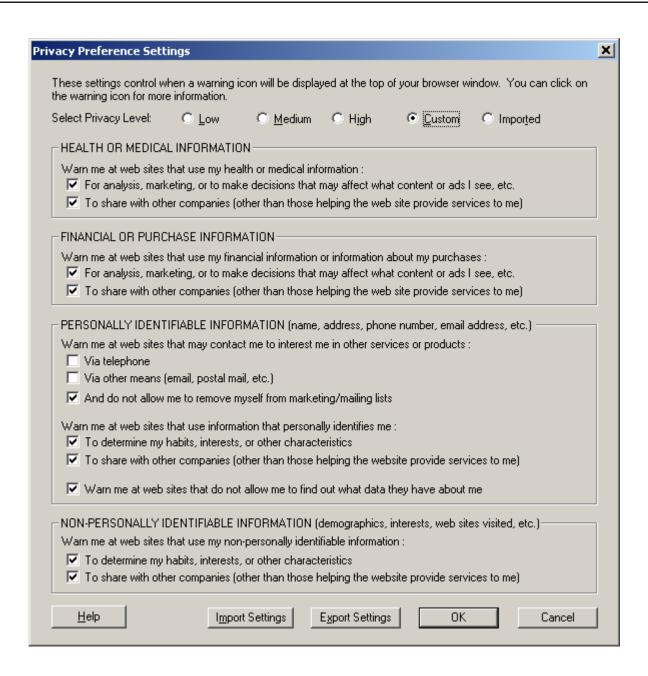
Click + for more -

Bird checks policies for embedded content



Privacy Bird icons





Evaluating P3P user agents

- Questions
 - Does P3P user agent perform useful function?
 - Can users use it effectively?
- Evaluation techniques
 - User survey
 - Laboratory study

Privacy Bird user survey

- ~20,000 downloads in first six months of beta trial
- Users asked whether they were willing to participate in survey when they downloaded software
- 2000 email addresses randomly selected from those willing to participate
- Sent invitation to fill out online 35-question survey

Demographics and Internet use

- Sample was older, more male, better educated, and had more Internet experience than random sample
- Most respondents from English speaking countries 70% from US, 14% from Australia, 6% from Canada
- US respondents had more Internet experience and more likely to have made purchases from web sites
- Are our skewed survey respondent demographics representative of Privacy Bird users?
- Are our demographics similar to demographics of users of other privacy software?

Attitudes about privacy

- 34% never heard of P3P
- 21% identified as "P3P experts"
- Most never or occasionally read privacy policies before installing Privacy Bird
- Level of privacy concern similar to other studies
- Our respondents more knowledgeable and concerned about cookies than typical Internet users
- Our respondents are not very knowledgeable about third-party cookies – 18% never heard of them, 41% heard of them but don't really know what they are
- P3P experts more knowledgeable about third-party cookies and less concerned about cookies

General evaluation of Privacy Bird

- Beta had some installation and stability problems that showed up on only some systems
- Frequent criticism: too many yellow birds!
 - In August 2002, E& Y reported 24% of to 100 domains visited by US Internet users were P3P enabled
- Average usefulness on 5 point scale (5=very useful)
 - Today: 2.9
 - If most web sites P3P-enabled: 4.0
 - If Privacy Bird could block cookies at sites with red bird: 4.1
- Women and non-US respondents found Privacy Bird most useful and more likely to recommend to a friend
- Average ease-of-use on 5 point scale (5=very easy)
 - Installation: 4.6
 - Changing privacy settings: 3.9
 - Understanding policy summary: 3.3

Policy summary

- Amount of information in policy summary
 - Right amount: 64%
 - Too much: 15%
 - Not enough: 20%
- No specific suggestions about what additional information to include
- How often did you look at policy summary?
 - Never: 15%
 - Once or twice: 34%
 - Several times: 36%
 - Ten or more times: 15%
- In beta 1.2 we reworded policy summary slightly and added expand/ collapse

Privacy settings

- How often did you change your privacy settings?
 - Never: 25%
 - Once or twice: 52%
 - Several times: 21%
 - Ten or more times: 2%
- P3P experts changed their settings more frequently
- A few comments that people did not fully understand what all the choices mean

