ROUGH DRAFT — NOT FOR CIRCULATION

Bridging the Gap between Computer Science and Legal Approaches to Privacy

Kobbi Nissim[1,2], Aaron Bembenek[2], Alexandra Wood[3], Mark Bun[2], Marco Gaboardi[4], Urs Gasser[3], David R. O'Brien[3], Thomas Steinke[2], and Salil Vadhan[2]

[1]Dept. of Computer Science, Ben-Gurion University.
[2]Center for Research on Computation and Society, Harvard University.
{kobbi|tsteinke|mbun|salil}@seas.harvard.edu, bembenek@g.harvard.edu.
[3]Berkman Center for Internet & Society, Harvard University.
{awood|ugasser|dobrien}@cyber.law.harvard.edu.
[4]The State University of New York at Buffalo. gaboardi@buffalo.edu.

June 8, 2016

This work is the product of a working group of the *Privacy Tools for Sharing Research Data* project at Harvard University.[1] Kobbi Nissim led the group discussion. Aaron Bembenek, Kobbi Nissim, and Alexandra Wood were the lead authors of this document. Working group members Mark Bun, Marco Gaboardi, Urs Gasser, Kobbi Nissim, David O'Brien, Thomas Steinke, Salil Vadhan, and Alexandra Wood contributed to the conception of this work. We thank Ann Kristin Glenster, Deborah Hurley, Georgios Kellaris, Michel Reymond, Or Sheffet, and Olga Slobodyanyuk for their insightful comments.

Work supported by NSF grant CNS-1237235, grants from the Sloan Foundation, and a Simons Investigator grant to Salil Vadhan.

[1] Privacy Tools for Sharing Research Data, http://privacytools.seas.harvard.edu.

Abstract

The fields of law and computer science incorporate contrasting notions of the privacy risks associated with the analysis and release of statistical data about individuals and groups of individuals. Emerging concepts from the theoretical computer science literature provide formal mathematical models for quantifying and mitigating privacy risks, where the set of risks they take into account is much broader than the privacy risks contemplated by many privacy laws. An example of such a model is *differential privacy*, which provides a provable guarantee of privacy against a wide range of potential attacks, including types of attacks currently unknown or unforeseen. The subject of much theoretical investigation, new privacy technologies based on formal models have recently been making significant strides towards practical implementation. For these tools to be used with sensitive personal information, it is important to demonstrate that they satisfy relevant legal requirements for privacy protection. However, making such an argument is challenging due to the conceptual gaps between the legal and technical approaches to defining privacy. Notably, information privacy laws are generally subject to interpretation and some degree of flexibility, which creates uncertainty for the implementation of more formal approaches.

This Article articulates the gaps between legal and technical approaches to privacy and presents a methodology for rigorously arguing that a technological method for privacy protection satisfies the requirements of a particular law. The proposed methodology has two main components: (i) extraction of a formal mathematical requirement of privacy based on a legal standard found in an information privacy law, and (ii) construction of a rigorous mathematical proof for establishing that a technological privacy solution satisfies the mathematical requirement derived from the law. To handle ambiguities that can lead to different interpretations of a legal standard, the methodology takes a conservative "worst-case" approach and attempts to extract a mathematical requirement that is robust to potential ambiguities. Under this approach, the mathematical proof demonstrates that a technological method satisfies a broad range of reasonable interpretations of a legal standard. The Article demonstrates the application of the proposed methodology with an example bridging between the requirements of the Family Educational Rights and Privacy Act of 1974 and differential privacy.

Keywords: privacy, privacy regulations, information privacy law, FERPA, de-identification, formal privacy models, differential privacy

Contents

1 Introduction

The fields of law and computer science incorporate contrasting notions of the privacy risks associated with the analysis and release of statistical data about individuals and groups of individuals. Many information privacy laws adopt conceptions of risk that are narrowly defined and subject to interpretation, compared to the formal privacy definitions described by the recent computer science literature.[2] This general approach is intended to support flexibility in allowing organizations to implement a variety of specific privacy measures that are appropriate to their varying institutional contexts, adaptable to evolving best practices, and able to address a wide range of privacy-related harms. In practice, however, such flexibility in the interpretation and application of such standards also creates uncertainty as to threshold standards for privacy protection. This uncertainty may pose a barrier to the adoption of new technologies, and lead organizations to implement measures that fail to protect against a sufficient range of data privacy risks.[3]

Emerging concepts from computer science provide formal mathematical models for quantifying and mitigating privacy risks that differ signficantly from traditional approaches to privacy such as de-identification. This creates conceptual challenges for the interpretation and application of existing legal standards, many of which implicitly or explicitly adopt concepts based on a de-identification approach to privacy. An example a formal privacy model is *differential privacy*, which provides a provable guarantee of privacy against a wide range of potential attacks, including types of attacks currently unknown or unforeseen.[4] The subject of much theoretical investigation, new technical methods providing differential privacy have recently been making significant strides towards practical implementation. Several first-generation real-world implementations of differential privacy currently exist, and researchers in industry and academia are now building and testing more tools for differentially private statistical analysis.[5]

