Data breaches and identity theft

Lorrie Faith Cranor
November 12, 2013

8-533 / 8-733 / 19-608 / 95-818:
Privacy Policy, Law, and Technology
Writing a research paper
Organizing a research paper

• Decide up front what the point of your paper is and stay focused as you write

• Once you have decided on the main point, pick a title

• Start with an outline

• Use multiple levels of headings (usually 2 or 3)

• Don’t ramble!
Typical paper organization

- Abstract - Short summary of paper
- Introduction - Motivation (why this work is interesting/important, not your personal motivation)
- Background and related work - Sometimes part of introduction, sometimes two sections
- Methods - What you did; in a systems paper you may have system design and evaluation sections instead
- Results - What you found out
- Discussion/Conclusions - May include conclusions, future work, discussion of implications, etc.
- References
- Appendix - Stuff not essential to understanding the paper, but useful, especially to those trying to reproduce your results - data tables, proofs, survey forms, etc.
Road map

• Papers longer than a few pages should have a “road map” so readers know where you are going

• Road map usually comes at the end of the introduction

• Tell them what you are going to say, then say it, (and then tell them what you said)

• Examples
  – In the next section I introduce X and discuss related work. In Section 3 I describe my research methodology. In Section 4 I present results. In Section 5 I present conclusions and possible directions for future work.
  – Waldman et al, 2001: “This article presents an architecture for robust Web publishing systems. We describe nine design goals for such systems, review several existing systems, and take an in-depth look at Publius, a system that meets these design goals.”
Use topic sentences

• (Almost) every paragraph should have a topic sentence
  – Usually the first sentence
  – Sometimes the last sentence
  – Topic sentence gives the main point of the paragraph

• First paragraph of each section and subsection should give the main point of that section

• Examples from Waldman et al, 2001
  – In this section we attempt to abstract the particular implementation details and describe the underlying components and architecture of a censorship-resistant system.
  – Anonymous publications have been used to help bring about change throughout history.
Avoid unsubstantiated claims

• Provide evidence for every claim you make
  – Related work
  – Results of your own experiments

• Conclusions should not come as a surprise
  – Analysis of related work, experimental results, etc. should support your conclusions
  – Conclusions should summarize, highlight, show relationships, raise questions for future work
  – Don’t introduce completely new ideas in discussion or conclusion section (other than ideas for future work)
  – Don’t reach conclusions not supported by the rest of your paper
Creating a research poster
December 4 Poster Fair

• During class in GHC 6115

• 32x40 inch foam core boards, 9x12 inch construction paper, glue sticks, and thumb tacks will be made available
  – You can get them from Tiffany Todd
    ttodd@cs.cmu.edu in Wean 4114

• Present your preliminary project results and get feedback you can use as you finish your paper
Creating a research poster

• Any word processor, drawing, or page design software will work
  – PowerPoint is well-suited for making posters

• Design poster as single panel or modular units
  – Single panel posters
    • Have a professional look (if well designed)
    • Should be printed on large format printers (SCS has one for student use, requires SCS account)
    • Other large printers on campus or local copy shops – some can also print on fabric
  – Modular units
    • Easier to design and transport
    • Print on letter paper (optionally, mounted on construction paper)
Research poster content

• Don’t try to present your whole paper
  – Convey the big picture
  – Don’t expect people to spend more than 3-5 minutes reading your poster
  – 500 words, maximum (can be a lot shorter!)
• Introduce problem, your approach, and results
• Provide necessary background or glossary
• A picture is worth 1000 words
  – Graphs, diagrams, etc.
• Use bullets and sentence fragments, similar to making slides
• Don’t forget to include title and author
Research poster design

• Use a modular design
• Each section of your poster can go in a box
• Use a large, easy-to-read font
  – Most text should be at least 20 point font
  – No text less than 14 point font
  – Headings should be larger and in bold
• Use color consistently
• Arrange elements for a sensible visual flow
Presenting your research poster

- Be prepared to give a 1-minute overview of your poster and answer questions
- Let people read your poster without interrupting them
- Consider bringing a laptop if you have software to demo or a video to show
- Consider making handouts available with abstract, web URL for obtaining your paper, and your contact information
Privacy Enhancing Technologies, and Policy
Weisi Dai, Carnegie Mellon University

Do you know using PETs may bring problems?
PETs are computer applications to help the user better control their information, including:

I KNOW THESE PETs

Cryptography as an arm?
Yes. Import/export controls and domestic controls exist.
In the U.S., strong cryptography needs a license to be exported.
Examples: 2 versions of Microsoft Internet Explorer 5, the PGP scanning project.
In China, using any item that implemented encryption without prior approval from the authorities is subject to fines and being sued.
Example: China restricts the use of TPM.

Categories of PETs
- Communication Anonymizers
- Cryptography
- Search Engines
- Digital Currencies

Browsing anonymously using a VPN at Amsterdam?
Maybe not a good idea, due to data retention laws.
In the European Union, VPN providers and privacy-friendly search engines including StartPage.com are required to log requests.
Startpage Web Search launched Startpage Australia.

Still want to use PETs to protect your privacy?
Of course, we are on the way.

Contact me: weisi@cmu.edu. Cite as:
Understanding Data Practices that Influence User Sharing Preferences for Online Behavioral Advertising

Ashwini Rao (arao@cmu.edu)

Background
- Free Internet services supported by online advertising
  E.g. search, social networking

Online Behavioral Advertising (OBA)
- Ads shown based on user behavior
  -670% increase in ad success rate

User Privacy and OBA
- Users find OBA “creepy and scary”, “embarrassing”
- Advertiser data practices vary

Purpose and Retention
- 1 Year vs. 6 Month vs. 1 Year

Collection and Sharing
- Collect Ads
- Use Ads

Access to User Profile
- Review only
- Review and edit

Survey
1. Collect your information on all sites and other websites you visit
2. Use the collected information only on facebook.com
3. Use the collected information for targeted ads and other purposes
4. If asked and use collected information for a maximum period of one year

Question
Would you allow Facebook to collect the following information...

Targeted ads only
- Vs.

