

# Internet Privacy and P3P

WWW10 Tutorial  
May 1, 2001

Marc Langheinrich  
ETH Zurich, Switzerland

[www.inf.ethz.ch/~langhein/](http://www.inf.ethz.ch/~langhein/)

# Outline – Part I

---

WWW10 Tutorial – May 1, 2001

- What is Privacy?
  - Definitions
  - Public Concern
- How do they get my Data?
  - Browser Chatter
  - Cookies
  - Ad Networks
  - Web Bugs
  - Spyware
- Solutions
  - Privacy Policies
  - Laws and Regulations
  - Privacy Tools
- Privacy Tools
  - Encryption
  - Anonymity
  - Management
  - Trust

# Outline – Part II

---

WWW10 Tutorial – May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# What is Privacy?

WWW10 Tutorial  
May 1, 2001

- What is Privacy?
  - Definitions
  - Public Concern
- How do they get my Data?
  - Browser Chatter
  - Cookies
  - Ad Networks
  - Web Bugs
  - Spyware
- Solutions
  - Privacy Policies
  - Laws and Regulations
  - Privacy Tools
- Privacy Tools
  - Encryption
  - Anonymity
  - Management
  - Trust

# What is Privacy?

---

WWW10 Tutorial – May 1, 2001

- „The right to be let alone.“
  - Louis Brandeis, 1890 (Harvard Law Review)
- Facets
  - Territorial Privacy
  - Behavioral Privacy / Media Privacy
  - Bodily Privacy
  - Privacy of Communications
  - Information Privacy
- „The desire of people to choose freely under what circumstances and to what extent they will expose themselves, their attitude and their behavior to others.“
  - Alan Westin, 1967 („Privacy and Freedom“)

# Do People Care?

---

WWW10 Tutorial – May 1, 2001

- March 2000 Business Week Poll
  - 63% not comfortable with anonymous online profiling
  - 89% not comfortable with identified online profiling
- August 2000 Pew Internet Poll
  - Most respondents want guarantee of privacy when they go online
  - Many consumers are unaware of how privacy invasions take place and are consequently unable to take advantage of available privacy-enhancing technologies.

# Preferences Vary

---

WWW10 Tutorial – May 1, 2001

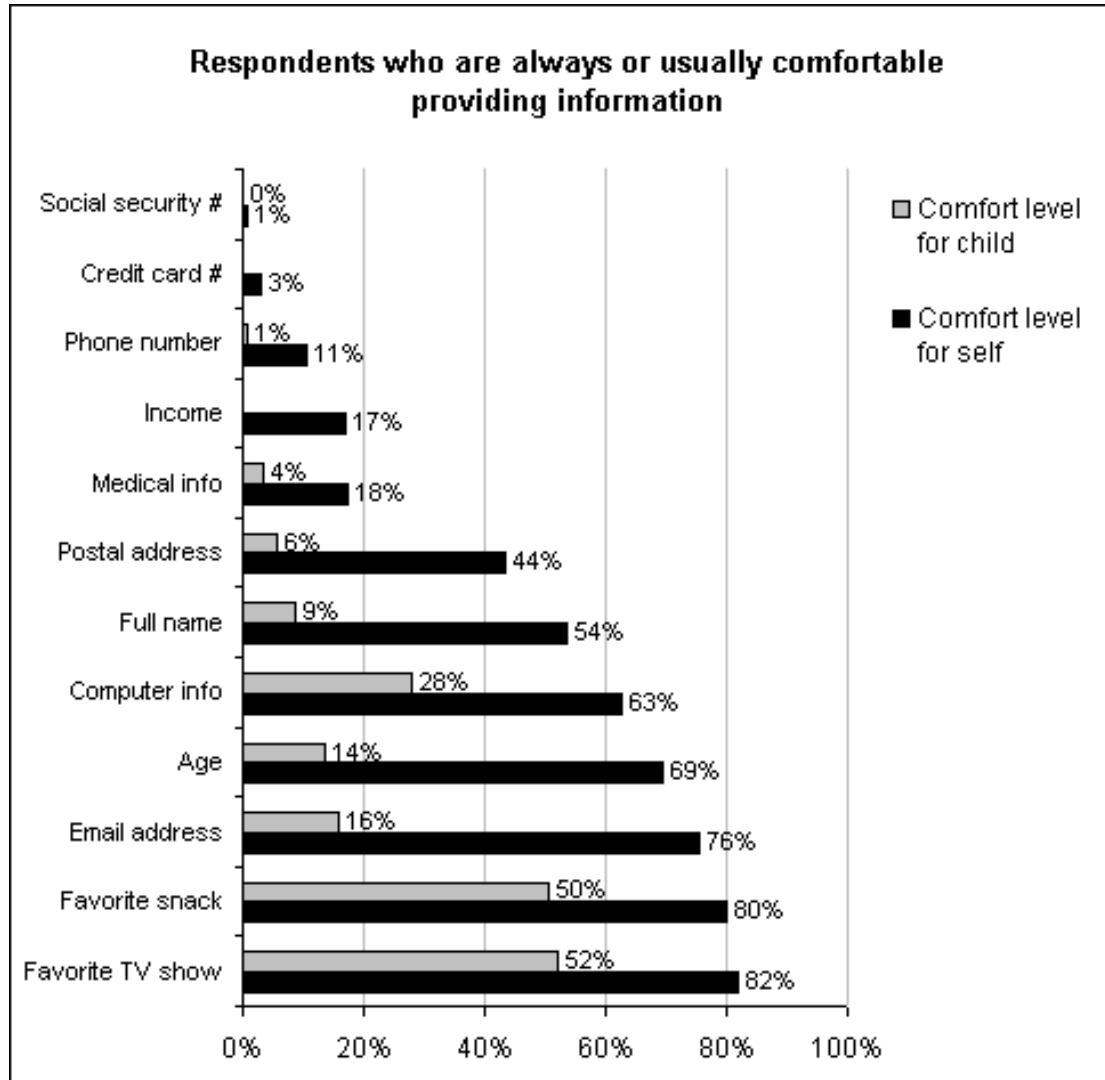
- April 1999 Study „Beyond Concern“
  - Internet users more likely to provide information when they are not identified
  - Acceptance of persistent identifiers (e.g. cookies) varies according to purpose
  - Some types of data more sensitive than others

<http://www.research.att.com/resources/trs/TRs/99/99.4/99.4.3/report.htm>

---

# What Data is Private?

WWW10 Tutorial – May 1, 2001



**Source:** Cranor, Reagle, Ackerman „Beyond Concern: Understanding Net Users' Attitudes About Online Privacy“

# Regional Differences

---

WWW10 Tutorial – May 1, 2001

- IBM-Harris multinational survey
  - Phone interviews with 1000+ adults in each of three countries: US, UK and Germany (10/1999)
  - US:
    - greatest trust in companies, but
    - most likely to actively protect privacy
  - Germany:
    - most comfortable with governmental privacy protection
- Japan's Ministry of Postal & Telecomm. survey
  - interviews with 968 adults, 1999
  - 70% have interest in privacy protection
  - 92% fear that personal information is used unknowingly

[http://www.ibm.com/services/files/privacy\\_survey\\_oct991.pdf](http://www.ibm.com/services/files/privacy_survey_oct991.pdf)

# Web Privacy Concerns

---

WWW10 Tutorial – May 1, 2001

- Data is often collected silently
  - Web allows large quantities of data collected cheaply & unobtrusively
- Data from multiple sources may be merged
  - Non-identifiable information can easily become identifiable when merged
- Users given no meaningful choice
  - Few sites offer alternatives

# How do they get my Data?

WWW10 Tutorial  
May 1, 2001

- What is Privacy?
  - Definitions
  - Public Concern
- How do they get my Data?
  - Browser Chatter
  - Cookies
  - Ad Networks
  - Web Bugs
  - Spyware
- Solutions
  - Privacy Policies
  - Laws and Regulations
  - Privacy Tools
- Privacy Tools
  - Encryption
  - Anonymity
  - Management
  - Trust

# Browsers like to Chatter

---

WWW10 Tutorial – May 1, 2001

- A typical HTTP request:

GET /index.html HTTP/1.0

**User-Agent:** Mozilla/3.01 (X11; I; SunOS 4.1.4 sun4m)

**Host:** www.amazon.com

**Referer:** http://www.alcoholics-anonymous.org/books.html

**Accept:** image/gif, image/x-xbitmap, image/jpeg, \*/\*

**Cookie:** session-id-time=868867200; session-id=6828-2461327-649945; group\_discount\_cookie=F

# Servers like to Record

---

WWW10 Tutorial – May 1, 2001

- A typical Logfile entry:

```
ppp109.bu.edu - - [09/Dec/1996:20:33:22 -500]
```

```
"Get /cgi-bin/wwwais?hemoglobin+gene HTTP/1.0" 200 527
```

- Stores date & time, requested URL, and optionally browser chatter
- Often allows some inference
  - Affiliation: Boston University
  - probably working from home
  - probably student or faculty in biology

# Cookies 101



WWW10 Tutorial – May 1, 2001

II. How do they get my Data?

- Set by the server using the `set_cookie` HTTP header
- Sent (replayed) by the browser with every subsequent request *to this domain only*
- Can be further restricted
  - Replay only to certain servers
  - Replay only at certain path
  - Replay only for certain time (t=0: only until browser is restarted)
- More Info: <http://www.cookiecentral.com/>

# Why Cookies?



WWW10 Tutorial – May 1, 2001

- Cookies can be useful
  - used like a staple to attach multiple parts of a form together [state management]
  - used to identify you when you return to a web site so you don't have to remember a password
  - used to help web sites understand how people use them [click-trails]
- Cookies can be harmful
  - used to profile users and track their activities, especially across web sites

# Ad Networks

WWW10 Tutorial – May 1, 2001



# Referer Log Problems

---

WWW10 Tutorial – May 1, 2001

- GET methods result in form values in URL
- These URLs are sent in the `referer:` header to next host
- Example:

```
http://www.merchant.com/cgi_bin/order?name=John+Do  
e&address=here+there&credit+card=234876923234&ex  
pires=0902& -> index.html
```

# Online and Offline Merging

WWW10 Tutorial – May 1, 2001

- In November 1999, DoubleClick purchased Abacus Direct, a company possessing detailed consumer profiles on more than 90% of US households.
- In mid-February 2000 DoubleClick announced plans to merge “anonymous” online data with personal information obtained from offline databases
- By the first week in March 2000 the plans were put on hold
  - Stock dropped from \$125 (12/99) to \$80 (03/00)

The logo for DoubleClick, with "Double" in black and "Click" in red.The logo for Abacus Direct, with the letters A, B, A, C, U, S each inside a red circle.

A division of DoubleClick Inc.

# Web Bugs



WWW10 Tutorial – May 1, 2001

- Invisible “images” (1-by-1 pixels, transparent color) embedded in Web pages that cause referrer info and cookies to be transferred
- Work just like banner ads from ad networks, but you can’t see them unless you look at the code behind a web page
- Also embedded in HTML formatted email messages

For more info on web bugs see:

<http://www.privacyfoundation.org/resources/webbug.asp>

Find your own bugs at:

<http://users.rcn.com/rms2000/privacy/wbfind.htm>

# Web Bugs + +



WWW10 Tutorial – May 1, 2001

II. How do they get my Data?

- Tracking Word Documents
  - Embedding Web bug as picture in Word
  - Everytime document is openend, Web server can log event
  - Allows even cookies to be set!
- Email Wiretapping
  - Small Javascript as part of HTML msg.
  - Sends BCC: of any forwarded message to wiretapper, *including full text of forwarder!*
  - Mostly illegal by law (constitutes wiretap)

<http://www.privacyfoundation.org/resources/docbug.asp>

# Spyware



WWW10 Tutorial – May 1, 2001

- Spyware: Any Software which employs a user's Internet connection, without their knowledge or explicit permission, to collect information.
  - Most products use pseudonymous, but unique ID.
- Over 800 known freeware and shareware products contain Spyware, for example:
  - Beeline Search Utility
  - GoZilla Download Manager
  - Comet Cursor
- Often hard, if not impossible to uninstall!
- Anti-Spyware Sites
  - <http://grc.com/oo/spyware.htm>
  - <http://www.adcop.org/smallfish>
  - <http://www.spychecker.com>
  - <http://cexx.org/adware.htm>

# Solutions

WWW10 Tutorial  
May 1, 2001

- What is Privacy?
  - Definitions
  - Public Concern
- How do they get my Data?
  - Browser Chatter
  - Cookies
  - Ad Networks
  - Web Bugs
  - Spyware
- Solutions
  - Privacy Policies
  - Laws and Regulations
  - Privacy Tools
- Privacy Tools
  - Encryption
  - Anonymity
  - Management
  - Trust

# Some Solutions

---

WWW10 Tutorial – May 1, 2001

- Privacy Policies
- Voluntary Guidelines and Codes of Conduct
- Seal Programs
- Laws and Regulations
- Privacy Tools