[2] For a previous discussion of the differences between legal and computer science definitions of privacy, see Felix T. Wu, *Defining Privacy and Utility in Data Sets*, 84 U. Colo. L. Rev. 1117 (2013). For a general introduction to a formal computer science definition of privacy called differential privacy, see Kobbi Nissim, Thomas Steinke, Alexandra Wood, Mark Bun, Marco Gaboardi, David O'Brien & Salil Vadhan, *Differential Privacy: An Introduction for Social Scientists*, Working Paper (forthcoming 2016); Cynthia Dwork, *A Firm Foundation for Private Data Analysis*, 54 COMM. ACM 86 (2011); Ori Heffetz & Katrina Ligett, *Privacy and Data-Based Research*, 28 J. Econ. Persp. 75 (2014); Erica Klarreich, *Privacy by the Numbers: A New Approach to Safeguarding Data*, Quanta Mag. (Dec. 10, 2012), https://www.quantamagazine.org/20121210-privacy-by-the-numbers-a-new-approach-to-safeguarding-data. Differential privacy was introduced in Cynthia Dwork, Frank McSherry, Kobbi Nissim & Adam Smith, *Calibrating Noise to Sensitivity in Private Data Analysis*, Proceedings of the 3rd Conference on the Theory of Cryptography 265 (2006).

[3] *See, e.g.,* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 U.C.L.A. L. Rev. 1701 (2010); Arvind Narayanan, Joanna Huey & Edward W. Felten, *A Precautionary Approach to Big Data Privacy* (Mar. 19, 2015), http://randomwalker.info/publications/precautionary.pdf.

[4] *See* Cynthia Dwork, Frank McSherry, Kobbi Nissim & Adam Smith, *Calibrating Noise to Sensitivity in Private Data Analysis*, Proceedings of the 3rd Conference on the Theory of Cryptography 265 (2006).

[5] *See* U.S. Census Bureau, OnTheMap, http://onthemap.ces.census.gov (last visited Jan. 22, 2016); Úlfar Erlingsson, Vasyl Pihur & Aleksandra Korolova, *RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response*, Proceedings of the 21st ACM Conference on Computer and Communications Security (2014); Andrew Eland, *Tackling Urban Mobility with Technology*, Google Europe Blog (Nov. 18, 2015), http://googlepolicyeurope.blogspot.com/2015/11/tackling-urban-mobility-with-technology.html; Privacy Tools for Sharing Research Data Project at Harvard University, Differentially Private Statistical Exploration, https://beta.dataverse.org/custom/DifferentialPrivacyPrototype (last visited Apr. 26, 2016); Microsoft Research, Privacy Integrated Queries (PINQ) Project, http://research.microsoft.com/en-us/projects/pinq (last visited Apr. 26, 2016); University of Pennsylvania, Putting Differential Privacy to Work Project, http://privacy.cis.upenn.edu/index.html (last visited Apr. 26, 2016); University of Texas, Airavat Project,

For these new technological tools to be used with sensitive personal information, it is important to demonstrate that they satisfy relevant legal requirements for privacy protection. However, making such an argument is challenging due to significant gaps between legal and mathematical approaches to defining privacy.[6] For instance, legal standards for privacy protection, and the definitions they employ, often vary according to industry sector, jurisdiction, institution, types of information involved, or other contextual factors.[7] Variations between laws create challenges for interpretation in the implementation of technological tools for privacy protection that are designed to be broadly-applicable. Legal approaches often, implicitly or explicitly, focus on a limited scope of attacks, such as re-identification by matching a named individual to a record in a database through linkage to information from other sources such as public records databases. This conceptualization of privacy risks, in part, leads many legal standards to turn on the presence of *personally identifiable information* in a data release. The concept of personally identifiable information is defined differently in various settings,[8] involves substantial ambiguity, and does not have a clear analog in mathematical definitions of privacy.[9]

In addition, standards and implementation guidance often emphasize techniques for protecting information released at the individual level but provide little guidance for releasing aggregate data, where the latter setting is particularly relevant to formal privacy models. In addition, limited or no guidance is available for cases in which personal information may be partly leaked, or where inferences about individuals can be made without full certainty. Mathematical models of privacy must make determinations within difficult gray areas, where the boundaries of cognizable legal harms may be uncertain. For these reasons, regulatory terminology and concepts do not directly apply to formal privacy models, creating challenges for interpretation.

