

*Privacy, Economics,
and Immediate Gratification:
Why Protecting Privacy is Easy,
But Selling It is Not*

Alessandro Acquisti

Heinz School, Carnegie Mellon University
PGuardian Technologies, Inc.

acquisti@andrew.cmu.edu

Why do we have great privacy enhancing technologies... that almost nobody uses?

Why do so many people claim to be concerned about privacy... and then do little to protect it?

It's the economy, stupid!

- Privacy is an economic problem...
- ... even when privacy issues do not have direct economic interpretation
- Privacy is about trade-offs: pros/cons of revealing/accessing personal information...
 - For individuals
 - For organizations
- ... and trade-offs are the realm of economics

Agenda

1. Privacy and the Economy / Economics and Privacy
2. So Many Open Questions, So Little Time
3. Privacy and Rationality: Alternative Explanations and Preliminary Survey Results

1. Privacy and the Economy

Economics and Privacy

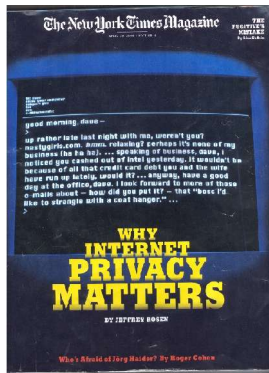
Privacy and the economy

“I am under no moral or other Obligation, to publish to the World, how much my Expenses and my Incomes amount to yearly. [...] Dissimulation is a branch of Wisdom.”

John Adams (1761)

- American census, 1799
- Warren and Brandeis, 1890
- “Franklin Mills Flour” girl, 1901 (Ellis Smith [2000])
- SSNs, 1935
- Retail Credit Co., TRW
- Equifax, Experian
- Amazon, Real Audio, eToys, 2000s

Privacy and the roaring dot com days



"It is surprising how recently changes in law and technology have been permitted to undermine sanctuaries of privacy that Americans have long taken for granted."

"Suddenly, shopping that had once seemed anonymous was being archived in personally identifiable dossiers"

NYT Magazine, April 30, 2000



"In the background, advertising services are building profiles of where people browse, what they buy, how they think, and who they are."

"For about 9 cents, some medical data sites will sell you your neighbor's history of urinary tract infections. It will get worse."

BusinessWeek, March 20, 2000

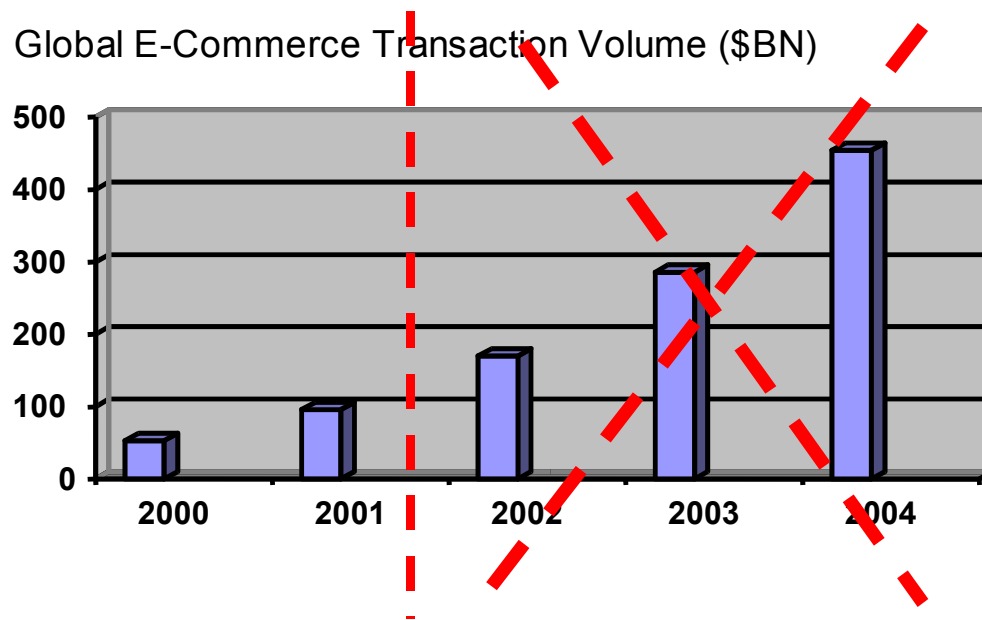


"Most commercial health-care Web sites lure consumers with free medical information, then sell data on them to third parties in ways that threaten the consumer's privacy."

Mark Smith, February 1, 2000

And then...

- *“The overall B2C market opportunity should reach \$450BN in transaction volume by 2004.”*
- Actually...