# Privacy Policies

---

WWW10 Tutorial – May 1, 2001

- Policies let consumers know about site's privacy practices
- Consumers can then decide whether or not practices are acceptable, when to opt-in or opt-out, and who to do business with
- The presence of privacy policies increases consumer trust

# Privacy Policy Drawbacks

---

WWW10 Tutorial – May 1, 2001

- BUT policies are often
  - difficult to understand
  - hard to find
  - take a long time to read
    - usually 3-4 pages!
  - changed without notice

# Voluntary Guidelines

---

WWW10 Tutorial – May 1, 2001

- Online Privacy Alliance  
<http://www.privacyalliance.org>
- Direct Marketing Association  
Privacy Promise  
<http://www.thedma.org/library/privacy/privacypromise.shtml>

# OECD Fair Information Principles

---

WWW10 Tutorial – May 1, 2001

- Collection limitation
- Data quality
- Purpose specification
- Use limitation
- Security safeguards
- Openness
- Individual participation
- Accountability

<http://www.oecd.org/dsti/sti/it/secur/prod/PRIV-en.HTM>

# Simplified Principles

---

WWW10 Tutorial – May 1, 2001

- Notice and disclosure
- Choice and consent
- Data security
- Data quality and access
- Recourse and remedies

# Seal Programs

WWW10 Tutorial – May 1, 2001

- TRUSTe – <http://www.truste.org>
- BBBOnline – <http://www.bbbonline.org>
- CPA WebTrust – <http://www.cpawebtrust.org/>
- Japanese Privacy Mark  
<http://www.jipdec.or.jp/security/privacy/>



# Seal Program Problems

---

WWW10 Tutorial – May 1, 2001

- Basic Principle:
  - Publish a policy (*any* policy) and follow it
- Only few require base-level standard
  - BBBOnline requires client in good standing with Better Business Bureau
- Effect:
  - Good notices of bad practices

# Laws and Regulations

---

WWW10 Tutorial – May 1, 2001

- Privacy laws and regulations vary widely throughout the world
- US has mostly sector-specific laws, with relatively minimal protections
  - Self-Regulation favored over comprehensive Privacy Laws
  - Fear that regulation hinders e-commerce
- Europe has long favoured strong privacy laws
  - First data protection law in the world: State of Hesse, Germany (1970)
  - Privacy commissions in each country (some countries have national and state commissions)

# Some US Privacy Laws

---

WWW10 Tutorial – May 1, 2001

- Bank Secrecy Act, 1970
- Fair Credit Reporting Act, 1971
- Privacy Act, 1974
- Right to Financial Privacy Act, 1978
- Cable TV Privacy Act, 1984
- Video Privacy Protection Act, 1988
- Family Educational Right to Privacy Act, 1993
- Electronic Communications Privacy Act, 1994
- Freedom of Information Act, 1966, 1991, 1996

# US Law – Recent Additions

---

WWW10 Tutorial – May 1, 2001

- HIPAA (Health Insurance Portability and Accountability Act, 1996)
  - Privacy Rule in effect 04/2001; allows until 04/2003 for implementation (changes probable)
  - Protects all medical records and other individually identifiable health information
- COPPA (Children's Online Privacy Protection Act, 1998)
  - Certain Web sites must obtain parental consent before collecting personal information from children (effective 04/2000)
- GLBA (Gramm-Leach-Bliley-Act, 1999)
  - requires privacy policy disclosure and opt-out mechanisms from financial service institutions

# EU Data Directive

---

WWW10 Tutorial – May 1, 2001

- 1995 Data Protection Directive 95/46/EC
  - sets a benchmark for national law for processing personal information in electronic and manual files
  - facilitates data-flow between member states and restricts export of personal data to „unsafe“ non-EU countries
- 1997 Telecommunications Directive
  - establishes specific protections covering telecommunications systems
  - July 2000 proposal to strengthen and extend directive to cover „electronic communications“
- Member states responsible for passing relevant national laws by 10/1998
  - 10 out of 15 member states have passed legislation, 5 are still pending (as of 04/2001)

# Safe Harbor

---

WWW10 Tutorial – May 1, 2001

- Membership
  - US companies self-certify adherence to requirements
  - Dept. of Commerce maintains signatory list  
<http://www.export.gov/safeharbor/SafeHarborInfo.htm>
- Signatories must provide
  - **notice** of data collected, purposes, and recipients
  - **choice** of opt-out of 3rd-party transfers, opt-in for sensitive data
  - **access** rights to delete or edit inaccurate information
  - **security** for storage of collected data
  - **enforcement** mechanisms for individual complaints
- Approved July 26, 2000 by EU
  - reserves right to renegotiate if remedies for EU citizens prove to be inadequate

# Privacy around the World

WWW10 Tutorial – May 1, 2001

- Australia\*
  - Proposed: Privacy Amendment (Private Sector) Bill in 2000
  - In talks with EU officials
- Brazil
  - Proposed: Bill No. 61 in 1996 (pending)
- Canada\*
  - Passed: Bill C-6 in 4/2000
  - Under review by EU
- Hong Kong\*
  - Passed: Personal Data (Privacy) Ordinance in 1995
- Japan
  - Currently: self-regulation & prefectural laws
  - In talks with EU officials
- Russia
  - Law on Information, Informatization, and Inform. Protect. 1995
  - In Progress: updated to comply with EU directive
- South Africa
  - Planned: Privacy and Data Protection Bill
- Switzerland\*
  - EU-certified safe third country for data transfers

<http://www.privacyinternational.org/survey/>

\* Has National Privacy Commissioner

# Data Protection Agencies

---

WWW10 Tutorial – May 1, 2001

- Australia: <http://www.privacy.gov.au/>
- Canada: <http://www.privcom.gc.ca/>
- France: <http://www.cnil.fr/>
- Germany: <http://www.bfd.bund.de/>
- Hong Kong: <http://www.pco.org.hk/>
- Italy: <http://www.privacy.it/>
- Spain: <http://www.ag-protecciondatos.es/>
- Switzerland: <http://www.edsb.ch/>
- UK: <http://www.dataprotection.gov.uk/>

... And many more

# Privacy Web Sites

---

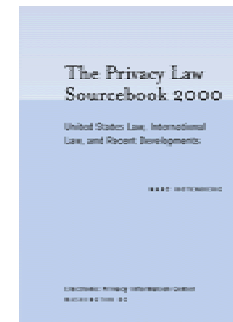
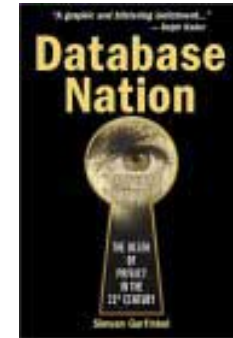
WWW10 Tutorial – May 1, 2001

- <http://www.privacyinternational.org>
- <http://www.privacyfoundation.org>
- <http://www.privacyexchange.org>
- <http://www.privacycouncil.com>
- <http://www.privacyplace.com>
- <http://www.junkbusters.com>
- <http://www.privacy.org>
- <http://www.pandab.org>
- <http://www.epic.org>
- <http://www.cdt.org>

# Books

WWW10 Tutorial – May 1, 2001

- Database Nation by Simson Garfinkel
- The Privacy Law Sourcebook 2000: United States Law, International Law, and Recent Developments by Marc Rotenberg



# Privacy Tools

WWW10 Tutorial  
May 1, 2001

- What is Privacy?
  - Definitions
  - Public Concern
- How do they get my Data?
  - Browser Chatter
  - Cookies
  - Ad Networks
  - Web Bugs
  - Spyware
- Solutions
  - Privacy Policies
  - Laws and Regulations
  - Privacy Tools
- Privacy Tools
  - Encryption
  - Anonymity
  - Management
  - Trust

# Privacy Tools

---

WWW10 Tutorial – May 1, 2001

- Encryption tools
  - Prevent others from listening in on your communications
- Anonymity tools
  - Prevent your actions from being linked to you
- Transparency tools
  - Make informed choices about how your information will be used
- Trust tools
  - Know that assurances about information practices are trust worthy

# Encryption Standards

---

WWW10 Tutorial – May 1, 2001

- Public Key Cryptography
  - Allows secure key exchange over insecure channel
- Applications & Protocols
  - IPSec – Secure IP
  - SSH – Secure Shell
  - SSL – Secure Socket Layer
  - SET – Secure Electronic Transactions
  - PGP – Pretty Good Privacy

# Anonymity – Low Tech

---

WWW10 Tutorial – May 1, 2001

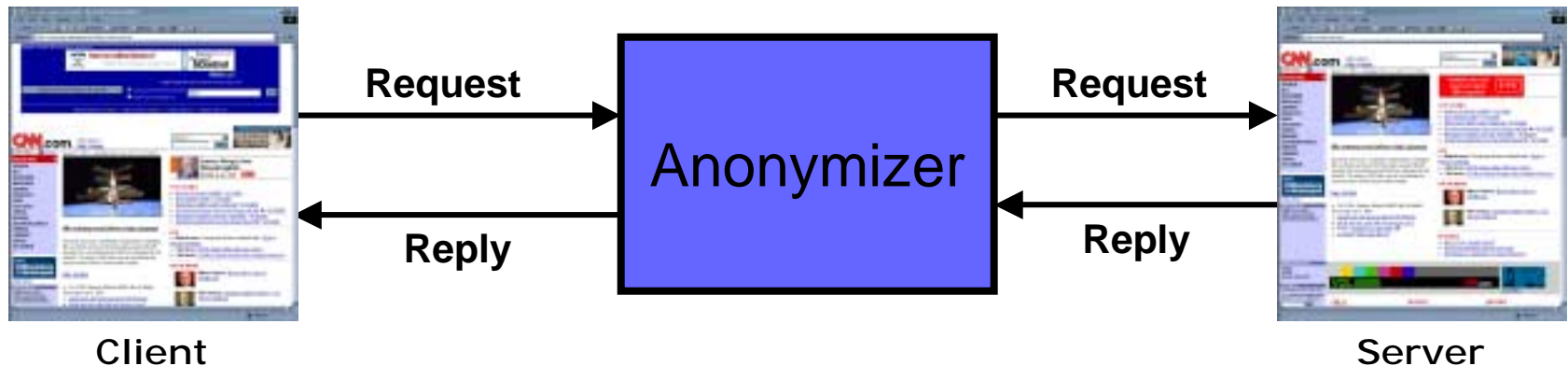
- Wander around cyber cafes
- Use free e-mail service instead of ISP
- Set up a pre-paid cash account with ISP
  - give all phony information
- Forge e-mail, spoof IP, etc.

. . . And don't give out any personally-identifiable data!

# The Anonymizer

WWW10 Tutorial – May 1, 2001

- Acts as a proxy for users
- Hides information from end servers



- Sees all web traffic
- Adds ads to pages (free service; subscription service also available)

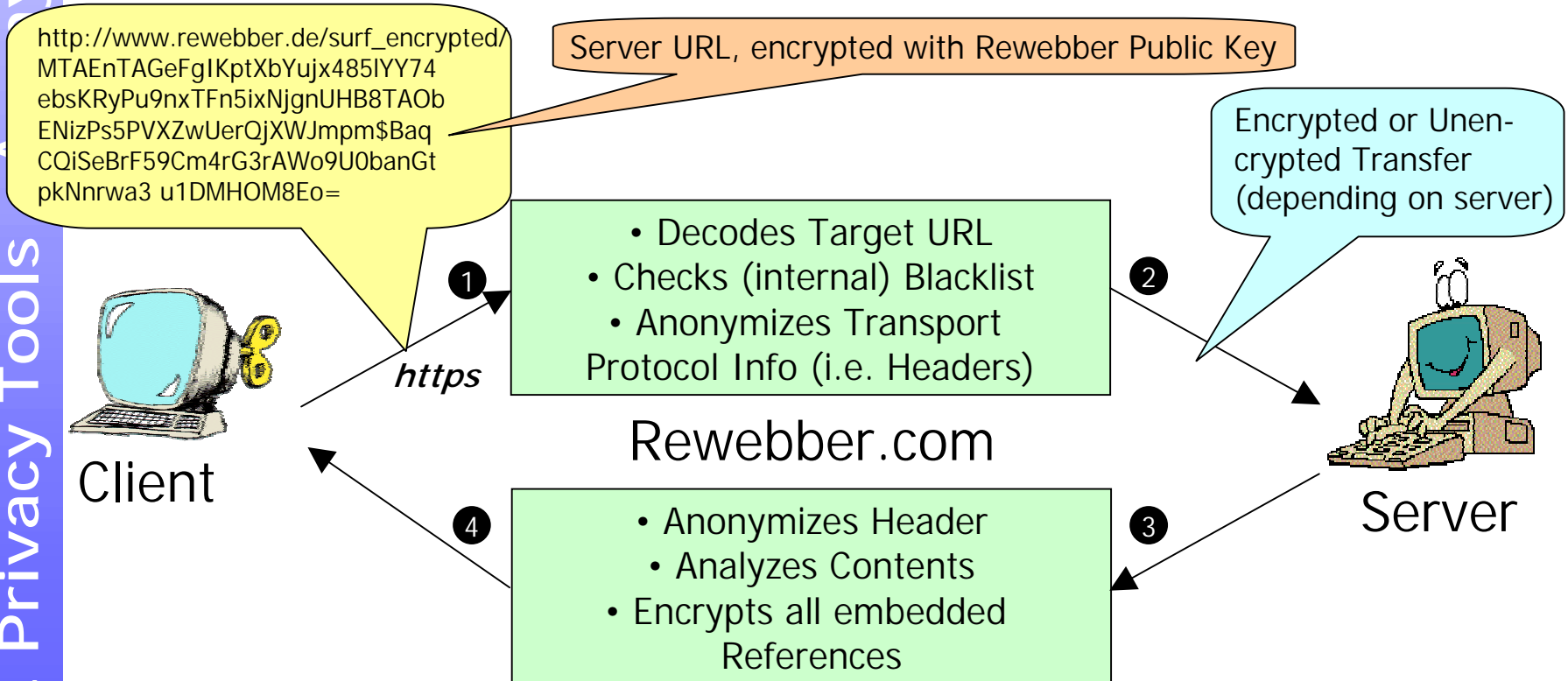
<http://www.anonymizer.com>

# Rewebber.com

WWW10 Tutorial – May 1, 2001

- Created at Hagen University, Germany
- Provides both Client- and Server-Anonymity
- Only as subscription service (\$5-\$15 per month)