References
1. J. Yen et al. How much can behavioral targeting help online advertising? WWW’20
2. B. S. et al. Smart, useful, scary, creepy perceptions of online behavioral advertising. SOUPS 2012
3. Against it et al. Does personalization matter? Measuring user’s perception for online tracking and advertising. SOSP 2013
4. F. G. E. et al. What matters to users? Measuring factors that affect users’ willingness to share data with an organization. SOUPS 2013
5. Amazon Mechanical Turk www.mturk.com
Do Teens Have a Right to Privacy? Parents’ and Teens’ Perspectives

Adam Durity, Abigail Marsh, Blase Ur

Motivation
- Legally, teens have few rights to privacy from their parents
- FERPA protects education records, but mandates sharing with parents/guardians
- COPPA protects children under age 13 from online third-party tracking
- No omnibus protections beyond age 12
- **Hypothesis:** Families believe teens have a de facto right to privacy from their parents
- Teens and parents have differing expectations of the boundaries
- Boundaries expand with age
- What do parents feel they have a right to know? Not to know? What is acceptable and ethical in their view?
- What do teens feel parents have a right to know? Does this differ from parents’ opinions?

Methodology
- Semi-structured interviews with teens in high school and parents of teens in high school
- 2 participants (Eventually 20 participants)
- Recruited participants from Pittsburgh, PA using Craigslist and flyers
- Selected only one participant per family
- In participant’s eyes, to what extent do teens have a right to privacy from their parents?

Area of Inquiry Examples

<table>
<thead>
<tr>
<th>Area of Inquiry</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy at home</td>
<td>Closing bedroom doors, areas that are off-limits, knocking</td>
</tr>
<tr>
<td>Social privacy</td>
<td>Knowing their friends, always knowing where they are</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Reading texts, monitoring computer, parental controls</td>
</tr>
</tbody>
</table>

Preliminary Results

<table>
<thead>
<tr>
<th>Theme</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respect for teen ➔ Privacy</td>
<td>P0, P1</td>
</tr>
<tr>
<td>A parent’s concerns override a teen’s right to privacy</td>
<td>P0, P1</td>
</tr>
<tr>
<td>Privacy as parent-teen negotiation</td>
<td>P0, P1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theme</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teens’ bedrooms are generally private</td>
<td>P0: “If they are actually in there and don’t want me in there … I respect their wishes.”</td>
</tr>
<tr>
<td></td>
<td>P1: “It’s his private [area], it’s his domain.”</td>
</tr>
<tr>
<td></td>
<td>However, P1 examines son’s room when he is not at home “just to make sure he’s not doing nothing he shouldn’t be doing.”</td>
</tr>
<tr>
<td>Some privacy attitudes varied</td>
<td>P0 tried to use parental controls, whereas P1 never tried to monitor technology usage</td>
</tr>
<tr>
<td></td>
<td>P1 knows most of son’s friends, whereas P0 knows only a handful</td>
</tr>
<tr>
<td>Teens’ right to privacy is not absolute</td>
<td>P0: “[they] have a right to privacy to some extent … but not overriding a parent’s need to know something.”</td>
</tr>
<tr>
<td></td>
<td>P1: “It’s my house and I’m gonna go in that room whenever I want to.”</td>
</tr>
<tr>
<td>Responsibility for teens’ actions vs. privacy</td>
<td>P0’s nephew was arrested for downloading child porn on grandparent’s computer and nearly liable</td>
</tr>
<tr>
<td></td>
<td>P1: “Hell, there could be a mad man living in the room, how would I know? I could see Dr Phil, ‘Well, you never went in your son’s room, huh?’”</td>
</tr>
<tr>
<td>Teen years are a privacy transition</td>
<td>P0: “By the time you’re done with it you have a right to privacy, when you start it you don’t.”</td>
</tr>
</tbody>
</table>

Carnegie Mellon University
CyLab
Get Me off Your Wearable Cameras

Yuan Tian
yuantian@cs.cmu.edu

Background
- Goal of the system: Usable notification for the video session
- Refine the privacy violation by the wearable cameras
- Easy and efficient opt-out notification scheme
- Techniques related:
  - Privacy concerns against wearable cameras
  - Information encoding in audio
  - Indoor localization

Motivation
- Wearable cameras are pervasive
- No usable notifications to individuals about the video session
- Individuals cannot opt-out conveniently
- When combined with social network and face recognition scheme, the privacy violation is even worse

Methods
- System design of privacy notification of wearable cameras
- Implementation overview
- Encoding information in audio
- Extracting the magnitude of the recorded video to get relative distances

Result
- Choice of transfer channel: why audio?
- Encoding and decoding information from audio: 1500-1800 Hz works best
- Extracting distance from the magnitude of collected video

Conclusion & Future Work
- Improve the accuracy of distance of devices so as to analyze the position of people with the device
- Evaluate the usable privacy of the notification
- Combine with social network service & provide meta data to opt-out individuals

Acknowledgments
We thank Professor Lorrie Cranor for her guidance on the project, and our peers: Nanya Shieh, Zheng Sun and Yasmin Kandissouin for their help with the project.

Carnegie Mellon University
Goals
- Facilitating Usable Privacy Policy Project (uebieprivacy.org) affiliated by:

- Identifying key policy features from Retail and News Entertainment sectors
- Extracting different types of information collected and their sharing targets for each sector

Towards Information Extraction
From Natural Language Privacy Policies In Retail & News Sectors

Aditya Mahela
Dilek Tukel O¨vrel¨ek
Poster Fair - December 6, 2013

Methodology
- Identify key features in each sector
- Build questionnaire to reflect key features
- Determine what each privacy policy says about each feature
- Collect terms used for information types, categories & sources, usage types, sharing targets
- Identify any patterns or anomalies in the privacy policies

News Entertainment

14 News Websites:
- 4 from top ten broadcast media
- 3 political websites
- 2 business websites
- 4 personal finance websites

Key Features
- News Entertainment
  - Services other than offering news?
  - Share behavioral data with other third parties?
  - Collection and usage of social media data incite the user to connect to the website using social media services

