

# Working Group on Current Technical Directions and Outlook

---

IBM Almaden Institute  
Privacy for Data Systems

Moderator:  
Johannes Gehrke  
Cornell University  
[johannes@cs.cornell.edu](mailto:johannes@cs.cornell.edu)

# Goals of the Working Group

---

- Formulate **existing capabilities** and **open problems**:
  - Identity management
  - Specification languages
  - Cryptographic primitives
  - Novel architectures
  - Privacy-preserving data mining
  - Technical definitions of privacy
  - ...
- Develop a set of challenge problems: Scenarios, benchmarks, challenge datasets

# Mechanics

---

- Existing capabilities versus open problems
- Each person wrote down top-5 existing capabilities and open problems
- Group arranged topics into categories and prioritized and discussed

# Existing Capabilities

---

- Policies languages/policy specification
  - EPAL
  - P3P
  - XACML
- Policy enforcement
  - Tivoli PM
  - Hippocratic databases
  - Access control (relational data, XML, ...)
  - Enterprise-mediated versus individual

## Existing Capabilities (Contd.)

---

- Database security and access control
- Inference control for statistical databases
- Identity management
  - IDEMIX (IBM Zurich)
  - Digital identities/pseudonyms and credentials
- Cryptographic primitives

# Existing Capabilities (Contd.)

---

- Data mining/analysis
  - Direct (centralized data collection) versus indirect (distributed protocols)
  - Randomized versus cryptographic protocols
- Data retrieval/access
  - PIR
  - Secure Multiparty Computation
- Anonymity
  - Voting
  - Anonymous remailers
  - Identity management

# Open Technical Problems

---

- Policy languages and specifications
  - Expressiveness versus ease of use: From natural to formal languages
  - Ontologies that define terms
  - Composability
  - Standards
- Policy enforcement
  - Goal: Automatic enforcement
  - Accountability/auditability

# Open Technical Problems (Contd.)

---

- Identity management
  - Linkage without identification, user-controlled linkability/reputation management
  - Location management
  - Data aggregation (pseudonyms)
- Cryptographic primitives
  - Power-efficiency and performance for resource-constrained environments
  - Efficiency of secure-multiparty computation protocols

# Open Technical Problems (Contd.)

---

- Database Issues

- Application of techniques from SMPC: Privacy-preserving SQL across distributed DBMS
- Support for enforcing privacy policies
- Structured versus un-structured data (lack of metadata)
- Performance overhead (Example: fine-grained access control)
- Architecture?
- Integration of legacy data that existed prior to privacy policies

# Open Technical Problems (Contd.)

---

- Context Management
  - How to define context
  - Context-dependent policy enforcement
- Scalability and performance
- Beyond relational data and text
- Economic models of data privacy
- Definitions of privacy

# Questions?

---