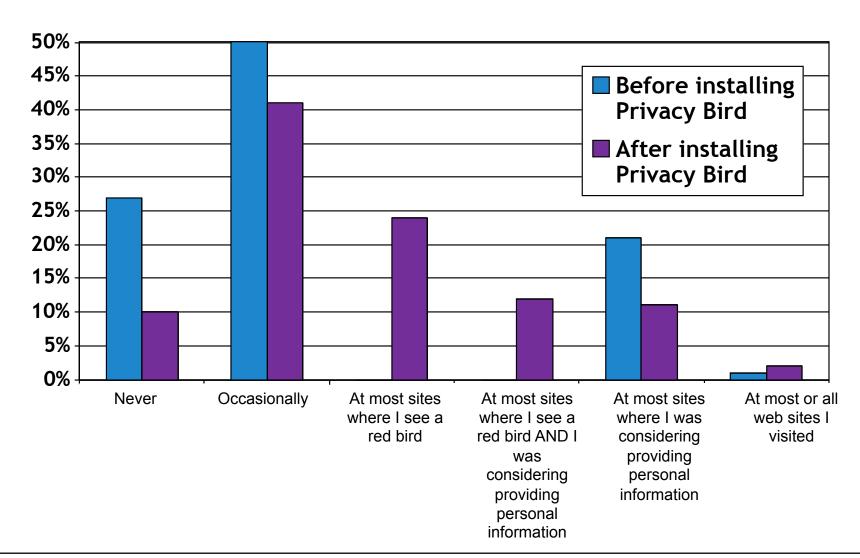
Icon and sounds

- What sound setting did you use?
 - Play sounds at all web sites: 19%
 - Play sounds with certain birds: 37%
 - No sounds: 45%
 - "Oh, how we love the squawking red crow"
 - "I was driven almost to a state of collapse, I used to jump when I heard the same bird call in my yard"
- Some complaints about location of bird in title bar
- In beta 1.2 we introduced a movable bird and a sound option that plays the sound only on the first visit to each site each day

Impact on online behavior

- 88% of respondents indicated some change in online behavior as a result of using Privacy Bird
 - Fill out fewer online forms: 37%
 - Take advantage of opt-outs: 37%
 - Stopped visiting some web sites: 29%
 - Comparing privacy policies at similar sites and frequenting sites with better policies: 18%
 - "Basically, I use Privacy Bird like a warning light. Whenever it's red
 I treat the website as hostile and am extra careful about the
 information I provide and activities I perform there"
 - "I told one mutual fund web site about Privacy Bird's findings, and they improved their pages because of it!"

Respondents who read privacy policies



Impact on online purchasing

- If you could find out before making an online purchase which of the websites that had the item you wanted had the best privacy policy, would you be likely to purchase the item form the site with the best privacy policy?
 - Almost always purchase from site with best privacy policy: 33%
 - Probably purchase from site with best privacy policy as long as price and services similar to other sites: 54%
 - Always purchase from site with best price: 6%
 - Do not plan to make online purchases: 7%

Privacy Bird laboratory study

- 12 experienced IE users (no P3P experience)
- Training on IE6 privacy features & Privacy Bird
- Asked to visit web sites and answer questions about their privacy policies by
 - Using IE6 privacy features
 - Using Privacy Bird
 - Reading privacy policy
- Order of tasks randomized
- Well-known sites with 2-3 page privacy policies and P3P with 2 "statement" elements
- L. Cranor, P. Guduru, and M. Arjula. User Interfaces for Privacy Agents. ToCHI, 2006. http://lorrie.cranor.org/pubs/privacy-bird-20050714.pdf

Questions about privacy policies

- Might site send unsolicited email?
- Might site send info to another company that might send unsolicited email?
- Does site use cookies?
- Does site offer opt-out or unsubscribe options?

Results

- Easier to find info with user agent than reading policies
- Find info fastest with Privacy Bird, slowest with IE6
- Some problems accurately answering questions with IE6 due to bugs in IE6 P3P implementation*
- L. Cranor and J. Reidenberg. Can user agents accurately represent privacy notices?. TPRC 2002 (September 2002). http://papers.ssrn.com/sol3/papers.cfm? abstract_id=32886

Rating of PB and IE6 (1-5)

	Privacy Bird	IE6
Usefulness	4.2	3.3
Likely to use in the future	4.6	3.5
Likely to recommend to a friend	4.6	2.8
Ease of understanding policy summary	4.0	2.7
Ease of finding information	4.2	2.8

Testing "ours" recipient terms

- IE6: Information may be used by this web site, entities for whom it is acting as an agent, and/or entities acting as its agent. An agent in this instance is defined as a third party that processes data only for the completion of the stated purpose, such as a shipping firm or printing service.
- Privacy Bird: Information may be used by this web site and the companies that help the site provide services to you (such companies must use your information only on behalf of this web site for the purposes stated in this policy).
- New alternative: Information may be used by this web site and the companies that help the site fulfill your requests (for example, shipping or billing companies -- such companies may not use your information for marketing or other purposes that go beyond fulfilling your request).
- P3P 1.1: With whom we may share your information: Companies that help us fulfill your requests (for example, shipping a product to you), but these companies must not use your information for any other purpose

