Online Privacy

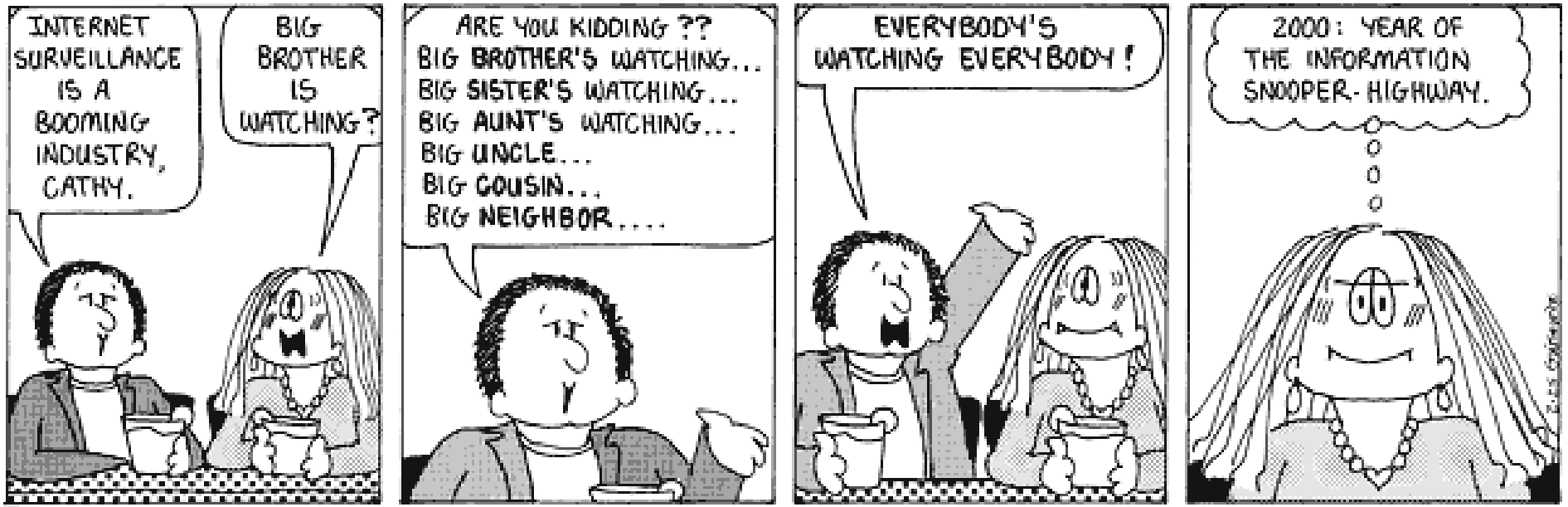
Promise or Peril?

Lorrie Faith Cranor
AT&T Labs-Research

http://lorrie.cranor.org/

Online privacy in the comics!

Cathy **February 25, 2000**



Why is Cathy concerned?

Cathy **March 1, 2000**

How did Irving find this out?

- He snooped her email
- He looked at the files on her computer
- He observed the "chatter" sent by her browser
- He set cookies through banner ads and "web bugs" that allowed him to track her activities across web sites

What do browsers chatter about?

- **Browsers chatter about**
  - IP address, domain name, organization,
  - Referring page
  - Platform: O/S, browser
  - What information is requested
    - URLs and search terms
  - Cookies

- **To anyone who might be listening**
  - End servers
  - System administrators
  - Internet Service Providers
  - Other third parties
    - Advertising networks
  - Anyone who might subpoena log files later

A typical HTTP request

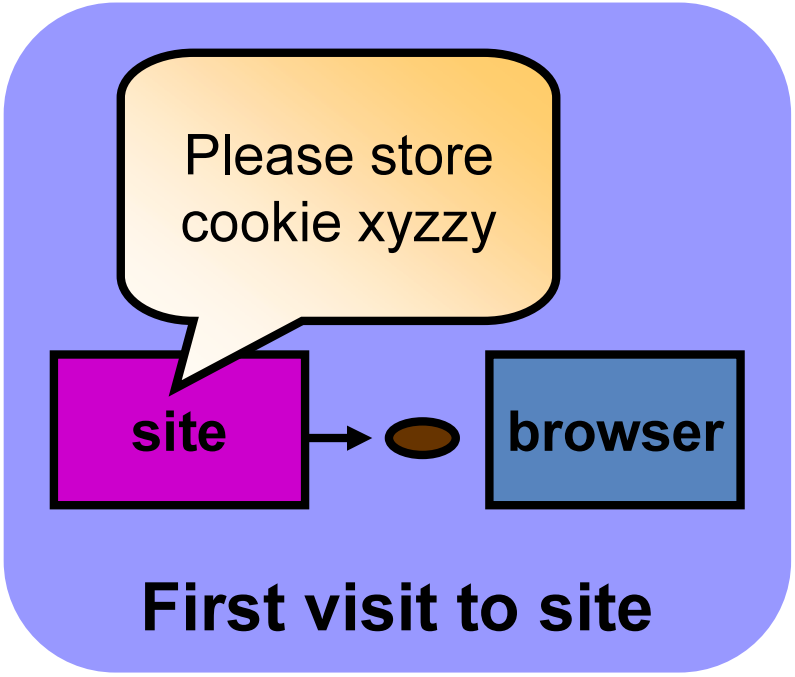
GET /retail/searchresults.asp?qu=beer HTTP/1.0
Referer: http://www.us.buy.com/default.asp
User-Agent: Mozilla/4.75 [en] (X11; U; NetBSD 1.5_ALPHA i386)
Host: www.us.buy.com
Accept: image/gif, image/jpeg, image/pjpeg, */*
Accept-Language: en
Cookie: buycountry=us; dcLocName=Basket; dcCatID=6773; dcLocID=6773; dcAd=buybasket; loc=; parentLocName=Basket; parentLoc=6773; ShopperManager%2F=ShopperManager%2F=66FUQULL0QBT8MMTVSC5MMNKBJFWDVH7; Store=107; Category=0