1.1 Contributions of this Article

In this Article, we present a methodology for bridging between legal and mathematical notions of privacy. This methodology seeks to overcome various conceptual differences between legal standards and formal mathematical models that make it difficult to translate from one notion to another. The proposed methodology has two main components: (i) extraction of a formal mathematical requirement of privacy based on a legal standard found in an information privacy law, and (ii) construction of a rigorous mathematical proof for establishing that a technological privacy solution satisfies the mathematical requirement derived from the law.

As a lens for demonstrating the application of this methodology, this Article discusses two specific examples of diverging privacy concepts: differential privacy and the Family Educational Rights and Privacy Act of 1974 (FERPA),[10] a federal law that protects the privacy of education records in the United States. Two arguments are made along the way. The first is a legal argument supported by a technical argument: the FERPA standard for privacy is relevant to analyses computed with differential privacy. The second is a technical argument supported by a legal argument: differential

http://z.cs.utexas.edu/users/osa/airavat (last visited Apr. 26, 2016); Prashanth Mohan et al., *GUPT: Privacy Preserving Data Analysis Made Easy*, Proceedings of SIGMOD '12 (2012).

[6] For an extended discussion of this argument, see Section 3.3 below.

[7] *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. Rev. 1814 (2011).

[8] *See id.*

[9] *See* Arvind Narayanan & Vitaly Shmatikov, *Myths and Fallacies of "Personally Identifiable Information,"* 53 Communications of the ACM 24, 26 (2010).

[10] Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g; 34 C.F.R. Part 99 (2013).

privacy satisfies the FERPA standard for privacy. To address ambiguities that can lead to different interpretations of the legal standard, the analysis takes a conservative, worst-case approach and extracts a mathematical requirement that is robust to potential ambiguities. Under this approach, the mathematical proof of privacy then demonstrates that differential privacy satisfies a large class of reasonable interpretations of the FERPA privacy standard.

While FERPA and differential privacy are used to illustrate the methodology, we believe it is a general approach that can be extended to bridge between technologies beyond differential privacy and privacy laws beyond FERPA. The degree of rigor employed enables us to make strong arguments about the privacy requirements of statutes and regulations. With the level of generalization afforded by this rigorous, conservative approach to modeling, differences between sector- and institution-specific standards are blurred, making the methodology broadly-applicable.

In this way, the methodology can help support the future adoption of emerging privacy-preserving technologies. Arguments that are rigorous from both a legal and technical standpoint can assure actors who release data that they are complying with the law, and also better inform data subjects of the privacy protection to which they are legally entitled. Furthermore, the process of formalizing privacy statutes and regulations can lead us to better understand those texts and help us identify areas of the law that seem ambiguous or insufficient. This approach can serve as a foundation for future extensions to address other problems in information privacy, an area that is both highly technical and highly regulated and therefore is well-suited to combined legal-technical solutions.[11]

1.2 Article structure

In the sections that follow, we establish various elements of a formal methodology for bridging between privacy concepts from the field of mathematics and the law, and provide examples demonstrating how the methodology can be applied in practice. Section 2 describes the setting in which the privacy issues relevant to this discussion arise, and provides a high-level overview of the computer science and legal approaches to privacy that have emerged to address these issues. Section 3 provides an introduction to two specific privacy concepts, differential privacy and FERPA, which are used as concrete examples for comparison throughout the paper. It also discusses the applicability of FERPA's requirements to differentially private computations and articulates the gaps between differential privacy and FERPA which create challenges for implementing differential privacy in practice. The later sections present a novel methodology for formally proving that a technological method for privacy protection satisfies the requirements of a particular privacy law.

Section 4 describes the process of extracting a formal mathematical requirement of privacy for FERPA. It uses the proposed methodology to construct a model of the attacker implicitly contemplated by FERPA, based on the text of the definitions found in FERPA and in documentation of the regulation's history. Section 5 provides a non-technical discussion of the mathematical proof to demonstrate that differential privacy satisfies the mathematical definition of privacy extracted from FERPA. The full proofs appear in the Appendix. Section 6 provides the general methodology

[11] For other research formalizing legal requirements for privacy protection using mathematical approaches, see, e.g., Omar Chowdhury, Andreas Gampe, Jianwei Niu, Jeffery von Ronne, Jared Bennatt, Anupam Datta & Limin Jia, *Privacy Promises That Can Be Kept: A Policy Analysis Method with Application to the HIPAA Privacy Rule*, Proceedings of the 18th ACM Symposium on Access Control Models and Technologies 3 (2013); Henry DeYoung, Deepak Garg, Dilsun Kaynar & Anupam Datta, *Logical Specification of the GLBA and HIPAA Privacy Laws*, Proceedings of 9th ACM Workshop on Privacy in the Electronic Society (2010).

for bridging between legal and technical approaches to privacy. Section 7 provides a concluding discussion describing how this new methodology can help further the real-world implementation of emerging formal privacy technologies, as well as the development of more robust privacy regulations. The Article concludes with a discussion of the policy implications of this approach, including gaps revealed in the privacy definitions used in FERPA and similar laws, as well as ways in which statutory definitions could be revised to incorporate a scientific understanding of privacy.