Sources: Forrester Research and IDC, circa 2001

How the *market* reacted

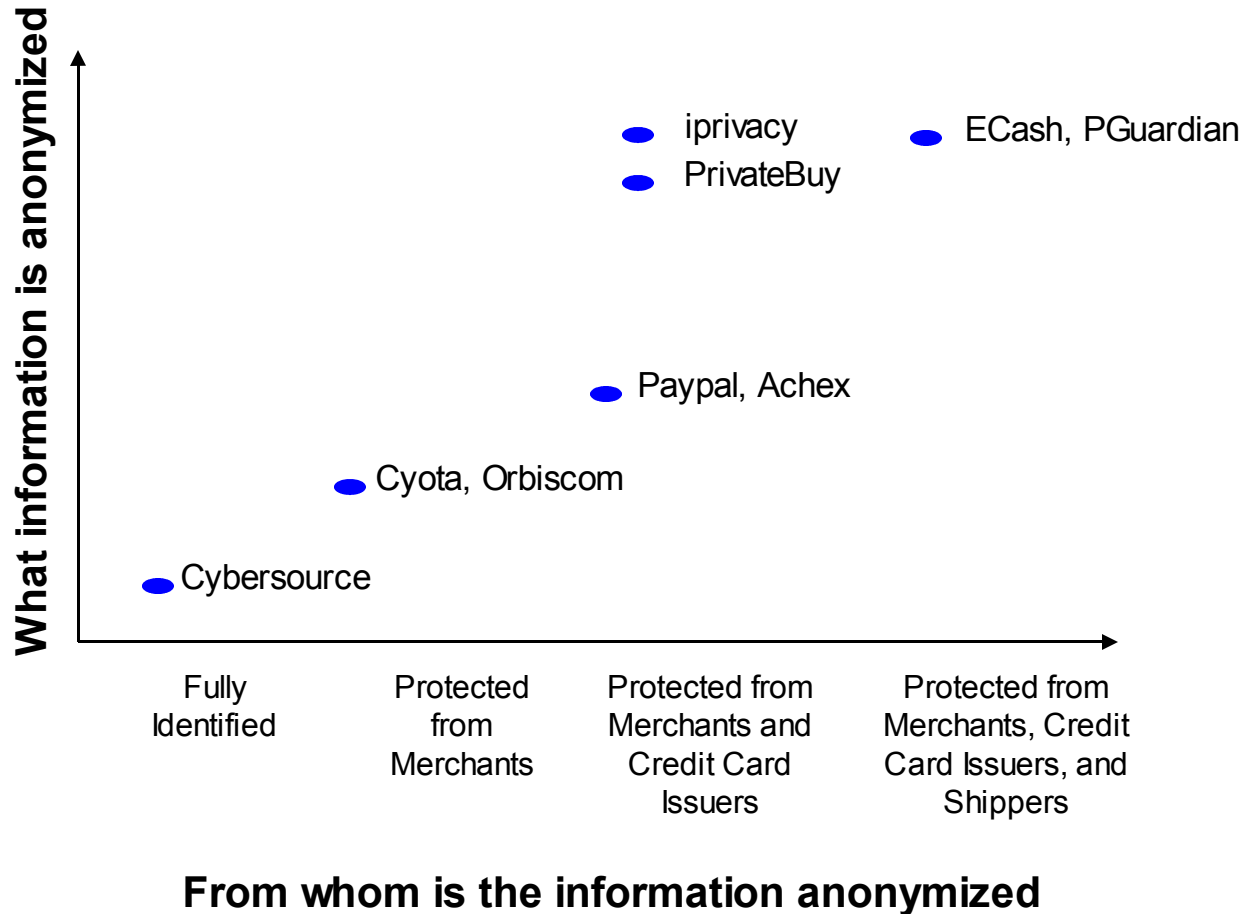
- Economic challenges pushed merchants towards more intrusive policies:

“This policy may change from time to time so please check back periodically.”

(Yahoo! Privacy Policy, circa 2001)



Technology: The case of the incredibly shrinking anonymous payments market



And the *Law*?

- National Zoo refuses to release a deceased giraffe's medical records on grounds that it would violate the animal's right to privacy (Politech, May 2002)
- *An Animal Privacy Entitlement Act?*
- No! A *Patriot Act* instead

What about economics?

- Posner (1980)
 - Privacy as concealment of information – *focus*
 - Privacy as quiet – *little economic relevance*
 - Privacy as freedom – *no economic relevance*

Economics and privacy

- Now:
 - Privacy as concealment of information - yes
 - Privacy as quiet - yes
 - Privacy as freedom – yes!
 - Even when privacy intrusions have no immediate economic relevance, immaterial dimensions of privacy still impact the well-being of the individual
 - Economics of happiness and well-being

The evolution of the *economics of privacy*

- Early 1980s
 - Chicago school vs. broader views of privacy
- Mid 1990s
 - The IT explosion: Varian, Noam, Laudon, Clarke
- After 2001
 - The Internet: personalization and dynamic behavior
 - Modeling: price discrimination, information and competition, costs of accessing customers
 - Empirical studies: surveys and experiments
 - Economics of (personal) information security (Anderson, Varian, WEIS)

The early days: Stigler

- Peculiar relation between “ownership” and privacy
 - Information about somebody may have been costly acquired by other people
- Free exchange of information will lead to desirable results regardless of ownership
 - If I am a good debtor, I want this information to be known; if I am a bad debtor, I want to keep it secret
 - Suppose I am a bad debtor: then, whether I do not reveal information or information about me is reported, I will pay higher rates

The early days: Posner

- Privacy as concealment of information
 - Individuals with bad traits (e.g., poor employees) have an interest in hiding them
 - Individuals with good traits have an interest in showing them
 - Reducing information available to “buyers” in this market (employers) reduces efficiency
- Extends argument to non-market behavior
 - E.g., marriage
- Costs of concealment borne by others
 - E.g., when privacy of sex-offenders is protected
- Privacy is re-distributive and reduces efficiency

