# The Need for Transparency in Privacy Statements & Practices

Annie I. Antón

North Carolina State University

Security and Liberty Forum
University of North Carolina
Chapel Hill, NC

April 14, 2007





### Context



- We've been studying privacy policy & healthcare privacy for 8 years
- Much of what we do in today's society involves information exchange and surveillance
  - e.g. government, commerce, healthcare
- □ For people to be comfortable and willing to take full advantage of IT services we need to:
  - Safeguard their privacy
  - Let them understand we're safeguarding their privacy and identity



### **Outline**



- Why do we care about transparency?
- Why do we care about privacy?
- Investigation framework
  - Analyzing privacy values vs. practices
  - Readability of privacy policies
  - User perception vs. comprehension of policies
  - To whom do we turn if we have questions?
- Recommendations





## Why do we care about transparency?

- Privacy policies are the standard to which organizations are held accountable.
- Transparency increases accountability.
- "Transparency is a clear, complete and readily available notice on an organization's privacy policies and practices." [www.nymity.com]
- "Experts agree that good privacy begins with effective transparency. Transparency requires privacy notices that are easy to understand, facilitate comparision, and are actionable. Privacy notices must also comply with legal requirements ...." [www.hunton.com]





## Medical Privacy in Particular



- Many medical conditions have some stigma associated with them
- Willingness to pay for treatment "out of pocket" or avoid life-saving treatment to keep information private
- Privacy concerns sometimes cause patients to withhold information that can prevent them from receiving the right kind of treatment
- In the U.S., HIPAA requires health care providers to secure medical records, etc.



### © Cartoonbank.com



"Your DNA doesn't match your credit history."

# Concerns About Healthcare Privacy in the U.S.





## InformationWeek BUSINESS INNOVATION POWERED BY TECHNOLOGY



NEWS

WINDOWS

SECURITY

OUTSOURCING

INTERNET

SOFTWARE

HARDWARE

MANAGEN

Security Tech Center: New! MyInformationWeek • Spyware • Security Blog • Viruses/Patches • Cybercrime • W • Dark Reading • Security Reviews • InformationWeek Download

## Government Report Finds Health Care Privacy Breaches Rampant

Agencies and contractors that experienced privacy breaches collectively have access to medical data for more than 100 million Americans.

» E-Mail

- » Print
- » Discuss
- » Del.icio.us
- » Digg

By Thomas Claburn InformationWeek

Sep 5, 2006 07:00 PM

Over 40% of federal health insurance contractors and state Medicaid agencies reported experiencing a privacy breach involving personal health information in the past two years, according to a Government Accounting Office (GAO) report released on Tuesday.



### Consider ....



- Data mining for anti-terrorism
- Airline passenger name record sharing w/ DHS
- ChoicePoint, Bank of America and other record thefts

All of these have been perceived as violations of privacy. And all have provoked public dismay and protest.



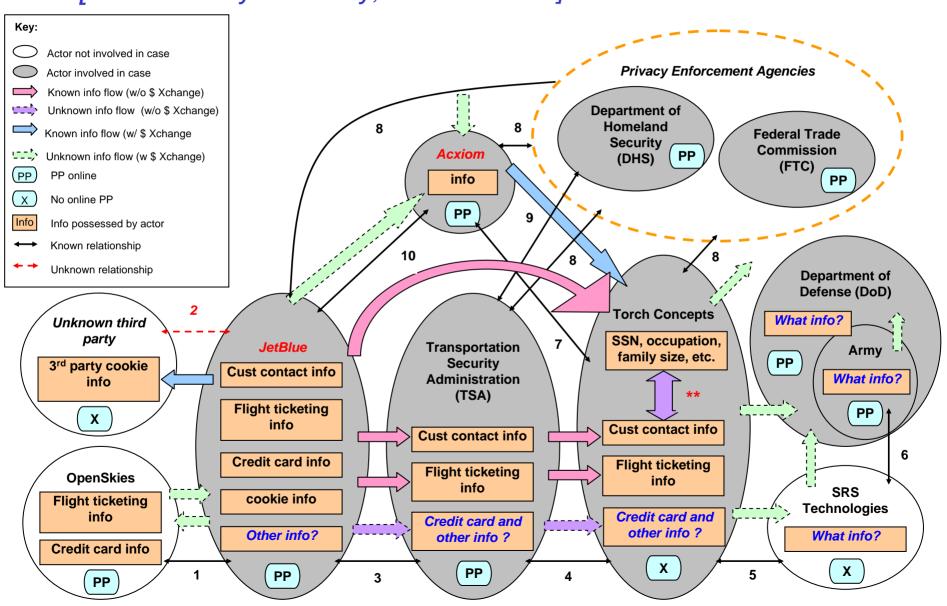
# An example of when things go wrong ...





