Overview of Today’s Lecture

- Some of the current problems in online privacy
- Tracking mechanisms
  - Cookies
  - Beacons
  - Browser fingerprinting
- Dangers of third-party tracking
- Ad ecosystem and user targeting
- Solutions for tracking prevention
- RePriv: combining personalization and privacy
Web privacy concerns

- Data is often collected silently
  - Web allows large quantities of data to be collected inexpensively and unobtrusively
- Data from multiple sources may be merged
  - Non-identifiable information can become identifiable when merged
- Data collected for business purposes may be used in civil and criminal proceedings
- Users are often given no explicit choice
HTTP Request + Cookie

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
Referer Logging Issues

- GET methods result in values in URL

- These URLs are sent in the referer header to next host

- Somewhat contrived example:

  http://www.ebay.com/cgi_bin/order?name=Bill+Clinton&address=here+there&credit+card=234876923234&PIN=1234& -> index.html
An HTTP cookie, originally invented by Lou Montulli and John Giannandrea at Netscape in 1994, is extremely useful for the web. Cookies are the easiest way to offer "stateful" user interfaces such as user accounts and logins, multi-page forms, or online shopping carts. Cookies also allow sites to store a unique ID in your browser, and to track you. Many people have learned to block, limit or delete their cookies.

Categories of cookies:
- Persistent cookie – cookie replayed until expiration date
- First-party cookie – cookie associated with the site the user requested
- Third-party cookie – cookie associated with an image, ad, frame, or other content from a site with a different domain name that is embedded in the site the user requested.
Tracking Mechanics: Beacons

- Often invisible 1x1 images
- 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, MS Word documents, etc.

Yahoo!’s Practices Regarding Web Beacons

Yahoo! may collect information through web beacons about your web browsing activities such as the address of the page you are visiting, the address of the referer page you previously visited, the time you are viewing the page, your browsing environment and your display settings. We may use the information we collect through web beacons:

- To understand traffic patterns and the number of visitors to the branded Yahoo! network of websites, websites within the Yahoo! Network Plus, and other non-Yahoo! websites that we partner with.
- To understand how you use and interact with Yahoo! products and services, including, but not limited to, the use of Yahoo! Mail outside of a browser-based experience.
- To improve Yahoo! products and services.
- To optimize your browsing experience.
- To provide anonymous individual and/or aggregate auditing, research, modeling and reporting for our advertisers and other partners. No personally identifiable information about you is shared with our advertisers and other partners as part of these services.
Tracking Mechanics: Fingerprinting

Panopticlick

How Unique — and Trackable — Is Your Browser?

Is your browser configuration rare or unique? If so, websites may be able to track you, even if you limit or disable cookies.

Panopticlick tests your browser to see how unique it is based on the information it will share with sites it visits. Click below and you will be given a uniqueness score, letting you see how easily identifiable you might be as you surf the web.

Only anonymous data will be collected by this site.

A paper reporting the statistical results of this experiment is now available: How Unique Is Your Browser?, Proceedings of the Privacy Enhancing Technologies Symposium (PETS 2013), Springer Lecture Notes in Computer Science.
## Panopticlick Results

Your browser fingerprint appears to be unique among the 1,865,596 tested so far. 
Currently, we estimate that your browser has a fingerprint that contains at least 20.83 bits of identifying information.

### Browser Characteristics

<table>
<thead>
<tr>
<th>Browser Characteristic</th>
<th>bits of identifying information</th>
<th>one in x browsers have this value</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Agent</td>
<td>17.51</td>
<td>108559.8</td>
</tr>
<tr>
<td>HTTP_ACCEPT Headers</td>
<td>3.83</td>
<td>14.25</td>
</tr>
</tbody>
</table>
| Plugin 0: BitDefender QuickScan: BitDefender QuickScan Web Necesso Plugin: mpqscanner.dll; mpqscanner; application/x-bitdefender-quickscanner; ), Plugin 1: Chrome PDF Viewer: pdf.dll; (Portable Document Format: application/pdf) (Portable Document Format: application/x-google-chrome-print-preview-pdf; pdf), plug-in 2: Default Plug-In: Plug-in functionality for installing third-party plug-in: default, plug-in, *, Plugin 3: Google Update: Google Update: no/GoogleUpdate3.dll; dpl; application/x-google-update-control), Plugin 4: Microsoft Office 2010: Office Authorization plug-in for FPAPI browsers: HFUITHDL.DLL; (4.0.7730.1000; application/vnd.msexcel), Plugin 5: Microsoft Office 2010: The plugin allows you to open and edit files using Microsoft Office applications: HFUITHDL.DLL; (SharePoint Plugin for Firefox: application/octet-stream; ), Plugin 6: Native Client: .ppGoogle{i}c{3}PluginChrome.dll; (Native Client Executables; application/octet-stream; native)), Plugin 7: Picasa: Picasa plugin: npPicasa3.dll; (3.1; application/octet-stream; plugin), Plugin 8: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plugin 7.7: The QuickTime Plugin allows you to view a wide variety of multimedia content in Web pages. For more information, visit the <A HREF=https://www.apple.com/quicktime/QuickTimeDownload Web site; npqplugplay.dll; (SDP stream descriptor; application/x-sdp) (RTSP stream descriptor; application/x-rtsp), Plug-in 10: QuickTime Plug...
Third-Party Tracking

