How private is your privacy?

Karen Sollins

MIT CSAIL

April 21, 2016

The increasing thunder

- In the news
  - Snowden
  - Apple vs. the FBI
- In our civil society
  - Warren/Brandeis, **The Right to Privacy**, 1890
  - Increasing tension between rights to privacy and expectation (need?) for surveillance
- In industry
  - Claims of Facebook/Google/etc.
  - Rise of examples such as Duck-duck-go, Bitcoin, Yik-Yak, etc.

Data at the core

- Privacy is about data
    - Access to data
    - Use of data
- Privacy is not binary
- Privacy is context sensitive
- Future privacy interests dependent on exposures inherent in future uses of data

The Data Life Cycle & privacy approaches

- Data collection
  - Notice and consent
  - Informed consent
  - Data restriction
    - Algorithms such as k-anonymity, l-diversity, t-closeness
    - Differential privacy
- Data access controls
  - Data use agreement
  - Tagging
  - DRM style management
  - Authentication/authorization protocols
  - Standard encryption
- Data processing (incl. fusion) and analytical methods
  - Individual insights vs. aggregate population insights: querying approaches including personal/private data stores, secure multi-party computation, homomorphic encryption
  - Aggregate population insights: statistical methods such as differential privacy, and synthetic datasets vs. baysian statistics
- Data compliance and audit
  - Legal policy compliance: Legalease and Grok (from Microsoft)
  - User access logging
  - Accountable systems
- Data destruction
  - Deletion or encryption
  - Inremental forgetting of bits of encryption keys (Garfinkel)

The Data Stakeholders

- Data subject(s): primary and secondary
- Decision makers
- Data collectors
- Data curators
- Data analysts
- Data platform providers
- Policy enforcers
- Auditors

Interests of each group:

Their own effectiveness

Their integration with others interests

The challenge: risk vs. trust

- In the human (individual, societal, and commercial) arenas how do we compose the risks and willingness to trust into a unified decision-making opportunity.
- The question we are often left with is: Should we take the risk to privacy by trusting stakeholders to provide some definition and degree of "privacy".

Looking forward in PrivSec

- Talk series: Metrics for Privacy
  - K-anonymity (Sweeney)
  - L-diversity (Machanavajihala et al,)
  - T-closeness (Li et al.)
  - Differential privacy: ε (Dwork, Vadhan, etc.)
  - Information theoretic approach to Privacy (Bezzi – SAP Labs)
  - Taxonomy for Information Privacy Metrics (Davarathna)
  - Discussions with Facebook, Thomson-Reuters and seeking others (banking and other financial industries, healthcare, etc.)
- Observation: risk metrics used in a) defining algorithms or b) evaluation
- Objective: understand role of metrics of privacy and possible composability.

For further information

Karen Sollins

sollins@csail.mit.edu