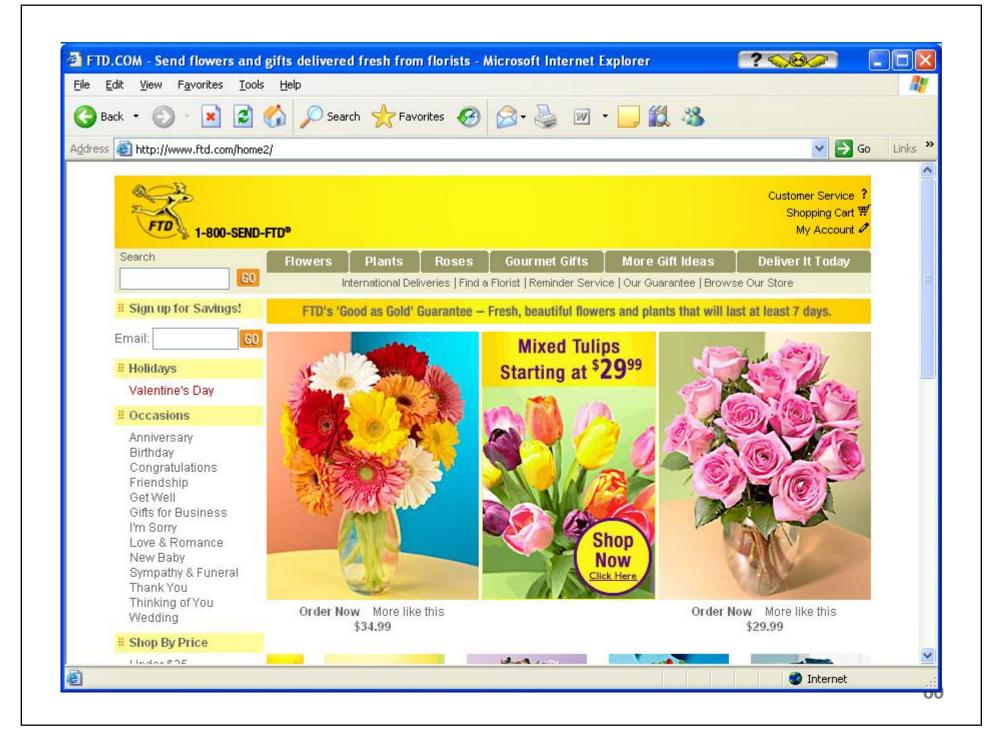
Privacy Bird icons

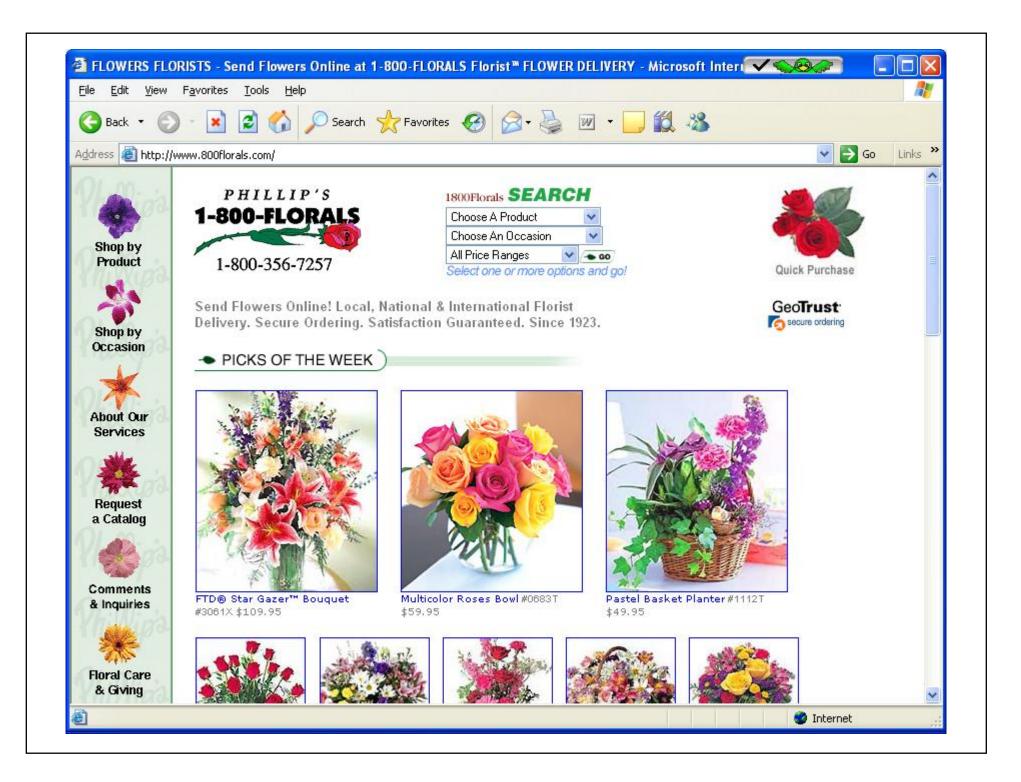


Privacy policy matches user's privacy preferences

Privacy policy
does not
match user's
privacy
preferences

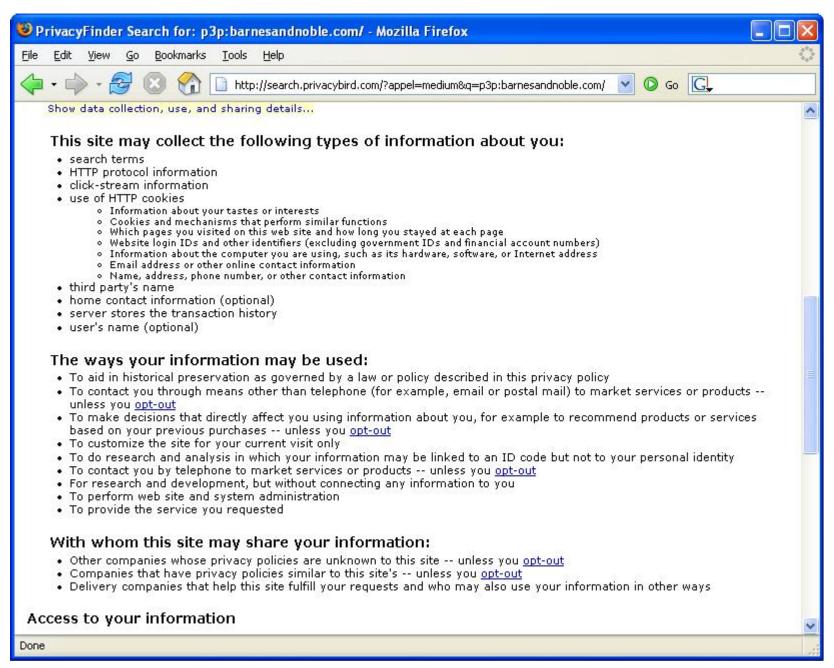
Example: Sending flowers





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



Is Privacy Finder useful?

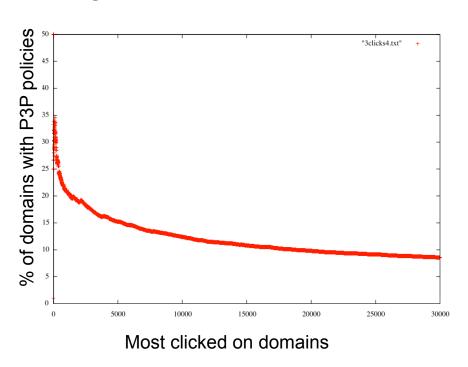
- Do users care about web site privacy?
- Have enough web sites adopted P3P that typical search results contain sites with P3P policies?
 - Do users have meaningful choices among privacy policies?
- Do users understand information provided by Privacy Finder?
- Does Privacy Finder influence online purchasing decisions?

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
- L. Cranor, S. Egelman, S. Sheng, A. McDonald, and A. Chowdhury. P3P Deployment on Websites. Electronic Commerce Research and Applications, 2008.

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
 - 5% of sites in our cache have
 P3P
 - 9% of 30K most clicked on domains have P3P
 - 17% of clicks to 30K most clicked on domains 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. Paper presented at the Workshop on the Economics of Information Security, June 7-8, 2007, Pittsburgh, PA.

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

One policy or many?

- P3P allows policies to be specified for individual URLs or cookies
- One policy for entire web site (all URLs and cookies) is easiest to manage
- Multiple policies can allow more specific declarations about particular parts of the site
- Multiple policies may be needed if different parts of the site have different owners or responsible parties (universities, CDNs, etc.)

Third-party content

- Third-party content should be P3P-enabled by the third-party
- If third-party content sets cookies, IE6 will block them by default unless they have P3P compact policy
- Your first-party cookies may become thirdparty cookies if your site is framed by another site, a page is sent via email, etc.