2 Background

Privacy is conceptualized very differently across a range of contexts, from surveillance to criminal procedure to public records releases to research ethics to medical decision making.[12] Because privacy law is quite broad in its reach and privacy measures can be designed to address a wide range of harms, it is important to define the scope of the analysis in this Article. Specifically, this Article focuses on a context related to the statistical analysis of data or the release of statistics derived from personal data. We refer to the relevant setting as *privacy in statistical computation*.

2.1 The setting: Privacy in statistical computation

Many government agencies, commercial entities, and research organizations collect data about individuals and groups of individuals. These entities frequently release data, or statistics based on the data, that they have collected, processed, and analyzed. If the data contain private information, law or policy likely restricts the degree to which the data can be released, including the formats that the release can take.

Federal and state statistical agencies, such as the Census Bureau, the Bureau of Labor Statistics, and the National Center for Education Statistics, release large quantities of statistics about individuals, households, and establishments, upon which policy and business investment decisions are based. For instance, the National Center for Education Statistics collects data from U.S. schools and colleges using surveys and drawing from administrative records and releases statistics to the public.[13] These statistics include total enrollment at public and private schools by grade level, class sizes, participation in activities such reading at home, access to computers and technology, standardized test scores by state and student demographics, and high school graduation rates, among other figures, and are used in shaping education policy and school practices.[14] The release of such statistics by federal statistical agencies is subject to strict rules for protecting the privacy of the

[12] For a broad survey of privacy concepts, see, e.g., Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. Penn. L. Rev. 477 (2006) (grouping privacy problems into categories such as surveillance, interrogation, aggregation, identification, insecurity, secondary use, exclusion, breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, distortion, intrusion, and decisional interference); Bert-Jaap Koops, Bryce Clayton Newell, Tjerk Timan, Ivan Škorvánek, Tomislav Chokrevski, & Maša Galič, *A Typology of Privacy*, 38 U. Pa. J. Int'l L. (forthcoming 2016) (surveying constitutional protections and privacy scholarship from nine North American and European countries, resulting in a classification of privacy based on "eight basic types of privacy (bodily, intellectual, spatial, decisional, communicational, associational, proprietary, and behavioral privacy), with an overlay of a ninth type (informational privacy) that overlaps, but does not coincide, with the eight basic types"); Alan F. Westin, Privacy and Freedom (1967) (identifying four states of privacy: solitude, intimacy, anonymity, and reserve).

[13] *See* National Center for Education Statistics, Digest of Education Statistics: 2013, NCES Pub. No. 2015-011 (2015).

[14] *See id.*

individuals in the data.[15]

In the commercial context, companies such as Google and Facebook collect personal information and use it to provide services to individual users and advertising customers. For instance, Facebook enables advertisers to target their ads on the Facebook platform by choosing the locations, demographics, interests, and behaviors of their target audience, and it provides an estimate of the potential number of users who would see the ad given the selections made.[16] The commercial collection, use, and release of data is regulated by the Federal Trade Commission,[17] as well as relevant sector-specific information privacy laws.[18]

Researchers and their institutions also release statistics to the public and make their data available to other researchers. These research activities are subject to oversight by an institutional review board in accordance with the Federal Policy for the Protection of Human Subjects.[19] Researchers are required to obtain consent from participants and to employ data privacy safeguards when collecting, storing, and sharing data about individuals.[20]

In this Article, we explore how statistical computations can be performed by government agencies, commercial entities, and researchers while providing privacy to individuals in the data. Additionally, we propose a method for demonstrating how a technology that preserves privacy in computation can be shown to satisfy legal requirements for privacy protection.

2.1.1 What is a computation?

A *computation* (alternatively referred to as an algorithm, mechanism, or analysis) is a procedure for producing an output given some input data, as illustrated in Figure 1.[21] This very general definition does not restrict the nature of the relationship between the input data and the output. For instance, a computation might output its input without any transformation, or a computation might completely ignore its input by producing an output that is independent of the input. Some computations are *deterministic* and others are *randomized*. A deterministic computation will always produce the same output given the same input, while a randomized computation does not have this guarantee. A computation that returns the mean age of participants in a dataset is an example of a deterministic computation. However, a similar computation that estimates the mean age in the dataset by sampling at random a small subset of the dataset records and returning the mean age

[15] *See* Confidential Information Protection and Statistical Efficiency Act, 44 U.S.C. § 3501 note (prohibiting the release of statistics in "identifiable form" and establishing specific confidentiality procedures to be followed by employees and contractors when collecting, processing, analyzing, and releasing data).