Online Retail
- Collection & sharing of sensitive information (credit card, credit history)
- Restrictions on sharing target’s privacy policies
- Use of SSL while transferring sensitive information
- Opt-out choices wrt advertising and promotional emails

Results: News Entertainment
- News websites not limited to “news”, 100% of the samples sell product and services, offer interactive services...
- If registered, all of them collect contact information
- 72.8% collect current location of a user
- 92% use cookies, beacons or other tracking technologies
- 78% use (OBA) to deliver targeted advertising

Results: Online Retail
- Contact Information
  - all of them collect contact information and addresses
  - 70% share for purposes other than provisioning core services
- Financial Information
  - all of them collect credit card information and
  - 20% collect credit history information
- SSL: 50% protect personal information; 30% protect only sensitive information; 20% do not mention SSL

Retail Sector

15 Retail Websites
- 4 popular online stores
- 3 not so popular stores
- 2 each: Health foods & Kid stores
- 2 each: Electronic & Home goods

Questionnaire
- 22 Questions for News Entertainment Sector
- 18 Questions for Retail Sector
- The questions are designed to be answered as:
  a) Yes
  b) No
  c) Not clear from the policy
  d) Policy does not answer the question

Results: collection of terms
- Personal Information: name, address, phone, email, age, etc., credit card information, social security number, personal description, photograph, location, device identifier, purchase information, redemption information, etc.
- Behavioral Information: purchase history, products viewed, products searched, session information, page response times, download errors, viewing duration, clicks, signs, mouse-over, page view information, search terms, etc.
- Technical Information: IP computer, browser, version, time zone, plug-ins, plug-in versions, OS platform, etc.
- Full spreadsheet is available on request
Assessment of Web Browser Privacy Features

Methodology
- Latest versions of browsers tested on Win 7
- Check lists prepared with the help of previous work from CQT (2009) and new privacy features
- Some list items were only checked for existence
- Some other list items checked for functionality
  - Fire Examine, read, write etc.
  - Some popular websites used for behavior analysis
  - Cookie tracking (Facebook, ...)

Browsers
- Firefox
- Chromium
- Opera

Five Main Areas of Privacy Features

Sample List Elements

Capability Scoring
- 6: privacy feature works perfectly, best among browsers
- 5: privacy feature works well, but not best by functionality
- 4: privacy feature works well, lacks functionality
- 3: privacy feature works poorly, lacks functionality
- 2: feature not reachible via traditional interface/reached via advanced mechanisms
- 0: feature non-existent, does not work

Overall Scores

Results
- No single champion
- For general privacy options: FF got the highest score
- Browser privacy modes all work similarly well
- Safari is the last in all areas (possibly because of Win version)
- Chrome is good at granular controls like site-by-site options, cookie controls
- Chrome is good at controlling plugins & extensions storage, functionality
I. Introduction
As privacy violations by social media services become more prevalent, it is essential to address the ethical and strategic implications of these violations. To what extent should they improve privacy controls? Social media providers must consider business ethics to determine how to protect their users and the privacy of data they generate.

Existing corporate social responsibility (CSR) research does not address the increasingly controversial issues of social media services' privacy practices.

This project seeks to evaluate whether social networking providers are bound by CSR to protect user privacy beyond the minimum legal requirements.

II. Overview
Goal: To develop a Privacy Decision-Making Framework for social media services to evaluate CSR obligations.

Method: Evaluate privacy-related CSR through 2 classic normative philosophical approaches (teleology, deontology) to CSR and a case study to understand the philosophies in practice.

<table>
<thead>
<tr>
<th>Teleological Approach</th>
<th>Deontological Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overwhelmed by choices</td>
<td>Duty to Shareholders</td>
</tr>
<tr>
<td>Focus on outcome</td>
<td>Duty to Shareholders</td>
</tr>
<tr>
<td>Greatest good for the greatest number</td>
<td>Duty to Shareholders</td>
</tr>
<tr>
<td>Consequentiion</td>
<td>Natural law theory, Natural law theory</td>
</tr>
<tr>
<td>Difficulty to forecast outcome</td>
<td>Contractualism, divinence theory</td>
</tr>
<tr>
<td>Often avoided in modern science except evolutionary biology</td>
<td>Flexibility and terms of duty are difficult to define</td>
</tr>
<tr>
<td>John Stuart Mill, Jeremy Bentham</td>
<td>Immanuel Kant, Thomas Nagel, Thomas Scanlon</td>
</tr>
</tbody>
</table>

III. Philosophical Approaches to CSR

Is Facebook Morally Obligated to Protect Privacy?

A teleological and deontological analysis of corporate social responsibility and privacy protection

Mia Wang
MSSFM
mwang@ecu.edu
45-41011

IV. Case Study - Facebook

A 2011 Wall Street Journal article titled "What the Mark Perez Secretly Got Right About Facebook" documented how Facebook's privacy practices were designed to promote user engagement. Facebook's (FB) corporate philosophy "notified" that they were sharing data with third parties.

The two students faced receiving notifications that their emails became accessible by Open: "Open Access" as a Facebook group after the news report.

Misalignment and public notification did not receive consent from the students.

V. Privacy Decision-Making Framework

Company S is considering whether or not to implement additional user privacy protections. Consider:

Teleological Approach

<table>
<thead>
<tr>
<th>Impact</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>(1, 4, +)</td>
</tr>
<tr>
<td>Shareholders</td>
<td>(1, 0, +)</td>
</tr>
</tbody>
</table>

Deontological Approach

<table>
<thead>
<tr>
<th>Impact</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>(1, 0, +)</td>
</tr>
<tr>
<td>Shareholders</td>
<td>(1, 0, +)</td>
</tr>
</tbody>
</table>

Total: 11

If Total 15, Company S should consider implementing additional user privacy protections.

VI. Conclusion

There are not definitive standards and perceptions of corporate social responsibility; as implementation will depend on an organization's leadership. A privacy-enhancing improvement is considered a social media service has the responsibility to reconcile strategic and ethical benefits, which may not be mutually exclusive: to implementing the privacy controls. The volume and scale at which social media services are collecting data, privacy decisions impact billions of users on a global level.