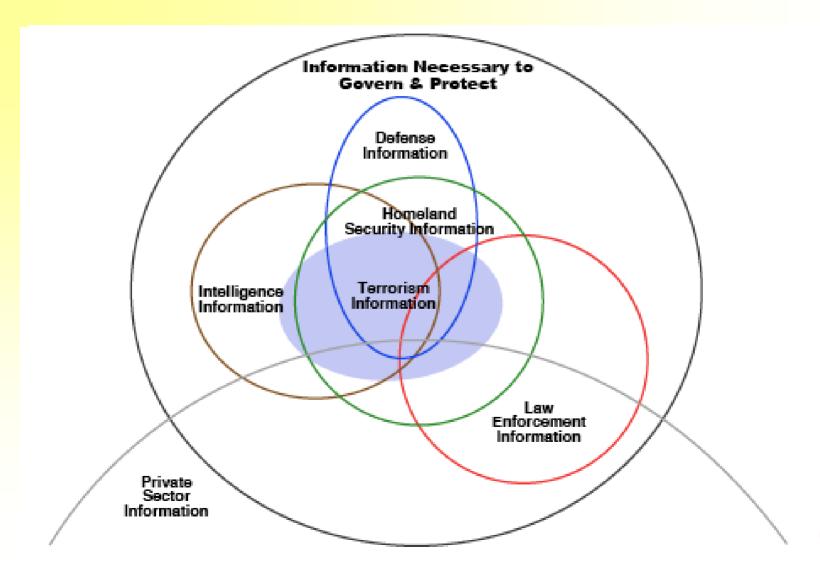
## jetBlue Analysis (Feb '04)

[IEEE Security & Privacy, Nov/Dec 2004]



### World of Information

[Rus05] Richard Russell. *Information Sharing and Collaboration Office Overview*, Department of Homeland Security Presentation, 2005.



## **Commercial Players**









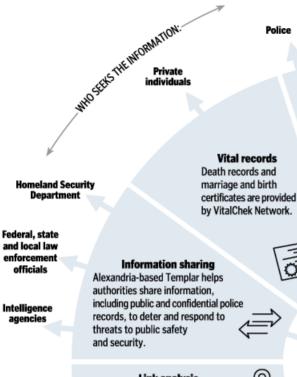






67 07





### Link analysis

12, based in Springfield and Britain, provides "visual investigative and link analysis" software that helps find patterns in databases in part to help authorities predict criminal activitiv.

U.S. intelligence analysts

**Police** 

Foreign law enforcement authorities



Intelligence analysts

**Public records information** Provides online access to more

than 17 billion public records

gathered from state and

local sources. Includes

addresses of neighbors

vehicle ownership.

and relatives, and

assets, names and

Reporters

### Corporations

### Parents hiring nannies

### Volunteer recruiters

Hiring executives

### http://www.washingtonpost.com/wpsrv/business/dailv/grap hics/choicepoint 012005.html1

Control

Information

### **Background checks**

Conducts checks on credit. employment history. education history, criminal records and military service. It also offers drug tests.



### Insurance claims

Gathers claim histories from automobile and home insurers to identify high-risk individuals and calibrate insurance premiums.





### Shoplifting

The Esteem system gets reports, including Social Security numbers, about employees convicted of shoplifting, those who have admitted shoplifting or those referred for prosecution.



Home

insurers

Auto

insurers

### Biometric service

Provides access to fingerprints of 46 million criminals.



ChoicePoint

The Springfield-based Bode Technology Group specializes in DNA identification analysis.