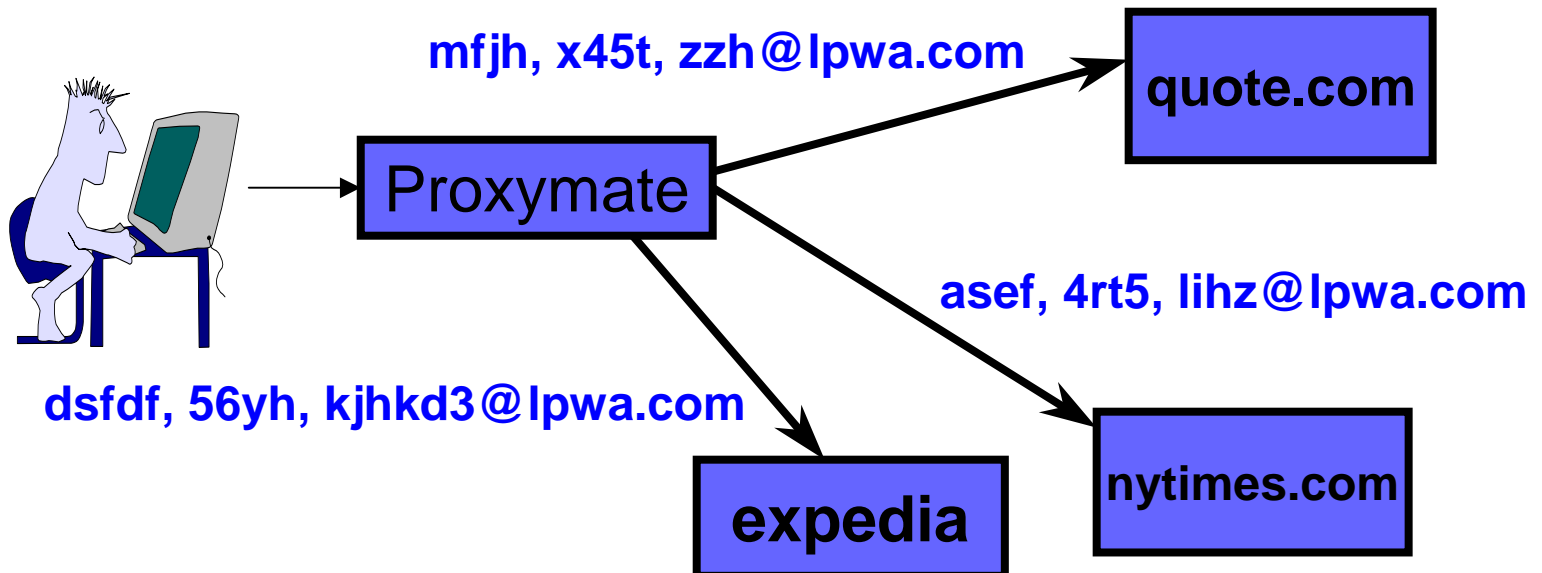
IV. Privacy Tools  
Anonymity



# Proxymate

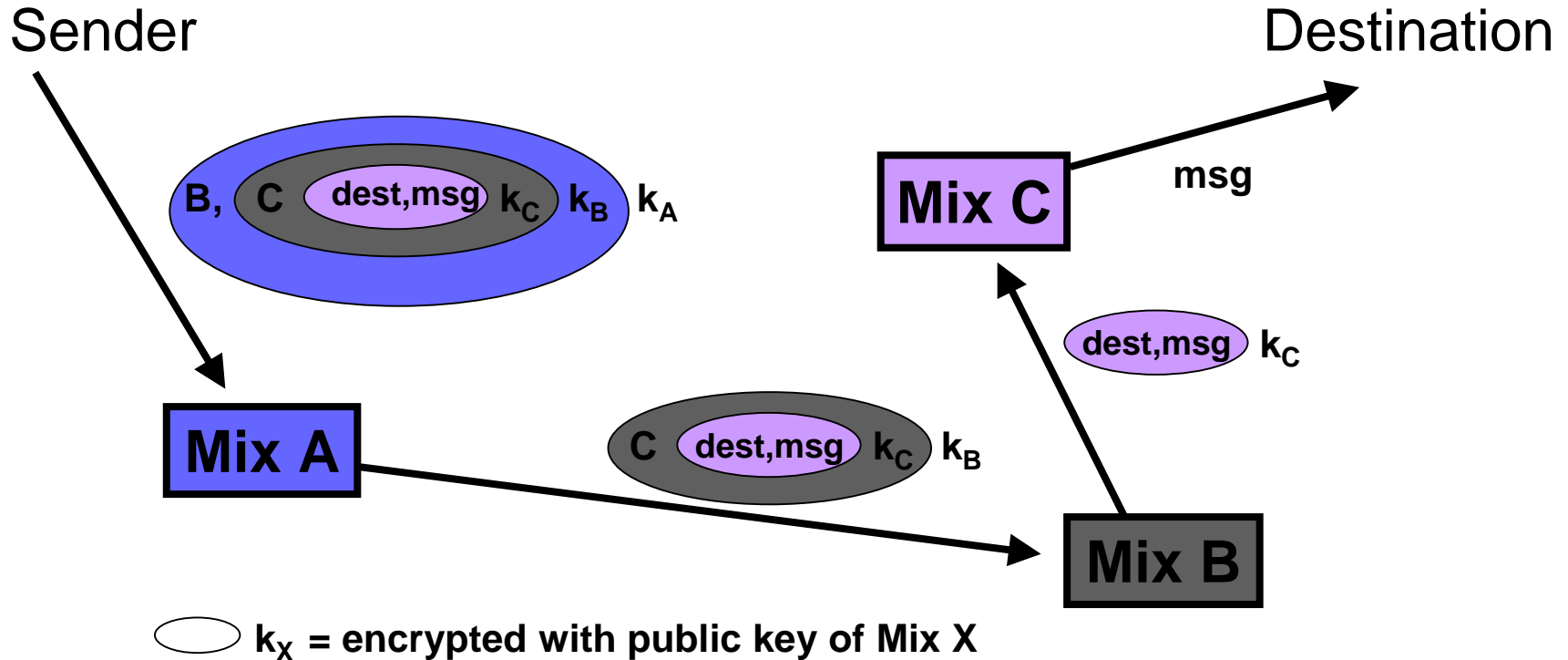
WWW10 Tutorial – May 1, 2001

- „Lucent Personal Web Assistant“ (LPWA) 1997
- Automatically generates user name, password and email address unique to each web site you visit
- Allows selective blocking of email aliases
- <http://www.proxymate.com/> (ended in July 2000)



# Mixes [Chaum81]



WWW10 Tutorial – May 1, 2001



Sender routes message randomly through network of "Mixes", using layered public-key encryption.

# Realization of Mixes

WWW10 Tutorial – May 1, 2001

- Onion Routing (Office of Naval Research) 
  - <http://www.onion-router.net>
  - service ended 01/2000
- Freedom (Zero-Knowledge Systems, Canada) zerøknowledge
  - <http://www.zeroknowledge.com>
- Java Anon Proxy (TU Dresden) 
  - <http://anon.inf.tu-dresden.de>

# Crowds

WWW10 Tutorial – May 1, 2001

- Users join a *Crowd* of other users
- Web requests from the crowd cannot be linked to any individual
- Protection from
  - end servers
  - other crowd members
  - system administrators
  - eavesdroppers
- First system to hide data shadow on the web without trusting a central authority



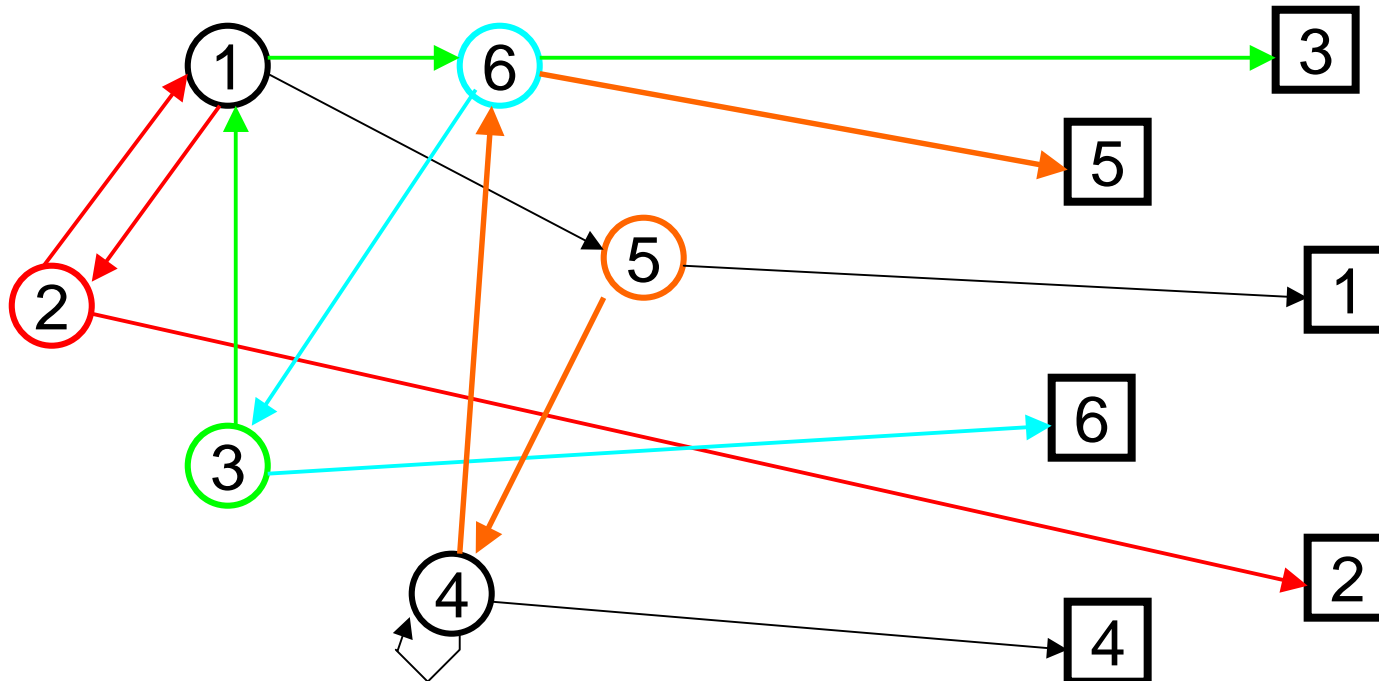
<http://www.research.att.com/projects/crowds/>

# Crowds Illustrated

WWW10 Tutorial – May 1, 2001

Crowd members

Web servers



# Anonymous Email

---

WWW10 Tutorial – May 1, 2001

- Anonymous remailers allow people to send email anonymously
- Similar to anonymous web proxies
- Some can be chained and work like mixes

<http://anon.efga.org/~rlist>

# Filters

---

WWW10 Tutorial – May 1, 2001

- Cookie Cutters

- Block cookies, allow for more fine-grained cookie control, etc.
- Some also filter ads, referer header, and browser chatter

<http://www.junkbusters.com/ht/en/links.html#measures>

- Child Protection Software

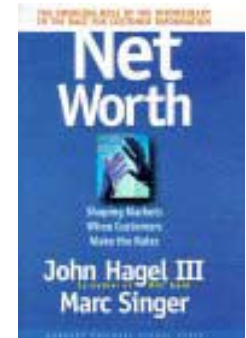
- Block the transmission of certain information via email, chat rooms, or web forms when child is using computer
- Limit who a child can email or chat with

<http://www.getnetwise.org/>

# Infomediaries

WWW10 Tutorial – May 1, 2001

- Hagel/Singer: „Net Worth“ 1997
- Services and tools that help people manage their online identities
  - Digitalme - <http://www.digitalme.com>
  - Jotter - <http://www.jotter.com>
  - Lumeria - <http://www.lumeria.com>
  - PrivacyBank - <http://www.privacybank.com>
  - Privaseek – <http://www.privaseek.com>