The framework developed in the project serves to help a service organize their priorities and responsibilities to assess whether or not they are ethically bound to ensuring a certain privacy control.

References

[1] Facebook's Privacy Policy
[3] The Ethics of Social Media

Cite the references if necessary.
Anonymous Dislike:
Users' Reaction to Anonymous Peer Reviews in Social Networks

Pranshu Kaivani
pok@andrew.cmu.edu

Chao Pan
chaop@andrew.cmu.edu

Introduction
This study explores the effects of anonymity on users' behavior and also tries to find out their response to anonymous comments. Its primary objective is to provide feedback so people can realize the error of their ways and thus make them more conscious with future posts.

Methodology
We have performed two surveys on Amazon's MTurk. Based on these surveys we have created an anonymous commenting system. It is a Google Chrome add-on for Facebook and we are currently conducting a user study to test the efficacy of our system.

Results (contd.)

<table>
<thead>
<tr>
<th>Replies to Inappropriate Posts</th>
<th>Will Reply</th>
<th>Will Not Reply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identified</td>
<td>63.33%</td>
<td>34.61%</td>
</tr>
<tr>
<td>Anonymous</td>
<td>73%</td>
<td></td>
</tr>
</tbody>
</table>

These were the important results from the first survey. However, we wanted to see if after receiving these comments any steps to improve privacy were taken.

Individual's Identification of Inappropriate Posts

<table>
<thead>
<tr>
<th>Prefer Help</th>
<th>Don't Want Help</th>
<th>Own Realization</th>
<th>Other's Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.31%</td>
<td>71.39%</td>
<td>14.30%</td>
<td></td>
</tr>
</tbody>
</table>

Results
Presence of Inappropriate Posts

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Partially Agree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.60%</td>
<td>50%</td>
<td>15.42%</td>
</tr>
</tbody>
</table>

System Architecture

Challenges & Future Work
- Due to Facebook changing its source code and DOM regularly our exclusion stops functioning. This leads to difficulties in conducting a user study.
- Getting large groups that provide meaningful data is one of the other sizeable challenges we face.
- We intend on adding more features to our system to make it more informative. A report dashboard is at the top of this list.
- Deleting the anonymous post content via peer review or natural language processing is another feature we intend on adding.
- There is still a lot more work possible in this area.
Web Application for Searching and Comparing Financial Companies' Privacy Practices

Gabriel Moreno
gabriel@cc.nmu.edu

Overview

- Comparing the privacy policies of financial institutions is a time-consuming task for consumers.
  - No centralized place to find the policies
- This web application allows users to:
  - Look at policies
  - Search for institutions with specific privacy practices and other criteria
  - Compare privacy practices of multiple institutions side-by-side

Motivation

- The Federal Trade Commission (FTC) envisioned that privacy notices would enable competition in a market where privacy practices would be part of the consumer's decision.
- Consumers are expected to compare and shop on privacy policies to protect their privacy.
- Doing this comparison is too much burden on consumers
  - It is time-consuming task

Limitations of Existing Tools

- Compare things other than privacy policies
  - Consumer products
    - Examples: pricegrabber.com, shopper.com
  - For banks, offered services, financial strength indicators, user reviews
    - Example: tradereview.com
  - Insurance policies (health, auto, homeowner's)
    - Example: insuranceinsight.com
  - Focus on the online practices of organizations
    - Example: privacygrade.com

Standard Privacy Notice for Financial Institutions

- Most financial institutions use the model privacy notice to comply with the requirements of federal regulators.
- Standardized privacy notices are easier to compare, but still involve a manual process for the consumer
  - Find the privacy notice
  - Compare them

Current Burden on Consumers

- Consumers must first obtain privacy notices from the different financial institutions and then compare them.
- What if a consumer wants to find a financial institution with specific privacy practices?
  - The consumer must then obtain all the privacy notices
  - Go one by one to select those that satisfy the specific criteria

Use Cases Supported by this Web Application

- Search for and view the privacy practices of a financial institution
  - No need to request it or find where it is on the web
- Compare two or more selected institutions side-by-side
- Search for financial institutions whose privacy policies match some specified characteristic
  - For example, institutions in Pennsylvania that do not share personal information for marketing purposes
Demonstration of financial companies’ privacy practices

Privacy Bank

Jie Chen, Ziwei Hu, Zhipeng Tian
5000 Forbes Ave, Pittsburgh, PA

Carnegie Mellon

WHY?
Have you ever wondered how your banks deal with your personal info?

Do they sell your personal info?
Do they share your personal info? with whom? for what?
OR do they keep your info secure and protected?

Traditional privacy policies have been difficult to read and understand. It also takes lots of time to read.

More importantly, it does not allow users compare privacy practices across different financial organization.

WHAT?
The Project is to design a website that allows users to search, compare, and review financial companies privacy policy.

Our Focus is not only on the main features, but also on communication and presentation.

HOW?
PrivacyBank=
User-friendly interface
+ Comprehensive and detailed data

MAIN FEATURES:
Search
Our database includes 729 financial companies info across the United States.

User Study
Interviewed with 10 people and ask them what they care about the most in the search results.

a) Does my bank share?
b) Can I opt out?
c) How to opt out?
d) Number of affiliates

Rate, Review, and Share
You can rate and review a bank after you search it. You can also share the search results on Facebook.

Are your banks selling your info?

Client
Web Browser
HTML CSS JS

Server
Apache+ PHP
MySQL Database

System Architecture Graph

Scan and Check out our website!
Revisiting Private E-mail
A review of anonymous remailers and similar technologies
Michael Kahn

Background
- Much of the discussion today surrounding online privacy concerns the collection of surfing habits and personal information.
- Many users have more specialized privacy needs. Among these is being able to send e-mail anonymously—specifically, to send an e-mail such that neither a recipient nor an attacker can identify the sender.
- Such technologies are important for whistleblowers and others who wish to protect their identities if their actions are revealed.
- The problem of anonymous e-mail has solutions that do not function effectively in the general case with all uses.
- Even members of the public who are sufficiently well informed about privacy may not be aware of ways to effectively send anonymous e-mail.