### **Criminal Locator and** Warrant System

Locates the most recent address of a wanted person.

**Police** 

Corporations

Private investigators



### 67 07

# Data Aggregators' Data Quality & Responsiveness

- Study examined the quality of data provided by ChoicePoint and Acxiom
- 100% of the reports given out by ChoicePoint had at least one error in them.
- □ Error rates for basic biographical data (including information people had to submit in order to receive their reports) fared almost as badly:
  - Acxiom had an error rate of 67%
  - ChoicePoint had an error rate of 73%.

The majority of participants had at least one significant error in their reported biographical data from each data broker.

http://www.privacyactivism.org/docs/DataAggregatorsStudy.html



## Transparency ...

□ Transparency increases accountability ...





# Online, how do you assure the public?







"Congratulations, Dave! I don't think I've read a more beautifully evasive and subtly misleading public statement in all my years in government."

## In particular, how ...



- do we assure that
  - policy complies with law?
  - system requirements comply with policy?
  - info. handling adheres to policy and system requirements?
- can policy be associated with data to ensure policies survive system boundaries?
- best do we support different social and legal definitions of privacy?





# Policy Content Analysis Using Antón & Earp Taxonomy ...

[Requirements Engineering Journal, 2004]

### Privacy Protection Goals

- Access/Participation
- Choice/Consent
- Enforcement/Redress
- Integrity/Security
- Notice/Awareness



### Privacy Vulnerabilities

- Aggregation of Information
- Collection of Information
- Monitoring of Information
- Personalization
- Solicitation
- Storage of Information
- Transfer of Information



# How does this relate to what users think?

What aspects of privacy are most important to on-line consumers?

Let's also ask:

Are consumers worried about the correct things?

Do institutions understand consumers' privacy concerns?





## **Privacy Survey Instrument**

### Privacy Values Baseline Reveals Misalignment

[IEEE Trans. On Eng. Mgmt, 2005]

- Data was collected from 1005 Internet users in 2002 to establish a privacy values baseline for correlation with our privacy protection goals and privacy vulnerabilities taxonomy.
- Consumers are most concerned with (in order):
  - information transfer
  - notice/awareness
  - information storage
- Privacy policies emphasize (in order)
  - data integrity/security
  - information collection, and
  - user choice/consent



Results Available: http://theprivacyplace.org/

# Are privacy statements readable?





## Readability Impact

[IEEE S&P, 2004 and 2006]



- Only 52.1% of general population has obtained at least 2 years of college
- Healthcare Industry
  - 24 online privacy documents at 9 healthcare websites
  - Flesch readability level = 14.2 (2 years of college)
- Financial Industry
  - 40 online privacy documents at 9 financial websites
  - Flesch readability range = 10.42 18.72
  - Flesch readability level = 14.1 (2 years of college)
- Difficult policies are time consuming to read



# Missing the intent of GLBA's clarity requirement



- The stark reality:
  - 8 require equivalent of a high school education or less
  - 13 require equivalent of some college education
  - 12 require 14-16 years of schooling
  - 1/6 require the equivalent of a postgraduate education (> 16 years).
  - 2/3 institutions had at least one policy document requiring the equivalent of a postgraduate education
- A full understanding of what 2/3 of these organizations are promising is only available to <15% of the adult Internet population





## Transparency ...

"Experts agree that good privacy begins with effective transparency. Transparency requires privacy notices that are easy to understand, facilitate comparision, and are actionable. Privacy notices must also comply with legal requirements ...."

[www.hunton.com]





# User perception vs. comprehension of policies...





## **2005 Privacy Policy Experiment**

[IEEE Trans. On Eng. Mgmt, 2007]

### Experiment

- Investigate user comprehension and perception of privacy policy expressions
- Theoretical framework based on the Privacy Requirements Taxonomy

### Objectives

- Gauge user perception of various alternatives to natural language privacy policies
- Measure user comprehension of the alternatives
- Compare user perception with user comprehension in order to determine whether they are in alignment with one another



## **Experimental Design: Blocking Factors**

- Drugstore.com
  - More vulnerable for consumers
- Healthcentral.com
  - Control
- Novartis.com
  - More protective for consumers