# Infomediaries - Examples

WWW10 Tutorial – May 1, 2001

- Jotter-Toolbar



Username and Passwords

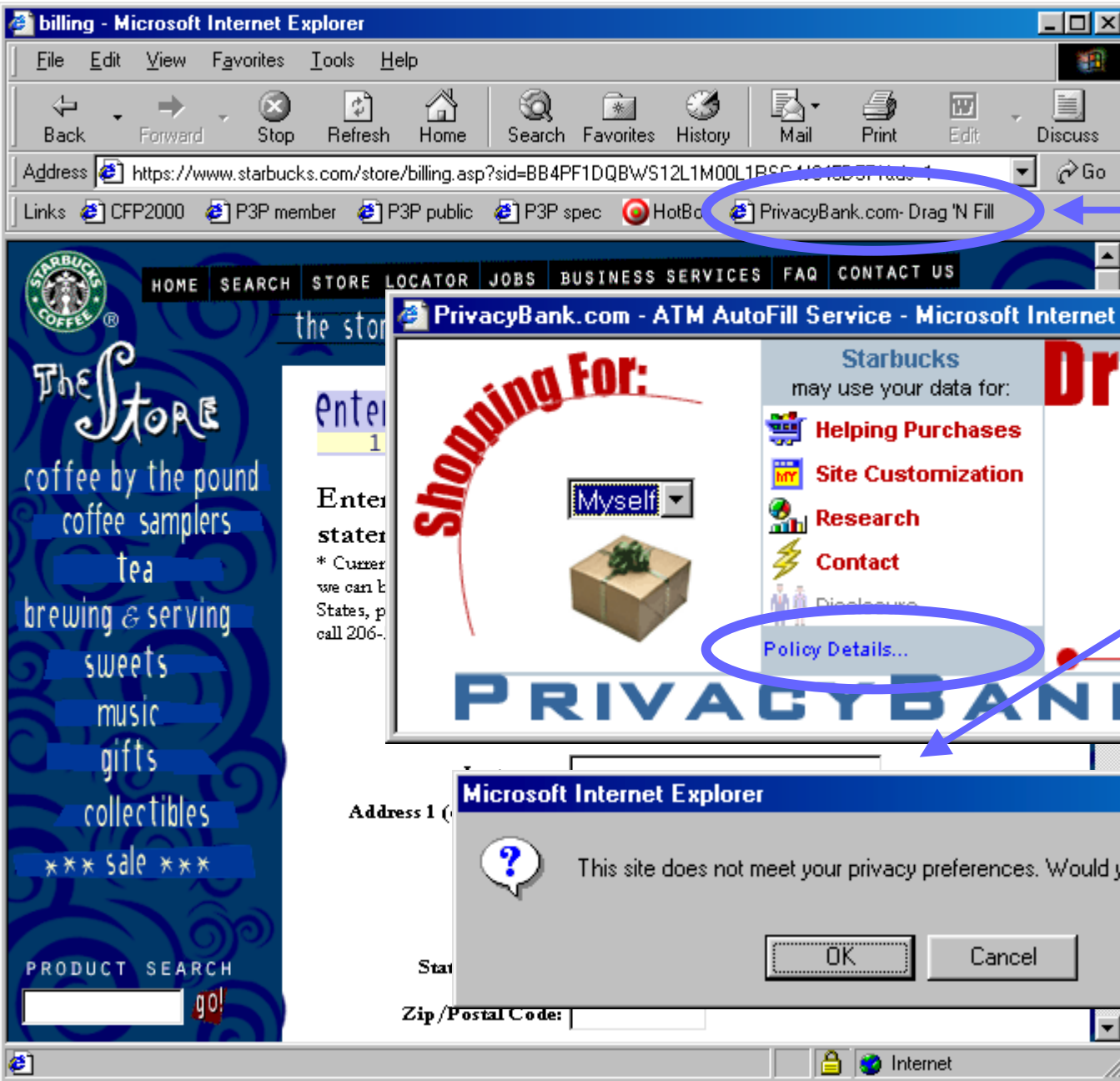
Shopping



Web-Formulare Auto-Fill

Show Privacy Policy

Ads



# Examples

Journal – May 1, 2001

PrivacyBank  
bookmark


billing - Microsoft Internet Explorer

Full Policy Details - Microsoft Internet Explorer



















































Back

Address

Links



**Privacy Policy of Starbucks**

Information Requested	Used For:	Required?
First Name	    	YES
Last Name	    	YES
Address 1	    	YES
Address 2	    	NO
City	    	YES
State/Province	    	YES
Zip/Postal Code	    	YES
Country	    	YES
E-mail	    	YES
Daytime Phone Number	    	YES

OK

Comments? Send mail to [comments@privacybank.com](mailto:comments@privacybank.com)

Copyright © 1998-1999 Millet Software, Inc.

Internet

# Summary – Part I

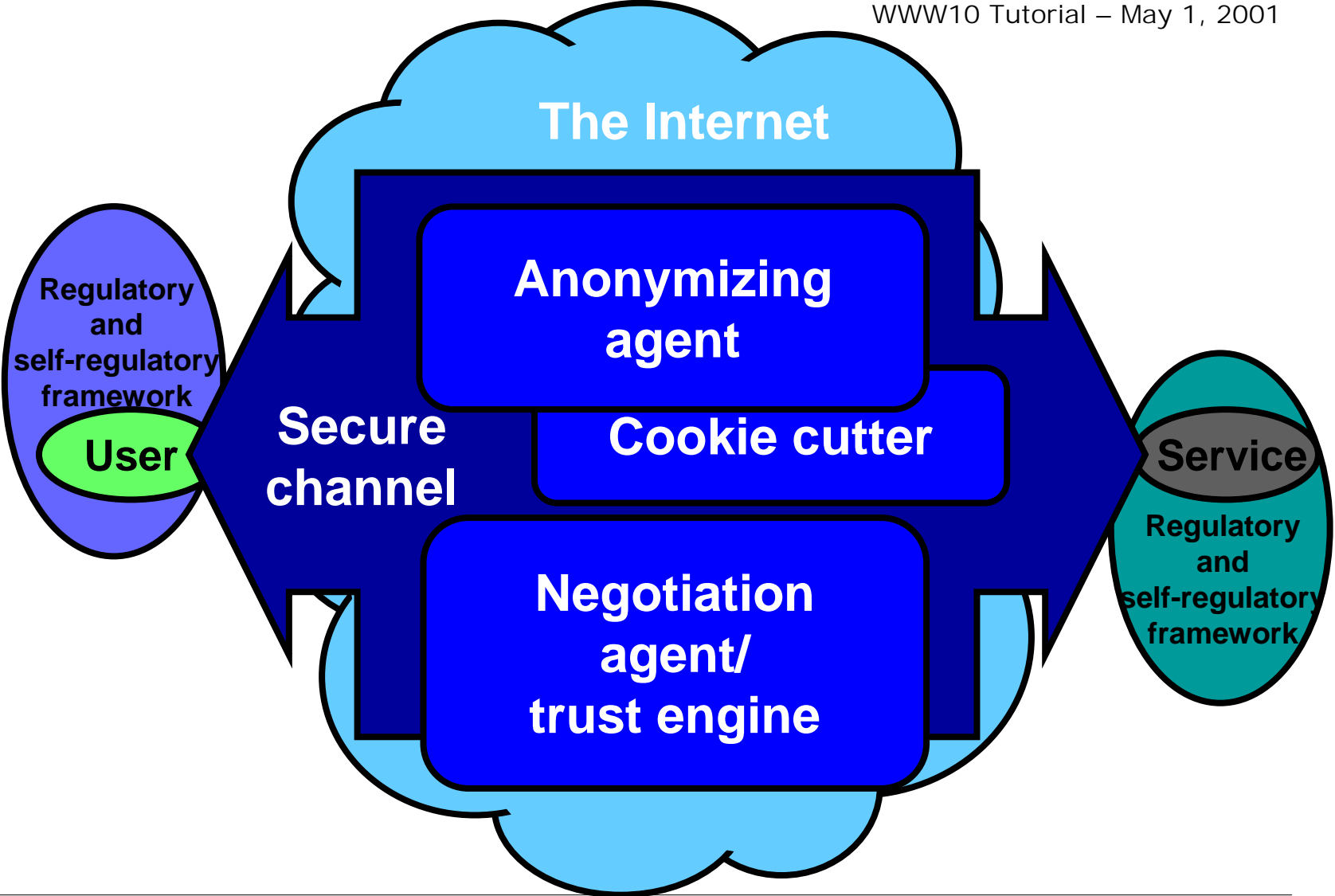
---

WWW10 Tutorial – May 1, 2001

- What is Privacy?
  - Definitions
  - Public Concern
- How do they get my Data?
  - Browser Chatter
  - Cookies
  - Ad Networks
  - Web Bugs
  - Spyware
- Solutions
  - Privacy Policies
  - Laws and Regulations
  - Privacy Tools
- Privacy Tools
  - Encryption
  - Anonymity
  - Management
  - Trust

# Privacy Tools

WWW10 Tutorial – May 1, 2001



# Outline – Part II

---

WWW10 Tutorial – May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# P3P Overview

WWW10 Tutorial  
May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# Original Idea behind P3P

---

WWW10 Tutorial – May 1, 2001

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

# P3P1.0 – A First Step

---

WWW10 Tutorial – May 1, 2001

- 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

# P3P1.0 Spec Defines

---

WWW10 Tutorial – May 1, 2001

- 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 mechanism** for transporting P3P policies over HTTP

# Future Versions of P3P

---

WWW10 Tutorial – May 1, 2001

- Allow web sites to offer a choice of policies
  - P3P 1.0 supports only one policy per resource
- Allow for “negotiation” and explicit agreements to be reached between user agent and web site
  - P3P 1.0 policies are “take-it-or-leave-it”
- Allow for non-repudiation of agreements, signatures from third-party seal providers, etc.
  - P3P 1.0 offers no mechanism to *prove* that certain communication took place
- Facilitate automated data transfer
  - P3P 1.0 requires external mechanisms (e.g., automatic form-fill) to transfer data