Goals
- Design criteria to test different types of e-mail anonymizing methods.
- Examine a variety of existing solutions and choose a test group.
- Evaluate solutions, and determine which options would be best for different user groups.
- Suggest improvements for existing tools.

Anonymity Services

<table>
<thead>
<tr>
<th>Mail-specific services</th>
<th>General anonymity services</th>
</tr>
</thead>
<tbody>
<tr>
<td>approx-plain (Type 0)</td>
<td>Tor</td>
</tr>
<tr>
<td>RabbitMail (Type 1)</td>
<td>Uses onion routing behind IP traffic</td>
</tr>
<tr>
<td>Cypherpunk (Type 2)</td>
<td>IonCerium (uses Anonymous Proxy)</td>
</tr>
<tr>
<td>Mixmaster (Type 3)</td>
<td>Uses mix networks composed of known relays</td>
</tr>
<tr>
<td>Mixmaster (Type 4)</td>
<td>I2P</td>
</tr>
<tr>
<td>Mixmaster (Type 5)</td>
<td>Anonymity layer available to other applications</td>
</tr>
</tbody>
</table>

Evaluation Criteria
- Effectiveness: How effective is the system at protecting senders from identification by third parties?
  - Is the system backed by a mathematical model for what the probabilities of identifying the sender?
  - How secure is the theory's implementation? Can an attacker use indirect attacks to identify the sender?
  - Does the system allow governments to impose usage?

Accessibility
- How accessible is the system?
- What system-wide technical requirements does it have for an end-user?
- What technical skill set is required to use the system?
- What kind of support is available for users?

Openness
- Is the system able to be examined by experts? Is it open-source?

Restrictiveness
- How restrictive is the system?
- Does it create a delay before the message can be received? If so, how long?
- Does it restrict the kind of traffic?
- Can attachments or other rich content be sent?

Results

<table>
<thead>
<tr>
<th>Remailers</th>
<th>General solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProtonMail</td>
<td>Tor</td>
</tr>
<tr>
<td>I2P</td>
<td>Mixmaster (Type 5)</td>
</tr>
<tr>
<td>Mixmaster (Type 3)</td>
<td>Mixmaster (Type 4)</td>
</tr>
<tr>
<td>Mixmaster (Type 4)</td>
<td>Mixmaster (Type 5)</td>
</tr>
<tr>
<td>Mixmaster (Type 5)</td>
<td>Mixmaster (Type 6)</td>
</tr>
</tbody>
</table>

Conclusion
- General anonymity services are consistently easier to use than currently available remailers. These remailers are typically used and maintained by a small group of researchers or enthusiasts, whereas Tor is widely used and has a large base of developers and support.
- For the majority of users, a general solution such as Tor or I2P provides ease of use and broader applications than are offered even by advanced remailer implementations like Mixmaster. However, these programs are more vulnerable to timing attacks than the higher latency remailers, and I2P offers greater anonymity.
- Anonymous remailers still have a place in the privacy landscape, but in order to remain feasible, they must provide the accessibility users expect from other, more popular services.
Motivation
- First P3P Search Engine by Byers, et al.
- Dual purpose engine:
  - Provide P3P-enabled search
  - Facilitate P3P deployment research
- Goals:
  - Improve scalability and usability
  - Add research functionality

Enhanced P3P-Enabled Search Engine
Damon Smith
15508 Privacy Policy, Law, and Technology

Policy Retrieval
- Problem: performance bottleneck
- Solution: parallel policy requests

Presentation
- Custom APPEL rule sets in addition to predefined low, medium, high rule sets
- Save rule set choice in cookie
- Privacy Bird Icons
  - Green: policy passes rule set
  - Yellow: no policy
  - Red: policy fails rule set

Cache Query
- P3P policy cache as a research tool
- Use special queries:
  - p3p:<website> display cached policy
  - p3psstat: stats about policies in cache
    - Total policy count
    - Percent deployment
    - Multiple policies/site count
  - appel<ruleset> test policies against rule set
    - Percent matching policy

Architectue

Proactive Caching

Scalability
- Google API allows 1000 queries per day
- Let users input their own Google API key
Instant Messenger Privacy Concerns & Remedies
Ryan Mahon
rmahon@andrew.cmu.edu

Concerns
- Exposed Information
  - Conversations
  - Social Networks
  - Internet Presences
- Exposed To
  - IMSP
  - ISP
  - Snoopers

Private Conversations With Existing Architecture
- Chaum’s Mix Nets [1981]
- Onion Routing via other IM Clients
- Advantages: Interoperability, Privacy
  - IMSP, ISP, snoopers
    - Cannot tell what is being said
    - Cannot tell who is being spoken to
- Disadvantages: Latency, Centralization

Private Conversations With Peer-To-Peer Infrastructure
- Content-Addressable Networks: Overlay network by Ratnasamy et al. [2001]
- Crowds: Anonymity tool by Reiter and Rubin [1999]
- Advantages: Decentralized, Better Latency-Privacy Tradeoff
- Disadvantages: Interoperability, Misbehavior-Detection

Preventing Presence Exposure
- Focus: AIM, YIM, & WinMessenger
- Three Main Problems (all solvable):
  - Poor Default Privacy Settings
  - Lack of Granularity in Configurations
  - No Notice of Presence Viewing

Default Privacy Settings
- WinMessenger (Top Left)
- YIM (Bottom Left)
- AIM (Bottom Right)

Conclusions
- Future Work
  - Implementation of Architectures
  - Evaluation: Fault Tolerance, Latency
  - Examination of Legal and Ethical Issues
- Privacy in Current Popular Instant Message Systems is Poor, But Fixable!
Analyzing Software Architectures for Privacy

Jeff Barnes (jmbarne@cs.cmu.edu)

1. Background
Organizations use privacy policies for many reasons:
- To demonstrate their privacy commitment to consumers, regulators, and industry groups.
- To protect against litigation.
- To assess their own compliance with relevant laws.
- To engender trust.