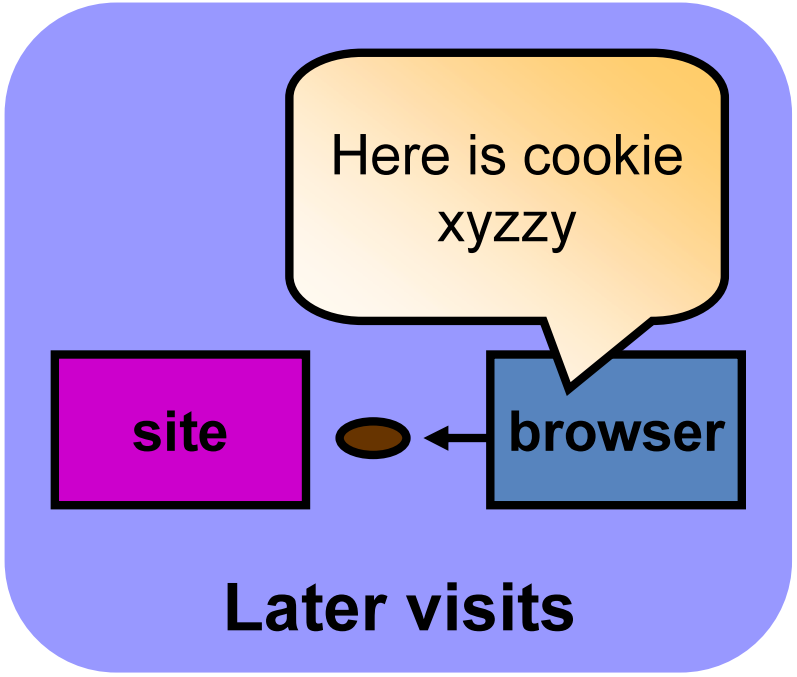
What about cookies?

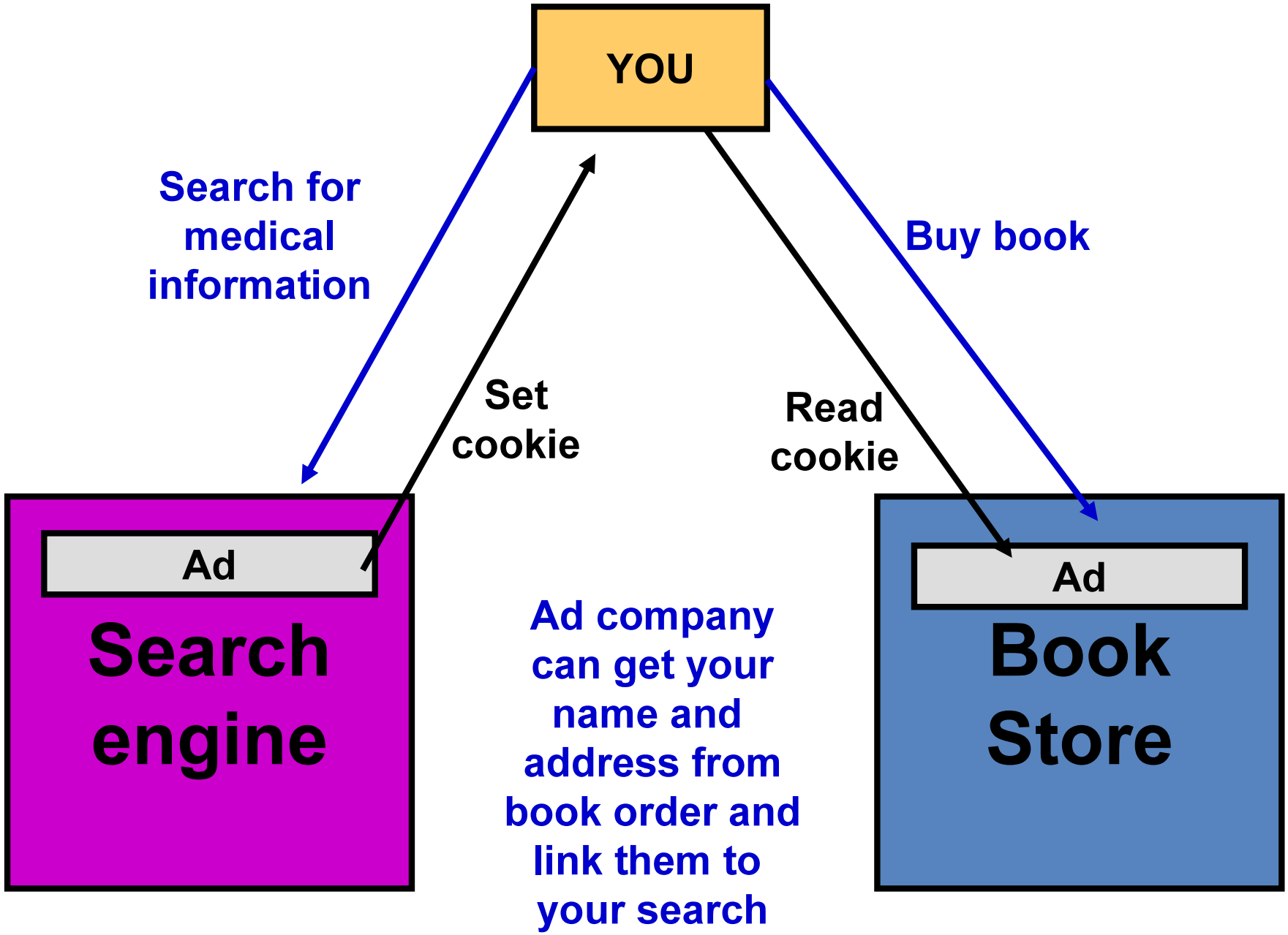
- **Cookies can be useful**
  - ★ used like a staple to attach multiple parts of a form together
  - ★ used to identify you when you return to a web site so you don't have to remember a password
  - ★ used to help web sites understand how people use them
- **Cookies can do unexpected things**
  - ★ used to profile users and track their activities, especially across web sites

How do cookies work?