[16] *See* Facebook, Facebook for Business: How to target Facebook ads, https://www.facebook.com/business/a/online-sales/ad-targeting-details (last visited Apr. 27, 2016). These audience-reach statistics are apparently rounded, in part, to protect the privacy of Facebook users. *See* Andrew Chin & Anne Klinefelter, *Differential Privacy as a Response to the Reidentification Threat: The Facebook Advertiser Case Study*, 90 N. Carolina L. Rev. 1417 (2012).

[17] *See* Federal Trade Commission, Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers, FTC Report (2012).

[18] *See, e.g.,* Children's Online Privacy Protection Act, 15 U.S.C. §§ 6501–6506 (requiring certain web site operators to provide notice and obtain consent when collecting personal information from children under 13 years of age); Cal. Civ. Code § 1798.82 (requiring businesses to disclose any data breach to California residents whose unencrypted personal information was acquired by an unauthorized person).

[19] *See* 45 C.F.R. Part 46.

[20] 45 C.F.R. § 46.111.

[21] This figure is reproduced from Kobbi Nissim, Thomas Steinke, Alexandra Wood, Mark Bun, Marco Gaboardi, David O'Brien & Salil Vadhan, *Differential Privacy: An Introduction for Social Scientists*, Working Paper (forthcoming 2016), in which an extended discussion of privacy in computation can also be found.

computed for these records is a randomized computation. As another example, a computation that outputs the mean age with the addition of some random noise is a randomized computation.

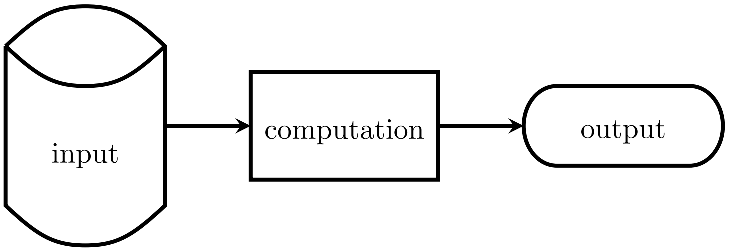


Figure 1: A computation produces an output given some input data.

Computations are used to create public statistics about private information. For instance, many government agencies use computations to generate and release statistical information from private datasets. The U.S. Census Bureau collects information about individual households, including sensitive personal information, and then releases statistics about the larger population. An example is the Census Bureau's use of computations to produce the median household income of a city or region after collecting household-level income information from a sample population in that area. Similarly, a school district, seeking to study and report on the performance of students in its schools, applies a set of computations to the raw data it has collected about individual students to generate more general statistics for publication. Releasing statistics about private data can benefit society by enabling research and informing policy decisions. However, the computations used to create the statistics also need to provide sufficient privacy protection for the participants in the original, private datasets. In the next subsection, we discuss how this is possible.

2.1.2 Privacy as a property of computation

In this Article, we consider whether a computation provides privacy protection to be a property of that computation. To see why this approach makes sense from an informational point of view, note that risks to privacy occur when the output of a computation carries information about its input. Consider, for example, a computation that receives as input a student's transcript and outputs the student's GPA. Seeing that the outcome of this computation is a GPA of 3.7, an observer can rule out certain transcripts for the student that would result in a different GPA. In some cases, the observer would be able to uniquely determine the student's grades (e.g., when the GPA is zero, or when the observer has prior information that, together with the computed GPA, uniquely determines the grades).[22]

When we make the claim that a certain computation provides privacy protection, we mean that the *informational relationship between input and output* of this computation satisfies the requirements of a particular definition of privacy. We emphasize that it is the computation that is private and not a particular output that is private. To see why distinguishing between a supposedly *private* or *non-private* output fails to capture a reasonable notion of privacy, consider a policy that states that statistics needs to be coarsened to the nearest ten (e.g., 0–9, 10–19, 20–29, etc.), with the

[22] While this example is using a deterministic computation of the GPA for simplicity, it can be also generalized to randomized computations.

hope that such a coarsening would hide the effect of individuals. Suppose a school releases statistics for the fall semester showing that 20–29 of its students have a disability, an output that seems innocuous. In the spring, it releases updated statistics showing that 30–39 of its students have a disability, again, an output that seems innocuous. However, reasoning about how the two statistics were computed reveals that the new student who joined between the fall and spring semesters has a disability. Although both *outputs* seem innocuous as each output only reveals aggregate statistics and does not directly identify any individual student, reasoning about how they *depend on the input data*—a dependency created by the computation of the statistics—reveals sensitive information.