# P3P is a Partial Solution

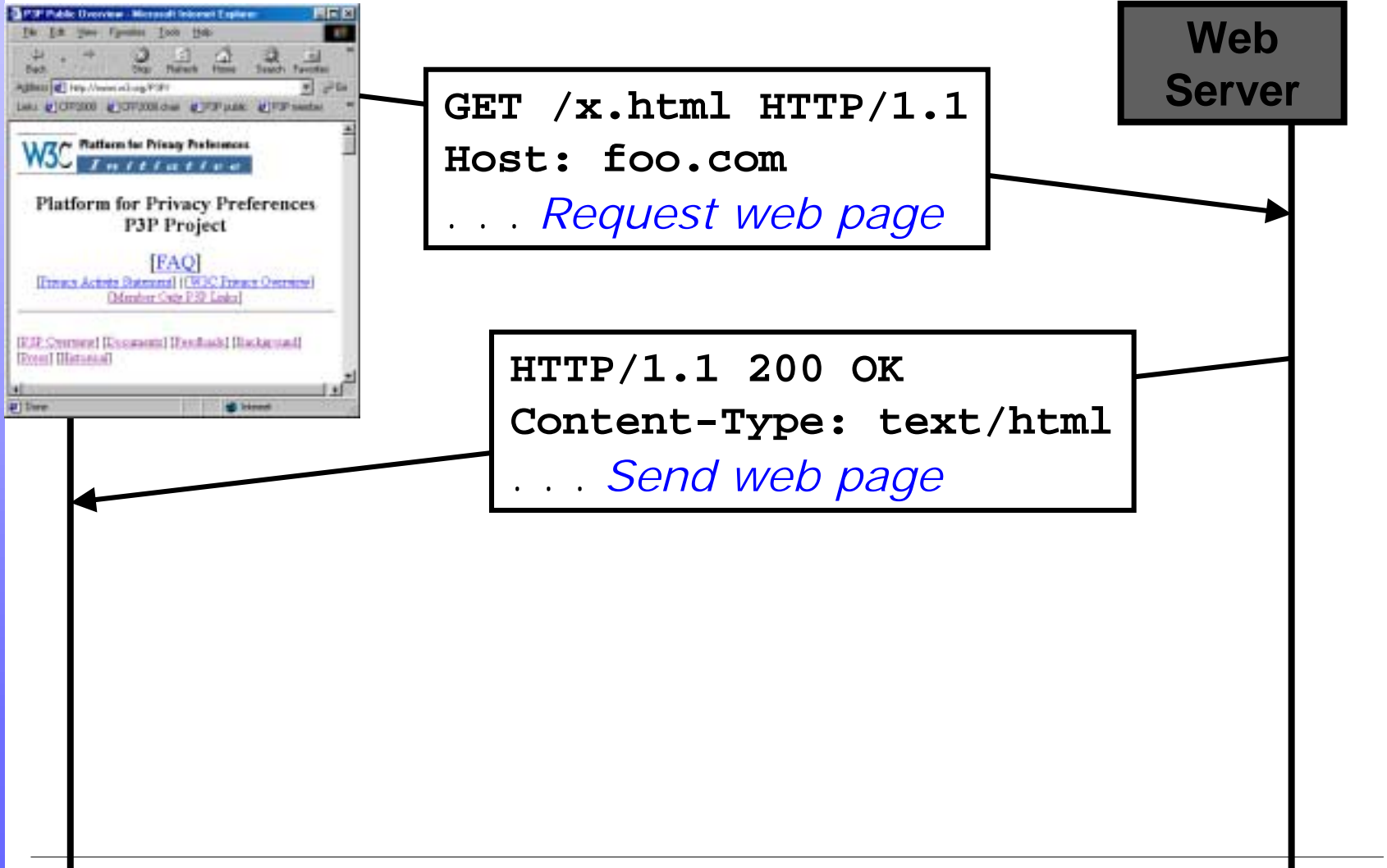
---

WWW10 Tutorial – May 1, 2001

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

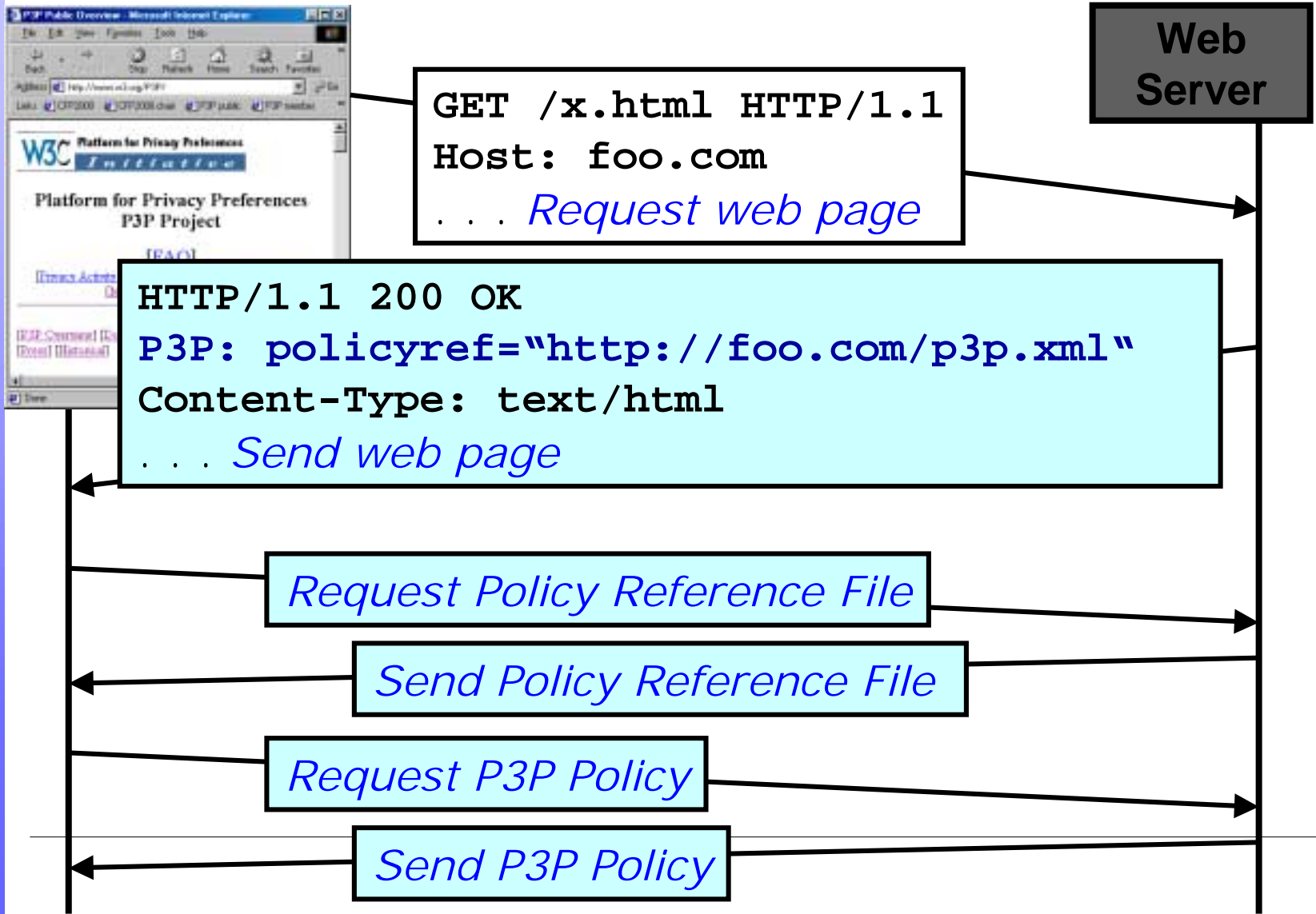
# A simple HTTP Transaction

WWW10 Tutorial – May 1, 2001



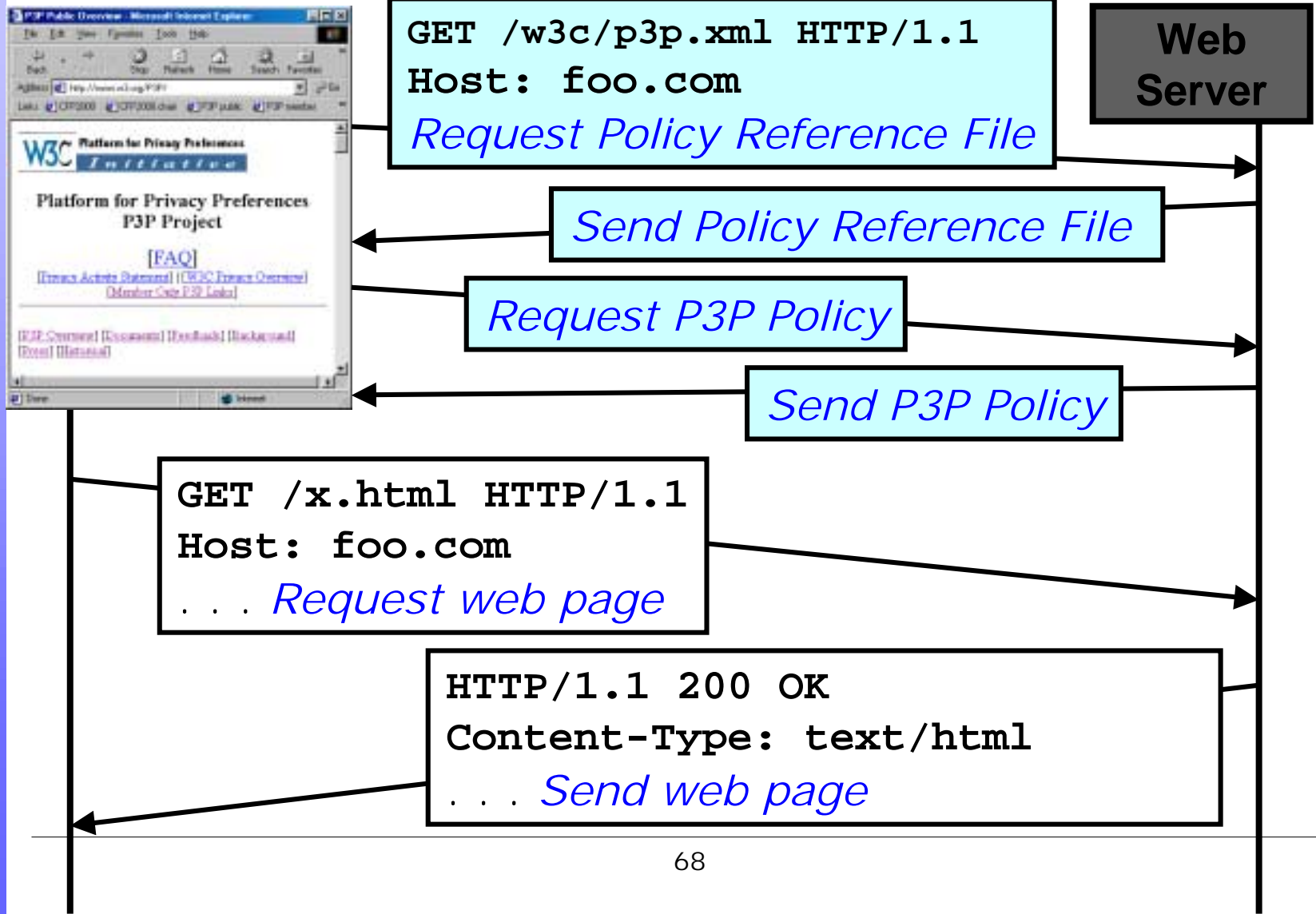
# P3P1.0 over HTTP

WWW10 Tutorial – May 1, 2001



# Or using p3p.xml File

WWW10 Tutorial – May 1, 2001



# P3P1.0 Clients

---

WWW10 Tutorial – May 1, 2001

- Client can be implemented as browser, proxy, plug-in, java applet, JavaScript, etc.
  - Can be entirely server side
  - Can be part of an infomediary service, shopping tool bar, automatic form filler, etc.
- Look for link to P3P policy and fetch policy with HTTP GET request
- Parse policy and take appropriate action
  - Display symbol, play sound, prompt user, etc.
  - Action can optionally be based on user preferences
  - Action can optionally allow data to be automatically filled into form or transferred from electronic wallet

# User Privacy Preferences

---

WWW10 Tutorial – May 1, 2001

- P3P 1.0 agents may (optionally) take action based on user preferences
  - Users should not have to trust privacy defaults set by software vendors
  - User agents that can read **APPEL** (A P3P Preference Exchange Language) files can offer users a number of canned choices developed by trusted organizations
  - Preference editors allow users to adapt existing preferences to suit own tastes, or create new preferences from scratch

# P3P Policies

---

WWW10 Tutorial – May 1, 2001

- Machine-readable (XML) version of web site privacy policies
- Use *P3P Vocabulary* to express data practices
- Use *P3P Base Data Set* to express type of data collected
- Capture common elements of privacy policies but may not express everything (sites may provide further explanation in human-readable policies)

# The P3P Vocabulary

---

WWW10 Tutorial – May 1, 2001

- Who is collecting data?
- What data is collected?
- For what purpose will data be used?
- Is there an ability to change preferences about (opt-in or opt-out) of some data uses?
- Who are the data recipients (anyone beyond the data collector)?
- To what information does the data collector provide access?
- What is the data retention policy?
- How will disputes about the policy be resolved?
- Where is the human-readable privacy policy?

# P3P Base Data Schema

---

WWW10 Tutorial – May 1, 2001

- A set of common data elements that all P3P implementations should know about
- Includes `user`, `thirdparty`, and `business` elements such as name, address, phone number, etc.
- Includes “Dynamic” elements such as indicators that a site collects click-stream, uses cookies, collects info of a certain category, etc.
- Extensible using custom data schemas

# Example Privacy Policy

---

WWW10 Tutorial – May 1, 2001

At CatalogExample, we care about your privacy. When you come to our site to look for an item, we will only use this information to improve our site and will not store it in an identifiable way. CatalogExample is a licensee of the PrivacySealExample Program.

...

