Bridging Privacy Definitions: Differential Privacy and Concepts from Privacy Law & Policy

Alexandra Wood
Berkman Klein Center for Internet & Society at Harvard University

DIMACS/Northeast Big Data Hub Workshop on Overcoming Barriers to Data Sharing including Privacy and Fairness
October 23 - 24, 2017

An Interdisciplinary Collaboration

This work is the product of an *interdisciplinary working group* bringing together computer scientists and legal scholars

Kobbi Nissim, Aaron Bembenek, Mark Bun, Marco Gaboardi, Thomas Steinke, Salil Vadhan

Alexandra Wood, David O'Brien, Urs Gasser

These opinions are my own. They are not the opinions of the Berkman Klein Center, any of our funders, nor (with the exception of co-authorship on previously published work) my collaborators.

Motivation

Formal privacy models like differential privacy offer a solution for providing wide access to statistical information with guarantees that individual-level information will not be leaked inadvertently or due to an attack.

- Formal mathematical privacy concept that addresses weaknesses of traditional schemes (and more).
- Supported by a rich theoretical literature and now in initial stages of implementation and testing by industry and statistical agencies.

Motivation

Formal privacy models like differential privacy offer a solution for providing wide access to statistical information with guarantees that individual-level information will not be leaked inadvertently or due to an attack.

- Formal mathematical privacy concept that addresses weaknesses of traditional schemes (and more).
- Supported by a rich theoretical literature and now in initial stages of implementation and testing by industry and statistical agencies.

However, these tools cannot be used to share sensitive data with the general public unless they satisfy legal standards with some certainty.

Introduction to the Legal Framework for Privacy

What Is Privacy?

“The claim of individuals, groups, or institutions, to determine for themselves when, how, and to what extent information about them is communicated to others.”

- Alan Westin

Broad Notions of Privacy

- A function of generally accepted social norms
- Access to information about the self – gradients between public and private
- Individuality, personhood, intimacy, dignity, reputation, and autonomy
- Freedom to inquire
- Enabler of creativity, counter-culture
- Control over information; power

Sources of Governance

- Constitutional Law (limits on government action)
  - Fourth Amendment
  - First Amendment
- Written law (statutes, regulations)
  - FERPA, HIPAA, etc.
  - Common Rule research regulations
  - Various state laws
- Common law (judicially developed)
  - Judicial opinions, precedent of statutes
  - Torts – civil injuries
  - Contracts

Relevance to Data Analysis and Sharing

- Various legal provisions restrict disclosures of identifiable or sensitive information about individuals.
    - e.g., FERPA generally prohibits the disclosure of personally identifiable information from education records, except with consent or pursuant to one of several narrow exceptions to the consent requirement. Notably, FERPA permits the disclosure of de-identified information.
- However, there is a lack of certainty around the use of terms like *personally identifiable information* and *de-identified information*, especially as the understanding of privacy risks continues to evolve over time.

Challenges

- De-identification standards are highly sector- and context-specific and vary widely depending on the setting. For example, some standards provide an objective for de-identification, while others prescribe a method for de-identification.
- Applicability is typically a binary determination that turns on the interpretation of terminology such as personal information, personally identifiable information, or individually identifiable information.
- Practices also vary, but generally are heuristic and focus on withholding, removing, or coarsening pieces of information considered to be identifying.

Variations in Standards: Selected Laws

- Family Educational Rights and Privacy Act
- HIPAA Privacy Rule
- Privacy Act
- OMB Guidance
- Title 13 (U.S. Census Bureau)
- Confidential Information Protection and Statistical Efficiency Act
- Massachusetts data security regulation

Overview of Selected Privacy Laws

FERPA: Family Educational Rights and Privacy Act

Protects personally identifiable information in education records maintained by educational agencies and institutions, including

"names, addresses, personal identifiers (e.g., SSNs, student numbers, biometric records), indirect identifiers (e.g., date of birth, place of birth, mother's maiden name), other information that, alone or in combination, is linked or linkable to a specific student that would allow a reasonable person in the school community, who does not have personal knowledge of the relevant circumstances, to identify the student with reasonable certainty, or information requested by a person who the educational agency or institution reasonably believes knows the identity of the student [in the requested record]."