The mid 1990s: Noam

- With no transaction costs in trading or negotiation, initial assignment of privacy rights is arbitrary from viewpoint of economic efficiency
 - Encryption
 - “The existence of encryption may largely determine who has to pay whom, not whether something will happen.”
 - Encryption makes other parties pay
 - Redistributes wealth to consumers
- Difficulties
 - Incomplete information
 - Human right
 - Burden on poor

The mid 1990s: Varian

- Consumers rationally want certain kinds of information to be available to producers, not other kinds
 - E.g., consumer wants seller to know what goods she likes, but not *how* much she likes them
- Annoyances comes from *too little* information being shared
 - E.g., tele-marketers offering products I do not want
- *Externalities* derive from secondary use of information
- Define property rights in private information in ways that allow consumers to retain control over how information about them is used
 - E.g., timed contracts
 - E.g., make it costly to access certain digital information

2001 and after: A new interest in the economics of privacy

- Calzolari and Pavan (2001)
- Taylor (2001)
- Acquisti and Varian (2001)
- ...

Optimal privacy policies

- Calzolari and Pavan (2001)
- Contracting environments where two “principals” (e.g., sellers) sequentially interact with a common “agent” (e.g., buyer)
 - First seller releases information that is correlated with agent’s type
- Welfare effects of privacy-protecting laws that prevent information disclosure on consumers’ shopping activity
 - Information transmission between two vendors may result in welfare increase
 - Reduces (expected) distortions

Customer privacy

- Taylor (2001)
- Value of customer information derives from ability of firms to identify individual consumers and charge them personalized prices
- Considers two settings: anonymity regime and recognition regime
- Welfare comparisons depend critically on whether consumers anticipate sale of the list
 - If consumers do not foresee sale of their data, firms have incentives to charge higher prices
 - If consumers anticipate sale of list, this results in lower prices than would prevail under the anonymity regime

Inducing customers to try new goods

- Acquisti and Varian (2001)
- Cookies-like technology vs. anonymizing technology
- Questions
 - Will cookies-like technology bring more profits?
 - Will buyers use the anonymizing technology?
- Results
 - No larger profits from cookies-like technology...
 - ... unless something more is offered
 - *Enhanced* services based on gathered information
 - Anonymizing technologies could make society worse off

Summarizing results

- Allowing firms to use cookies *can* make customers and society better off
- Sharing information between sellers reduces “distortions”
- With “strategic” customers, firms better off respecting customer’s privacy
- *So, where is the problem?*

Off-line vs. on-line identities

- *On-line identity*
 - Carries information about an individual's tastes, her purchase history, etc. (e.g., Amazon account)
- *Off-line identity*
 - The persistent identity of an individual, as revealed by identifiers such as credit card numbers and social security numbers
- *The problem: Linked on-line/off-line identities*
 - Different needs
 - Externalities
 - *Technology can separate them. Why is this not happening?*

2. So Many Open Questions, So Little Time

Open questions

1. Is too much privacy bad for you?
2. Do you really have zero privacy?
3. What are the costs of privacy?
4. Who should protect your privacy?
5. Do people really care about privacy?

Is too much privacy bad for you?

or, too much privacy can act against the interests of society or the individual

- Economics says:
 - More sharing of on-line identity information is good: market laws can allow the right amount of information to be shared
 - But, this is *not* in contradiction with protection of privacy (off-line identity)
 - *Problem: linkages and trails shrinks the anonymity set (Danezis and Serjantov [2002]). Then:*

Do you really have zero privacy?

“Get over it: You already have zero privacy.”

or, the loss of control on personal information is simply unavoidable in our networked society

- But information technology can also:
 - Either link *or* unlink on-line and off-line identities
 - Or make linkages difficult (e.g., Sweeney [2002])

Do you really have zero privacy?

- In almost every conceivable on-line and off-line scenario, we have developed tools and methods to adequately protect privacy
 - Anonymous payments (e.g., Chaum [1982])
 - Anonymous browsing (e.g., Goldschlag *et al* [1999])
 - Private preferences (e.g., Canny [2002])
 - Re-mailers (e.g., Chaum [1981])
 - (Good) electronic voting (e.g., Juels and Jakobsson [2002])
 - ...

And yet....

- Economic arguments show that trade-offs between *sharing* and *protecting* personal information may be reconciled
- Technology *could* do it
- So, why econ & technology did *not* do it?
- Solve the following equation:

Find a *privacy* combination *convenient* for customers (e.g. Bob), *profitable* for vendors (e.g. Amazon.com), *advantageous* for other existing players (e.g. credit card networks), *non replicable* by competitors

Who should protect your privacy?

- Self-regulation?
 - Fails under pressure
- Policy/legislation?
 - EU vs. US
 - Samuelson (2003): The social costs of confusing privacy policies
- Individual responsibility?
 - Can individuals protect themselves?
 - Should they?

Phrasing the policy debate?

