18734: Foundations of Privacy

Privacy: Conceptual Frameworks

Anupam Datta

Fall 2014

Today

- Fair Information Practices Principles
  - …and case studies of real privacy policies
- Philosophical and Legal Studies of Privacy
  - …and some connections to technical work

Fair Information Principles

- Task: Critically read the entire privacy policy of a Web services company.
  1. Try to find one example of a piece of the policy that maps to each principle.
  2. Can you find examples of principles that are not reflected in the policy?
  3. Can you find examples of policy clauses that reflect a principle that is not included in these principles?
  4. Are there policy clauses that could be more restrictive or less restrictive with respect to information use in order to better adhere to the principles?
  5. Are there parts of the policy that are too vague? If so, suggest alternatives.

Fair Information Principles

- **Collection Limitation.** There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

- **Data quality principle.** Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

Fair Information Principles

- **Purpose specification.** The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.

- **Use limitation principle.** Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except: (a) with the consent of the data subject; or (b) by the authority of law.

Fair Information Principles

- **Security safeguards principle.** Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.

- **Openness principle.** There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity about usual residence of the data controller.

Fair Information Principles

- **Individual participation principle.** An individual should have the right: (a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him;

(b) to have communicated to him, data relating to him

  - within a reasonable time;
  - at a charge, if any, that is not excessive;
  - in a reasonable manner; and
  - in a form that is readily intelligible to him;

(c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and

Fair Information Principles

(d) to challenge data relating to him and, if the challenge is successful, to have the data erased; rectified, completed or amended.

- **Accountability principle.** A data controller should be accountable for complying with measures which give effect to the principles stated above.

Fair Information Principles

- Task: Critically read the entire privacy policy of a Web services company.
  1. Try to find one example of a piece of the policy that maps to each principle.
  2. Can you find examples of principles that are not reflected in the policy?
  3. Can you find examples of policy clauses that reflect a principle that is not included in these principles?
  4. Are there policy clauses that could be more restrictive or less restrictive with respect to information use in order to better adhere to the principles?
  5. Are there parts of the policy that are too vague? If so, suggest alternatives.

Failure of FIPPs

By Fred Cate

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156972

- Reduced to only procedural principles (like notice and consent)

Consumer Privacy Bill of Rights 2012

http://www.whitehouse.gov/sites/default/files/privacy-final.pdf

- Respect for Context: Consumers have a right to expect that companies will collect, use, and disclose personal data in ways that are consistent with the context in which consumers provide the data.

- Accountability: Consumers have a right to have personal data handled by companies with appropriate measures in place to assure they adhere to the Consumer Privacy Bill of Rights.

Philosophical studies on privacy

- Reading
  - Overview article in Stanford Encyclopedia of Philosophy

    http://plato.stanford.edu/entries/privacy/

- Alan Westin, Privacy and Freedom, 1967
- Ruth Gavison, Privacy and the Limits of Law, 1980
- Helen Nissenbaum, Privacy as Contextual Integrity, 2004

Westin 1967

- Privacy and control over information

"Privacy is the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others"

- Relevant when you give personal information to a web site; agree to privacy policy posted on web site
- May not apply to your personal health information

Contextual Integrity [Nissenbaum 2004]

- Philosophical framework for privacy
- Central concept: *Context*
  - Examples: Healthcare, banking, education
- What is a context?
  - *Set of interacting agents in roles*
    - Roles in healthcare: doctor, patient, …
  - *Informational norms*
    - Doctors should share patient health information as per the HIPAA rules
    - Norms have a specific structure (descriptive theory)
  - *Purpose*
    - Improve health
    - Some interactions should happen - patients should share personal health information with doctors

Informational Norms

"In a context, the flow of information of a certain type about a subject (acting in a particular capacity/role) from one actor (could be the subject) to another actor (in a particular capacity/role) is governed by a particular transmission principle."

