CPSC156: The Internet Co-Evolution of Technology and Society

Lecture 15: March 8, 2007

Identity, Anonymity, and Accountability in Information Systems

Reading Assignment

"Identity and Anonymity: Some Conceptual Distinctions and Issues for Research," by G. T. Marx

web.mit.edu/gtmarx/www/identity.html

Posits 7 types of "identity knowledge."

1. A Person's Legal Name

- Usually an answer to the question "Who are you?"
- Involves connection to biological and social lineage.
- Many people may have the same name, but the assumption is often made that there is at most one person of each name born to particular parents at a given time and place.

2. A Person's Address

- Usually an answer to the question "Where are you?"
- Involves location and reachability in actual space or cyberspace.
- Need not involve knowing the person's name or even a pseudonym.
- Note that a person may be unreachable even if his name and address are known; this was true of, *e.g.*, Robert Vesco when he was a fugitive in Cuba.

3. Unique ID: Linkable

- Unique alphanumeric strings, biometric patterns, or pseudonyms that *can* be linked back to actual people *but need not be*.
- Involves trusted intermediaries and conditions under which they should link IDs to people.
- Social security numbers could be used in this way if we had widespread agreement on how they should be used and when they should be linked to names and addresses.

4. Unique ID: Unlinkable

- Unique alphanumeric strings, biometric patterns, or pseudonyms that *cannot* be linked back to actual people.
- Provides a means of discerning information about people without identifying them; someone tested for AIDS may be given a number that he can use to call for results but never have to reveal his name or address.
- Spies, undercover operatives, and con artists may use fraudulent IDs and never reveal their real names to those they deal with.

5. Distinct Appearance or Behavior Patterns

- Some information is necessarily revealed when one interacts with others.
- "Being unnamed is not necessarily the same as being unknown." To a limited extent, you "know" the person you see at 8:15 a.m. on the M23 bus every day.
- Leakage of identifying information is a condition of social existence and has been greatly expanded by new technologies.

6. Social Categorization

- Forms of identification that do not distinguish among members of a group; the group may be defined by gender, ethnicity, religion, age, economic class, *etc.*
- Number of categories has exploded with new technology and expanded bureaucracy.
- New categories (credit scores, IQs, life-style categories used in marketing, *etc.*) may or may not be known by the people in them; this was not true of traditional social categories.

7. Certification: Proof of the possession of knowledge or skill

- Knowing a secret password, being able to swim, *etc.* are ways to prove that one is entitled to certain privileges or is a member of a certain group.
- These proofs may be linkable to individual people (as passwords often are) but need not be.
- Provides essential balance between the need to control sensitive personal information and the need to restrict access to and prevent abuse of systems.

"Policies" for the Handling of Sensitive Data (not from Marx)

- Collection
- Retention, destruction
- Use, mining
- Sharing, selling
- Updating, cleaning, correcting
- De-identifying, scrubbing, re-identifying
- . . .

Basic Questions (1)

- What are the best tools for expressing and analyzing policies?
- How can an organization ensure that it is following its own data-management policies?
- How can those who transmit data to an organization ensure themselves that the organization is following its data-management policies?

Basic Questions (2)

- What recourse does one have when an organization that handles one's data violates a policy?
- Are there "implicit policies" or, more generally, when should one be held accountable for actions *not* clearly governed by a specific policy?

Who is Accountable to Whom?

- Individuals
- Organizations
- Governments
- Technology vendors
- Network operators
- . . .

When is it ok NOT to be Accountable?

- Anonymous activity?
- Unobservable activity?
- "Pseudonymous" = Unidentifiable but accountable?
- Offline analogs
- . . .