- A third party is typically an advertiser or ad network

- Their content is placed alongside primary (first-party) content

- Requests go to their site and result in
  - Referred often containing the URL and user identifying information to be sent to the site
  - An ID that is stored in the cookie for cross-correlation
  - Date, time, etc.
Clickstreams

- In the language of computer science, clickstreams – browsing histories that companies collect – are not anonymous at all; rather, they are pseudonymous.

- The latter term is not only more technically appropriate, it is much more reflective of the fact that at any point after the data has been collected, the tracking company might try to attach an identity to the pseudonym (unique ID) that your data is labeled with.

- Thus, identification of a user affects not only future tracking, but also retroactively affects the data that's already been collected. Identification needs to happen only once, ever, per user.

Arvind Narayanan, Stanford
Magnitude of the Problem

- Recorded interactions with 120 popular sites for information leakage to third parties
- Found that
  - 56% leaked some form of private information
  - 48% leaked a user identifier
Suppose you find the same username on different online services, what is the probability that these usernames refer to the same physical person?

Our experiments, based on crawls of real web services, show that a significant portion of the users' profiles can be linked using their usernames.

To the best of our knowledge, this is the first time that usernames are considered as a source of information when profiling users on the Internet.
Recent Stanford Experiments

- Picked 185 popular sites
- Used FourthParty web measurement platform to create an account and interact with the site
- Explored content that dealt with user identity, such as profile and settings pages
- After collecting data, searched Request-URIs and Referer headers for known personal information

- User name/ID leaked in 113 websites or 61%

![Bar chart showing leaked information]

- facebook.com
- doubleclick.net
- quantserve.com
- google-analytics.com
- scorecardresearch.com

http://donottrack.us/blogs/
More Results from the Stanford Study

- Viewing a local ad on the Home Depot website sent the user's first name and email address to 13 companies
- Entering the wrong password on the Wall Street Journal website sent the user's email address to 7 companies
- Changing user settings on the video sharing site Metacafe sent first name, last name, birthday, email address, physical address, and phone numbers to 2 companies
- Signing up on the NBC website sent the user's email address to 7 companies
- Signing up on Weather Underground sent the user's email address to 22 companies.
- The mandatory mailing list page during CNBC signup sent the user's email address to 2 companies.
- Clicking the validation link in the Reuters signup email sent the user's email address to 5 companies.
- Interacting with Bleacher Report sent the user's first and last names to 15 companies.
- Interacting with classmates.com sent the user's first and last names to 22 companies.
Many first-party websites make what would appear to be incorrect, or at minimum misleading, representations about not sharing PII. Here are some examples:

- **The Home Depot:**
  - Personal Information Disclosure: The Home Depot will not trade, rent or sell your personal information, without your prior consent, except as otherwise set out herein. [Does not describe sharing with third-parties for advertising or analytics.]

- **The Wall Street Journal:**
  - We will not sell, rent, or share your Personal Information with these third parties for such parties' own marketing purposes, unless you choose in advance to have your Personal Information shared for this purpose. Information about your activities on our Online Services and other non-personally identifiable information about you may be used to limit the online ads you encounter to those we believe are consistent with your interests. Third-party advertising networks and advertisers may also use cookies and similar technologies to collect and track non-personally identifiable information such as demographic information, aggregated information, and Internet activity to assist them in delivering advertising on our Online Services that is more relevant to your interests.
Players in the Online Space: Ad Scenario

- Ad networks
- Hosts – sites on which ads are placed
- Users – some are concerned about their privacy
Ad Targeting

- The better (more relevant) ads are, the more they appeal to the user
- The more they appeal to the user, the higher the click-through rates (CTR) become
- The more click the advertising network gets, the more they get paid (pay-per-click)

- How do we create more relevant ads?
- Need to know what the user finds relevant
- How can we find that out?
- One option is to do user profiling/modeling
- Followed by ad targeting
Tracking Prevention Solutions