## **Experimental Design: Four variants (treatments)**

- Original natural language policy
- List of privacy goals and privacy vulnerabilities
- Categorical representation based on the taxonomy
- Original natural language policy supplemented with highlighted privacy goals and vulnerabilities





## Experimental Design (n=993)

	Original NL Policy	Goals & Vulnerabilities	Categorical	G/V & Policy
More Vulnerable	76	80	77	93
Control	86	87	81	83
More Protective	77	81	83	89



## Respondent Demographics

(n=993)

- 41.33% engage in online healthcare research at least once a month
- 31.7% have made a healthcare related purchase online
- □ 53.95% have beyond a college degree
- □ 29.74% female
- □ 66.87% male





## "I feel secure sharing my personal info with BrandX after viewing their privacy practices"

Policy Variant	Average Response
Goals/Vulnerabilities & Original Policy	2.85
Original Policy	2.72
Categorical	2.58
Goals/Vulnerabilities	2.44

Users perceive websites that express privacy policy with G/V highlights to be more secure.





## "I believe BrandX will protect my personal information more than other companies"

Variant	Average Response
Goals/Vulnerabilities & Original Policy	2.80
Original Policy	2.71
Categorical	2.60
Goals/Vulnerabilities	2.60
Average	2.68

Users perceive websites that express privacy policy with G/V highlights as more protective.





## "I believe BrandX will protect my personal information more than other companies"

Originating Website	Average Response
HealthCentral (control)	2.74
Novartis (more protective)	2.73
Drugstore.com (more vulnerable)	2.57

Users are able to distinguish between most vulnerable website & more protected ones.





## "I feel that BrandX's privacy practices are explained thoroughly in the policy I read"

Variant	Average Response
Original Policy	3.34
Goals/Vulnerabilities & Original Policy	3.33
Categorical	2.88
Goals/Vulnerabilities	2.77
Average	3.06

**Most explained:** 

GV & Policy variant Drugstore.com Least explained:

👉 Categorical variant ★

Novartis.com





# "I feel that BrandX's privacy practices are explained thoroughly in the policy I read"

Originating Website	Average Response		
Drugstore.com (more vulnerable)	3.25		
HealthCentral (control)	2.99		
Novartis (more protective)	2.96		

Users perceive length as reflecting thoroughness.





# "I feel confident in my understanding of what I read of BrandX's privacy policy"

Policy	Average Response
HealthCentral (control)	3.27
Novartis (more protective)	3.24
Drugstore.com (more vulnerable)	3.05
Average	3.19

User confidence in understanding raises interesting questions about perception of thoroughness.





### **Example Quiz Question**

### Which statement is TRUE regarding BrandX's information collection practices?

- BrandX collects Personal Health Information (PHI) in order to fill a prescription
- BrandX collects information about your browser type
- BrandX collects information from customer emails sent to their institution
- BrandX collects information from forms filled out on their site
- The policy does not address how information is collected



### **Average Comprehension Score**

All respondents (p<.0001)

Variant	Average
Categorical **	52.14 **
Goals/Vulnerabilities	43.82
Goals/Vulnerabilities & Original Policy	43.27
Original Policy	35.70
Average	43.74

User perception and comprehension are misaligned.





### **Average Comprehension Score**

(only respondents who read the entire policy)

Variant	Average
Categorical	65.67
Goals/Vulnerabilities	55.46
Goals/Vulnerabilities & Original Policy	49.38
Original Policy	40.00
Average	52.88

User perception and comprehension are misaligned.





# "I read the entire set of privacy policies of the website"

□ Categories 62%

□ Policy 56%

■ Goals/Vulnerabilities & Policy 50%

■ Goals/Vulnerabilities 44%





### **Comprehension Summary**

[IEEE Trans. On Eng. Mgmt, 2007]

- Users comprehend natural language policies the least
- Simply adding goals to natural language policies increased comprehension
- Significant increase in comprehension when goals are categorized
- Scores alarmingly low, despite which variant used





# What about demographics?



- No correlation between demographic factors and comprehension/perception exists.
- Exception
  - Respondents age 57 and higher scored <u>lower</u> on comprehension questions



### **Summary of User Perceptions**

[IEEE Trans. On Eng. Mgmt, 2007]

- Users believe:
  - they are more secure sharing PII with websites that display
     NL policies that highlight the goals and vulnerabilities
  - the companies that display NL policies that highlight goals and vulnerabilities will protect their information the most
  - the two NL variants are explained more thoroughly than alternative expressions
- User perception and comprehension are misaligned!!
  - Users feel most secure and protected by natural language policies, but comprehend them the worst





# What if we have questions about an organization's privacy policy?





# U.S. Airlines Study May 2006

[Otto & Antón -- Unpublished Sneak Preview!]