(20 C.F.R. § 99.3)

FERPA: Family Educational Rights and Privacy Act

Permits the release of de-identified information, without consent,

"after the removal of all personally identifiable information provided that the educational agency or institution or other party has made a reasonable determination that a student's identity is not personally identifiable, whether through single or multiple releases, and taking into account other reasonably available information."

(20 C.F.R. § 99.31(b)(1))

HIPAA Privacy Rule

HIPAA establishes rules governing protected health information held by covered entities.

Protected health information is information, including demographic information, which relates to:

- the individual's past, present, or future physical or mental health or condition,
- the provision of health care to the individual, or
- the past, present, or future payment for the provision of health care to the individual, and

that identifies the individual or for which there is a reasonable basis to believe can be used to identify the individual.

HIPAA Privacy Rule

Method #1 for de-identifying data: Expert determination

A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:

> (i) Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and
>
> (ii) Documents the methods and results of the analysis that justify such determination

HIPAA Privacy Rule

Method #2 for de-identifying data: Safe harbor

(i) Categories of information from a list of 18 identifiers (e.g., names, geographic units containing 20,000 or fewer people, dates (except year), telephone numbers, Social Security numbers, etc.) are removed, and

(ii) The covered entity does not have actual knowledge that the information could be used alone or in combination with other information to identify an individual who is a subject of the information.

(45 C.F.R. § 164.514)

Privacy Act of 1974

- Generally prohibits federal executive agencies from disclosing personal information about U.S. citizens and legal permanent residents maintained in a system of records, except as authorized by the data subject.
    - A system of records contains information that is retrieved by an individual's name or other unique identifier.
- Establishes a code of fair information practices that governs the collection, maintenance, use, and dissemination of information about individuals that is maintained in systems of records by federal agencies.

OMB Guidance

Breach notification policies and guidance for federal agencies: "The term PII refers to information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual. Because there are many different types of information that can be used to distinguish or trace an individual's identity, the term PII is necessarily broad. To determine whether information is PII, the agency shall perform an assessment of the specific risk that an individual can be identified using the information with other information that is linked or linkable to the individual. In performing this assessment, it is important to recognize that information that is not PII can become PII whenever additional information becomes available--in any medium or from any source--that would make it possible to identify an individual."

OMB Memorandum M-17-12, "Preparing for and Responding to a Breach of Personally Identifiable Information," Jan. 3, 2017.

Title 13 (U.S. Census Bureau)

- Authorizes the Census Bureau to conduct the census and supplemental surveys
- Provides that the information collected by the Census Bureau from individual persons, households, or establishments be kept strictly confidential and be used only for statistical purposes.
- Prohibits Census Bureau employees from "mak[ing] any publication whereby the data furnished by any particular establishment or individual under this title can be identified" 13 U.S.C. § 9(a)(2).

CIPSEA: Confidential Information Protection and Statistical Efficiency Act

- Protects information collected by any federal agency directly from respondents under a pledge of confidentiality for exclusively statistical purposes.

- Protects data in identifiable form, meaning "any representation of information that permits the identity of the respondent to whom the information applies to be reasonably inferred by either direct or indirect means."

Pub. L. No. 107-347, tit. V, § 502 (4) (2002).

Massachusetts data security regulation

Protects personal information, defined as the combination of

(1) a Massachusetts resident's first name (or first initial) and last name, and

(2) any one or more of the following:

  (a) Social Security number,

  (b) Driver's license number or state-issued identification card number, or

  (c) Financial account number, or credit or debit card number.

This definition explicitly excludes publicly available information.

(201 Mass. Code Regs. § 17.00)

Gaps between Differential Privacy and Legal Standards for Privacy

Challenges for Formal Privacy Models

Demonstrating that formal privacy models satisfy applicable legal requirements is challenging due to the conceptual gaps between legal and technical approaches to defining privacy.

Notably, information privacy laws are generally:

- context-specific,
- subject to interpretation,
- allow for some degree of flexibility, and
- rely on traditional, often heuristic, conceptions of privacy,

which creates uncertainty for the implementation of more formal approaches.

Example Points of Mismatch

FERPA

- Applies to highly sector- and context-specific settings
- Contemplates a small set of specific types of privacy attacks
- Protects a small set of information (non-directory PII)
- Refers to the obvious extreme cases, not to more difficult "gray areas"
- Applies to releases of microdata and tabulations
- Imprecise, not rigorous/formal from a technical standpoint

Differential Privacy

- Offers general privacy protection
- Addresses a very large class of potential data misuses
- Protects any information contributed by an individual
- Applies to all analyses, does not leave "gray areas"
- Not limited to releases of microdata and tabulations
- A mathematically rigorous definition

Is it possible to bridge these very different languages?