We argue that defining privacy as a property of a computation makes sense from both a computer science and legal perspective. Computer scientists seek to reason about the properties of computations, and treating privacy as a computational property fits naturally in this world view. This approach has successful precedents in established areas of computer science such as cryptography.[23] We argue that this approach is also applicable to regulatory notions of privacy. While legal texts might not explicitly refer to computation, we observe that they often attempt to implicitly define privacy as a property that certain computations possess. For instance, consider the safe harbor de-identification standard found in the HIPAA Privacy Rule, which allows the disclosure of data from which certain pieces of information deemed to be identifying have been removed.[24] This provision is in effect specifying a computation; that is, a computation that produces an output in which the identifiers specified by the safe harbor provision have been redacted is considered to provide sufficient privacy protection. Similarly, regulatory requirements or guidance prescribing minimum cell counts for aggregate data tables produced using personal information,[25] or recommending applying k-anonymization techniques,[26] are effectively treating certain computations as providing privacy.[27]

A benefit of viewing privacy as a property of computation is that computations are formal mathematical objects, and as such can be reasoned about with a high degree of mathematical

[23] For instance, a cryptographic computation might be considered to provide privacy protection if any hypothetical adversary given an encrypted message can do no better at guessing a property of that message than another hypothetical adversary that is not given the encrypted message. *See* Shafi Goldwasser & Silvio Micali, *Probabilistic Encryption and How to Play Mental Poker Keeping Secret All Partial Information*, Proceedings of ACM Symposium on Theory of Computing 36 (1982).

[24] 45 C.F.R. § 164.514

[25] For example, in accordance with the Elementary and Secondary Education Act of 1965, states must define minimum cell counts for the publication of student achievement results "[b]ased on sound statistical methodology" that "[y]ields statistically reliable information for each purpose for which disaggregated data are used" and does not "reveal personally identifiable information about an individual student." 34 C.F.R. § 200.7.

[26] "A release provides k-anonymity protection if the information for each person contained in the release cannot be distinguished from at least $k - 1$ individuals whose information also appears in the release." *See* Latanya Sweeney, *k-Anonymity: A Model for Protecting Privacy*, 10 International Journal on Uncertainty, Fuzziness and Knowledge-based Systems 557 (2002). Guidance from the Department of Health and Human Services covers the application of k-anonymity as one approach to protecting health records subject to the HIPAA Privacy Rule. *See* Office for Civil Rights, Department of Health and Human Services, *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule* (2012) (providing guidance on applying k-anonymity and noting that "[t]he value for k should be set at a level that is appropriate to mitigate risk of identification by the anticipated recipient of the data set," while declining to "designate a universal value for k that covered entities should apply to protect health information in accordance with the de-identification standard").

[27] Note that by specifying properties of the output, such requirements restrict the computation, but do not necessarily successfully restrict the informational relationship between input and output. We include these examples here only to support the claim that conceptualizing the problem in terms of restrictions on computation is compatible with some existing legal and regulatory approaches to privacy protection.

rigor. In this Article we attempt to use this view to make rigorous arguments about the privacy requirements of laws and regulations. In turn, these arguments can assure actors who release data that they are following the law, and may give data subjects a better understanding of the privacy to which they are legally entitled. Additionally, the process itself of formalizing legal requirements as computational objects can lead us to better understand those requirements and help identify areas of the law that seem ambiguous or potentially insufficient.

2.1.3 Privacy risks in computation

A number of approaches have been developed and widely implemented to limit the disclosure of personal information when sharing statistical data about individuals. Traditional approaches include obtaining consent from data subjects, entering into data use agreements restricting the use and re-disclosure of data, and applying various techniques for de-identifying data prior to release.[28] Statistics about individuals or groups of individuals are generally made available after de-identification techniques have transformed the data, by removing, generalizing, aggregating, and adding noise to pieces of information deemed to be identifiable. At the core of this approach is a concept of *personally identifiable information* (PII), which is based in a belief that privacy risks lurk in tables of individual-level information, and that the removal of PII, such as names, addresses, and Social Security numbers, provides adequate privacy protection.[29]

Releases of statistical data about individuals are under increasing scrutiny. Frequent data privacy breaches have demonstrated that privacy risks can be discovered even in releases of data that have been redacted of PII. For example, in the late 1990s, Latanya Sweeney famously demonstrated that the medical record of Massachusetts Governor William Weld could be identified in a release of data on state employee hospital visits that had been stripped of the names and addresses of hospital patients in order to protect their privacy.[30] Using Governor Weld's date of birth, ZIP code, and gender, which could be found within public records, she was able to locate his record in the released data, as it was the only record that matched all three attributes. Indeed, well over 50% of the U.S. population can be uniquely identified using these three pieces of information.[31]

[28] *See, e.g.,* Department of Education, Data-Sharing Tool Kit for Communities: How to Leverage Community Relationships While Protecting Student Privacy (March 2016), http://www2.ed.gov/programs/promiseneighborhoods/datasharingtool.pdf (providing best practice guidance on sharing education data while protecting privacy, by de-identifying data, obtaining written consent, or entering into a written data-sharing agreement with the recipient).