1. Browser privacy modes
2. Opting out of cookie-based tracking
3. "Do Not Track (DNT)"
4. Tracking Protection Lists (TPLs)
Browser Privacy Modes

- Prevent access to persistent user data
- Prevent storing persistent data
- Cleanse referers

<table>
<thead>
<tr>
<th>Privacy Mode Comparison</th>
<th>Chrome's Incognito</th>
<th>IEB's InPrivate Browsing</th>
<th>Firefox 3.5's Private Browsing</th>
<th>Safari's Private Browsing²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visited sites are not stored in the browser history</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Downloaded files are not stored in the download history</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Form field data (including passwords) is not stored</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Addresses typed into the address bar are not stored</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Visited links are not stored</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Search queries are not stored in the browser</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cached files are deleted at the end of the browsing session</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Existing third-party cookies cannot be read</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Privacy Mode Comparison</th>
<th>Chrome's Incognito</th>
<th>IEB's InPrivate Browsing</th>
<th>Firefox 3.5's Private Browsing</th>
<th>Safari's Private Browsing²</th>
</tr>
</thead>
<tbody>
<tr>
<td>New cookies are deleted at the end of the session</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Blocks referring URL from being sent</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mode can operate on a per-window basis</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mode can persist even when user quits and re-starts browser</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Controlling Cookie Access
InPrivate Filtering in IE8/IE9

When you visit multiple websites that contain content from the same provider, such as a map, advertisement, or web measurement tools, some information about your visits might be shared with the content provider. If you choose to block content, portions of the websites you visit might not be available.

- **Automatically block**
- **Choose content to block or allow**
- **Off**

### Content Provider

<table>
<thead>
<tr>
<th>Content provider</th>
<th>Status</th>
<th>Used by</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>2mdn.net</td>
<td>Blocked</td>
<td>3 sites</td>
<td>flashwrite_1_2.js</td>
</tr>
<tr>
<td>atdmt.com</td>
<td>Blocked</td>
<td>3 sites</td>
<td>pixel.gif</td>
</tr>
</tbody>
</table>

More information from this content provider

- **Allow**
- **Block**

3 Show content from providers used by this number of websites you’ve visited. (3-30)

Advanced settings

Learn more about InPrivate Filtering

OK Cancel
Opting out of Cookie-based Tracking

- Instead of preventing cookie access, explicitly set opt-out cookies
- Many ad networks provide mechanisms for this
- There are tools to help you set the right cookie: SelectOut.org
Manipulating Opt-Out Cookies
"Do Not Track (DNT)

- The Do Not Track proposal is to include a simple, machine-readable header indicating that you don't want to be tracked. The header that would be inserted is **DNT : 1**

- Because this signal is a header, and not a cookie, users will be able to clear their cookies at will without disrupting the functionality of the Do Not Track flag.

- It’s important to note that there is no "list" that consumers need to sign up for. Early discussion of Do Not Track included proposals about a list-based registry of users, similar to the Do Not Call Registry. This proposal does not collect data on consumers in a central list.
DNT: Fear, Uncertainty, and Doubt

**Recent Updates**

EFF IN THE NEWS | November 13, 2011
Advertisers Can’t Be Trusted to Self-regulate on Data Collection, Says EFF

DEEPLINKS BLOG | November 14, 2011
The DAA’s Self-Regulatory Principles Fall Far Short of Do Not Track

DEEPLINKS BLOG | October 21, 2011
An EFF Guide to the Silicon Valley Human Rights Summit

EFF IN THE NEWS | October 6, 2011
Kindle Fire’s Silk browser raises issues over website tracking history of users

DEEPLINKS BLOG | September 22, 2011
EFF Advocates for User Privacy in W3C Workshop on Do Not Track

EFF IN THE NEWS | August 18, 2011
Do you know about your digital fingerprint?

EFF IN THE NEWS | August 18, 2011
Tracking Your Every Move on the Internet

DEEPLINKS BLOG | July 7, 2011
EFF Urges Senators to Recognize Need for Updated Privacy Laws

EFF IN THE NEWS | May 27, 2011
'Like' Button Follows Web Users

EFF IN THE NEWS | May 27, 2011
US Senate Sinks Its Teeth Into Online Privacy Reform
Tracking Protection Lists (TPLs)

Abine, The Online Privacy Company, is the leading provider of online privacy solutions for consumers. Abine’s products and services allow regular people to regain control over their personal information while continue browsing, interact and shop online.

Abine’s Tracking Protection List blocks many online advertising and marketing technologies that can track and profile you as you browse the Web. This list is updated weekly to keep you safer and more private.