*It is true that there are potential costs of using Gmail for email storage [...] The question is **whether consumers should have the right to make that choice and balance the tradeoffs**, or whether it will be preemptively denied to them by privacy fundamentalists out to deny consumers that choice.*

-- Declan McCullagh (2004)

- Can consumers really make the choice that best serves their own interests?

Privacy attitudes...

- Attitudes: usage
 - Top reason for not going online (Harris)
 - 78% would increase Internet usage given more privacy (Harris)
- Attitudes: shopping
 - \$18 billion in lost e-tail sales (Jupiter)
 - Reason for 61% of Internet users to avoid ECommerce (P&AB)
 - 73% would shop more online with guarantee for privacy (Harris)
- (most of the above is 2001 data...)

... versus privacy behavior

- Behavior
 - Anecdotic evidence
 - DNA for BigMac
 - Experiments
 - Spiekermann, Grossklags, and Berendt (2001): privacy “advocates” & cameras
 - Everyday examples
 - Dot com deathbed
 - Abundance of information sharing

Explanations

- Syverson (2003)
 - “Rational, after all” explanation
- Shostack (2003)
 - “When it matters” explanation
- Huberman (2004)
 - “Privacy and deviance” explanation
- *Are there other explanations?*

3. Privacy and Rationality: Alternative Explanations and Preliminary Survey Results

Personal information is a very peculiar economic good

- Asymmetric information
 - Individual does not know how, how often, for how long her information will be used
 - Intrusions invisible and ubiquitous
 - Externalities and moral hazard
- Ex-post
 - Value uncertainty
 - Keeps on affecting individual after transaction
 - Imagine: lump sum vs. negative annuity

Personal information is a very peculiar economic good

- Context-dependent (states of the world)
 - Anonymity sets
 - Recombinant growth
 - Sweeney (2002): 87% of Americans uniquely identifiable from ZIP code, birth date, and sex
- Subjective
 - “Willingness to pay” affected by considerations beyond traditional market reasoning

Personal information is a very peculiar economic good

- Both private and public good aspects
 - As information, it is non rival and non excludable
 - Yet the more other parties use that personal information, the higher the risks for original data owner
- Buy vs. sell
 - Individuals value differently *protection* and *sale* of same piece of information
 - Like insurance, but...

... maybe because...

- ... privacy issues actually originate from *two* different markets
 - Market for personal information
 - Market for privacy
- Related, but not identical
- Confusion leads to inconsistencies
 - Different rules, attitudes, considerations
 - Public vs. private
 - Selling vs. buying
 - Specific vs. generic
 - Value for other people vs. damage to oneself
 - Lump sum vs. negative annuity

Privacy and rationality

- Traditional economic view: forward looking agent, utility maximizer, bayesian updater, perfectly informed
 - Both in theoretical works on privacy
 - And in empirical studies