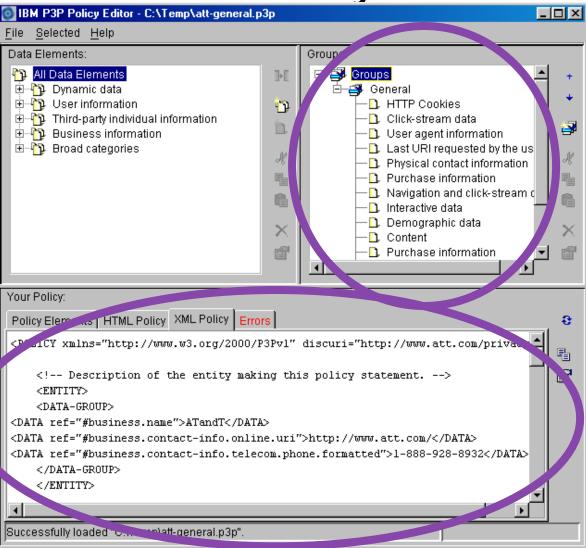
Cookies and P3P

- P3P policies must declare all the data stored in a cookie as well as any data linked via the cookie
- P3P policies must declare all uses of stored and linked cookie data
- Sites should not declare cookie-specific policies unless they are sure they know where their cookies are going!
 - Watch out for domain-level cookies
 - Most sites will declare broad policy that covers both URLs and cookies

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

Locating the policy reference file

- Place policy reference file in "well known location" /w3c/p3p.xml
 - Most sites will do this
- Use special P3P HTTP header
 - Recommended only for sites with unusual circumstances, such as those with many P3P policies
- Embed link tags in HTML files
 - Recommended only for sites that exist as a directory on somebody else's server (for example, a personal home page)

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
- Procedure depends on server see P3P Deployment Guide appendix http://www.w3.org/TR/p3pdeployment or Appendix B of Web Privacy with P3P

Don't forget to test!

- Make sure you use the P3P validator to check for syntax errors and make sure files are in the right place http://www.w3.org/P3P/validator/ or http://validator.privacyfinder.org/
 - But validator can't tell whether your policy is accurate
- Use P3P user agents to view your policy and read their policy summaries carefully
- Test multiple pages on your site

XML syntax basics

Element opening tag

```
Attribute
Element that <BIG-ELEMENT>
doesn't contain <element name="value" />
(ending slash) </BIG-ELEMENT>
                                        Comment
Element
          <!-- This is a comment
closing tag
(beginning
slash)
          <ELEMENT>Sometimes data goes
          between opening and closing
Element
that contains
          tags</ELEMENT>
character
data
```