Visit the Abine website for more information about this Tracking Protection List.

EasyPrivacy Tracking Protection List is based on the popular EasyPrivacy subscription for Adblock Plus and is managed by the well known EasyList project, which serves nearly ten million daily users and has a large support forum with dozens of experienced members able to assist resolving any issues that may arise.

Visit the EasyList website for more information about this Tracking Protection List.

PrivacyChoice maintains a comprehensive database of tracking companies, including domains used by nearly 300 ad networks and platforms, tracking methods, summaries of key policies, oversight, and opt-out and opt-in processes.

PrivacyChoice has created Tracking Protection Lists based on this data. You have the option of installing two lists. The first list blocks companies that are not subject to oversight by the NAI and the second list blocks all tracking company domains in the PrivacyChoice database. These lists will be automatically updated with new tracking domains discovered through continuous website scanning and user panels.

Visit the PrivacyChoice website for more information about this Tracking Protection List.

TRUSTe is the leading online privacy certification and services provider. TRUSTe’s TRUSTe Certified Tracking Protection List enables relevant and targeted ads from companies that demonstrate respectful consumer privacy practices and comply with TRUSTe’s high standards and direct oversight. TRUSTe helps users get good ads, without compromising personal privacy.

Visit the TRUSTe website for more information about this Tracking Protection List.
Tracking Protection Lists (TPLs)

How do they work?

- The websites you visit often contain content from third parties. In order to load this content, certain information about your computer, including your IP address and the address of the webpage you’re viewing, is sent to each of the third parties. If a site is listed as a “do not call” site on a TPL, Internet Explorer 9 will block third-party content from that site, unless you visit the site directly by clicking on a link or typing its web address. By limiting “calls” to third-party websites, Internet Explorer 9 limits the information these third-party sites can collect about you.

Do TPLs only block third-party calls?

- TPLs can include “do not call” or “OK to call” entries that permit calls to specific third-party sites. Please be aware that if there are conflicts between “do not call” and “Ok to call” TPLs, the “Ok to call” rules will govern. You should review carefully the TPLs that you choose to download to ensure that you want to allow calls to each of the sites included in any “Ok to call” list.
Privacy in the News

• Concerns about tracking
• Personal data siloed away
• Browser features help
• Legislative pressure
What are some of the reasons for the outrage caused by third-party tracking?
RePriv

Re-Envisioning In-Browser Personalization & Privacy

[Oakland S&P 2011]

Ben Livshits
Microsoft Research
users want a highly personalized web experience
Share data to get personalized results

Privacy concerns
Browser: Personalization & Privacy

• Broad applications:
  – Site personalization
  – Personalized search
  – Ads

• User data in browser
  • Browsing history
  • User interest profile
  • Distill

• Control information release

Top: Computers: Security
Top: Arts: Movies
Top: Sports: Hockey
Top: Science: Math
Top: Recreation: Outdoors

Distill

Browsing history

Your browser

User interest profile
bn.com would like to learn your top interests. We will let them know you are interested in:

- Science
- Technology
- Outdoors

Accept  Decline
RePriv Protocol

GET /index.html HTTP 1.1
Host: www.example.com
Accept: repriv ...

HTTP/1.1 300 Multiple Choices
index.html
index.html?top-n&level=m

POST /index.html HTTP 1.1
Host: www.example.com
Content-Length: x
category1=c1&...

HTTP/1.1 200 OK

Personalized page content
Would you like to install an extension called “Bing Personalizer” that will:

- Watch mouse clicks on bing.com
- Modify appearance of bing.com
- Store personal data in browser
Contributions of RePriv

RePriv
• An in-browser framework for collecting & managing personal data to facilitate personalization.

Core Behavior Mining
• Efficient in-browser behavior mining & controlled dissemination of personal data.

RePriv miners
• A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation
• Evaluation of above mechanisms on real browsing histories & two in-depth case studies.
RePriv Architecture

Browser equipped with RePriv

- Core mining
- Miners
- Personal store
- RePriv APIs

User consent and policies

3rd party providers

1st party providers
Core Mining

- Taxonomy from first two levels of ODP taxonomy
  - ~450 categories total
  - 20 top-level categories
  - Overlap exists

- Naïve Bayes
  - All categories equally likely
  - Training: min(3000, # pages) sites per category
  - Attribute words occur in at least 15% of docs for ≥1 category

- Classification is fast enough: $O(c \cdot n)$
  - $n$ is # words in document
  - $c$ is # document categories
Global Mining Convergence

Interest profiles are fast to build
RePriv

- An in-browser framework for collecting & managing personal data to facilitate personalization.