Questions regarding this statement should be directed to:  
CatalogExample 1-248-392-6753

When you browse through our site we collect:

The basic information about your computer and connection to make sure that we can get you the proper information and for security purposes

Aggregate information on what pages consumers access or visit to improve our site

We purge the browsing information that we collect regularly

# P3P/XML Encoding

WWW10 Tutorial – May 1, 2001

```
<POLICY xmlns="http://www.w3.org/2000/12/P3Pv1"
  discuri="http://www.catalog.example.com/Privacy.html">
  <ENTITY><DATA-GROUP><DATA ref="#business.name">CatalogExample</DATA>
    <DATA ref="#business.contact-info.telecom.telephonenumber.intcode">1</DATA>
    <DATA ref="#business.contact-info.telecom.telephonenumber.loccode">
      248</DATA>
    <DATA ref="#business.contact-info.telecom.telephonenumber.number">
      3926753</DATA>
  </DATA-GROUP></ENTITY>
  <ACCESS><nonident/></ACCESS>
  <DISPUTES-GROUP> <DISPUTES resolution-type="independent"
    service="http://www.PrivacySeal.example.org"
    short-description="PrivacySeal.exampleorg"
    <REMEDIES><correct/></REMEDIES>
    <IMG src="http://www.PrivacySeal.example.org/Logo.gif"/>
  </DISPUTES></DISPUTES-GROUP>
  <STATEMENT>
    <PURPOSE><admin/><develop/></PURPOSE>
    <RECIPIENT><ours/></RECIPIENT>
    <RETENTION><stated-purpose/></RETENTION>
    <DATA-GROUP>
      <DATA ref="#dynamic.clickstream"/>
      <DATA ref="#dynamic.http"/>
    </DATA-GROUP>
  </STATEMENT>
</POLICY>
```

# Referencing P3P Policies

WWW10 Tutorial  
May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# Policy References

---

WWW10 Tutorial – May 1, 2001

- Allows web sites to indicate what policy applies to what resource
- Allows user agents to determine what policy applies to what resource
- Performance optimization
  - Send reference rather than full policy with each response
  - Only parse and process each policy once as long as results are cached

# Policy Reference Files (PRF)

---

WWW10 Tutorial – May 1, 2001

- Allow specification of which policy applies to which resources on a site
  - **<EXPIRY>**
    - Determines how long PRF is valid
  - **<POLICY-REF>**
    - URI of policy
  - **<INCLUDE>, <EXCLUDE>**
    - URI prefixes (local) to which policy applies / doesn't apply
  - **<EMBEDDED-INCLUDE>, <EMBEDDED-EXCLUDE>**
    - Absolute URI to 3<sup>rd</sup> party content to which policy applies / does not apply
  - **<COOKIE-INCLUDE>, <COOKIE-EXCLUDE>**
    - Associates / disassociates cookies with policy
  - **<METHOD>**
    - Methods to which policy applies

# PRF Example

WWW10 Tutorial – May 1, 2001

```
<META xmlns="http://www.w3.org/2000/P3Pv1">  
<POLICY-REFERENCES>  
  <EXPIRY max-age="172800" />  <!-- relative expiry: 2 days -->
```

```
<POLICY-REF about="/P3P/Policy1.xml">  
  <INCLUDE>/*</INCLUDE>  
  <EXCLUDE>/catalog/*</EXCLUDE>  
  <EXCLUDE>/cgi-bin/*</EXCLUDE>  
  <EXCLUDE>/servlet/*</EXCLUDE>  
</POLICY-REF>
```

```
<POLICY-REF about="/P3P/Policy2.xml">  
  <INCLUDE>/catalog/*</INCLUDE>  
</POLICY-REF>
```

```
<POLICY-REF about="/P3P/Policy3.xml">  
  <INCLUDE>/cgi-bin/*</INCLUDE>  
  <INCLUDE>/servlet/*</INCLUDE>  
  <EXCLUDE>/servlet/unknown</EXCLUDE>  
</POLICY-REF>
```

```
</POLICY-REFERENCES>  
</META>
```

# EXPIRY

---

WWW10 Tutorial – May 1, 2001

- States how long policy reference file (or policy) stays valid
- Relative time (in seconds)
  - Denotes time before a new policy can replace existing one
  - Minimum: 24 hours (86400 seconds)
- Absolute time (GMT/UTC)
  - Used to phase out policies
  - If date is in the past: no policy!

# METHOD

---

WWW10 Tutorial – May 1, 2001

- Allows different P3P policies for the *same* resource when accessed through *different methods*.
- E.g., Web publishing systems might only collect clickstream data for GET requests, but collects login information for PUT and DELETE methods.
- Notice: GET and HEAD requests must use same policies!

# PRF example

WWW10 Tutorial – May 1, 2001

```
<META xmlns="http://www.w3.org/2000/P3Pv1">
<POLICY-REFERENCES>
  <EXPIRY max-age="172800" />  <!-- relative expiry: 2 days -->

  <POLICY-REF about="/P3P/Policy1.xml">
    <INCLUDE>/docs/*</INCLUDE>
    <METHOD>HEAD</METHOD>
    <METHOD>GET</METHOD>
  </POLICY-REF>

  <POLICY-REF about="/P3P/Policy2.xml">
    <INCLUDE>/docs/*</INCLUDE>
    <METHOD>PUT</METHOD>
    <METHOD>DELETE</METHOD>
  </POLICY-REF>

</POLICY-REFERENCES>
</META>
```

# Embedded Content

---

WWW10 Tutorial – May 1, 2001

- User agents should check for policies on all embedded content (images, frames, etc.)
  - Good use of policy reference files should reduce need for extra round trips
- `<EMBEDDED-INCLUDE/EXCLUDE>`
  - Performance optimization: allows declaration of 3<sup>rd</sup> party contents (`<INCLUDE>` allows only local URIs)
  - Specified policy only applies when accessed from site making declaration
    - avoids „sticky“ misdeclarations from rogue sites

# PRF Example

WWW10 Tutorial – May 1, 2001

- Example policy at [www.example.org](http://www.example.org):

```
<META xmlns="http://www.w3.org/2000/12/P3Pv1">
<POLICY-REFERENCES>

<POLICY-REF about="/P3P/Policy1.xml">
  <INCLUDE>/docs/*</INCLUDE>
  <INCLUDE>/other/index.html</INCLUDE>
  <EMBEDDED-INCLUDE>
    http://*.adserver.example.com/ads/*
  </EMBEDDED-INCLUDE>
  <EMBEDDED-EXCLUDE>
    http://*.adserver.example.com/ads/network/*
  </EMBEDDED-EXCLUDE>
</POLICY-REF>

</POLICY-REFERENCES>
</META>
```

- Policy1.xml only applies to [adserver.example.com/ads](http://adserver.example.com/ads) if accessed from [www.example.org](http://www.example.org) pages!

# Forms

---

WWW10 Tutorial – May 1, 2001

- Forms are special kind of embedded content („ACTION“ URL)
  - User agents should be especially careful not to unknowingly submit data when no policy is available
  - Check well-known location **before** submitting form data, if policy is unknown

# Cookies

WWW10 Tutorial – May 1, 2001

- P3P policy only applies to *resource*, not its associated cookies!
- <COOKIE-INCLUDE/EXCLUDE>
  - Associates P3P policy to (named) cookie
- „cookie“-policy must cover
  - Data stored in, or linked via, the cookie
  - All purposes associated with stored or linked data
  - If data collection done via HTTP, then separate policy must also cover that data transfer

# Cookies Example

WWW10 Tutorial – May 1, 2001

Entrance Page. Sets **unique\_id** for session tracking.



covers

policy1

Declares only **clickstream data** logging.



covers

policy2

Declares collection of **contact info** (optional, only required for „ACTION“ URL handling the POST of the data)

Set\_cookie

Contact page. Sets **unique\_id** associated with customer data.

Set\_cookie



Assigns **unique id** for **state management**, but also allows **linking** to **contact information**.

covers

policy3

Declares **contact info** and **state management**

# PRF example

WWW10 Tutorial – May 1, 2001

```
<META xmlns="http://www.w3.org/2000/12/P3Pv1">
<POLICY-REFERENCES>

  <POLICY-REF about="/P3P/Policy1.xml">
    <COOKIE-INCLUDE>* * *</COOKIE-INCLUDE>
    <COOKIE-EXCLUDE>obnoxious-cookie .example.com </COOKIE-EXCLUDE>
  </POLICY-REF>

  <POLICY-REF about="/P3P/Policy2.xml">
    <COOKIE-INCLUDE>obnoxious-cookie .example.com </COOKIE-INCLUDE>
  </POLICY-REF>

</POLICY-REFERENCES>
</META>
```

# Locating a PRF

---

WWW10 Tutorial – May 1, 2001

- Well-known file
  - `/w3c/p3p.xml` is standard location for policy reference file
- HTTP Header
  - References appear in response headers
- LINK tags
  - References appear in LINK tags

# HTTP Header

---

WWW10 Tutorial – May 1, 2001

- Example (2.1)

Client request:

```
GET /index.html HTTP/1.1
Host: catalog.example.com
Accept: */*
Accept-Language: de, en
User-Agent: WonderBrowser/5.2 (RT-11)
```

Server response:

```
HTTP/1.1 200 OK
P3P: policyref="http://www.example.com/P3P/p1.xml "
Content-Type: text/html
Content-Length: 7413
Server: CC-Galaxy/1.3.18
```

# LINK Tags

WWW10 Tutorial – May 1, 2001

- LINK tag embedded in an HTML document encodes the information that could be expressed using the P3P PolicyRef header
- Most useful for entities that wish to supply P3P policies but can't put file in well-known location or change headers (Geocities homesteaders, for example)
- Example

```
<link rel="P3Pv1" ref="http://www.example.com/P3P/p1.xml">
```

# Safe Zone

---

WWW10 Tutorial – May 1, 2001

- User agents should ensure that minimal data collection takes place while fetching a P3P policy
  - Suppress transmission of unnecessary data
  - Try to fetch policy reference file from well-known location

# P3P Vocabulary

WWW10 Tutorial  
May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# The POLICY Element

---

WWW10 Tutorial – May 1, 2001

- Contains a complete P3P policy
- Takes mandatory `discuri` attribute
  - indicates location of human-readable privacy policy
- Sub-Elements
  - `<ENTITY>`, `<DISPUTES-GROUP>`, `<ACCESS>`, `<STATEMENT>`,  
`<TEST>`, `<EXTENSION>`, `<EXPIRY>`
- Example:

```
<POLICY xmlns= "http://www.w3.org/2000/12/P3Pv1 "  
  discuri= "http://www.catalog.example.com/Privacy.html"/>
```

# The ENTITY Element

WWW10 Tutorial – May 1, 2001

- Mandatory
- 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
- Example

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

# The DISPUTES Element

WWW10 Tutorial – May 1, 2001

- Describes a dispute resolution procedure
  - may be followed for disputes about a service's privacy practices
- Part of a **<DISPUTES-GROUP>**
  - allows several dispute resolution procedures to be listed
- Attributes:
  - resolution-type\*
    - customer service
    - independent org.
    - court
    - applicable law
  - service\* (URI)
  - short-description
  - verification (URI)
- Sub-Elements
  - <IMAGE>
  - <LONG-DESCRIPTION>
  - <REMEDIES>

\* Mandatory Attribute

# The REMEDIES Element

WWW10 Tutorial – May 1, 2001

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

```
<DISPUTES-GROUP>
  <DISPUTES
    resolution-type="independent"
    service="http://www.PrivacySeal.org"
    description="PrivacySeal.org"
    image=http://www.PrivacySeal.org/Logo.gif>
    <REMEDIES><correct/></REMEDIES>
  </DISPUTES>
</DISPUTES-GROUP>
```

# The ACCESS Element

---

WWW10 Tutorial – May 1, 2001

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

- Example:

```
<ACCESS><nonident/></ACCESS>
```

# The STATEMENT Element

---

WWW10 Tutorial – May 1, 2001

- Data practices applied to data elements
  - mostly serves as a grouping mechanism
- Contains the following sub-elements:
  - <CONSEQUENCE>
  - <NON-IDENTIFIABLE>
  - <PURPOSE> \*
  - <RECIPIENT> \*
  - <RETENTION> \*
  - <DATA-GROUP> \*

\* Mandatory Elements

# The CONSEQUENCE Element

---

WWW10 Tutorial – May 1, 2001

- 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>A site with clothes you would appreciate</CONSEQUENCE>
```

# The NON-IDENTIFIABLE Element

---

WWW10 Tutorial – May 1, 2001

- 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
- Must have a human readable explanation how this is done at the **discuri**
- No attributes, no sub-elements:

<NON-IDENTIFIABLE />

# The PURPOSE Element

WWW10 Tutorial – May 1, 2001

- Purposes of data collection, or uses of data
  - <current/>
  - <admin/>
  - <develop/>
  - <customization/>
  - <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>
  <admin/>
  <develop
    required="opt-out" />
</PURPOSE>
```

# The RECIPIENT Element

WWW10 Tutorial – May 1, 2001

- Recipients of the collected data
  - <ours>
  - <delivery>
  - <same>
  - <other-recipient>
  - <unrelated>
  - <public>
- Note:
  - <delivery> only used if delivery service does NOT agree to use data only for completion of delivery.
- Optional attribute (all but <ours>):
  - **required**
    - always (default)
    - opt-in
    - opt-out
- Optional sub-element:
  - <recipient-description>
- Example:

```
<RECIPIENT>
  <ours />
  <delivery
    required="opt-out" />
</PURPOSE>
```

# The RETENTION Element

WWW10 Tutorial – May 1, 2001

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

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

```
<RETENTION><indefinitely/></RETENTION>
```

# The DATA Element

WWW10 Tutorial – May 1, 2001

- Describes the data to be transferred or inferred
- Contained in a DATA-GROUP
- Attributes:
  - `ref*`
  - `optional`
- Sub-Elements:
  - `<CATEGORIES>`
- Example:

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

\* Mandatory Attribute

# The CATEGORIES Element

---

WWW10 Tutorial – May 1, 2001

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

# The TEST Element

---

WWW10 Tutorial – May 1, 2001

- Used for testing purposes
  - Presence (anywhere in policy) indicates that policy is just an example and **MUST** be ignored
- Prevents misunderstandings during initial deployment
- No attributes, no sub-elements:

<TEST />

# Extension Mechanism

WWW10 Tutorial – May 1, 2001

- `<EXTENSION>` describes extension to P3P syntax
- `optional` attribute indicates whether the extension is mandatory or optional (default is `optional="yes"`)
- Example:
  - „This set of data elements is only collected from users living in USA, Canada or Mexico“ (optional extension)

```
<DATA-GROUP>
. . .
  <EXTENSION>
    <COLLECTION-GEOGRAPHY type = "include"
      xmlns="http://www.TheCoolCatalog.com/P3P/region">
      <USA/><Canada/><Mexico/>
    </COLLECTION-GEOGRAPHY>
  </EXTENSION>
</DATA-GROUP>
```

# Compact Policies (CP)

---

WWW10 Tutorial – May 1, 2001

- *Optional* performance optimization
- Summary of (full) P3P policies
- Only apply to cookies
  - Allows quick decision whether to accept or reject cookie
  - If not enough information, full policy should be fetched
  - Must declare both data **stored** and **linked to** cookie
  - Only for cookies set in **current response**

# CP Syntax

---

WWW10 Tutorial – May 1, 2001

- Part of P3P Header
  - **P3P: policyref="...", CP="NON NID DSP NAV CUR"**
- Supports 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)

# P3P Data Schemas

WWW10 Tutorial  
May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# Base Data Schema

---

WWW10 Tutorial – May 1, 2001

- 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
- Dynamic
  - clickstream, http, clientevents, cookies, miscdata, searchtext, interactionrecord

# dynamic.miscdata

WWW10 Tutorial – May 1, 2001

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

```
<POLICY ...>
  . . .
  <DATA ref = " #dynamic.miscdata" >
    <CATEGORIES><online/></CATEGORIES>
  </DATA>
  . . .
</POLICY>
```

# Custom Data Schemas

---

WWW10 Tutorial – May 1, 2001

- Use the `<DATASHEMA>` element
  - Embedded in a policy or in a stand-alone XML file
  - Use `<DATA-DEF>` and `<DATA-TYPE>` elements to define data elements and data types respectively
- Updates in referenced XML schema files must either be **backwards-compatible**, or a new name (URI) must be used!