Assertions in a P3P policy

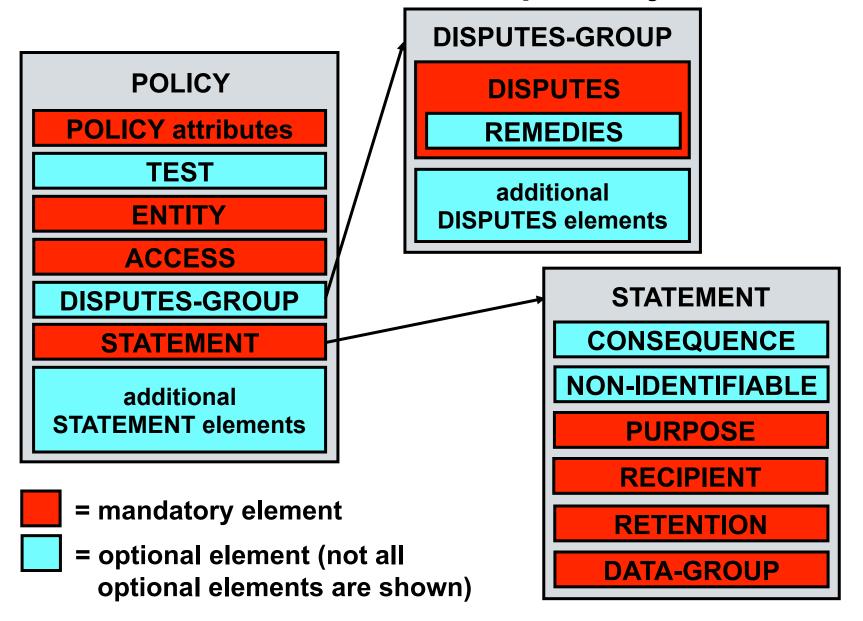
General assertions

- 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>

Structure of a P3P policy



Example privacy policy

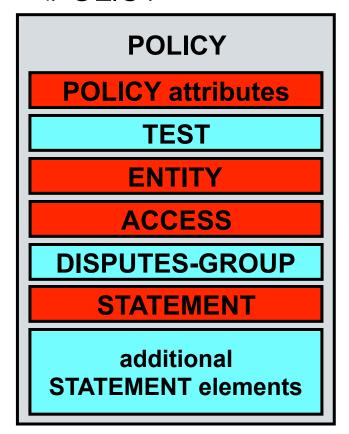
• We do not currently collect any information from visitors to this site except the information contained in standard web server logs (your IP address, referer, information about your web browser, information about your HTTP requests, etc.). The information in these logs will be used only by us and the server administrators for website and system administration, and for improving this site. It will not be disclosed unless required by law. We may retain these log files indefinitely. Please direct questions about this privacy policy to privacy@p3pbook.com.

```
P3P/XML encoding
                                                                      P3P version
      <POLICIES xmlns="http://www.w3.org/2002/01/P3Pv1"> *
      <POLICY discuri="http://p3pbook.com/privacy.html"
                                                                    Location of
              name="policy">
                                         P3P policy name
                                                                    human-readable
        <ENTITY>
                                                                    privacy policy
         <DATA-GROUP>
          <DATA
Site's
            ref="#business.contact-info.online.email">privacy@p3pbook.com
name
          </DATA>
and
          <DATA
contact
            ref="#business.contact-info.online.uri">http://p3pbook.com/
info
          </DATA>
          <DATA ref="#business.name">Web Privacy With P3P</DATA>
        </DATA-GROUP>
                                            Access disclosure
         </ENTITY>
                                                                Human-readable
        <access><nonident/></access>
                                                                explanation
        <STATEMENT>
          <CONSEQUENCE>We keep standard web server logs.</CONSEQUENCE>
          <PURPOSE><admin/><current/><develop/></PURPOSE> ← How data may
   Statement
          <RECIPIENT><ours/></RECIPIENT>
                                                                    be used
          <RETENTION><indefinitely/></RETENTION>
                                                             Data recipients
          <DATA-GROUP>
             <DATA ref="#dynamic.clickstream"/>
                                                            Data retention policy
             <DATA ref="#dynamic.http"/>
          </DATA-GROUP>
                                                      Types of data collected
         </STATEMENT>
      </POLICY>
      </POLICIES>
                                                                              83
```

The POLICY element

- Contains a complete P3P policy
- Takes mandatory discuri attribute
 - indicates location of human-readable privacy policy
- Takes opturi attribute (mandatory for sites with opt-in or opt-out)
 - Indicates location of opt-in/opt-out policy
- Takes mandatory name attribute
- Sub-Elements
 - <EXTENSION>, <TEST>, <EXPIRY>,<DATASCHEMA>, <ENTITY>,<ACCESS>, <DISPUTES-GROUP>,<STATEMENT>, <EXTENSION>

- Example
 - < POLICY



The TEST element

- Used for testing purposes
 - Presence indicates that policy is for testing purposes and MUST be ignored
- Prevents misunderstandings during initial P3P deployment
- <TEST/>

The ENTITY element

- Identifies the legal entity making the representation of the privacy practices contained in the policy
- Uses the business.name data element and (optionally) other fields in the business data set (at least one piece of contact info required)
- Example

```
• <ENTITY> <DATA-GROUP> <DATA ref="#business.name">CatalogExample</DATA> <DATA ref="#business.contact-info.telecom.telephone. intcode">1</DATA> <DATA ref="#business.contact-info.telecom.telephone. loccode">248</DATA> <DATA ref="#business.contact-info.telecom.telephone. number">3926753</DATA> </DATA> </DATA-GROUP> </FNTITY>
```

The ACCESS Element

- Indicates the ability of individuals to access their data
 - <nonident/>
 - < all />
 - <contact-and-other/>
 - <ident-contact/>
 - <other-ident/>
 - <none/>
- Example
 - <ACCESS><nonident/></ACCESS>

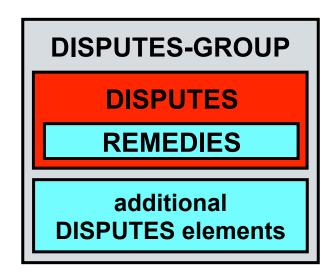
The DISPUTES Element

- Describes a dispute resolution procedure
 - may be followed for disputes about a service's privacy practices
- Part of a <DISPUTES-GROUP>
 - allows multiple dispute resolution procedures to be listed

- Attributes:
 - resolution-type
 - customer service
 - independent organization
 - court
 - applicable law
 - service
 - short-description (optional)
 - Verification (optional)
- Sub-Flements
 - <IMAGE> (optional)
 - <LONG-DESCRIPTION> (optional)
 - <REMEDIES> (optional)

The REMEDIES element

- Sub element of DISPUTES element
- Specifies possible remedies in case a policy breach occurs
 - <correct/>, <money/>, <law/>
- Example of DISPUTES and REMEDIES