Contextual Integrity [Nissenbaum2004]

Privacy Regulation Example (GLB Act)

Sender role

Subject role

Financial institutions must notify consumers if they share their non-public personal information with non-affiliated companies, *but the notification may occur either before or after the information sharing occurs*

Attribute

Recipient role

Transmission principle

Exactly as **CI** says!

Privacy Laws in the US

- HIPAA (Health Insurance Portability and Accountability Act, 1996)
  - Protecting personal health information
- GLBA (Gramm-Leach-Bliley-Act, 1999)
  - Protecting personal information held by financial service institutions
- COPPA (Children's Online Privacy Protection Act, 1998)
  - Protecting information posted online by children under 13

- *More details in later lecture about these laws and a formal logic of privacy that captures concepts from contextual integrity*
- *"Respect for Context" included in Consumer Privacy Bill of Rights 2012*

Gavison 1980

- Privacy as limited access to self

"A loss of privacy occurs as others obtain information about an individual, pay attention to him, or gain access to him. These three elements of *secrecy*, *anonymity*, and *solitude* are distinct and independent, but interrelated, and the complex concept of privacy is richer than any definition centered around only one of them."

- Basis for database privacy definition discussed later

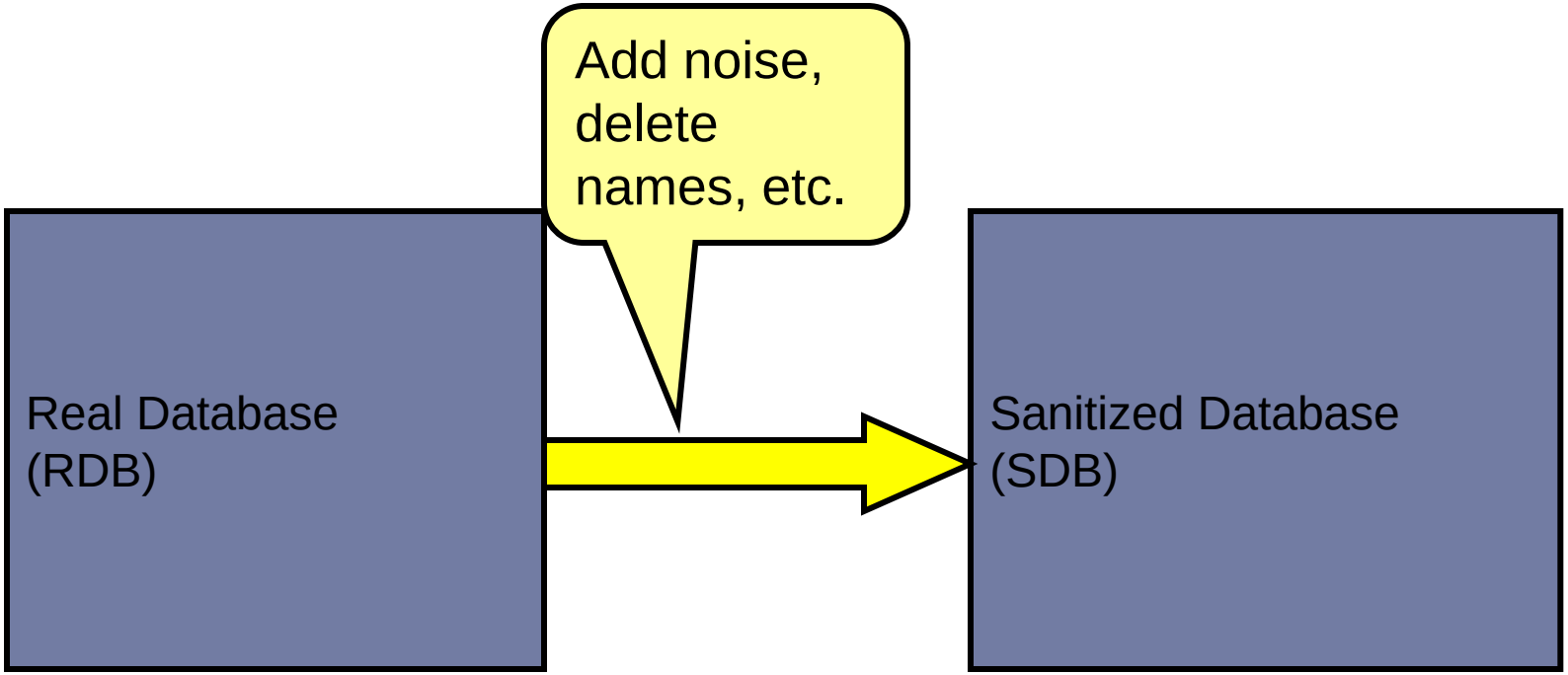
Gavison 1980

- On utility

"We start from the obvious fact that both perfect privacy and total loss of privacy are undesirable. Individuals must be in some intermediate state – a balance between privacy and interaction …Privacy thus cannot be said to be a value in the sense that the more people have of it, the better."