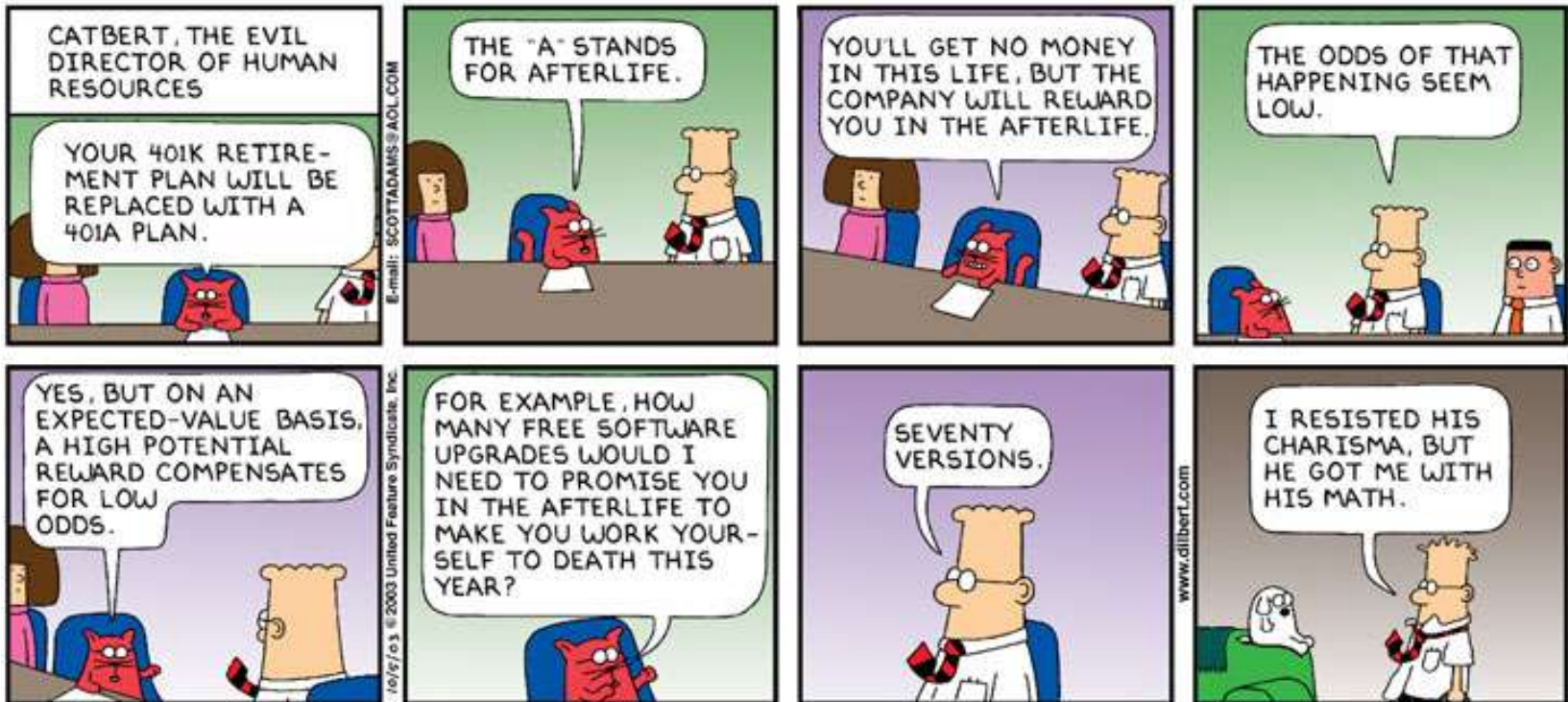
Yet: privacy trade-offs

- Protect:
 - Immediate costs or loss of immediate benefits
 - Future (uncertain) benefits
- Do not protect:
 - Immediate benefits
 - Future (uncertain) costs

Why is this problematic?

- Incomplete information
- Bounded rationality
- Psychological/behavioral distortions
 - Complacency towards large risks
 - Inability to deal with prolonged accumulation of small risks
 - Coherent arbitrariness
 - Hyperbolic discounting
- *Theory: Acquisti ACM EC 04*
- *Empirical approach: Acquisti and Grossklags WEIS 04*

Immediate gratification...



Hyperbolic discounting

$$U^t = \sum_{\tau=t}^T \delta^\tau u_\tau,$$

$$U_t(u_t, u_{t+1}, \dots, u_T) = \delta^t u_t + \beta \sum_{\tau=t+1}^T \delta^\tau u_\tau$$

Survey time vs. decision time

$$\min_{\text{wrt } x} DU_1 = \beta[(E(c_p^s)\delta^s x) + (E(c_i^{s+n})\delta^{s+n}(1-x))]$$

$$\min_{\text{wrt } x} DU_s = \delta E(c_p^s)x + \beta E(c_i^n)\delta^n(1-x)]$$

Time consistency vs. time inconsistency

	Period 1	Period 2	Period 3	Period 4
Benefits from selling period 1	2	0	0	0
Costs from selling period 1	0	1	1	1
Benefits from selling period 3	0	0	2	0
Costs from selling period 3	0	0	0	1

Sophisticated vs. naïve

	Period 1	Period 2	Period 3	Period 4
Costs	-5	-7	-9	-16

Consequences

- Rationality model not appropriate to describe individual privacy behavior
- Time inconsistencies lead to under protection and over release of personal information
- Genuinely privacy concerned individuals may end up not protecting their privacy
- Also sophisticated users will not protect themselves against risks
- Large risks accumulate through small steps
- Not knowing the risk is *not* the issue

Survey and experiment

- Phase One: pilot
- Phase Two: ~100 questions, 119 subjects from CMU list
- Paid, online survey (CMU Berkman Fund)
- Goals
 - Contrast three sets of data
 - *Privacy attitudes*
 - *Privacy behavior*
 - *Market characteristics and psychological distortions*
 - Test rationality assumption
 - Explain behavior and dichotomy
- Phase Three: experiment

Questions

1. Demographics and IT usage
2. Knowledge of privacy risks
3. Knowledge of protection
4. Attitudes towards privacy (generic)
5. Attitudes towards privacy (specific)
6. Risk neutrality/aversion (unframed)
7. Strategic/unstrategic behavior (unframed)
8. Hyperbolic discounting (unframed)
9. Buy and sell value for same piece of information
10. Behavior, past: “Sell” behavior (i.e., give away information)
11. Behavior, past: “Buy” behavior (i.e., protect information)

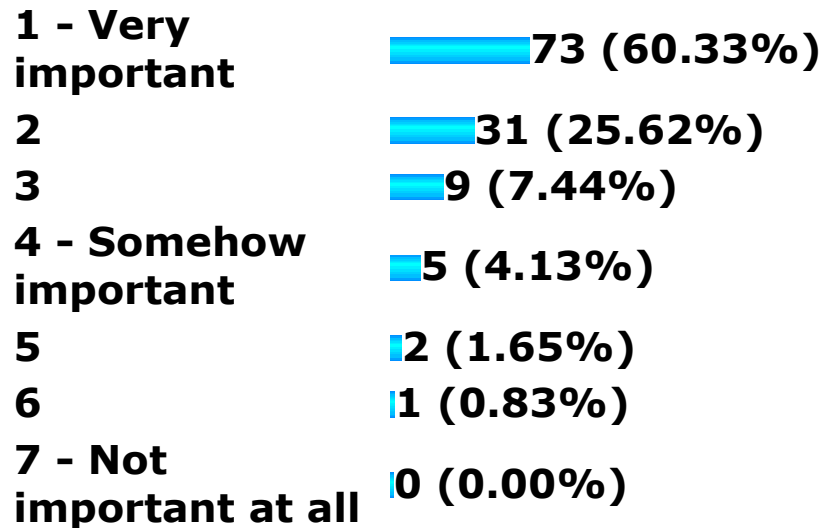
Demographics

- Age: 19-55 (average: 24)
- Education: mostly college educated
- Household income: from <15k to 120k+
- Nationalities: USA 83%
- Jobs: full-time students 41.32%, the rest in full time/part time jobs or unemployed

Privacy attitudes

(excerpts)



How important is privacy to you?




