Sensitive Information in a Wired World

CPSC 457/557, Fall 2013

Lecture 4, September 10, 2013

1:00-2:15 pm; AKW 400

http://zoo.cs.yale.edu/classes/cs457/fall13/

OECD fair information principles

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

Simplified principles

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

Laws and regulations

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

Retail Shopping on the Internet

- Consumer can complete the purchase
  - Without leaving his home
  - Without having to face or talk to another person
- Each purchase leaves a trail of electronic evidence
  - Retailer logs the transaction both for order fulfillment and for customer profiling.
  - Retailer sends the transaction data to other organizations in order to complete the transaction (credit card, shipper, warehouse, factory, *etc.*).
  - Retailer gives or sells these transaction data to business partners and others.
  - Retailer and advertisers put cookies on consumers' machines.
  - Internet traffic is carried by many routers owned by many ISPs.

Retail Shopping in a B&M Store

- Consumer can make the purchase
  - In a store that he has never been to before, where he is unlikely to know anyone.
  - With cash (and not have to identify himself).
- But he may leave a trail of evidence anyway.
  - There may be a surveillance camera in the store.
  - Someone in the store may recognize him, even if he's never been there before and doesn't recognize the observer.
  - A check-out clerk or inventory system may record the purchase, particularly if he buys an unusual item.

Discussion Point:
Which Scenario is More Private?

- Bottom line: Neither is private!

  "You have no privacy. Get over it."

  - Scott McNeely, SUN Microsystems CEO

- However, the B&M-store purchase with cash is, at this time, more likely not to create a searchable, linkable, profilable record.

“Public Records” in the Internet Age

Depending on State and Federal law, “public records” can include:

- Birth, death, marriage, and divorce records
- Court documents and arrest warrants (including those of people who were acquitted)
- Property ownership and tax-compliance records
- Driver’s license information
- Occupational certification

They are, by definition, “open to inspection by any person.”

How "Public" are They?

Traditionally: Many public records were "practically obscure."

- Stored at the local level on hard-to-search media, *e.g.*, paper, microfiche, or offline computer disks.
- Not often accurately and usefully indexed.

Now: More and more public records, especially Federal records, are being put on public web pages in standard, searchable formats.

What are “Public Records” Used For?

In addition to straightforward, known uses (such as credential checks by employers and title searches by home buyers), they’re used for:

- Commercial profiling and marketing
- Dossier compilation
- Identity theft and “pretexting”
- Private investigation

Discussion point: Will “reinventing oneself” and “social forgiveness” be things of the past?

Do We Need a More Nuanced Approach?

Can we distinguish among

- Private information
  - Only the "data subject" has a right to it.
  - ? Example: Legal activity in a private home.
- Public information
  - Everyone has a right to it.
  - ? Example: Government contracts with businesses
- Nonpublic personal information
  - Only parties with a legitimate reason to use it have a right to it.
  - Example: Certain financial information (see, *e.g.*, the Graham-Leach-Bliley Act)

Discussion point: Should some Internet-accessible "public records" be only conditionally accessible? Should data subjects have more control?

Further Reading on These and Related Topics

Electronic Privacy Information Center:

http://www.epic.org/

In particular, EPIC's material on

Public records:

http://www.epic.org/privacy/publicrecords/

Spam:

http://www.epic.org/privacy/junk_mail/spam/

Profiling:

http://www.epic.org/privacy/profiling/

Daniel Solove's Privacy Taxonomy

- Motivation
- Solove's two-level, 16-part taxonomy
- Definitions and discussion of five of the parts
- Questions about this taxonomy and its relationship to other taxonomies we have considered

Motivation for this Work

"Under the **secrecy paradigm**, privacy is tantamount to complete secrecy, and a privacy violation occurs when concealed data [are] revealed to others. If the information is not previously hidden, then no privacy interest is implicated by the collection or dissemination of the information."

Solove's thesis in this article is that the secrecy paradigm has strongly influenced court decisions but is a thoroughly inadequate organizing principle for privacy law.

A. Information Collection

1. Surveillance
2. Interrogation

B. Information Processing

1. Aggregation
2. Identification
3. Insecurity
4. Secondary Use
5. Exclusion

C. Information Dissemination

1. Breach of Confidentiality
2. Disclosure
3. Exposure
4. Increased Accessibility
5. Blackmail
6. Appropriation
7. Distortion

D. Invasion

1. Intrusion
2. Decisional Interference

Aggregation

- The gathering of information about a person
- When combined, disparate pieces of information begin to form a portrait of a person.
- When analyzed, aggregated information can reveal facts that the person did not expect to be revealed when the isolated pieces of information were collected.

Identification

- The linking of information to “the person in the flesh”
- Identification enables us to attempt to verify that, e.g., a person who wants to access a data record is the owner of the account or the subject of the record.
- Identification can attach “informational baggage” to people and inhibit their ability to change.

Insecurity

- Technical glitches, security lapses, carelessness, and abuse or illicit use of information about people
- Information insecurity can lead to identity theft, unauthorized access to bank accounts, and other serious harms.
- Many privacy laws require that information be kept secure. However, courts have been reluctant to award damages to victims, because harm is hard to measure.

Exclusion

- Failure to provide people with notice and input about their records
- Exclusion reduces accountability of organizations that maintain records, violates Fair Information Practices, and often goes hand-in-hand with inadequate security.
- Organizations often claim that requirements to provide notice and input are too costly or, in the case of law enforcement and intelligence, that notice can tip off people under investigation and render the data records useless.

Disclosure

- Revelation to others of true information about a person
- Various statutes restrict disclosure of information from government records, health records, motor vehicle records, school records, and even records of certain commercial activities (such as cable viewing and video rental). The information is deemed to be private and not of interest to the general public.
- Some critics contend that restrictions on disclosure are restrictions on free speech.

Question: General Approach

Is this a useful taxonomy? Are these 16 categories truly distinct, and are they collectively exhaustive of all privacy violations?

Is this highly particularized, incremental approach the right one, or would a broad, simply stated "right to privacy" be more effective?

Question: Relationship to Fair Information Practices

Consider Solove's privacy taxonomy in the context of Fair Information Practices.

Are these two conceptual frameworks consistent? Taken together, do they address essentially all legitimate concerns about "cyber rights"?