- A cookie stores a small string of characters
- A web site asks your browser to "set" a cookie
- Whenever you return to that site your browser sends the cookie back automatically
- Cookies are only sent back to the site that set them






YOU
Search for medical information
Buy book
Set cookie
Read cookie
Ad
Search engine
Ad
Book Store
Ad company can get your name and address from book order and link them to your search

Web bugs

- Invisible "images" embedded in web pages that cause 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/education/webbug.html

Referer log problems

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

```
http://www.merchant.com/cgi_bin/order?name=Tom+Jones&address=here+there&credit+card=23487692323 4&PIN=1234& -> index.html
```

What DoubleClick knows…

… about Richard M. Smith

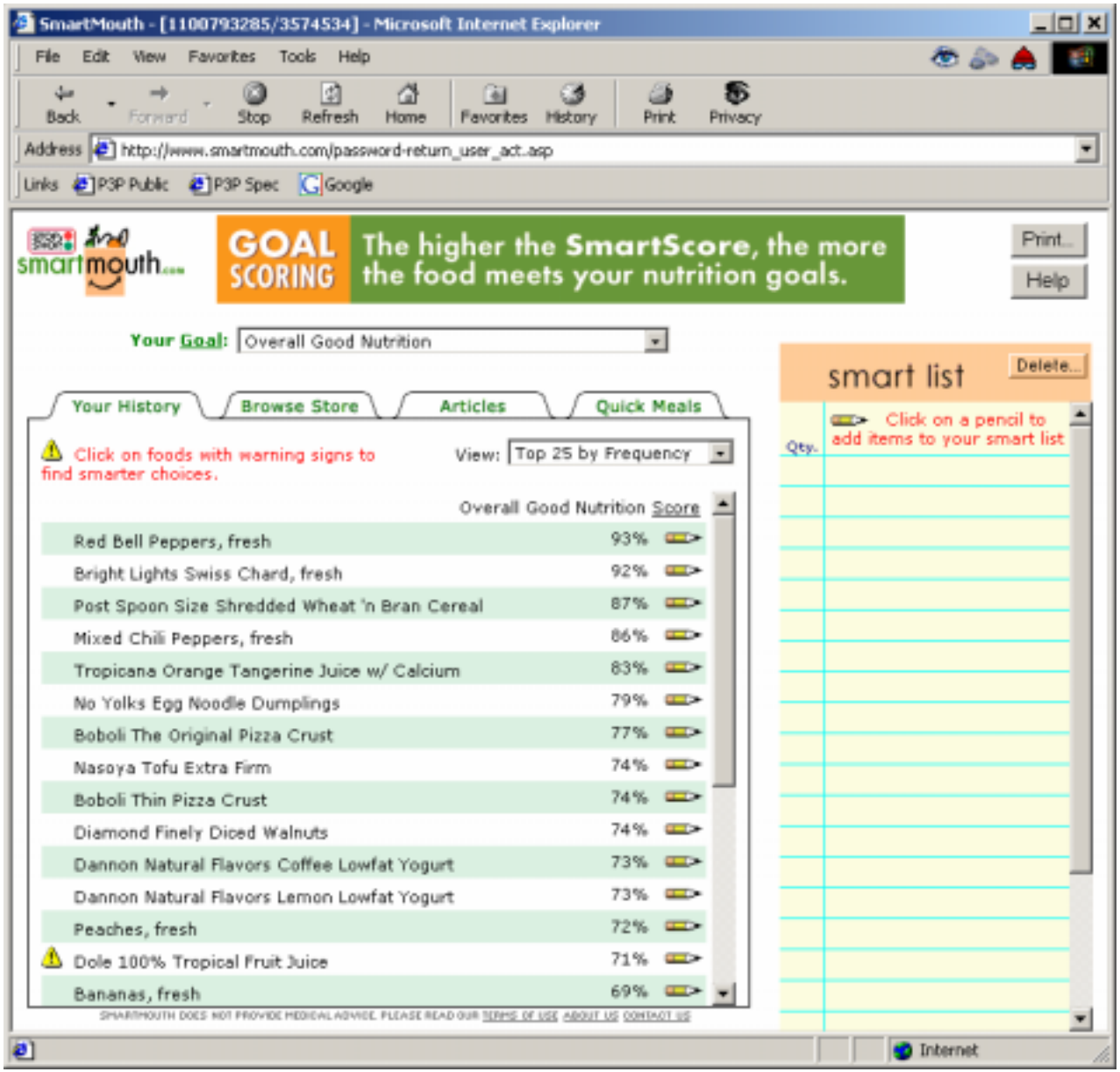
- **Personal data:**
  - ★My Email address
  - ★My full name
  - ★My mailing address (street, city, state, and Zip code)
  - ★My phone number
- **Transactional data:**
  - ★Names of VHS movies I am interesting in buying
  - ★Details of a plane trip
  - ★Search phrases used at search engines
  - ★Health conditions

No clicks required

"It was not necessary for me to click on the banner ads for information to be sent to DoubleClick servers."

- *Richard M. Smith*

Offline data goes online…
My 25 most frequent grocery purchases



My purchase patterns have changed recently

Public concern

- **April 1997 Louis Harris Poll of Internet users**
  - ★5% say they have been the victim of an invasion of privacy while on the Internet
  - ★53% say they are concerned that information about which sites they visit will be linked to their email address and disclosed without their knowledge

Beyond concern

- **April 1999 Study: *Beyond Concern: Understanding Net Users' Attitudes About Online Privacy* by Cranor, Ackerman and Reagle (US panel results reported)**

  http://www.research.att.com/projects/privacystudy/

  - ★ Internet users more likely to provide info when they are not identified
  - ★ Some types of data more sensitive than others
  - ★ Many factors important in decisions about information disclosure
  - ★ Acceptance of persistent identifiers varies according to purpose
  - ★ Internet users dislike automatic data transfer