But what happens when the privacy policy is wrong?
- 2000: Chase Manhattan Bank violated its own privacy policy by selling personal information about 18 million customers to marketers. Chase agreed to correct its privacy practices and pay the New York attorney general $101,900.
- 2002: Due to a software error, a subsidiary of Sony Music transmitted personal e-mail addresses to marketers in violation of its privacy policy. The company agreed to take measures including hiring an independent auditor and paid $75,000.
- 2004: The FTC fined Gateway Learning Corp. for renting the personal information of users of its flagship product, Hooked on Phonics.
- 2006: New York sued the Internet for selling personal information to a marketer in violation of its own privacy policy. A $1.1M settlement was reached.

Companies misunderstand their own privacy practices and consequently misrepresent themselves, underestimate their legal culpability, and damage their reputations.

Why is this a hard problem?
Part of the problem is human misunderstanding or ignorance of organization privacy policies. But another problem is the complexity of the software systems that manage and store personal information.

Even the developers of a software system may find it difficult to make statements about its privacy characteristics, because its complexity makes it difficult to infer how privacy-sensitive information travels through the system as a whole.

My approach is to use software architecture to confront this problem directly.

2. Software Architecture
Software architecture views software systems as comprising, at a high level, software components that communicate with each other through connectors.

Primary uses of software architecture include:
- Engineering a new system
- Reverse-engineering an existing system

Analysis techniques can be applied to both uses: analyzing the properties of proposed designs for a new system and analyzing the properties of an existing system.

Such properties include performance, security, etc. Privacy can be analyzed in this way too.

3. Conceptual Overview
Key idea: Rather than trying to determine the privacy properties of a software system holistically, evaluate the privacy behaviors of its constituent elements and model the flow of privacy-sensitive data through the system.

Why does this make sense?
Figuring out the privacy properties of an entire system is hard. But figuring out those of a small software component should be easy for the software engineers responsible for a project.

Then, we can apply our analysis to infer how privacy-sensitive information travels through the system from those of the constituent elements.

4. Theoretical Framework
Graph theory provides a mathematical model of our approach.

We can view a software architecture as a directed graph where the vertices are components and the edges are connectors.

Model the set of privacy-sensitive information as a set of labels, like contact for contact information.

Annotate each vertex with a set of labels indicating the privacy-sensitive information that enters the system at that component.

Annotate each edge with a set of labels indicating the privacy-sensitive information that may pass through that connector.

Finally, use these annotations to model how different types of data flow through the system.

5. Example
Consider a company that collects sensitive user information (name, contact information, Social Security Number) through a web interface. All of this information is stored in a secure database of user information. Individuals’ names and contact information are periodically extracted from this database and sent to a marketing database to be shared with marketing partners, in accordance with the privacy policy. SSNs are not supposed to be sent to the marketing database.

6. Implementation
An architecture description language (ADL) is used to describe software architectures in a clear and unambiguous way.

A typical ADL provides a way to describe components and connectors and how they are hooked up.

Allows elements of an architecture (components and connectors) to be annotated with user-defined properties such as performance attributes.

Allows definition of architectural styles—classes of software architectures. An architectural style is characterized by a vocabulary of architectural elements and a set of constraints on how they may be assembled.

Acme is an ADL developed at CMU. I picked it in part because of its GUI, AcmeStudio, which allows easy usage of Acme, provides graphical representation of architectures, and supports extensions for analysis.

I implemented my privacy analysis in Acme by:
1. Developing a style to accommodate the expression of privacy-relevant information.
2. Developing an external privacy analysis for systems of that style.

This is a screen shot of our example in AcmeStudio:

7. Future Work
- Basic improvements: better UI, more sophisticated definition of data types
- Sophisticated description and analysis of where and how data exit the system
- Model information that is anonymous or pseudonymous but privacy-sensitive.
- Check conformance between an implemented system and its described privacy characteristics.
- Check conformance with a privacy statement.
### Legal Background
- Fourth Amendment: Protects against unreasonable search and seizure. Requires probable cause and particular description of what will be searched.
-保護品: artery to unreasonable search and seizure. Requires probable cause and particular description of what will be searched.
- Courts limited applicability to cars. Carcass moved to new jurisdiction while law enforcement seeks warrant on car. 
- Racial profiles: plan sight, expectation to privacy limited.
- Expectations limited to privacy in car, therefore no reasonable expectation in glove compartment.

### Black Boxes
- What they are: Similar to concept as airplane black boxes.
- How do they do it: Record and store crash data; seat belt use, ejection deployment, speed at impact.
- Initial use: Car companies improved at bags.
- New uses: Understanding causes of accidents
- Court cases: determining fault, improving safety.
- Monitoring: real time feedback & logging. Insurance companies: time & speed.

### Traffic Cameras
- What they are: Multiple functions: red light, speed, hit patterns, automated speed, tolls.
- What do they do: Monitor violations will issue plate squeal cars.
- Initial uses: Deter unruly behavior.
- New Uses: Optimize post-accident traffic, enhance safety.
- Problem: increase accidents. Solution: timing.

### GPS transponders
- What they are: Determination of elevation and location through GPS signals.
- What do they do: Transponder sends signal to satellites, translates location. Provides MR, Carriage + owners with enhanced location.
- Initial use: Department of Defense. Detect missile launches.
- New Uses: May systems in cars.
- Problem: increase accidents. Solution: timing.

### Potential Privacy Invasions

<table>
<thead>
<tr>
<th>Privacy Invasions</th>
<th>Past</th>
<th>Present</th>
<th>Car</th>
<th>GPS</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved security</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Increased efficiency</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reduced crime</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Improved public safety</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

### Highway Use Tax Proposals
- What they are: Electronic toll collection.
- What do they do: E-ZPass data
- New uses: E-ZPass data without judicial oversight for PATRIOT re-authorization.
Digital Enhancement of the Female Figure: Harmful Fallacy of Perfection or Fair Marketing Tool?

Background

What happens when women are societal body ideals of perfection on television and in advertising? How do they feel about their bodies and how does it affect their confidence and self-esteem? This digital enhancement of women in the media is a prevalent issue that has been studied extensively in the field of psychology.

Issue

The problem arises when women are constantly exposed to these societal body ideals. This can lead to feelings of inadequacy and a desire to change one's body to meet these unrealistic standards. This can result in disordered eating, body dysmorphic disorder, and other mental health issues.