$M: X^n \rightarrow T$ satisfies ϵ-differential privacy if

$\forall x, x' \in X^n$ s.t. $dist_H(x, x') = 1$ $\forall S \subseteq T$,

$$\Pr_{\mathbf{M}}[M(x) \in S] \leq e^{\epsilon} \Pr_{\mathbf{M}}[M(x') \in S].$$

Opportunities for Bridging Privacy Definitions

Approach #1: Formal Modeling

Approach #1: Formal Modeling

We seek a methodology for rigorously arguing that a technological privacy solution satisfies the requirements of a particular law.

The proposed approach has two components:

1. Extraction of a formal mathematical requirement of privacy based on a legal standard found in an information privacy law, and
2. Construction of a rigorous mathematical proof for establishing that a technological privacy solution satisfies the mathematical requirement derived from the law.

Illustration: Formally Modeling FERPA

Goal: To extract a formal model of the Department of Education's privacy desiderata for FERPA, in the form of a game-based privacy definition:

- Provides a concise and fairly intuitive abstraction of the requirements in FERPA.
- Enables us to prove that if a formal model, such as differential privacy, satisfies the game-based definition, then we have a strong argument that it satisfies the requirements of FERPA.

Although FERPA is not written with a privacy game framework in mind, we claim (and demonstrate) that it is possible to extract a game that is based on its requirements.

Extracting a Formal Definition from FERPA

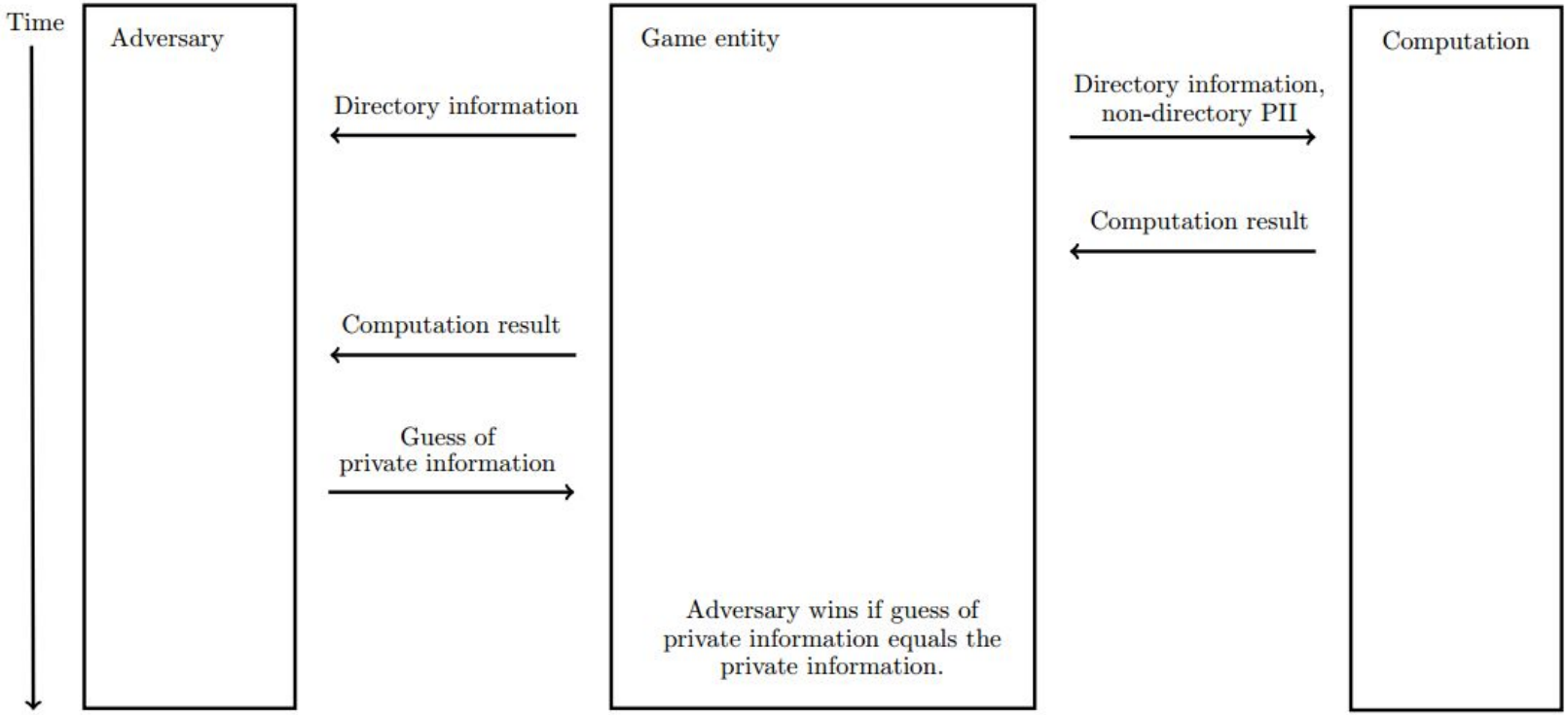
FERPA allows the release of de-identified information and directory information from education records.