March 2000 BusinessWeek poll

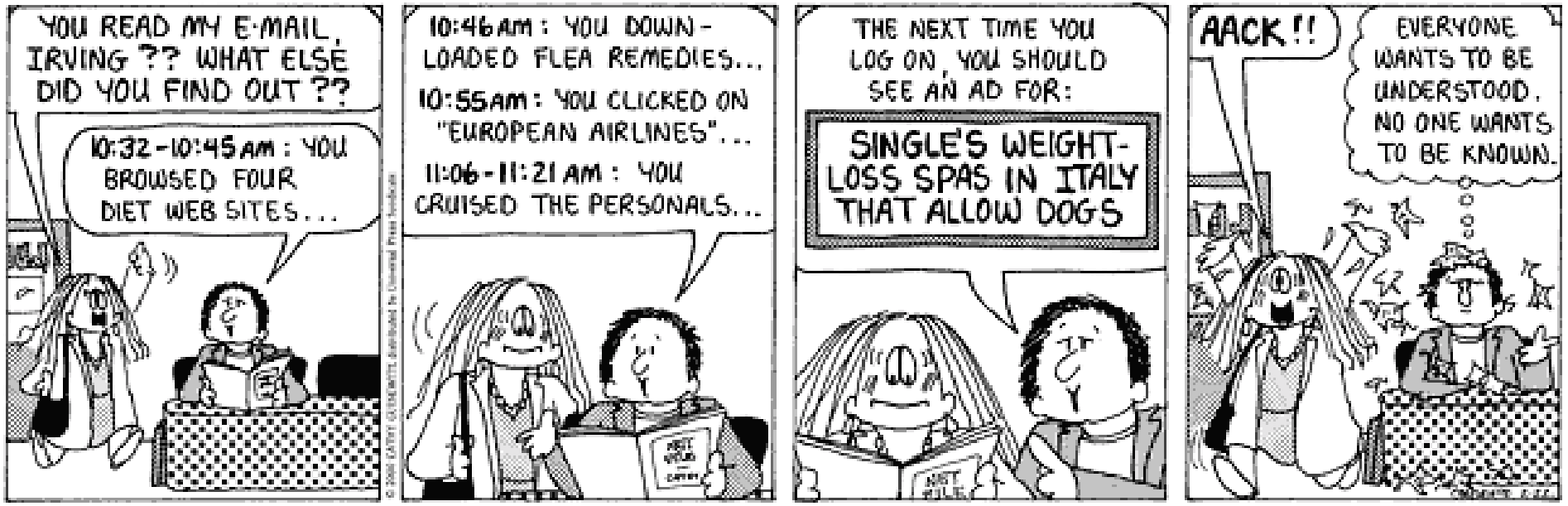
- **Telephone survey of 1,014 US adults by Harris Interactive**

  http://businessweek.com/2000/00_12/b3673006.htm

  - ★63% not comfortable with anonymous online profiling
  - ★89% not comfortable with identified online profiling
  - ★95% not comfortable with identified online profiling that includes sensitive information
  - ★91% not comfortable with web sites sharing their info to track them across multiple sites

No one wants to be known

Cathy **February 22, 2000**



IBM-Harris multi-national survey

- **Telephone interviews with 1000+ adults in each of three countries: US, UK, Germany**

  http://www.ibm.com/services/e-business/priwkshop.html

  - ★ Americans profess the greatest degree of confidence in the way companies handle their personal information, but Americans also are the most likely among the three groups of citizens to take steps to protect their privacy.
  - ★ Americans appear to be motivated to take privacy protection measures, not so much from a set of specific concerns, but by a general sense that their personal information may be misused.

International issues

- **European Union Data Directive prohibits secondary uses of data without informed consent**
  - ★Creating personally-identifiable online profiles will have to be **opt-in** in most cases
  - ★Upfront notice must be given when data is collected – no web bugs
  - ★No transfer of data to non-EU countries unless there is adequate privacy protection

Children's issues

Children's Online Privacy Protection Act (COPPA) requires parental consent before collecting personally-identifiable data from children online

Subpoenas

- Data on online activities is increasingly of interest in civil and criminal cases
- The only way to avoid subpoenas is to not have data
- Your files on your computer in your home have much greater legal protection that your files stored on a server on the network

Online privacy – key concerns

- **Data is often collected silently**
  - ★ Web allows lots of data to be collected easily, cheaply, unobtrusively and automatically
  - ★ Individuals not given meaningful choice
- **Data from many sources may be merged**
  - ★ Even non-identifiable data can become identifiable when merged
- **Data collected for business purposes may be used in civil and criminal proceedings**

Some solutions

- Privacy policies
- Voluntary guidelines and codes of conduct
- Seal programs
- Chief privacy officers
- Laws and regulations
- Software tools

Privacy policies

- 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 or privacy policies increases consumer trust
- BUT policies are often difficult to understand, hard to find, and take a long time to read
- Many policies are changed frequently without notice

Voluntary guidelines

- Online Privacy Alliance
  http://www.privacyalliance.org
- Direct Marketing Association Privacy Promise
  http://www.thedma.org/library/privacy/privacypromise.shtml
- Network Advertising Initiative Principles
  http://www.networkadvertising.org/