Experiments Concerning Women and Advertising

Digital Enhancement Programs

There are many programs that enhance photos of women. Some of these programs include Adobe Photoshop, FaceTune, and BodyMorph. These programs allow for the alteration of photographs, including the enhancement of facial features and body proportions.

Experiment 1:37

One study found that women who saw images of themselves that were digitally enhanced were more likely to feel self-conscious and dissatisfied with their bodies. This suggests that digital enhancement can have a negative impact on women's self-esteem.

Experiment 2:

Another study found that women who saw images of themselves that were digitally enhanced were more likely to compare themselves to these images and feel inferior. This suggests that digital enhancement can lead to feelings of inadequacy and a desire to change one's body.

Conclusion

It is difficult to determine whether digital enhancement of the female figure is beneficial or detrimental to women. Future research is needed to better understand the impact of digital enhancement on women's self-esteem and body image.

Sources

Cookie Taster: Using P3P Compact Policies to Protect Your Privacy on the Internet
Dave Gordon, Hanan Hibshi, and Pedro Leon

Background and Motivation
- Cookies are used extensively to collect, correlate, and share user information on the Internet.
- Compact Policies (CPs) are part of the Platform for Privacy Preferences (P3P) and can be used to communicate privacy policies with respect to cookies.
- CPs are three-character and four-character tokens transmitted in HTTP headers that help protect user privacy by expressing the following with regards to user information:
  - What information is being collected?
  - How is it used?
  - Who has access to it?
  - How long is it stored?
  - What information can the user access?
- Research has shown that CPs are being misused by websites and that the only user agent (embedded in the IE browser) that utilizes P3P CPs has been ineffective at using them properly to protect user privacy.

Our Goals
- Design an effective, non-intrusive, and easy-to-use user agent for Firefox that will:
  - Increase user awareness about how cookies are used and help them protect their privacy according to their expectations.
  - Encourage websites to comply with the P3P standard, avoid deceptive practices, and refrain from collecting excessive information from users.
  - Help to better understand privacy preferences through user feedback, and use that feedback to improve agent design.

Our Architecture

Configuring User Preferences

Showing Evaluation Results

Conclusions
- Our prototype can be extended for implementation in other popular web browsers.
- User agents can help to improve both user privacy awareness and protection.
- Accurate and effective user agents can encourage proper use of P3P compact policies.
- Design decisions are critical; they have implications on both usage and policy.
A Survey and Review of Privacy-Related Extensions for Mozilla Firefox

Aaron J. Couch
Carnegie Mellon University
Helen College
aaronjcouch@cmu.edu

Introduction
“Privacy software” is available to users to address the concerns and problems associated with the distribution of personal information online.

Fears of identity theft, the annoyance of unwanted marketing, and the general desire to be left alone are the greatest drivers of the market for privacy software.

This project is intended to survey and review extensions for Mozilla Firefox that offer privacy-related functionality.

Firefox extensions offer a means of altering the web-browsing experience to protect personal and private data. With some extensions, users can regain control over their online interactions and privacy.

NoScript
- 16 million downloads
- Provides an easy-to-use interface for blocking or allowing content to be displayed
- Provides a significant level of security for the user
- Requires manual configuration by the user

Add-on Plus
- 2 million downloads
- Facilitates easy installation and management of add-ons
- Provides a significant level of security for the user
- Requires manual configuration by the user

Conclusions
Empowering users with the ability to control their online privacy is crucial in a political and legal landscape which offers negligible safeguards or regulations for privacy-invading practices.

Extensions frequently serve as front-line defenses against new or previously undiscovered privacy threats, like session hijacking.

As web developers get trickier with obfuscating tracking mechanisms, extension developers have to stay on top of the game.

Increasing awareness of extension options is critical for all users to protect their privacy.

Recommendations:
- NoScript
- Add-on Plus
- Ghostery
- BetterPrivacy
- TrackMeNot

A note on proxy-enabling extensions
A variety of extensions are available for Firefox to enable anonymized web browsing, such as proxies. Proxy servers can act as intermediaries for Internet requests, effectively anonymizing users. While these extensions are not specifically explored here, users may want to investigate popular proxy extensions such as Torbutton, FxAProxy, AutoProxy, and Quickproxy.

Work referenced
- D. T. Wright, “How to use the Tor Bridge,” in The Tor Project, 2015.

For more info
Look at my draft paper.
Contact me at aaronjcouch@cmu.edu
### What is P3P anyway?
**The Platform for Privacy Preferences**
A machine-readable description of website privacy practices in XML format.

### Privacy Label Editor
**A simple, graphical web-based tool for creating P3P policies**

Yousifah Mahyshah
Taye Ogunyemi
Katheryn Sa

Privacy Policy, Law, Technology - Fall 2010

### New Approach: A Graphical Editing Tool
**Benefits:**
- Web-based
- Point to learn
- Point and click interface
- Easy update of existing P3P policies
- No need for use to be a P3P expert
- Easy to compare privacy policy with P3P policy to discover inconsistencies
- Easy to visualize differences between privacy policies

### Privacy Policies on the Web Today
- Long
- College age reading level
- Full of legalese

What does this mean?

### Sample Graphical P3P Policy

#### Traditional Tools
**Confusing?**

### Why are P3P and the Privacy Label Editor needed?
**P3P** - A standardized description of privacy policies

But what if the P3P policy for a website is created incorrectly?
- Capital One's P3P policy has numerous inconsistencies with its privacy policy

**Privacy Label Editor** - Visual representation for viewing and editing P3P policy

### Sample Graphical P3P Policy

#### Using the Privacy Label Editor
Data breaches and identity theft
Data breach

• Personal data lost or stolen
  – How?
• Data breach may lead to identity theft (but not always, and not for all people involved)
• Many states have notification statutes
• What can organizations do to prevent?
Identity theft

- Fraudulent acquisition and use of a person’s identifying information, usually for financial gain

- Range of offenses
  - Making purchases on someone else’s credit card
  - Opening credit in someone else’s name
  - Providing someone else’s identity to get a job
  - Providing someone else’s identity to avoid arrest, or to have someone else arrested