De-identification can be thought of in terms of a computation; e.g., requiring the removal of identifying attributes can be seen as requiring a computation to redact those identifiers from the input data.

- ➢ This framing is useful for modeling a law's requirements using the formal language used in computer science. This modeling allows us to extract a mathematical definition for determining whether a computation meets the FERPA privacy standard.

But how do we know whether a given computation provides a sufficient level of privacy protection to meet the requirements of FERPA?

Components of a FERPA Privacy Game



Modeling FERPA: Directory Information

The regulatory language is **ambiguous**, so we interpret the language as conservatively as reasonably possible. In other words, where there is ambiguity, we err on the side that is most beneficial for the adversary.

- For example, the definition of **directory information** (i.e., information that can be disclosed because it is not considered harmful) is ambiguous (e.g., the definition varies between schools).

We could make assumptions in defining directory information in our model. However, new interpretations could call these assumptions into question.

- Instead, **we let the attacker to choose** what constitutes directory information.

Modeling FERPA: The Adversary

Personally identifiable information: "information that, alone or in combination, is linked or linkable to a specific student that would allow a reasonable person in the school community, who does not have personal knowledge of the relevant circumstances, to identify the student with reasonable certainty."

This is FERPA's **implicit adversary**. Key points from guidance:

- We should not assume anything about the skill level of the adversary.
- Standard is based on the knowledge of a member of the school community, which is stronger than one based on the knowledge of any reasonable person.
- The adversary can have both high-level knowledge (e.g., demographics of school) and "insider" knowledge about specific individuals in local community.

Modeling FERPA: Adversary's Knowledge

The adversary clearly has (potentially a lot of) knowledge, but by definition does not have "personal knowledge of the relevant circumstances."

In our model, the adversary has access to any information that is publicly available, but has some uncertainty about private student information.

We model the adversary's knowledge via probability distributions. Adversary associates with each student a probability distribution that represents her knowledge about the private information of that student. We allow the adversary to choose these statistics.

Example: If Alice comes from a school where 50% of the students failed the state math proficiency exam, then adversary might associate with Alice a distribution that has her failing the exam with a probability of 0.5.

Proving Differential Privacy Satisfies FERPA

Developing a formal definition of privacy protection based on the requirements of FERPA allows us to reason, with high confidence, about whether the use of a privacy technology satisfies FERPA.

For instance, we can prove mathematically that any computation that is differentially private meets this definition, and (since the requirements of this definition are likely stricter than that of FERPA) thus satisfies the privacy requirements of FERPA.

Opportunities for Bridging Privacy Definitions

Approach #2: Interpreting the Differential Privacy Guarantee

Approach #2: Interpreting the Differential Privacy Guarantee

- Legal requirements relevant to issues of privacy in computation rely on an understanding of a range of different privacy concepts.
- While none of the privacy concepts that appear in the law refer directly to differential privacy, the differential privacy guarantee can be interpreted in reference to these concepts—while accommodating differences in how these concepts are defined across contexts.

Common Privacy Concepts in the Law

- Personally identifiable information
- De-identification
- Linkage
- Inference
- Identification risk
- Consent and opting out
- Purpose and access restrictions

These concepts are interpreted differently across laws. They also appear in the technical literature, often with different definitions and interpretations.