# Custom Schema Example

WWW10 Tutorial – May 1, 2001

```
<POLICY>
  [...]
  <!-- Custom data elements defined by this policy. -->
  <DATASHEMA>
    <DATA-DEF name="example" short-description="Example Data">
      <LONG-DESCRIPTION>Custom data elements by example.com</LONG-DESCRIPTION>
      <CATEGORIES><uniqueid/></CATEGORIES>
    </DATA-DEF>
    <DATA-DEF name="example.registration"
      short-description="Registration information">
      <CATEGORIES><uniqueid/></CATEGORIES>
    </DATA-DEF>
    <DATA-DEF name="example.registration.userid" short-description="User ID">
      <LONG-DESCRIPTION>User ID created by registering
        at our site.</LONG-DESCRIPTION>
      <CATEGORIES><uniqueid/></CATEGORIES>
    </DATA-DEF>
    <DATA-DEF name="example.registration.password" short-description="Password">
      <LONG-DESCRIPTION>Password created by the user
        when registering at our site.</LONG-DESCRIPTION>
      <CATEGORIES><uniqueid/></CATEGORIES>
    </DATA-DEF>
  </DATASHEMA>
  [...]
</POLICY>
```

# P3P Deployment

WWW10 Tutorial  
May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# Deployment Issues

---

WWW10 Tutorial – May 1, 2001

- Project Timeline
  - History
  - Outreach
- Site Deployment
  - Planning
  - Sample Installation Apache Server
- Client Examples
  - Prototypes

# Project Timeline

---

WWW10 Tutorial – May 1, 2001

- June 1997 – W3C P3P kickoff meeting
- 1997-1999 – Many working drafts published
- October 1999 – W3C patent analysis published  
<http://www.w3.org/TR/P3P-analysis>
- November 1999 – “Last call” working draft
- June 21, 2000 – P3P “Interop” event, New York
- December 15, 2000 – P3P becomes „W3C Candidate Recommendation“
- Required for „Proposed Recommendation“
  - implementations (2 user agents, 2 tools)
  - at least 10 P3P-enabled Web sites

# Outreach and Deployment

---

WWW10 Tutorial – May 1, 2001

- P3P Policy Outreach Working Group convened in October 1999
- P3P European Workshop
- Many prototype/demo implementations
- P3P/WAP Workshop in December 2000
- Currently looking for user agent implementations and commitments from web sites to use P3P
- P3P “Interop” events
  - June 21, 2000 in New York City
  - November 2, 2000 in Palo Alto
  - Possible European event in 2001

# Interest from Europe

---

WWW10 Tutorial – May 1, 2001

- Several meetings with European Commission working party (Article 29 WG, DG13, DG15, etc.)
- Interest in using P3P to complement and help enforce EU laws
- Plan to work together to create APPEL files corresponding to national laws
  - Process should help identify remaining holes in P3P vocabulary
- Several European demonstration projects plan to include P3P

# Site Deployment

---

WWW10 Tutorial – May 1, 2001

- Creating 1 or more policy statements
- Creating a policy reference file (PRF)
- Creating a human readable policy
- Publish policies and PRF
- Tell browsers where to find PRF

# Planning

---

WWW10 Tutorial – May 1, 2001

- How many policies?
- What method to use?
  - well-known location (/w3c/p3p.xml)
  - HTTP header
  - HTML LINK tag
- Should compact policies be used?
- Should cookie-policies be created?
- How will policy updates be handled?

# How many Policies?

---

WWW10 Tutorial – May 1, 2001

- One human readable policy
- P3P policies as specific as possible
  - otherwise: shop needs credit card info if I want to view homepage? (no – only at checkout!)
- striking a balance
  - the more policies, the more specific
  - the fewer policies, the easier to administer
- Usually about 10 policies per site

# Locating the PRF

---

WWW10 Tutorial – May 1, 2001

- Fast & Easy: well-known location
  - browsers look here first
  - site might not have access to server root
- Flexible: HTTP header based
  - relatively easy to administer (based on server)
  - server might not support custom headers
- Last Resort: HTML LINK-tag based
  - can be used on *any* site
  - without server-side include (SSI, PHP, ASP) almost impossible to administer on larger sites!

# Apache Installation

---

WWW10 Tutorial – May 1, 2001

- including mod\_headers module
  - LoadModule
  - AddModule
- specifying headers
  - httpd.conf
  - .htaccess
- Example:

```
<Location/>  
Header append P3P "policyref=\"http://www.example.com/P3P/policy1.xml\""  
</Location>
```

# Policy Updates

---

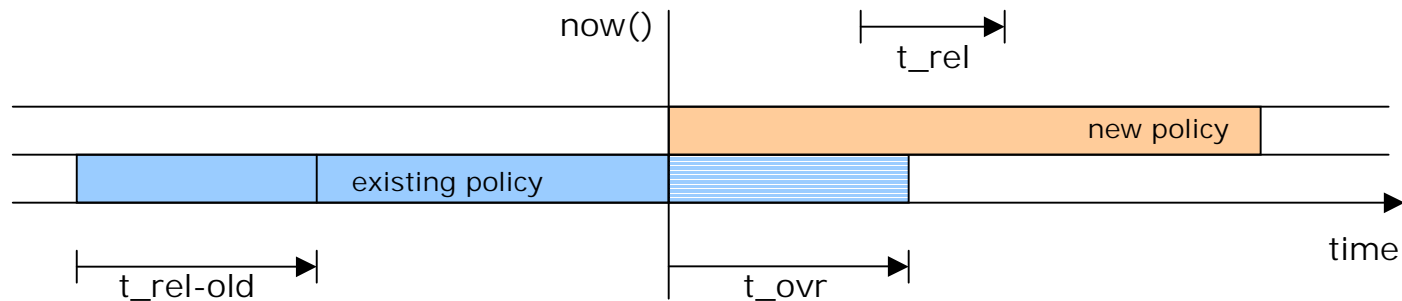
WWW10 Tutorial – May 1, 2001

- Always keep track of policy rollout time
  - collected data must be synchronized with policy in effect at that time
- Overlap
  - Simple to install
  - needs to honor two policies at same time
- Seamless
  - almost no overlap time ( $\leq 1$  second)
  - more effort to setup and perform

# Overlapping Update

WWW10 Tutorial – May 1, 2001

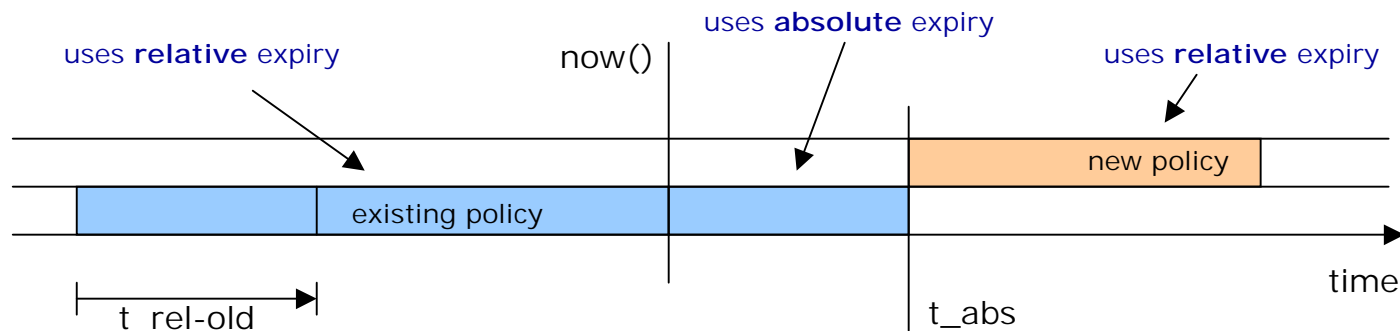
- Replace policy reference file (and/or policies) at any time
- Needs to honor *both* policies / reference files until overlap time
  - $t_{\text{ovr}} = \text{now}() + t_{\text{rel-old}}$



# Seamless Update

WWW10 Tutorial – May 1, 2001

- Update policy reference file (and/or policies) with absolute expiry time
  - $t_{abs} \geq \text{now}() + t_{rel}$
- Batch process:
  - cron-activated @ update time
  - copy new policy reference file (contains relative expiry time)
  - copy new policies (contain relative expiry time)



# P3P Prototypes

---

WWW10 Tutorial – May 1, 2001

- Early Prototypes
  - AT&T Privacy Minder
  - AT&T P3P Proposal Generator
  - ENC Privacy Information Management System
  - Microsoft Privacy Wizard
  - NEC P3P for Perl
  - IBM P3P Library
  - NCR P3P user agent demo
- Commercial Prototypes
  - Microsoft Internet Explorer 6 (Public Preview)
  - YouPowered Orby
  - IDecide Privacy Companion
- Recent Work
  - W3C P3P client prototype
  - IBM P3P Policy Editor

<http://www.w3.org/P3P/implementations>

---

# W3C P3P Client Prototype

WWW10 Tutorial – May 1, 2001

The screenshot displays the W3C P3P Client Prototype interface. On the left, a 'P3P Warning' dialog box asks: 'You are going to access to a WWW page without P3P policy. Do you want to proceed?' with a 'Proceed' button and a 'Cancel' button. The main window shows a tree view of a policy document for 'localhost/inform.xml'. The tree structure is as follows:

- POLICY
  - DISCLOSURE STATEMENT
    - IDENTIFIABLE (no)
      - RECIPIENT
      - PURPOSE
      - DATA-GROUP (DATA)

Below the tree is an 'Attributes' table:

name	value
name	dynamic.cookies

On the right, the 'User Preference' section shows the current status: 'inform'. Below it, there are radio buttons for 'ACCEPT', 'INFORM', 'WARN', and 'REJECT'. A 'Modify' button is at the bottom right.

Annotations with arrows point to various parts of the interface:

- 'View Details of Site Policy' points to the tree view.
- 'Warning to the user' points to the P3P Warning dialog.
- 'Matching Status of Site Policy and User's Preference' points to the 'inform' status.
- 'Change User's Preference' points to the radio buttons.