OECD fair information principles

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

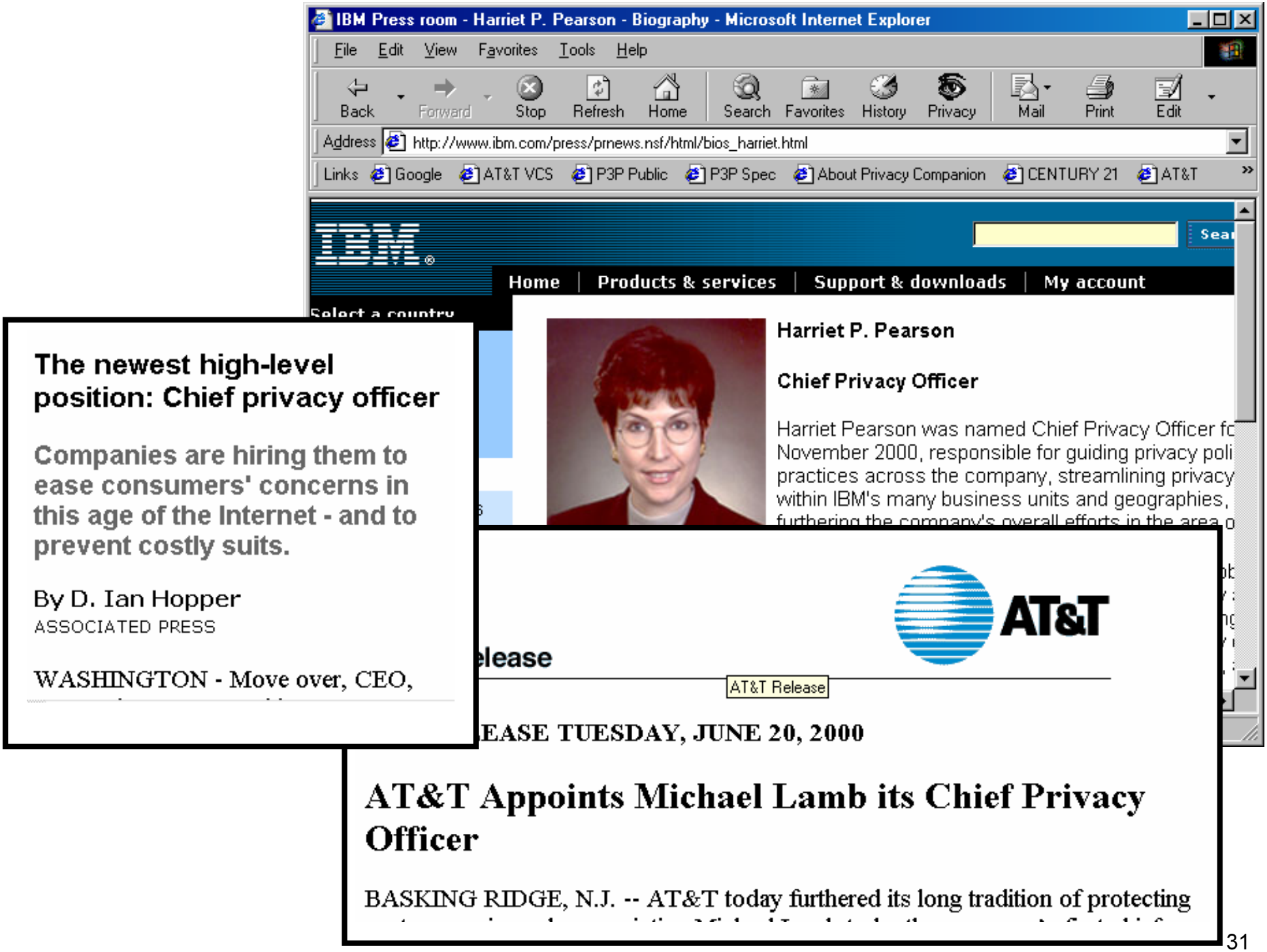
- Collection limitation
- Data quality
- Purpose specification
- Use limitation
- Security safeguards
- Openness
- Individual participation
- Accountability

Simplified principles

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

Seal Programs

- 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/


IBM Press room - Harriet P. Pearson - Biography - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Forward Stop Refresh Home Search Favorites History Privacy Mail Print Edit
Address http://www.ibm.com/press/prnews.nsf/html/bios_harriet.html
Links Google AT&T VCS P3P Public P3P Spec About Privacy Companion CENTURY 21 AT&T
IBM
Home | Products & services | Support & downloads | My account
Harriet P. Pearson
Chief Privacy Officer
Harriet Pearson was named Chief Privacy Officer fo
November 2000, responsible for guiding privacy poli
practices across the company, streamlining privacy
within IBM's many business units and geographies,
The newest high-level position: Chief privacy officer
Companies are hiring them to ease consumers' concerns in this age of the Internet - and to prevent costly suits.
By D. Ian Hopper
ASSOCIATED PRESS
WASHINGTON - Move over, CEO,
AT&T
AT&T Release
EASE TUESDAY, JUNE 20, 2000
AT&T Appoints Michael Lamb its Chief Privacy Officer
BASKING RIDGE, N.J. -- AT&T today furthered its long tradition of protecting

Chief Privacy Officers

- **Companies are increasingly appointing CPOs to have a central point of contact for privacy concerns**
- **Role of CPO varies in each company**
  - ★ Draft privacy policy
  - ★ Respond to customer concerns
  - ★ Educate employees about company privacy policy
  - ★ Review new products and services for compliance with privacy policy
  - ★ Develop new initiatives to keep company out front on privacy issue
  - ★ Monitor pending privacy legislation

Laws and regulations

- **Privacy laws and regulations vary widely throughout the world**
- **US has mostly sector-specific laws, with relatively minimal protections**
  - ★ Federal Trade Commission has jurisdiction over fraud and deceptive practices
  - ★ Federal Communications Commission regulates telecommunications
- **European Data Protection Directive requires all European Union countries to adopt similar comprehensive privacy laws**
  - ★ Privacy commissions in each country (some countries have national and state commissions)

Software tools

- **Anonymity and pseudonymity tools**
  - ★ Anonymizing proxies
  - ★ Mix Networks and similar web anonymity tools
    - Onion routing
    - Crowds
    - Freedom
  - ★ Anonymous email
- **Encryption tools**
  - ★ File encryption
  - ★ Email encryption
  - ★ Encrypted network connections
- **Filters**
  - ★ Cookie cutters
  - ★ Child protection software
- **Information and transparency tools**
  - ★ Identity management tools
  - ★ P3P
- **Other tools**
  - ★ Privacy-friendly search engines
  - ★ Computer "cleaners"
  - ★ Tools to facilitate access

Platform for Privacy Preferences Project (P3P)

- **Developed by the World Wide Web Consortium (W3C)**
  http://www.w3.org/p3p/
- **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 (built into browsers or separate applications) 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

Using P3P on your Web site

1. **Formulate privacy policy**
2. **Translate privacy policy into P3P format**
   - ★ Use a policy generator tool
3. **Place P3P policy on web site**
   - ★ One policy for entire site or multiple policies for different parts of the site
4. **Associate policy with web resources:**
   - ★ Place P3P policy reference file (which identifies location of relevant policy file) at well-known location on server;
   - ★ Configure server to insert P3P header with link to P3P policy reference file; or
   - ★ Insert link to P3P policy reference file in HTML content

The P3P vocabulary

- Who is collecting data?
- What data is collected?
- For what purpose will data be used?
- Is there an ability to 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?