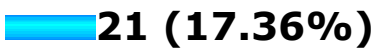
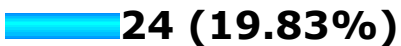
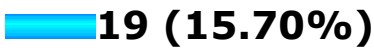
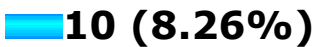


Privacy attitudes

(excerpts)

Do you think you have enough privacy in today's society?

Yes	
No	

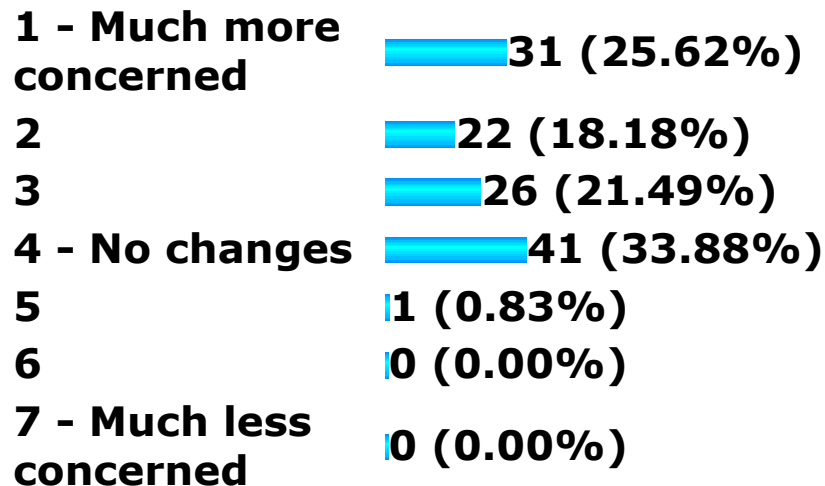
How concerned are you about threats to your personal privacy in today's information society?

1 - Very much	
2	
3	
4 - Somehow	
5	
6	
7 - Not at all	

Privacy attitudes

(excerpts)

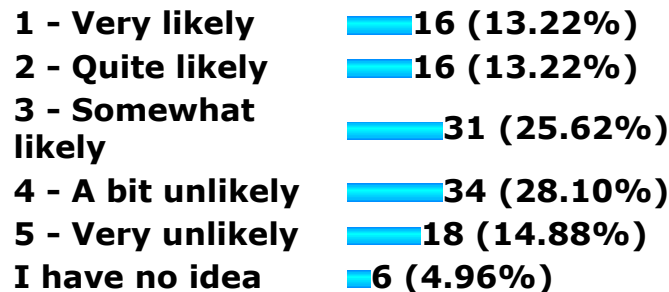
Has your concern about threats to your personal privacy changed in the last 24 months?



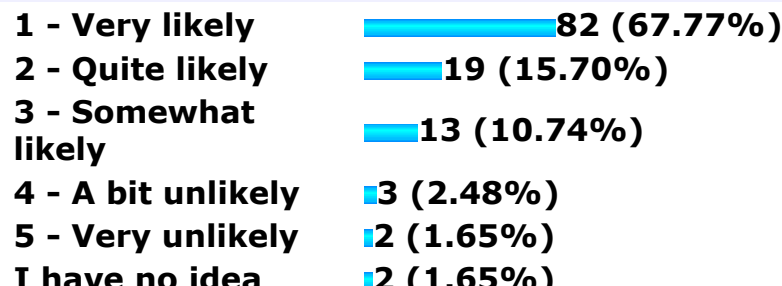
Knowledge of privacy risks (excerpts)

When you are releasing personal information during an ecommerce transaction, how likely do you consider the following outcomes?

Attempts to vary price during your next purchase based on your collected data



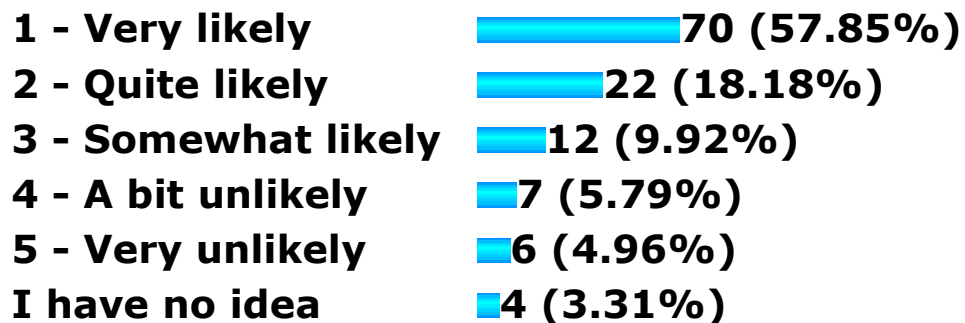
Use for marketing purposes (e.g., advertising emails)