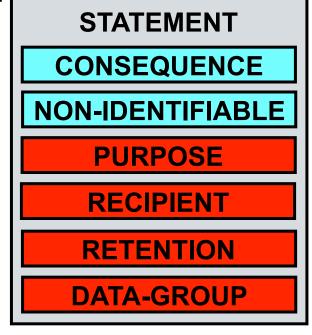
```
<DISPUTES-GROUP>
  <DISPUTES resolution-type="law"
service="http://www.ftc.gov/bcp/conline/edcams/kidzprivacy/" short-
description="Children's Online Privacy Protection Act of 1998, and Federal
Trade Commission Rule">
      <REMEDIES><law/></REMEDIES>
      </DISPUTES>
  </DISPUTES-GROUP>
```

The STATEMENT element

- Data practices applied to data elements
 - mostly serves as a grouping mechanism

Contains the following sub-elements

- <CONSEQUENCE> (optional)
- <NON-IDENTIFIABLE> (optional)
- <PURPOSE>
- <RECIPIENT>
- <RETENTION>
- <DATA-GROUP>



The CONSEQUENCE element

- Consequences that can be shown to a human user to explain why the suggested practice may be valuable in a particular instance, even if the user would not normally allow the practice
- Example
- CONSEQUENCE>We offer a 10% discount to all individuals who join our Cool Deals Club and allow us to send them information about cool deals that they might be interested in.
 CONSEQUENCE>

The NON-IDENTIFIABLE element

- Can optionally be used to declare that no data or no identifiable data is collected
 - non-identifiable: there is no reasonable way to attach collected data to identity of a natural person, even with assistance from a third-party
 - Stronger requirements than non-identified
- Must have a human readable explanation how this is done at the discuri
- Other STATEMENT elements are optinal when NON-IDENTIFIABLE is present
- <NON-IDENTIFIABLE/>

The PURPOSE element

- Purposes of data collection, or uses of data
 - <current/>
 - <admin/>
 - <develop/>
 - <tailoring/>
 - <pseudo-analysis/>
 - <pseudo-decision/>
 - <individual-analysis/>
 - <individual-decision/>
 - <contact/>
 - <historical/>
 - <telemarketing/>
 - <other-purpose/>

- Optional attribute:
 - required
 - always (default)
 - opt-in
 - opt-out
- Example
- <PURPOSE>
 <current/><admin/><develop
 required="opt-out"/></PURPOSE>

Customization purposes

Purpose	Does this involve creating a profile of the user?	How is the user identified?	Does this result in a decision that directly affects the user?
Research and development	No	user is not identified	No
One-time tailoring	No	user may not be identified at all, or may be identified with a pseudonym or with personally-identifiable information	Yes
Pseudonymous analysis	Yes	pseudonym	No
Pseudonymous decision	Yes	pseudonym	Yes
Individual analysis	Yes	personally- identifiable information	No
Individual decision	Yes	personally- identifiable information	Yes 94

The RECIPIENT element

- Recipients of the collected data
 - <ours>
 - <delivery>
 - <same>
 - <other-recipient>
 - <unrelated>
 - <public>
- Optional attribute
 - required
 - always (default)
 - opt-in
 - opt-out
- Optional sub-element
 - <recipient-description>

- Example
- <RECIPIENT>
 - <ours/>
 - <same required=
 - "opt-out"/>
 - <delivery>
 - <recipient-description>
 - FedEx
 - </recipient-description>
 - </delivery>
 - </RECIPIENT>

The RETENTION element

- Indicates the kind or retention policy that applies to the referenced data
 - <no-retention/>
 - <stated-purpose/>
 - <legal-requirement/>
 - <business-practices/>
 - <indefinitely/>

Requires publishing of destruction timetable linked from human-readable privacy policy

- Example
- <RETENTION><indefinitely/></RETENTION>

The DATA element

- Describes the data to be transferred or inferred
- Contained in a DATA-GROUP
- Attributes:
 - ref
 - optional (optional, default is no, not optional=required)
- Sub-Elements:
 - <CATEGORIES>
- Example

```
    <DATA-GROUP>
        <DATA ref="#dynamic.miscdata">
            <CATEGORIES>
            <preference/><political/>
            </CATEGORIES>
            </DATA>
            <DATA ref="#user.home-info" optional="yes"/>
            </DATA-GROUP>
```

The CATEGORIES element

Provides hints to user agents as to the intended uses of the data

- Physical contact information
- Online contact information
- Unique identifiers
- Purchase information
- Financial information
- Computer information
- Navigation and clickstream data
- Interactive data

- Demographic and socio-economic data
- Content
- State management mechanisms
- Political information
- Health information
- Preference data
- Government-issued identifiers
- Location information
- other

Base Data Schema

- User data user
 - name, bdate, cert, gender, employer, department, jobtitle, home-info, business-info
- Third party data thirdparty
 - Same as user
- Business data business
 - name, department, cert, contact-info
- Dynamically generated Dynamic
 - clickstream, http, clientevents, cookies, miscdata, searchtext, interactionrecord

dynamic.miscdata

- Used to represent data described only by category (without any other specific data element name)
- Must list applicable categories
- Example