Personally Identifiable Information

Personally identifiable information (also *personal information*, *individually identifiable information*) is a central concept appearing in information privacy law.

- Legal protections typically extend only to PII, and information not considered personally identifiable is not protected.
- Examples: FERPA, HIPAA Privacy Rule, Massachusetts data security regulation, OMB memorandum, among many others
- Although definitions of personally identifiable information vary significantly, they are generally understood to refer to the presence of pieces of information that are linkable to the identity of an individual or to an individual's personal attributes.

PII: Interpretation of DP Guarantee

The term PII does not have a precise technical meaning, and in practice it can be difficult to determine whether information is personal, identifying, or likely to be considered identifying in the future.

Further, the meaning of PII in releases that are not in a microdata or tabular format, such as statistical models or outputs of a machine learning system, is unclear.

Regardless of the definition of PII that is used, differential privacy can be interpreted as (essentially) ensuring that using an individual's data will not reveal any PII that is specific to her.

- Here, *specific* is used to refer to information that cannot be inferred unless the individual's information is used in the analysis.

De-identification

The term de-identification refers to a collection of techniques that aim to transform identifiable information into non-identifiable information, while also preserving some utility of the data. In principle, it is intended that de-identification, if performed successfully, can be used as a tool for removing PII, or transforming PII into non-PII.

- Examples: FERPA, HIPAA Privacy Rule

Any algorithm that satisfies the requirements of differential privacy has the property that using an individual's data will (essentially) not reveal PII that is specific to him or her.

- Because the output of a differentially private computation does not reveal such information, any differentially private algorithm should be considered sufficient for de-identification.

Linkage

One of the most common modes of privacy loss recognized by privacy regulations, implicitly or explicitly, is a successful record linkage.

- Linkage typically refers to the matching of information in a database to a specific individual, often by leveraging auxiliary data sources.
- Example: FERPA defines personally identifiable information in terms of information "linked or linkable to a specific student." (34 C.F.R. § 99.3)

Linkage attacks have a concrete meaning when data is published as a collection of individual-level records, often referred to as microdata. However, what is considered a successful linkage when a publication is made in other formats (including, e.g., statistical models and synthetic data) is open to interpretation.

Linkage: Interpretation of DP Guarantee

Despite these conceptual gaps, it can be argued that differential privacy addresses reasonable interpretations of record linkage.

- Microdata or contingency tables that allow the identification of population uniques cannot be created using statistics produced by a differentially private tool.
- Differential privacy masks the contribution of a single individual, making it impossible to infer any information specific to an individual, including whether an individual's information was used.
- Differentially private statistics provably hide the influence of every individual, and even groups of individuals, providing protection not only against releasing exact records but also approximate statistics that could leak individual-level information.

Inference

Some information privacy laws, or interpretations of these laws, refer to modes of privacy loss involving inference.

- Example #1: CIPSEA protects ``any representation of information that permits the identity of the respondent to whom the information applies to be reasonably inferred by either direct or indirect means.'' Pub. L. 107-347 § 502(4) (emphasis added).
- Example #2: FERPA defines personally identifiable information, in part, in terms of information that would allow one to identify a student "with reasonable certainty." 34 C.F.R. § 99.3.

Inference: Interpretation of DP Guarantee

It is important to distinguish between two types of inferences: inferences about individuals and inferences about large groups of individuals.

Differentiating between these two categories of inference is key to enabling socially beneficial uses of data, such as research investigating the relationship between smoking and lung cancer, while protecting individuals from disclosures of information specific to them.

Differential privacy rules out inferences about individuals, thereby protecting individuals from inferences about values or attributes that are specific to them. To achieve this goal, differential privacy adds a small amount of uncertainty, similar to traditional SDL techniques.

Identification Risk

Some information privacy laws refer to an acceptable level of risk of identification of a record in a data release. Similarly, other laws often acknowledge, implicitly or explicitly, that any disclosure of information carries privacy risks, and therefore the goal is to minimize rather than eliminate such risks.

- In guidance, the Dept. of Education refers to the goal of FERPA's de-identification requirements in terms of "minimiz[ing] the risk of disclosing personally identifiable information in redacted records or statistical information." 73 Fed. Reg. 74,806, 74,835 (Dec. 9, 2008).
- The HIPAA Privacy Rule requires covered entities to use de-identification techniques prior to releasing data in order to create a dataset with only a "very small" risk of identification. 45 C.F.R. § 164.514(b)(1).

