Transparency and Accountability
Motivating the Policy Aware Web

Daniel J. Weitzner djweitzner@csail.mit.edu
MIT Computer Science and Artificial Intelligence Laboratory
Decentralized Information Group
Motivating the Policy Aware Web

- Privacy and Security Challenges
  - Case studies
  - Larger trends
- Inadequacy of current legal and technical tools
- Way forward: Embrace transparency and require accountability
- What is the Policy Aware Web?
Privacy, Security and Transparency Challenges

- Law Enforcement
- Large-scale Data mining
- Database Security
“Authorities in the Washington region spotted the same faded blue 1990 Chevrolet Caprice and recorded its New Jersey tags on at least 10 different occasions this month….

Each time, however, they let the driver go after finding no record that it had been stolen or that its occupants were wanted for any crimes.


‘We were looking for a white van with white people, and we ended up with a blue car with black people,’ said D.C. Police Chief Charles H. Ramsey”
Large-scale datamining (for airline security)

- **TTIC**
  - Aliases: Alfred Marx (132 Main St)
  - = Aldo Moro
  - Watchlist: Alfred Marx

- **TSA**
  - PNR: Aldo Moro on fl 302 DCA->BOS
  - Inference: Aldo Moro = Alfred Marx => secondary screening

- **INS**
  - Out of Status list: Alfred Marx
  - Under suspicion: Aldo Moro
  - Possible Inference: Aldo Moro should be investigated
Database Leaks

- **Choicepoint**: access allowed for unauthorized persons
- **Lexis-Nexis**: data loss
- **Iron Mountain**: lost tape of plain-text personnel records
Larger technical trends toward transparency

- The end of ‘stovepipes’
- Cheap query
- Location-aware sensor nets
- Cost of storage = zero
1. The End of ‘Stovepipes’
2. Cheap query: institution-wide and Web-wide
3. Location-aware Sensor Nets

Motes attached to patients collect vital signs (pulse ox, heart rate, etc.)

Ambulance system makes triage decisions, relays to EMTs

PDAs carried by EMTs receive vital signs and enter into field report

Correlate with patient records at hospital
4. Cost of Storage approaches zero
Avoiding Technological Determinism
And Avoiding Policy Formalism
Challenge to existing privacy model: from evidence to inference

Most intrusive practices are from inferences drawn, not evidence collected:

- Credit card transactions → profiling
- Web logs → Web search patterns
- Instantaneous location → travel patterns
Current Legal Tools to address inferencing power

- US Const. Fourth amendment
- Wiretap laws
- EU Data Protection Directive

However, neither the Fourth Amendment nor wiretapping statutes impose any limits on what the government can do with information once it has been collected.
Current technical tools to address inferencing power

- Anonymity and de-identification: but see Sweeney
- “Privacy-preserving” data mining through Secure Multiparty Computation: but consider leaks in real-world application
Managing Transparency With Accountability

- Apply accountability techniques to:
  - Law enforcement data collection
  - Large scale data mining
  - Security policy audit
- Policy Aware Technology is *Necessary but not Sufficient*
- Public Policy agenda: facilitate accountability
  - from collection limits to usage rules
  - Enforcement of agreements
Living with the Transparency Paradox

"Assume the (unauthorized) user knows all ciphering procedures"
-Kerckhoffs’ Principle (1835-1903)

The transparency paradox: that we may have to embrace greater exposure of personal information in order to advance fundamental privacy values.