- AirTran Airways
- Alaska Airlines
- American Airlines
- ATA
- Continental Airlines
- Delta Airlines
- JetBlue Airways
- Northwest Airlines
- Southwest Airlines
- United Airlines
- US Airways

#### Objective:

 Contact each airline via the contact provided on their Privacy Policy to specifically ask questions regarding their privacy policy to determine each airline's preparedness to respond to such inquiries

#### Mechanisms:

- Telephone
- Web Forms
- Email



### **Phone Script**

Hello, my name is Annie Antón/Paul Otto. I've read your privacy policy and have a few questions about it, so I'm calling the number that you provided should we have additional questions ...

- With whom specifically does \_\_\_\_\_ share customer information?
  - Do you share it with:
    - DHS?
    - TSA?
    - Data Brokers like Axciom & ChoicePoint?
    - Direct Marketers?
- Who are your third parties?
  - What do you get in exchange for this sharing?
- What specific information do you share with these companies?



### **Airline Responses to Inquiry #1**

	original contact	Airline Reply to 1st Inquiry				
	date type	date	lapsed time	response quality	response author	notes
U.S.						
AirTran Airways	05/17/10 web form	5/17/10	same day (auto)	pending	auto	will reply w/in 72 business hrs; if request is urgent call cust serv: 800.airtran;
		5/17/10	same day	pending	Theda Moody (cust serv team)	your email has been forwarded to website development group; someone will respond promptly
Alaska Airlines	06/09/10 web form	6/9/10	same day (auto)	pending	auto	Please allow one business day for an agent to respond to your email.
American Airlines	05/17/10 web form	5/17/09	same day	pending	Mr. M. Exter (web services)	please view PP; then, if you require additional assistance, contact web services at: 800.222.2377
Continental Airlines	05/17/10 email	5/24/10	7 days	no answer	senior attorney: Abby Bried (priv. compliance contact)	please review privacy policy;
<del>Delta Air Lines</del>	05/17/10 web form	5/17/10	same day (auto)	pending	auto	auto reply acknowledged our email; asked for patience due to volume; but response may take longer
JetBlue Airways	05/17/10 email	5/26/10	9 days	no answer	cust. commitment crew member: Nancy	please review privacy policy; if you have questions please email or call; happy to respond, but more efficient if you first read the PP
Northwest Airlines	06/09/10 web form	6/9/10	same day (auto)	pending	auto	We respond to questions about online bookings within 3 hours, 7 days a week, 24 hours a day. All messages are personally read and handled to give you the attention you deserve.
		6/10/10	1 day	referral	Martez Franklin (cust serv)	I have researched your inquiry. In order to assist you, please call colleagues in Customer care at: 701.420.6282 (also provided mailing address and fax number). Thank you for writing, we appreciate your patience, I trust that this will assist you with y
Southwest Airlines	phone					
United Airlines	05/17/10 web form	5/17/10	same day (auto)	pending	auto	a member of our customer relations team will personally review your correspondence and respond as soon as possible; due to current volumes, response time is longer than normal
		5/19/10	2 days	referral	Peter Vohra (cust relations)	given your situation, you can call our cust. Relation team: 877.228.1327