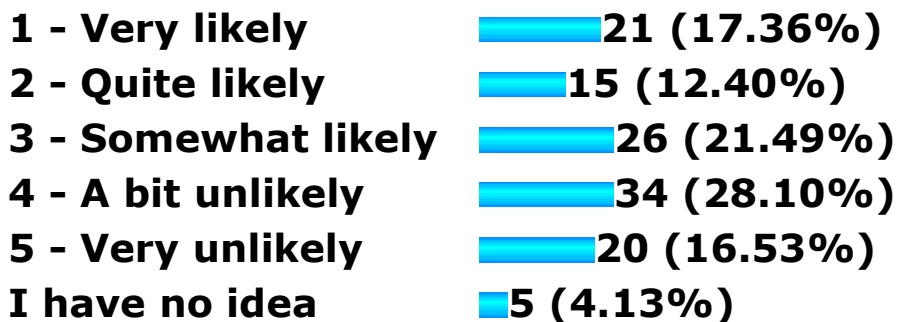
Knowledge of privacy risks (excerpts)

How likely do you consider the possibility that a 3rd party *can* monitor some details of the following activities you may engage in?

Using a file sharing client (e.g., Kazaa)



Writing a text memo to yourself on a computer connected to the Internet in your organization

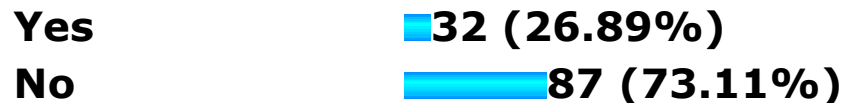


Knowledge of privacy risks (excerpts)

Do you know what Echelon is?



Do you know what Carnivore is?



Do you know what Total Information Awareness is?



Knowledge of privacy risks (excerpts)

You have completed a credit card purchase with an online merchant. Besides you and the merchant website, who has data about *parts* of your transaction?

Nobody: 36.4%

Credit card company: 18.7%

Hackers: 15%

"Nobody, assuming an SSL transaction, without which I would not commit an online transaction using my credit card"

Privacy knowledge and overconfidence (excerpts)

Can you estimate an interval for which you are 95% sure that it contains the number that correctly answers the following questions?

Example: Occurrences of identity theft in the US in 2003

Solution: lower bound 0.5 Million (complaints with FTC), less conservative estimates: 10 Million

Rational
31.9%

Overconfident
56.3%

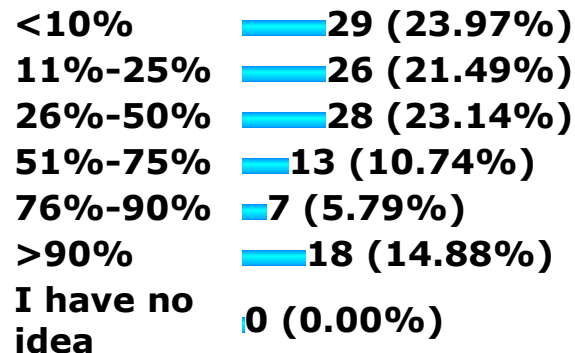
Missing
10.9%

Privacy risks and bundles (excerpts)

Privacy concern	<i>Data about offline identity</i>	<i>Bundled data about offline identity</i>
<i>Low concern</i>	27.30%	6.70%
<i>Medium concern</i>	26.70%	27.20%
<i>High concern</i>	20.00%	39.50%
<i>Missing data</i>	26.00%	26.60%

Privacy risks and bundles (excerpts)

Imagine that somebody does not know you but knows what your date of birth is, what your sex is, and the zip code where you live. What do you think is the probability that this person can uniquely identify you based on those data?



- Sweeney (2002): 87% of the population of the United States is likely to be uniquely identified by 5-digit ZIP code, birth date, and sex

Knowledge of privacy protection (excerpts)

- Privacy law:
 - *54% cannot quote a law or describe it*
- OECD Fair information principles:
 - *38% believe they include "litigation against wrongful behavior"*
- Goal: browse anonymously
 - *51% would not know how*
- Goal: browse the Internet with warnings if a website has an incompatible privacy policy
 - *67% would not know how (but most use IE6)*

Knowledge of privacy risks and attitude (excerpts)

Are you informed about the policy regarding monitoring activities of employees/students in your organization?

	Yes, I am informed	I don't know how such monitoring could take place	I somewhat know ... but don't know the details	There is a policy, but I don't know its details
Low concern	0.00%	0.00%	6.70%	0.00%
Medium concern	0.00%	13.40%	19.60%	6.70%
High concern	19.50%	0.00%	6.70%	20.00%

“Buy” behavior

- 74% adopted some *strategy* or *technology* or otherwise took some particular *action* to protect their privacy
 - Encryption, PGP
 - Do-not-call list
 - Interrupt purchase
 - Provide fake information
 - [...]
- However, when you look at details, percentages go down...
 - 8% encrypt emails regularly
 - Similar results for shredders, do-not-call lists, caller-IDs, etc.

“Sell” behavior

When interacting with any party except family and friends (e.g., a merchant or institution), have you ever given away the following pieces of information for a discount or bonus? Or did you receive a better service or recommendation for releasing this information?