Custom data schemas

- You can define your own data elements
- Not required you can always use categories
- May be useful to make specific disclosures, interface with back-end databases, etc.
- Use the <DATASCHEMA> element
 - Embedded in a policy file or in a stand-alone
 XML file

Extension mechanism

- <EXTENSION> describes extension to P3P syntax
- optional attribute indicates whether the extension is mandatory or optional (default is optional="yes")
 - Optional extensions may be safely ignored by user agents that don't understand them
- Only useful if user agents or other P3P tools know what to do with them
- Example (IBM GROUP-INFO extension used to add name attribute to STATEMENT elements)

```
    <STATEMENT>
    <EXTENSION optional="yes">
         <GROUP-INFO xmlns= "http://www.software.ibm.com/P3P/editor/extension-1.0.html"
         name="Site management"/>
         </EXTENSION>
         . . .
    </STATEMENT>
```

Compact policy syntax

- Part of P3P Header
 - P3P: CP="NON NID DSP NAV CUR"
- Represents subset of P3P vocabulary
 - ACCESS (NOI ALL CAO IDC OTI NON)
 - CATEGORIES (PHY ONL UNI PUR ... OTC)
 - DISPUTES (DSP)
 - NON-IDENTIFIABLE (NID)
 - PURPOSE (CUR ADM DEV CUS ... OTP) aio
 - RECIPIENT (OUR DEL SAM UNR PUB OTR) aio
 - REMEDIES (COR MON LAW)
 - RETENTION (NOR STP LEG BUS IND)
 - TEST (TST)

Policy reference files (PRF)

- Allows web sites to indicate which policy applies to each resource (URL or cookie)
 - Every resource (HTML page, image, sound, form action URL, etc.) can have its own policy
- User agents can cache PRFs (as long as permitted by EXPIRY) so they don't have to fetch a new PRF every time a user clicks

PRF elements

- <EXPIRY>
 - Determines how long PRF is valid default is 24 hours
- <POLICY-REF>
 - Provides URL of policy in about attribute
- <INCLUDE>, <EXCLUDE>
 - URL prefixes (local) to which policy applies/doesn't apply
- <COOKIE-INCLUDE>, <COOKIE-EXCLUDE>
 - Associates / disassociates cookies with policy if you want a policy to apply to a cookie, you must use <COOKIE-INCLUDE>!
- <METHOD>
 - HTTP methods to which policy applies
- <HINT>
 - Provides URLs of PRFs for third-party content

PRF example

```
<META xmlns="http://www.w3.org/2002/01/P3Pv1" xml:lang="en">
  <POLICY-REFERENCES>
    <EXPIRY max-age="172800"/>
    <POLICY-REF about="http://www.example.com/privacy.xml#policy1">
      <INCLUDE>/</INCLUDE>
      <INCLUDE>/news/*</INCLUDE>
      <EXCLUDE>/news/top/*</EXCLUDE>
    </POLICY-REF>
    <POLICY-REF about="http://www.example.net/pp.xml#policy2">
      <INCLUDE>/news/top/*</INCLUDE>
    </POLICY-REF>
    <POLICY-REF about="/P3P/policies.xml#policy3">
      <INCLUDE>/photos/*</INCLUDE>
      <INCLUDE>/ads/*</INCLUDE>
      <COOKIE-INCLUDE/>
    </POLICY-REF>
    <HINT scope="http://www.example.org"</pre>
      path="/mypolicy/p3.xml"/>
  </POLICY-REFERENCES>
</META>
```

Policy updates

- Changing your P3P policy is difficult, but possible
- New policy applies only to new data (old policy applies to old data unless you have informed consent to apply new policy)
- Technically you can indicate exact moment when old policy will cease to apply and new policy will apply
- But, generally it's easiest to have a policy phase-in period where your practices are consistent with both policies
- Default policy life time is 24 hours, so phase-in period would be just one day for most sites

P3P policy validation

- http://www.w3.org/P3P/validator.html
- http://validator.privacyfinder.org

APPEL

- A P3P Preference Exchange Language
- Working draft, never became an official recommendation
- Allows users to store their preferences and import them into another user agent
- Allows organizations to distribute canned settings files
- Not a very well designed language

Class exercise

- Create a P3P policy for a web site that has a fairly complete privacy policy but no P3P policy
 - For example, http://www.target.com/
 - What questions do you need to ask someone from that company?
 - How will you group data into statements?
 - Where will you put the PRF?



Carnegie Mellon UniversityCyLab



Engineering & Public Policy