Risk: Interpretation of DP Guarantee

Differential privacy enables a formal quantification of risk, and the privacy loss parameter epsilon can be tuned to different legal requirements for minimizing risk.

Regardless of how identification risk—or privacy risk, more generally—is defined, differential privacy guarantees that the risk to an individual is almost the same with or without her participation in the dataset.

- In this way, differential privacy can be interpreted to guarantee that the risk to an individual is minimal or very small.

Consent and Opting Out

Some information privacy laws include consent provisions, or opt out provisions, by which individuals can choose to allow, or not to allow, their information to be used by or redisclosed to a third party, respectively.

- For example, FERPA, the HIPAA Privacy Rule, and the Privacy Act of 1974 generally prohibit the disclosure of certain records containing personal information, absent the consent of the individuals involved.

- FERPA also includes a provision requiring educational agencies and institutions to offer students an opportunity to opt out of the disclosure of their personal information in school directories. 34 C.F.R. § 99.37.

Opt Out: Interpreting the DP Guarantee

Differential privacy can be viewed as automatically providing all individuals in the data with the protection that opting out is intended to provide.

- When differential privacy is used, the consequences for an individual's privacy are almost the same whether or not an individual's information is included in an analysis.
- Moreover, differential privacy provides all individuals with this privacy guarantee, thereby avoiding the possibility that individuals who choose to opt out would, by doing so, inadvertently reveal a sensitive attribute about themselves or attract attention as individuals who are potentially hiding sensitive facts about themselves.

Purpose Restrictions

Information privacy laws often permit the nonconsensual disclosure of data for certain purposes, such as "statistical purposes."

- Title 13 restricts the use of confidential information from respondents, prohibiting uses "for any purpose other than the statistical purposes for which it is supplied." 13 U.S.C. § 9(a)(1).
- CIPSEA: "Data or information acquired by an agency under a pledge of confidentiality and for exclusively statistical purposes shall be used by officers, employees, or agents of the agency exclusively for statistical purposes." § 512(a). A statistical purpose is "the description, estimation, or analysis of the characteristics of groups, without identifying the individuals or organizations that comprise such groups;" and "includes the development, implementation, or maintenance of methods, technical or administrative procedures, or information resources that support the purposes described [in this definition]." § 502(9).

Purpose Restrictions

Legal requirements reflecting purpose and access restrictions such as these can be divided into two categories. Restrictions limiting use to statistical purposes, including statistical purposes involving population-level rather than individual-level analyses or statistical computations, are consistent with the use of differential privacy.

- Tools that satisfy differential privacy can be understood to restrict uses to only those that are for statistical purposes.

Other use and access restrictions such as restrictions limiting access to individuals with "legitimate educational interests" are orthogonal to differential privacy.

Public data releases, free from use restrictions, may even be viewed as demanding the use of formal approaches such as differential privacy that preserve privacy in post-processing.

Conclusion

With the emergence of new technologies based on formal privacy models, can we claim they satisfy existing regulatory requirements?

- ➢ It is challenging due to conceptual gaps between formal privacy models and concepts used in law and policy.
- ➢ There are at least two promising approaches to this problem:
  1. Extracting a formal, conservative model of the regulation to make a combined mathematical-legal formal claim that differential privacy satisfies a legal requirement.
  2. Interpreting the differential privacy guarantee in terms of the specific language of a relevant law or policy in order to argue that the use of differential privacy is sufficient to satisfy requirements, such as protecting personally identifiable information from disclosure.

Related Work

Kobbi Nissim, Aaron Bembenek, Alexandra Wood, Mark Bun, Marco Gaboardi, Urs Gasser, David R. O'Brien, and Salil Vadhan, Bridging the Gap between Computer Science and Legal Approaches to Privacy, 31 Harvard Journal of Law & Technology ___ (forthcoming 2018), https://privacytools.seas.harvard.edu/publications/bridging-gap-between-computer-science-and-legal-approaches-privacy.

Kobbi Nissim, Thomas Steinke, Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, David O'Brien, and Salil Vadhan, Differential Privacy: A Primer for a Non-technical Audience (Preliminary Version) (2017), https://privacytools.seas.harvard.edu/publications/differential-privacy-primer-non-technical-audience-preliminary-version.