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

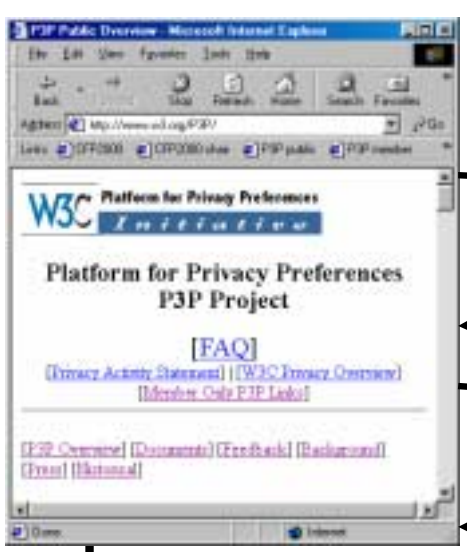
A simple HTTP transaction


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

Web Server

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

… with P3P 1.0 added




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

Web Server

Send Policy Reference File

Request P3P Policy

Send P3P Policy

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

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

User preferences

- **P3P spec does not specify how users should configure their preferences or what user agent should do**
  - ★Some guidelines are offered in Guiding Principles
- **A separate W3C specification - A P3P Preference Exchange Language (APPEL) provides a standard format for encoding preferences**
  - ★Not required for P3P user agent implementations

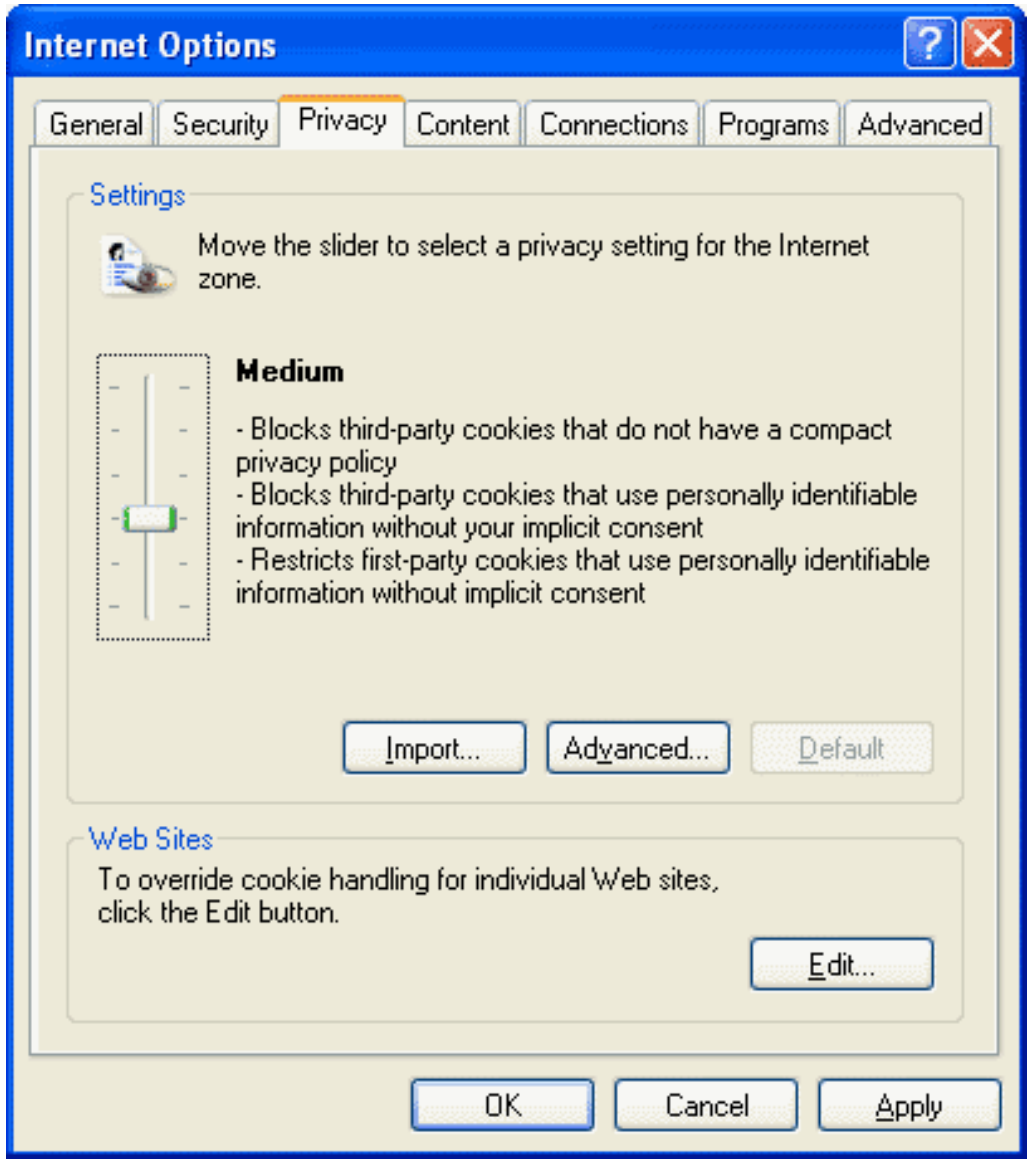
Types of P3P user agent tools

- **On-demand or continuous**
  - ★ Some tools only check for P3P policies when the user requests, others check automatically at every site
- **Generic or customized**
  - ★ Some tools simply describe a site's policy in some user friendly format – others are customizable and can compare the policy with a user's preferences
- **Information-only or automatic action**
  - ★ Some tools simply inform users about site policies, while others may actively block cookies, referrers, etc. or take other actions at sites that don't match user's preferences
- **Built-in, add-on, or service**
  - ★ Some tools may be built into web browsers or other software, others are designed as plug-ins or other add-ons, and others may be provided as part of an ISP or other service

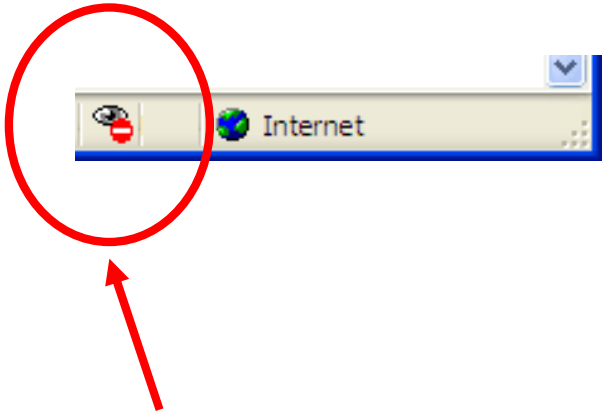
Other types of P3P tools

- **P3P validators**
  - ★ Check a site's P3P policy for valid syntax
- **Policy generators**
  - ★ Generate P3P policies and policy reference files for web sites
- **Web site management tools**
  - ★ Assist sites in deploying P3P across the site, making sure forms are consistent with P3P policy, etc.
- **Search and comparison tools**
  - ★ Compare privacy policies across multiple web sites – perhaps built into search engines