### **Airline Responses to Inquiry #1**

	original	original contact		Airline Reply to 1st Inquiry			
	date	type	date	lapsed time	response	response author	notes
					quality		
U.S.							
AirTran Airways	05/17/10	web form	5/17/10	same day (auto)	pending	auto	will reply w/in 72 business hrs; if request is urgent call cust serv: 800.airtran;
			5/17/10	same day	pending	Theda Moody (cust serv team)	your email has been forwarded to website development group; someone will respond promptly
Alaska Airlines	06/09/10	web form	6/9/10	same day (auto)	pending	auto	Please allow one business day for an agent to respond to your email.
American Airlines	05/17/10	web form	5/17/09	same day	pending	Mr. M. Exter (web services)	please view PP; then, if you require additional assistance, contact web services at: 800.222.2377
ATA						,	
Continental Airlines	05/17/10	email	5/24/10	7 days	no answer	senior attorney: Abby Bried (priv. compliance contact)	please review privacy policy;
Delta Air Lines	05/17/10	web form	5/17/10	same day (auto)	pending	auto	auto reply acknowledged our email; asked for patience due to volume; but
JetBlue Airways	05/17/10	email	5/26/10	9 days	no answer	cust. commitment crew member: Nancy	please review privacy policy; if you have questions please email or call; happy to respond, but more efficient if you first read the PP
Northwest Airlines	<del></del>	web form	6/10/10	same day (auto)  1 day	referral	auto  Martez Franklin (cust	week, 24 hours a day. All messages are personally read and handled to give you the attention you deserve.  I have researched your inquiry. In order to assist you, please call
				,		serv)	colleagues in Customer care at: 701.420.6282 (also provided mailing address and fax number). Thank you for writing, we appreciate your patience, I trust that this will assist you with y
Southwest Airlines		phone					
United Airlines	05/17/10	web form	5/17/10	same day (auto)	pending	auto	a member of our customer relations team will personally review your correspondence and respond as soon as possible; due to current volumes, response time is longer than normal
			5/19/10	2 days	referral	Peter Vohra (cust	given your situation, you can call our cust. Relation team: 877.228.1327



# The Good, the Bad and the Ugly ...





# The Good: Alaska Airlines

- □ Call #1: 8 mins (on hold 3 mins)
  - very helpful, answered what she could, gave us number for corporate office to follow up on our unanswered questions
  - Answer: She explained that BoA provides their visa cards and is the only partner that receives PII;
  - Question: pertinent PII?
  - Answer: she was not sure, but told us what was associated with the Mileage Plan status (#, mailing info, #miles)
- □ Call #2: to corporate office
  - transferred to legal dept voicemail ... call to legal dept:
     directed to contracted employee who works with PP
  - Answer: Info used in reservations: name, travel itinerary; don't sell PII, and share info only in the case of customers registering for promotions/sweepstakes.
- Total Length of Call: 8 minutes w/ contracted employee on 6/8/2006



# The Bad: Continental 3rd Party Data Retention

- "Continental contracts with reputable third parties to facilitate the collection and aggregation of Web usage and transaction information; the distribution of advertisements and electronic mail; and the operation of some survey, promotional and sweepstakes programs. Such third parties are subject to confidentiality and any necessary privacy requirements and, upon completion of these services, all information is returned to Continental."
- □ Question: How do you verify that these third parties actually destroy all copies of that information?
- Answer:
  - Wasn't sure at first (on hold for 3 minutes).
  - He had no info on this
  - Did not have information on a customer service manager
  - Said there was no place to go for an answer
  - Was not aware of any compliance officer with whom we could speak



# Continental: Phone + Email Personal Data for Marketing

- "If you object to Continental maintaining and using your personal data for marketing purposes, OnePass members may opt out of such purpose by using continental.com (sign in is required)."
- Question: What specific personal information is used for marketing purposes?
- Answer:
  - would have to find out (put on hold for 7 minutes)
  - Belinda tells us that Larry Lettick (sp?) is not at his desk, but he is who we need to talk to
    - would not give us Larry's email or phone number, saying she wasn't sure if he wanted to be directly reached
    - offered to take our email and phone # to pass on to Larry Annie gave her @purdue email address and @purdue phone #\* call ended
- □ Total Length of Call: 10 minutes (7 minutes on hold) on 6/8/2006
- Follow up response:
- Email received (6/9/2006 @ 11:55pm from onepass@coair.com). Belinda Harris: "After speaking with the Corporate Security contact and our Marketing contact, I have been advised that we do not provide a specific list of all the data collected for marketing purposes. If you refer back to the PP on Continental.com, each section advises what information is collected. ..."