[29] Many privacy laws explicitly or implicitly endorse the practice of removing personal information considered to be identifying prior to release. *See, e.g,* Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g; 34 C.F.R. Part 99 (permitting educational agencies and institutions to broadly release, without the consent of students or their parents, information from education records that have been de-identified through the removal of "personally identifiable information"); Health Insurance Portability and Accountability Act (HIPAA) Privacy Act, 45 C.F.R. Part 160 and Subparts A and E of Part 164 (providing a safe harbor permitting the release of health information that has been de-identified through the removal of information from a list of eighteen "identifiers"). For an extended discussion of various legal approaches to de-identification, see Section 2.3 below. Such approaches also appear in a wide range of guidance materials on privacy and confidentiality in data management. *See, e.g.,* Federal Committee on Statistical Methodology, Report on Statistical Disclosure Limitation Methodology, Statistical Policy Working Paper 22 (2005), http://www.hhs.gov/sites/default/files/spwp22.pdf.

[30] *See* Latanya Sweeney, *Weaving Technology and Policy Together to Maintain Confidentiality*, 25 J. L., Med., & Ethics 98 (1997).

[31] *See* Latanya Sweeney, *Simple Demographics Often Identify People Uniquely*, Data Privacy Lab Technical Report (2000); Philippe Golle, *Revisiting the uniqueness of simple demographics in the US population*, Proceedings of the 2006 ACM Workshop on Privacy in the Electronic Society (WPES) 77 (2006).

Repeated demonstrations across many types of data have confirmed that this type of privacy breach is not merely anecdotal but is in fact widespread. For instance, it has been shown that individuals can be identified in releases of Netflix viewing records and AOL search query histories, despite the companies' efforts to remove identifying information from the data prior to release.[32] Other research has demonstrated that just four data points providing an individual's location at a point in time can be sufficient to identify 95% of individuals in mobile phone data and 90% of individuals in credit card purchase data.[33] More generally, it is now understood that "[a]ny information that distinguishes one person from another can be used for re-identifying anonymous data."[34]

Privacy attacks have also exposed vulnerabilities in releases of aggregate data, as well as inference risks that are distinct from the risk of re-identification.[35] Many successful attacks have focused not on discovering the identities of individuals but rather on learning or inferring sensitive details about them.[36] For example, researchers have discovered privacy risks in databases containing information about mixtures of genomic DNA from hundreds of people.[37] Although the data were believed to be sufficiently aggregated so as to pose little risk to the privacy of individuals, it was shown that an individual's participation in a study could be confirmed using the data, thereby revealing that an individual suffers from the medical condition being studied. In another demonstration, vulnerabilities were uncovered in the Israel Central Bureau of Statistics' web-based portal, which allowed the public to make queries of aggregate statistics based on information from an anonymized survey.[38] Researchers were able to retrieve the records of more than one thousand individuals through queries to the system; moreover, they showed it was possible to link the records to identifiable individuals.[39] Privacy risks have also been uncovered in online recommendation systems used by web sites such as Amazon, Netflix, and Last.fm, which employ algorithms to recommend similar products based on an analysis of data generated by the behavior of millions of users.[40] In these attacks, it has been revealed that recommendation systems can leak

[32] *See* Arvind Narayanan & Vitaly Shmatikov, *Robust De-anonymization of Large Sparse Datasets*, Proceedings of the 2008 IEEE Symposium on Research in Security and Privacy 111 (2008); Michael Barbaro & Tom Zeller, Jr., *A Face Is Exposed for AOL Searcher No. 4417749*, N.Y. Times, Aug. 9, 2006.

[33] *See* Yves-Alexandre de Montjoye et al., *Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata*, 347 Science 536 (2015); Yves-Alexandre de Montjoye et al., *Unique in the Crowd: The Privacy Bounds of Human Mobility*, 3 Nature Sci. Reps. 1376 (2013).

[34] Arvind Narayanan & Vitaly Shmatikov, *Myths and Fallacies of "Personally Identifiable Information,"* 53 Communications of the ACM 24, 26 (2010).

[35] For a recent survey of different classes of privacy attacks, including both re-identification attacks and tracing attacks (i.e., attacks that aim to determine whether information about a target individual is in a database), see the discussion in Cynthia Dwork, Adam Smith, Thomas Steinke & Jonathan Ullman, *Hiding in Plain Sight: A Survey of Attacks on Private Data* (forthcoming 2017).

[36] For example, Dinur and Nissim showed that publishing just statistical estimates can lead to a massive leakage of individual information. *See* Irit Dinur & Kobbi Nissim, *Revealing Information While Preserving Privacy*, Proceedings of the 22nd ACM PODS 202 (2003).

[37] *See* Nils Homer et al., *Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-density SNP Genotyping Microarrays*, 4 PLoS Genetics 8 (2008). This attack has been strengthened and generalized in several works, such as Cynthia Dwork et al., *Robust Traceability from Trace Amounts*, IEEE Symposium on Foundations of Computer Science (2015), http://privacytools.seas.harvard.edu/files/robust.pdf.

