

Integrating Approaches to Privacy across the Research Lifecycle

September 24-25, 2013 — Harvard University

Use case discussion questions

-- for discussion in the morning, during the use case breakouts

1. *Characterization*. Are there key additional characteristics of the use case that should be noted? How do these characteristics change the analysis and treatment of privacy in these cases?
2. *Current approaches*. How is this use case treated now -- what's the state of the art & practice? How is success measured?
3. *Enhancing approaches*. Are any of the approaches discussed yesterday used? How could the tools and approaches mentioned earlier or other existing tools be used at particular stages of the research lifecycle to enhance utility and privacy?
4. *Integrating approaches*. Are approaches that have been developed and used in different communities compatible with each other? How should legal, computational, policy, and statistical tools be integrated so as to be most effective?
5. *Utility*. What things would stakeholders like to do with the data that the toolset doesn't restrict or obstruct? Where is social benefit sub-optimal? How is utility measured/perceived by the stakeholders?
6. *Privacy*. What sorts of data/outputs are considered particularly sensitive? What are the most important real and perceived risks -- what harms could occur if data is released and reidentified, how severe are these harms and how likely?
7. *Methodological Barriers*. What are technical, methodological, computational or infrastructural barriers to improving privacy and utility in the management of this data. What particular characteristics of the use case contribute barriers?
8. *Incentives*. If better tools already exist, why aren't they used? What are barriers to adoption of new tools and methods? What are the specific "market failures" in this area -- such as perverse incentives, lack/asymmetry of information, lack of well-developed market, irrational behavior, transaction cost, network effects, etc.? What particular characteristics of the use case most influence incentives?
9. *Future*. How is this use case likely to evolve over time? What are threats to stability/scalability/robustness/resilience of the proposed/current solutions?
10. *Prior work*. Are there key additional examples of the use case that should be noted? Are there additional key references or writings that should be noted?

Cross-cutting Questions

-- for discussion mid-afternoon, after use case report-out:

1. *Commonalities and differences in use cases*. Are there common themes in the use-case group findings? Are there common challenges across the use cases discussed, common tools in use, common ways of measuring privacy or utility? Where are needs

distinctly different – where is a concept, method or tool used in one area ignored or rejected in another? Where there are commonalities or differences what attributes of the use case could explain these?

2. *Gaps in use case coverage.* Are key use cases missing? If so, what are they, and in what important attributes do they differ from the use cases already discussed? What are key examples of such use cases and cogent analyses/discussions of them?
3. *Gaps in tool coverage.* What legal, computational, policy, and statistical tools for addressing data privacy have not been discussed? In what contexts are these tools most often used or applicable? What are good discussions of them?

Disciplinary Specific Questions

-- for topical breakouts late afternoon.

1. *Requirements analysis.* What aspects, if any, of the use cases discussed have not been well-understood in the discipline? What else would need to be understood about them to apply existing disciplinary theories, tools and methods effectively?
2. *Prior work.* What are key surveys or core foundational works in the discipline that provide analysis, methods or applications relevant to research in data privacy?
3. *Gaps in disciplinary basic, applied and translational research.* How might the discipline contribute to solving problems of managing research data, as described? What basic, applied and translational research questions would need to be answered?
4. *Near term Opportunities.* What gaps are most important? What are most tractable in 2 years, 5 years? How would we know if such research were successful?
5. *Grand challenges.* A grand challenge is a fundamental unresolved research problem with broad applications, large scale and deep impact, and long-term benefits for the progress of science and society, that is potentially solvable in the next ten-fifteen years. Are there grand challenge problems in the discipline that would also have a large impact on research data privacy?