- Implemented as an IE5 Helper Object
- Warn user before accessing *dangerous* page
- Enables user to:
  - View the site policy's detail
  - Change user's preference interactively

# Idcide Privacy Companion

---

WWW10 Tutorial – May 1, 2001

- Browser plug-in for Netscape or Internet Explorer (4.0-5.01) browsers
  - Includes icons to let users know that sites use first- and/or third-party cookies
  - Offers different privacy level that controls the cookie types allowed (1<sup>st</sup> or 3<sup>rd</sup> party)
  - Prevents data spills to 3<sup>rd</sup> parties through “referer” header
  - Lets users view tracking history
- Prototype P3P-enabled Privacy Companion allows for more fine-grained automatic decision making based on P3P policies

<http://www.idcide.com>

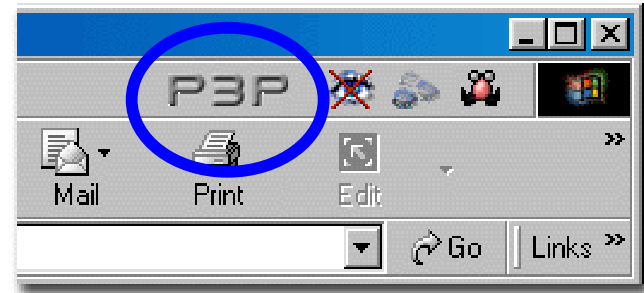
# I Decide P3P Indicator

WWW10 Tutorial – May 1, 2001



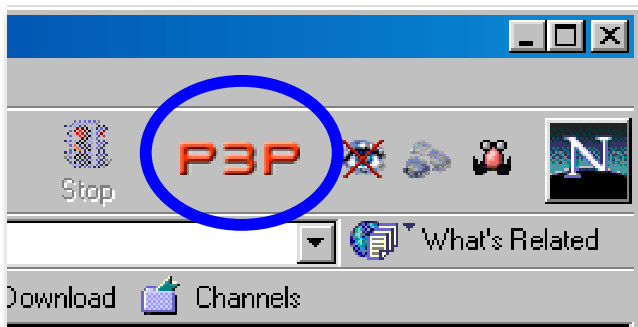
[www.idcide.com](http://www.idcide.com)

Searching for a P3P policy



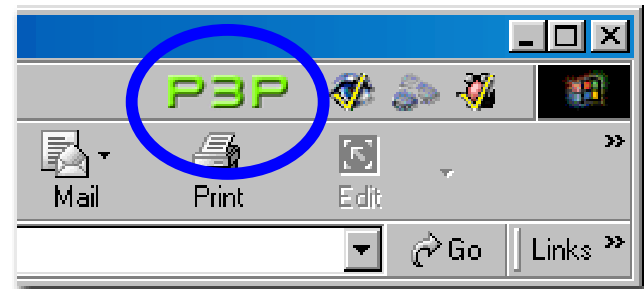
[www.idcide.com](http://www.idcide.com)

No P3P policy found



[www.idcide.com](http://www.idcide.com)

P3P policy is NOT acceptable



[www.idcide.com](http://www.idcide.com)

P3P policy is acceptable

# YOUpowered Orby

---

WWW10 Tutorial – May 1, 2001

- A tool bar that sits at the top of a user's desktop and allows a user to
  - Accept or deny cookies while surfing
  - Decide how, when and where to share personal information
  - Store website passwords
  - Enjoy the convenience of "one-click" form-fill
- P3P features in prototype automatically rate web sites based on their P3P policies

<http://www.youpowered.com>

# YOUpowered Orby

WWW10 Tutorial – May 1, 2001



The screenshot shows a Microsoft Internet Explorer window with the following components:

- Site Information Table:**

Site	Trust	Flags
a1828.ms.a.microsoft.com	██████████	
ads.web.aol.com	██████████	
c.microsoft.com	██████████	
images-eu.amazon.com	██████████	
p3ptestbed-1.w3.org	██████████	P3P
rt2.alt-idns.com	██████████	P3P
specialoffers.aol.com	██████████	
stats.klsoft.com	██████████	
www.amazon.de	██████████	
www.aol.com	██████████	P3P
www.cluborby.com	██████████	P3P
www.ibm.com	██████████	P3P
www.microsoft.com	██████████	
www.research.att.com	██████████	P3P
www.w3.org	██████████	P3P
www.whitehouse.gov	██████████	
www.youpowered.com	██████████	P3P
- Site Information Flags:**
  - OrbyPositives: P3P icon
  - OrbyNegatives: P3P icon
- Detailed P3P Policy Information:**

http://www.w3.org/2000/
- Browser Status Bar:**
  - YOUpowered logo
  - Anonymous profile selected
  - Orby Trust: private (red indicator)
  - Trust Meter: 10 bars (5 green, 5 yellow)
  - P3P indicator: P3P icon
  - Trusted sites: green checkmark

**MULTIPLE PROFILES**  
User can select from multiple, custom defined profiles.

**PRIVACY PREFERENCES**  
Three predefines security settings: Open, Trusting, Cautious, Private.

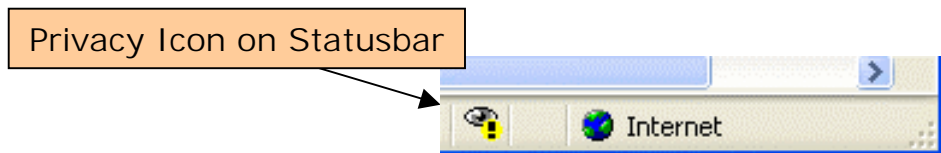
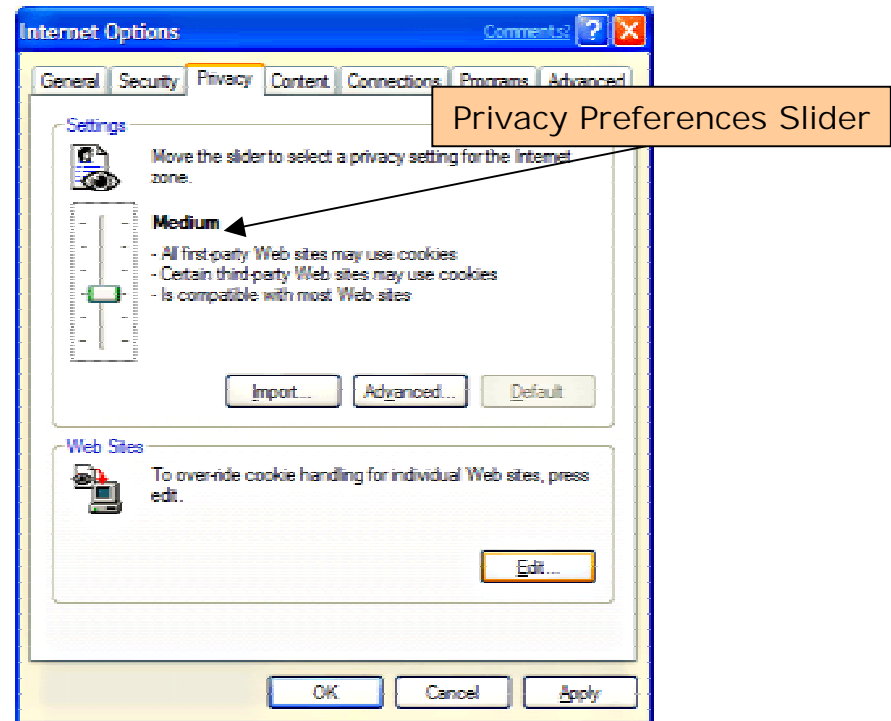
**TRUST METER**  
Shows how well privacy preferences match current site policy.

**P3P INDICATOR**  
Shows if site offers P3P policy.

# MS Internet Explorer 6

WWW10 Tutorial – May 1, 2001

- Uses P3P for advanced cookie filtering
- Implements only subset of P3P
  - Compact policies only
  - Only certain recipients, purposes and categories
- Public Preview available 4/2001
  - Limited subset of above functionality



<http://msdn.microsoft.com/workshop/security/privacy/ie6privacyfeature.asp>

# IBM P3P Policy Editor

---

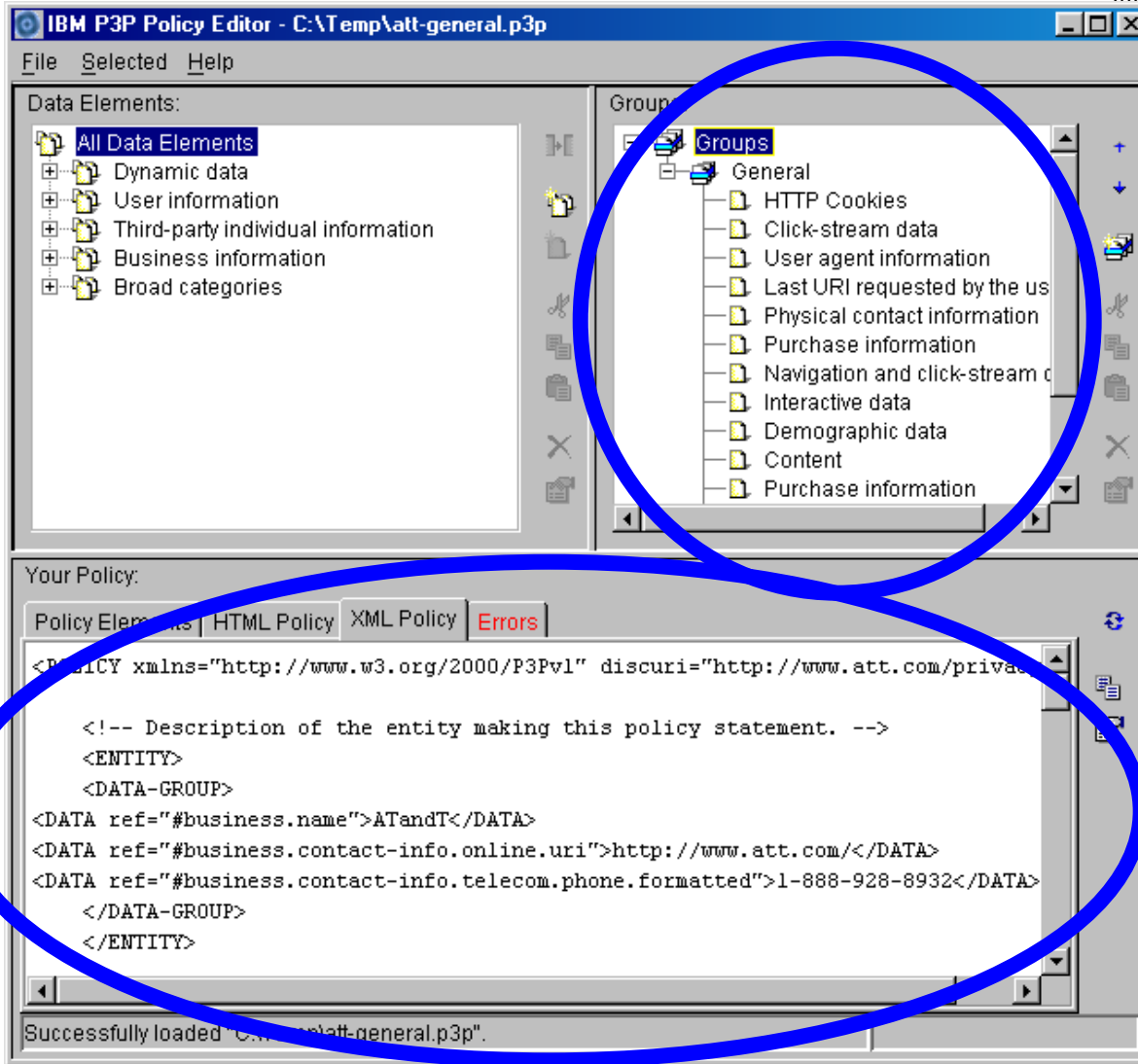
WWW10 Tutorial – May 1, 2001

- Allows web sites to create privacy policies in P3P and human-readable format
- Drag and drop interface
- Available from IBM AlphaWorks site:

<http://www.alphaworks.ibm.com/tech/p3peditor>

# IBM P3P Policy Editor

WWW10 Tutorial – May 1, 2001



Sites can list the types of data they collect

And view the corresponding P3P policy

# Summary & Outlook

WWW10 Tutorial  
May 1, 2001

- P3P
  - Overview
  - Referencing Policies
  - Vocabulary
  - Base Data Set
- P3P Deployment
  - Site Installation
  - Client Examples
- Summary & Outlook

# Topics Discussed

---

WWW10 Tutorial – May 1, 2001

- What is Privacy?
- How do they get my Data?
- Solutions
- Privacy Tools
- P3P (and APPEL)
- Deployment

# Internet Privacy

---

WWW10 Tutorial – May 1, 2001

- Data is often collected silently
  - Web allows large quantities of data collected cheaply & **unobtrusively**
- Data from multiple sources may be merged
  - Non-identifiable information can easily become identifiable when merged
- Solutions exist that
  - provide anonymity
  - ensure private communications
  - provide base-level of trust
  - help manage personal data
- No single tool does it all!

# P3P

---

WWW10 Tutorial – May 1, 2001

- Is ...
  - a user empowerment tool
  - is not a solution in itself
  - a first step (1.0), aimed at ease of deployment
- Provides ...
  - a vocabulary & base data set to express privacy practices
  - a protocol for publishing privacy practices
- Needs ...
  - no special software on server side
  - P3P-aware client software, tools
  - industry support

# A Glimpse of the Future?

WWW10 Tutorial – May 1, 2001



**Creative Labs Nomad JukeBox**  
Music transfer software reports all uploads to Creative Labs.

<http://www.nomadworld.com/welcome.asp>



**Sportbrain**  
Monitors daily workout. Custom phone cradle uploads data to company Web site for analysis.

<http://www.sportbrain.com/>



**:CueCat**  
Keeps personal log of advertisements you're interested in.

<http://www.crq.com/cuecat.html>



**Sony eMarker**  
Lets you figure out the artist and title of songs you hear on the radio. And keeps a personal log of all the music you like on the emarker Web site.

<http://www.emarker.com>



See <http://www.privacyfoundation.org/>

# Resources and Feedback

---

WWW10 Tutorial – May 1, 2001

For further info on P3P see  
**<http://www.w3.org/P3P/>**

Send comments to  
**[www-p3p-public-comments@w3.org](mailto:www-p3p-public-comments@w3.org)**