P3P in IE6



Initial focus is on P3P policies for cookies

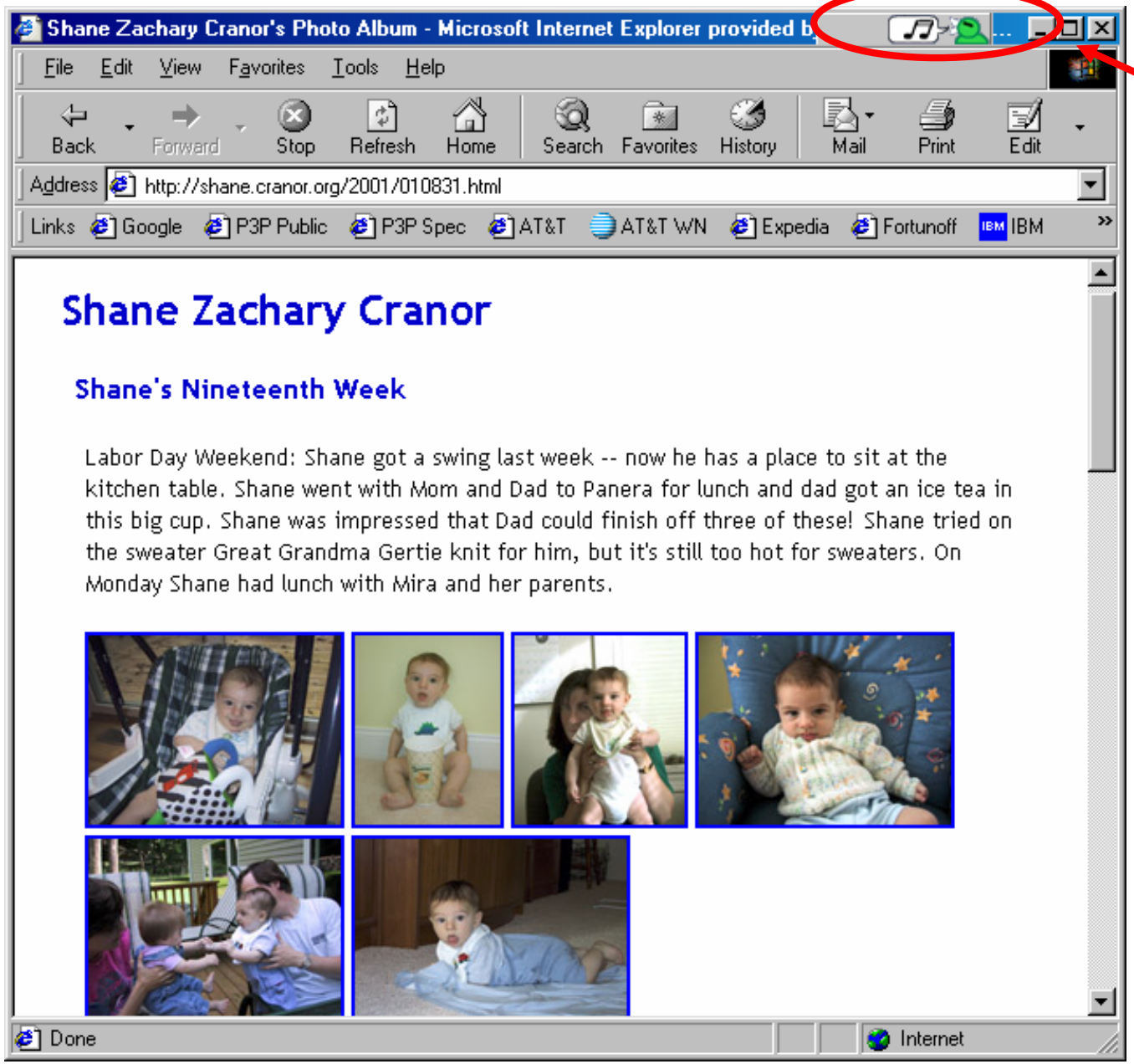


Privacy icon on status bar

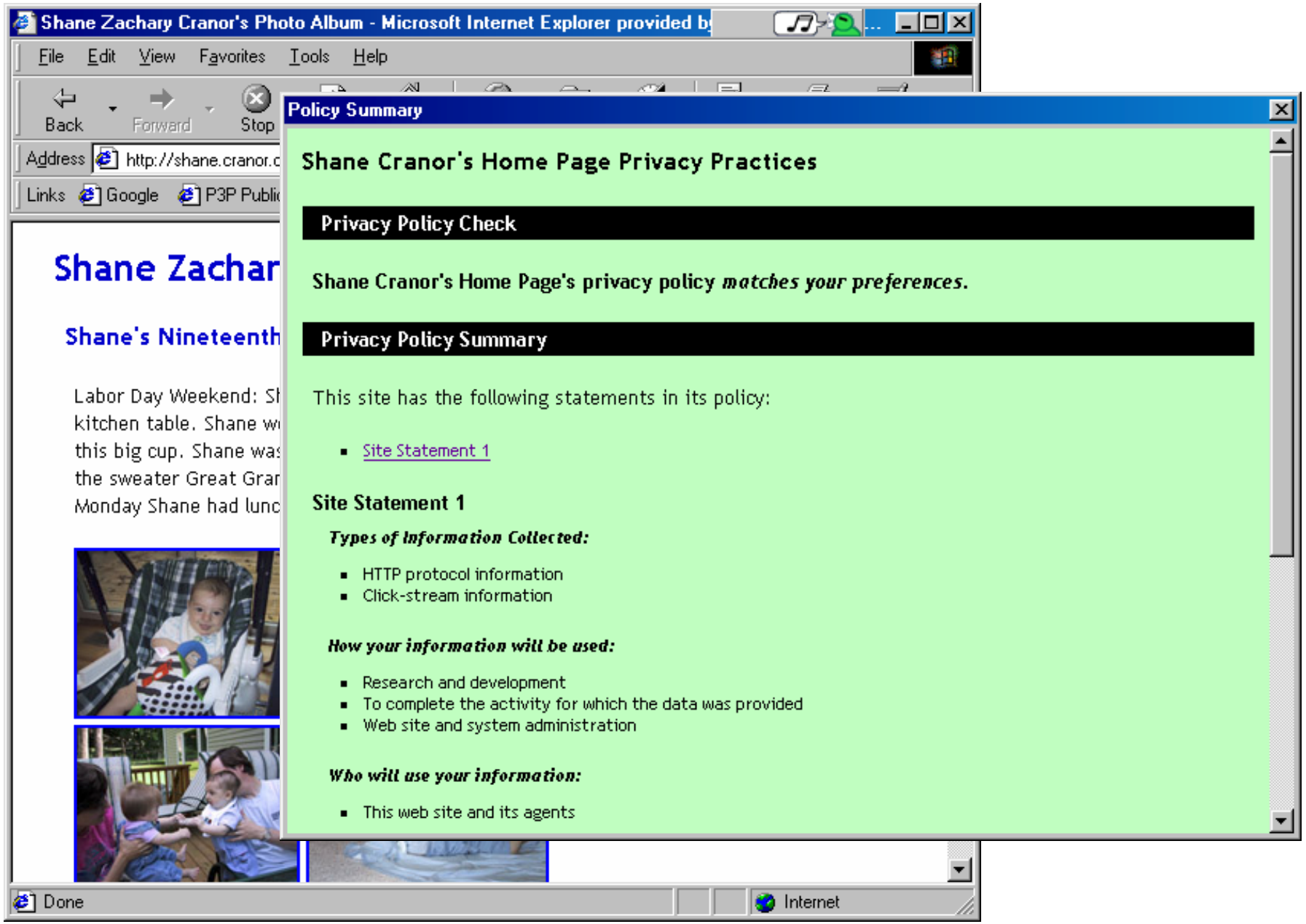
AT&T WorldNet Privacy Tool

- Testing in WorldNet Beta club later this month
- Future FREE public release
- `http://privacy.research.att.com/`

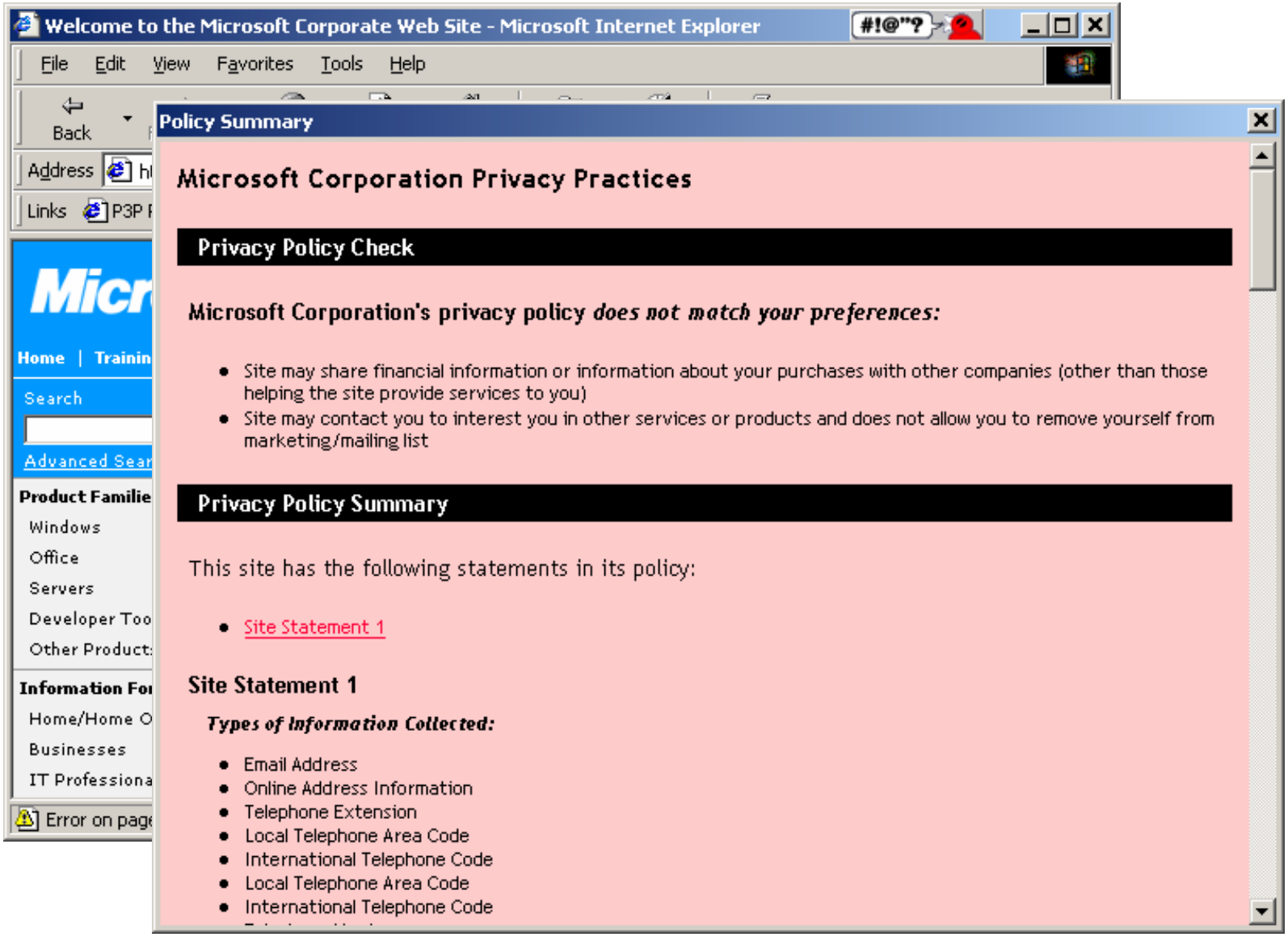
Chirping bird is privacy indicator



Click on the bird for more info



Privacy policy summary - mismatch



P3P deployment

- Look for P3P browsers and plug-ins to be available by the end of the year
- P3P tools for web site developers already available
- Web sites operators should start P3P-enabling their sites now

`http://www.w3.org/p3p/`

Cathy

January 21, 2001