- This balance between privacy and utility will show up in data privacy as well as in privacy policy languages, e.g. health data could be shared with medical researchers

A Connection to Database Privacy



Health records

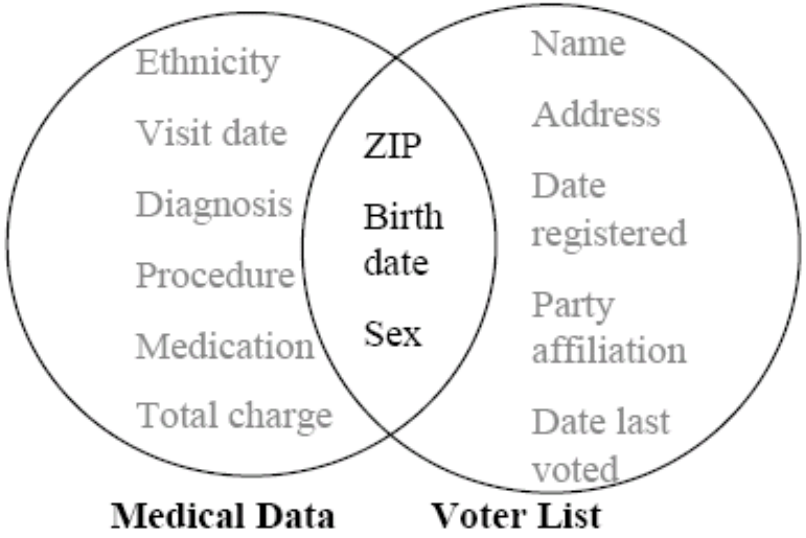
Census data

Protect privacy

Provide useful information (utility)

Re-identification by linking

Linking two sets of data on shared attributes may uniquely identify some individuals:



Example [Sweeney] : De-identified medical data was released, purchased Voter Registration List of MA, re-identified Governor

87 % of US population uniquely identifiable by 5-digit ZIP, sex, dob

K-anonymity

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	13053	28	Russian	Heart Disease
2	13068	29	American	Heart Disease
3	13068	21	Japanese	Viral Infection
4	13053	23	American	Viral Infection
5	14853	50	Indian	Cancer
6	14853	55	Russian	Heart Disease
7	14850	47	American	Viral Infection
8	14850	49	American	Viral Infection
9	13053	31	American	Cancer
10	13053	37	Indian	Cancer
11	13068	36	Japanese	Cancer
12	13068	35	American	Cancer

Figure 1. Inpatient Microdata

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	130**	< 30	*	Heart Disease
2	130**	< 30	*	Heart Disease
3	130**	< 30	*	Viral Infection
4	130**	< 30	*	Viral Infection
5	1485*	≥ 40	*	Cancer
6	1485*	≥ 40	*	Heart Disease
7	1485*	≥ 40	*	Viral Infection
8	1485*	≥ 40	*	Viral Infection
9	130**	3*	*	Cancer
10	130**	3*	*	Cancer
11	130**	3*	*	Cancer
12	130**	3*	*	Cancer

Figure 2. 4-anonymous Inpatient Microdata

Task before next class

- Read the HIPAA Privacy Rule

http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/