Core Behavior Mining

- Efficient in-browser behavior mining & controlled dissemination of personal data.

RePriv miners

- A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation

- Evaluation of above mechanisms on real browsing histories & two in-depth case studies.
Verifying Miners

- Untrusted miners are written in Fine
- API wrappers for RePriv functionality written in Fine
- Refined types on security to reflect critical arguments to reflect policy needs
- All Miners state policy at top of source code
- Won't compile unless code follows policy

<table>
<thead>
<tr>
<th>Miner Name</th>
<th>C# LoC</th>
<th>Fine LoC</th>
<th>Verif. Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>TwitterMiner</td>
<td>89</td>
<td>36</td>
<td>6.4</td>
</tr>
<tr>
<td>BingMiner</td>
<td>78</td>
<td>35</td>
<td>6.8</td>
</tr>
<tr>
<td>NetflixMiner</td>
<td>112</td>
<td>110</td>
<td>7.7</td>
</tr>
<tr>
<td>GlueMiner</td>
<td>213</td>
<td>101</td>
<td>9.5</td>
</tr>
</tbody>
</table>
assume ExtensionId "twitterminer"
assume CanCommunicateXHR "twitter.com" Nil
assume CanUpdateStore("twitter.com" "twitterminer")
Netflix Example

- Update interest profile based on Netflix.com interactions
  - Watches clicks on rating links, updates store
  - Reads store to find recently viewed movies by genre
- Can provide this information on request to
  - fandango.com
  - amazon.com
  - metacritic.com

14 lines of Fine code

```plaintext
let doGetMovies genre cdom =
  ...
  let flixEnts = GetStoreEntriesByTopic myprov "movie" in
  let genreFlix = bind myprov flixEnts (filterByGenre genre) in
  ExtensionReturn cdom myprov genreFlix
```

```plaintext
assume CanReadDOMClass "netflix.com" "rv5"
assume CanCaptureEvents "onclick" (P "netflix.com" "netflixminer")
assume CanServeInformation "fandango.com" (P "netflix.com" "netflixminer")
assume CanServeInformation "amazon.com" (P "netflix.com" "netflixminer")
assume CanServeInformation "metacritic.com" (P "netflix.com" "netflixminer")
assume CanHandleSites "netflix.com"
assume CanReadStore (P "netflix.com" "netflixminer")
assume CanReadLocalFile "moviegenres.txt"
```
RePriv

- An in-browser framework for collecting & managing personal data to facilitate personalization.

Core Behavior Mining

- Efficient in-browser behavior mining & controlled dissemination of personal data.

RePriv miners

- A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation

- Evaluation of above mechanisms on real browsing histories & two in-depth case studies.
Privacy-Aware News Personalization

Map RePriv interest taxonomy to del.icio.us topics

Query personal store for top interests

Ask del.icio.us API for “hot” stories in appropriate topic areas from nytimes.com

Replace nytimes.com front page with del.icio.us stories
Privacy Policy

Change “href” attribute of anchor elements on nytimes.com

Change TextContent of selected anchor and div elements on nytimes.com
Do Video Games Equal Less Crime?

That's one theory for the continuing fall in crime, despite the recession.

Gamers Finally Get Their Wheaties Box ...sort of

Dr Pepper is featuring the Halo 3 player Tom Taylor, who goes by Tsquared, on the labels, which will appear on about 175 million 20-ounce bottles from January to April.

Sony To Shut its SF Metreon PlayStation store

Sony is closing down its one-and-only U.S. PlayStation store at the Metreon mall in San Francisco. The recession is clearly to blame, but it's happening at time when Microsoft - which opened and shut its own Microsoft store at the Metreon - is going to open a chain of its own stores.

Microsoft Takes on Cable With Xbox Streaming Video

If talks with Disney work out, the game console could stream ESPN content, making it that much easier to watch TV without cable.

Some Video Gamers Leery of Obama's Views

Gamers are worried that the president-elect's positions on video games may signal new regulations or restrictions on the industry.
News Personalization: Effectiveness

Most responses rated highly!

Most responses rated poorly
RePriv Summary

- Existing solutions require privacy sacrifice

- RePriv is a browser-based solution
  - User retains control of personal information
  - High-quality information mined from browser use
  - General-purpose mining useful & performant
  - Flexibility with rigorous guarantees of privacy

- Personalized content & privacy can coexist

- See our Oakland papers and W2SP papers
Some of the current problems in online privacy

- Tracking mechanisms
  - Cookies
  - Beacons
  - Fingerprinting

- Dangers of third-party tracking

- Ad ecosystem and user targeting

- Solutions for tracking prevention

- RePriv: combining personalization and privacy