# The Ugly: United Airlines

- 3 total calls, plus 3 total transfers
- Call #1: (customer relations): on phone 9 mins (on hold 6 mins)
  - refused to give name & number, said to call reservations to ask these questions
- □ Call #2: on phone 3 mins (on hold 2)
  - told to call web support w/ questions, had no info on compliance officer or legal dept #s
- □ Call #3: on phone 25 mins (on hold 10)
  - person said she wasn't the right person to talk to for any of the questions (in India);
  - asked to speak to supervisor (put us on hold for 2 mins) came back & said "we don't share with gov't agencies or offices because you have a SSN"
  - can't give # for corporate office or legal dept
  - got very agitated at our questions and even yelled at us.



### **Transparency?**

"Experts agree that good privacy begins with effective transparency.
Transparency requires privacy notices that are easy to understand, facilitate comparision, and are actionable. Privacy notices must also comply with legal requirements ...." [www.hunton.com]



### In Summary ...





### Recommendations



- Construct meaningful, readable policies for your websites
- Consider alternative presentations of your policies
- Test your policies against focus groups
- Legalese is ambiguous and difficult to support
  - Compliance is easier to support with formalized methods
  - Analysis exposes ambiguities and choices to thwart potential abuses or privacy breaches
- Ensure people who may be contacted about your privacy policy are trained to respond in a meaningful way



# Something we might consider here in the U.S. ...





#### 2006 Top Privacy Policy in Canada

In 2006, the awards for transparency go to:

#### Banking

- TD Bank Financial Group
- Scotiabank
- CIBC
- Laurentian Bank
- RBC Royal Bank

#### Telcos

- TELUS
- Telebec
- Bell
- Rogers
- Aliant

#### Retail

- Chapters Indigo
- Holt Renfrew
- Sears Canada
- Costco Canada
- Roots Canada

#### Financial Services

(outside of banking)

- Vancity
- MCAP
- ATB Financial
- Financialinx
- GMAC Financial Services

#### Insurance

- Aviva
- First Canadian Title
- LAWPRO
- belairdirect
- Manulife Financial

#### Media

- CanWest Global Communications Corp
- Thomson Nelson
- Alliance Atlantis
- Transcontinental
- Chum Limited

#### **Consumer Services**

- Workopolis
- AIR MILES Reward Program
- Equifax
- TransUnion
- eBay Canada

#### **Business Services**

(outsourcing sector)

- Ceridian Canada
- Ipsos Canada
- Aon Canada
- ADP Canada
- Kelly Services

http://www.nymity.com/about\_us/ Top\_Privacy\_Policies.asp

Ranking is based on transparency, not on content. For example, an organization that stated several uses of personal information for secondary purposes would rank equal to an organization that stated they didn't use personal information for secondary purposes. The key is providing notice on this matter.



### Collaborators

- Faculty
  - Linda Aiman-Smith
  - David Baumer
  - Julie Earp
- Graduate Students
  - Davide Bolchini
  - Travis Breaux
  - Jack Frink
  - Qingfeng He
  - Neha Jain
  - Paul Otto
  - William Stufflebeam
  - Matt Vail
- Undergraduate Students
  - Carrie Gheen





### This work supported in part by ...

- National Science Foundation
  - Medium ITR Grant (2003-2007)
     Antón, Earp, Baumer, Aiman-Smith, Stufflebeam, Jain @ NCSU Colin Potts @ Georgia Tech
  - Digital Government Grant (2003-2005)
     Antón, He @ NCSU
     José Fortes @ University of Florida
  - Cyber Trust Grant (2004-2007)
     Antón, Yu, Baumer, Rappa @ NCSU
     Elisa Bertino & Ninghui Li @ Purdue University
  - ITR Grant (2004-2007)
     Antón, Earp @ NCSU
     Elisa Bertino @ Purdue University
- Computing Research Association DMP (`00,'01,'02,'04,'05)
- NCSU Digital Enterprise
- IBM in RTP
- Purdue University CERIAS



### **For More Information:**

http://www.csc.ncsu.edu/faculty/anton

and

http://theprivacyplace.org/





### Thank you!