- How it happens
  - Physical theft, phishing, malware, computer security breaches, acquaintances, hospitals and nursing homes, ….
Data breach laws

• First enacted in CA in 2002 – SB 1386

• Most states in the US now have them
  – 47 states, DC, Guam, Puerto Rico, Virgin Islands
  – Alabama, New Mexico, and South Dakota do not

• Require notifying customers of PII data breaches

• Who must comply, definitions of PII, definitions of breach, types of notification, exemptions, etc. vary

Pennsylvania Statutes  
Title 73: Trade and Commerce  
Chapter 43: Breach of Personal Information Notification Act  
Effective: June 20, 2006

§ 2301. Short title.  
§ 2302. Definitions.  
§ 2303. Notification of Breach.  
§ 2304. Exceptions.  
§ 2305. Notification to Consumer Reporting Agencies.  
§ 2306. Preemption.  
§ 2307. Notice exemption.  
§ 2308. Civil relief.  
§ 2329. Applicability.

§ 2301. Short title. This act shall be known and may be cited as the Breach of Personal Information Notification Act.

§ 2302. Definitions. The following words and phrases when used in this act shall have the meanings given to them in this section unless the context clearly indicates otherwise:

"Breach of the security of the system." The unauthorized access and acquisition of computerized data that materially compromises the security or confidentiality of personal information maintained by the entity as part of a database of personal information regarding multiple individuals and that causes or the entity reasonably believes has caused or will cause loss or injury to any resident of this Commonwealth. Good faith acquisition of personal information by an employee or agent of the entity for the purposes of the entity is not a breach of the security of the system if the personal information is not used for a purpose other than the lawful purpose of the entity and is not subject to further unauthorized disclosure.

"Business." A sole proprietorship, partnership, corporation, association or other group, however organized and whether or not organized to operate at a profit, including a financial institution organized, chartered or holding a license or authorization certificate under the laws of this Commonwealth, any other state, the United States or any other country, or the parent or the subsidiary of a financial institution. The term includes an entity that destroys records.

"Encryption." The use of an algorithmic process to transform data into a form in which there is a low probability of assigning meaning without use of a confidential process or key.

"Entity." A State agency, a political subdivision of the Commonwealth or an individual or a business doing business in this Commonwealth.

"Individual." A natural person.
"Notice." May be provided by any of the following methods of notification:

(1) Written notice to the last known home address for the individual.

(2) Telephonic notice, if the customer can be reasonably expected to receive it and the notice is given in a clear and conspicuous manner, describes the incident in general terms and verifies personal information but does not require the customer to provide personal information and the customer is provided with a telephone number to call or Internet website to visit for further information or assistance.

(3) E-mail notice, if a prior business relationship exists and the person or entity has a valid e-mail address for the individual.

(4) (i) Substitute notice, if the entity demonstrates one of the following:
   (A) The cost of providing notice would exceed $100,000.
   (B) The affected class of subject persons to be notified exceeds 175,000.
   (C) The entity does not have sufficient contact information.

   (ii) Substitute notice shall consist of all of the following:
       (A) E-mail notice when the entity has an e-mail address for the subject persons.
       (B) Conspicuous posting of the notice on the entity's Internet website if the entity maintains one.
       (C) Notification to major Statewide media.

"Personal information."

(1) An individual's first name or first initial and last name in combination with and linked to any one or more of the following data elements when the data elements are not encrypted or redacted:
   (i) Social Security number.
   (ii) Driver's license number or a State identification card number issued in lieu of a driver's license.
   (iii) Financial account number, credit or debit card number, in combination with any required security code, access code or password that would permit access to an individual's financial account.

(2) The term does not include publicly available information that is lawfully made available to the general public from Federal, State or local government records.

"Records." Any material, regardless of the physical form, on which information is recorded or preserved by any means, including in written or spoken words, graphically depicted, printed or electromagnetically transmitted. The term does not include publicly available directories containing information an individual has voluntarily consented to have publicly disseminated or listed, such as name, address or telephone number.
"Redact." The term includes, but is not limited to, alteration or truncation such that no more than the last four digits of a Social Security number, driver's license number, State identification card number or account number is accessible as part of the data.

"State agency." Any agency, board, commission, authority or department of the Commonwealth and the General Assembly.

§ 2303. General rule.

(a) General rule.--An entity that maintains, stores or manages computerized data that includes personal information shall provide notice of any breach of the security of the system following discovery of the breach of the security of the system to any resident of this Commonwealth whose unencrypted and unredacted personal information was or is reasonably believed to have been accessed and acquired by an unauthorized person. Except as provided in section 4 [FN1] or in order to take any measures necessary to determine the scope of the breach and to restore the reasonable integrity of the data system, the notice shall be made without unreasonable delay. For the purpose of this section, a resident of this Commonwealth may be determined to be an individual whose principal mailing address, as reflected in the computerized data which is maintained, stored or managed by the entity, is in this Commonwealth.

(b) Encrypted information.--An entity must provide notice of the breach if encrypted information is accessed and acquired in an unencrypted form, if the security breach is linked to a breach of the security of the encryption or if the security breach involves a person with access to the encryption key.

(c) Vendor notification.--A vendor that maintains, stores or manages computerized data on behalf of another entity shall provide notice of any breach of the security system following discovery by the vendor to the entity on whose behalf the vendor maintains, stores or manages the data. The entity shall be responsible for making the determinations and discharging any remaining duties under this act.

§ 2304. Exceptions. The notification required by this act may be delayed if a law enforcement agency determines and advises the entity in writing specifically referencing this section that the notification will impede a criminal or civil investigation. The notification required by this act shall be made after the law enforcement agency determines that it will not compromise the investigation or national or homeland security.

§ 2305. Notification to Consumer Reporting Agencies. When an entity provides notification under this act to more than 1,000 persons at one time, the entity shall also notify, without unreasonable delay, all consumer reporting agencies that compile and maintain files on consumers on a nationwide basis, as defined in section 603 of the Fair Credit Reporting Act (Public Law 91-508, 15 U.S.C. § 1681a), of the timing, distribution and number of notices.