YES (in decreasing order):

- Full name
- Email address
- Home address
- Phone number
- Job title
- Personal interests
- SSN
- Health history

Attitudes/behavior dichotomy (excerpts)

	Have a loyalty card, gave correct info	Have a card, gave fake info	Don't have a card
Low concern	0.00%	6.70%	6.70%
Medium concern	6.70%	13.30%	0.00%
High concern	26.70%	13.30%	0.00%


Recall of past behavior (excerpts)

	Have a loyalty card, gave correct info	Have a card, gave fake info	Don't have a card
Yes, have given identity data	20.00%	0.00%	0.00%
No, have not given identity data	26.70%	26.00%	13.30%

Password for chocolate?

- InfoSec Europe 2004 experiment
 - 71% of office workers at Liverpool Street Station were willing to reveal their password for a chocolate bar

Imagine that a person on the street asks you for your access password to your work computer *in exchange* for a chocolate bar. You believe the person does not know you or your workplace. Would you accept the deal?

Yes  2 (1.67%)

No  118 (98.33%)

- Loewenstein “hot/cold” theory

“Buy” vs. “sell” price

Is "sell" price higher or lower than "buy" price?

	Social Security Number	Most favorite online user name	Interests outside work/university
Sell > buy	90.00%	76.67%	75.83%
Sell = buy	5.83%	10.00%	10.83%
Sell < buy	0.00%	10.00%	10.00%
Missing	4.17%	3.33%	3.33%

“Buy” vs. “sell” price

Is "sell" price higher or lower than expected loss?

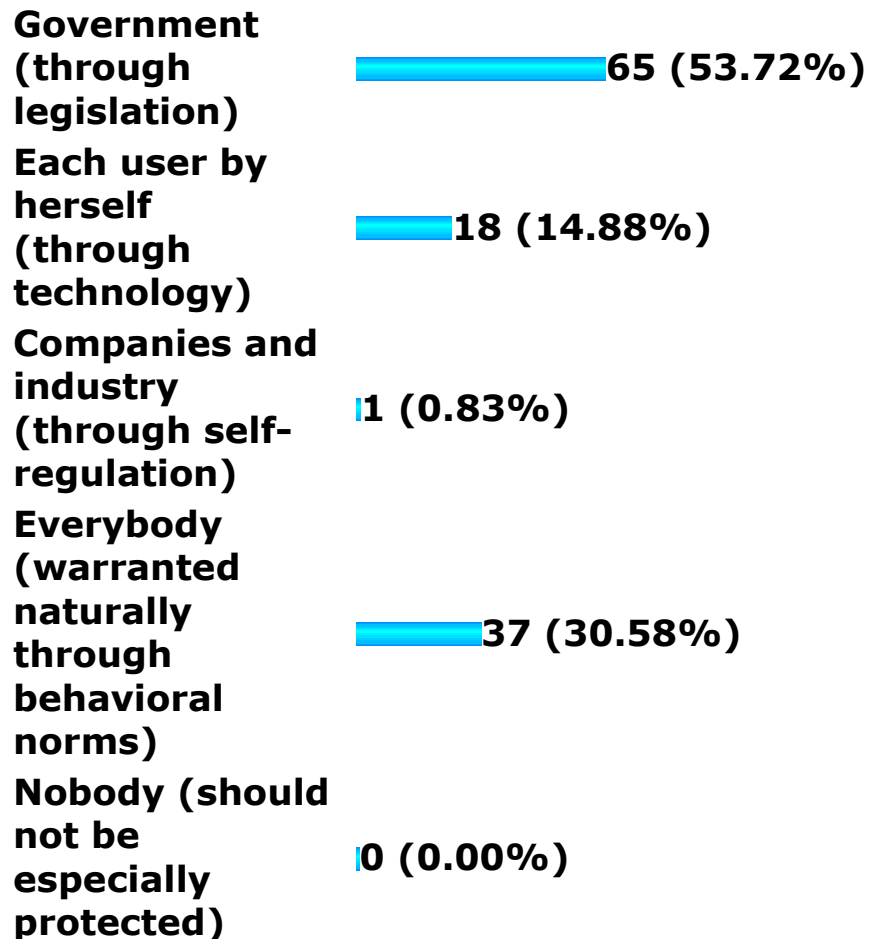
Sell > expected loss	71.43%
Expected loss > sell	7.56%
Missing	21.01%

Is "buy" price higher or lower than expected loss?

Buy > expected loss	39.50%
Expected loss > Buy	36.13%
Missing	24.37%

So... who should protect your privacy?

Do you think that privacy should be protected by:



Assorted conclusions

- Theory
 - Time inconsistencies may lead to under-protection and over-release of personal information
 - Genuinely privacy concerned individuals may end up not protecting their privacy
- Evidence
 - Evidence of overconfidence, incorrect assessment of own behavior, incomplete information about risks and protection, buy/sell dichotomy
 - Rationality model not appropriate to describe individual privacy behavior
- Implications
 - Privacy easier to protect than to sell
 - Self-regulation alone, or reliance on technology and user responsibility alone, will *not* work
 - Economics can show what to protect, what to share
 - Law can send appropriate signals to the market