[38] *See* Amitai Ziv, *Israel's "Anonymous" Statistics Surveys Aren't So Anonymous*, Haaretz (Jan. 7, 2013), http://www.haaretz.com/news/national/israel-s-anonymous-statistics-surveys-aren-t-so-anonymous-1.492256.

[39] *See id.*

[40] *See* Joseph A. Calandrino et al., *"You Might Also Like:" Privacy Risks of Collaborative Filtering*, IEEE Symposium on Security and Privacy (2011).

information about the transactions made by individuals. In light of these findings, privacy risks should be taken into account when releasing not just individual-level records but also aggregate statistics, and a conception of privacy risks should take into account inference risks in addition to re-identification risks.

Successful privacy attacks demonstrate how techniques for learning about individuals in a data release are rapidly advancing and exposing vulnerabilities in many commonly used measures for protecting privacy. Although traditional approaches to privacy—including those that are currently hard-wired into privacy laws—may have been sufficient at the time they were developed, they are becoming less and less suitable for the protection of information in the Internet age. It is likely that privacy risks will continue to evolve and even increase over time, enabled by rapid advances in analytical capabilities and the growing availability of personal data from different sources, and motivated by the potential misuses of the revealed data. In addition, the application of de-identification techniques produces data of reduced quality.[41] These realizations have led computer scientists in search of approaches to data privacy that will be robust against a wide range of attacks, including types of attacks that are currently unknown.

2.2 An introduction to the computer science approach to defining privacy

In computer science, privacy is often formalized as a game in which an adversary attempts to exploit a computational system to learn protected information. A system is considered to provide privacy protection if it can be demonstrated, via a mathematical proof, that no adversary can win the game with a probability that is "too high." Every aspect of a game framework must be carefully and formally defined, including any constraints on the adversary, the mechanics of the game, what it means for the adversary to win the game, and with what probability the adversary can win the game. This degree of formalization allows us to prove that a given system is private up to the explicit assumptions of the model.

The formal privacy games we will be considering have the following components: an adversary, a computation, and the game mechanics. Each of these components is discussed in turn below.

An *adversary* attempts to exploit the system to learn private information. The adversary is not defined in terms of how he attempts to do so, but rather by the computational power, access to the system, and background knowledge he can bring. Consequently, the adversary does not represent a uniquely specified attack, but rather a whole class of attacks, *including attacks that have not been conceived by the system designer*. This means that a system cannot be tested for its privacy, as testing its resilience to known attacks would not rule out its vulnerability to other attacks. Rather, privacy needs to be *proved* mathematically. By proving that the system provides privacy against such an adversary, we have made a strong claim: the system is private no matter the particular attack or attacker, provided that the assumptions we have made in our model are not violated.

A *computation* takes as input a dataset (potentially including private information) and produces some output. For instance, one could envision a computation that takes as input a spreadsheet of student grades and returns the average grade (or an approximation of it). Unlike the adversary, the

[41] *See, e.g.,* Jon P. Daries et al., *Privacy, Anonymity, and Big Data in the Social Sciences*, 12 ACM Queue (2014), http://queue.acm.org/detail.cfm?id=2661641 (discussing the significantly different results in analyses based on massive open online course data before and after applying de-identification techniques: "For example, the original analysis found that approximately 5 percent of course registrants earned certificates. Some methods of de-identification cut that percentage in half.").

computation needs to be uniquely specified[42] and, furthermore, known to the adversary.[43] We use the game framework to prove that a computation (or a class of computations) provides privacy given the assumptions of the game. For example, we might prove that in a particular game framework the computation reports the average grade in such a way as to maintain the individual privacy of the students in the original dataset.

The *game mechanics* acts as an intermediary between the adversary and the private data.[44] The adversary does not have direct access to non-public data, and instead receives information via the game mechanics. The game mechanics enforce a specific protocol and determines the ways in which the adversary can interact with the computation. In addition, the game mechanics defines when the adversary is deemed to have won the game and when a system is deemed to provide a sufficient level of privacy. For instance, the game mechanics might specify the following protocol: the previously described computation is used to calculate the average grade on a test, and then the computation result is released to the adversary. The adversary responds with a guess of the score of a particular student, and is considered to have won the game if his guess is within a certain range of that student's true grade.

A privacy game provides us with a privacy definition: A computation satisfies the definition of privacy if no adversary can win the game (with the computation) "too often." Note that we do not require that an adversary must never win the game. Such a requirement, even if intuitive, would not be achievable, as there is always some probability that an adversary could win the game by chance, and this could also be the case without the adversary getting access to an outcome of the computation. In other words, the standard is that no adversary should be able to win the game with a probability that is significantly greater than some baseline probability, which we would generally take to be the probability of winning without access to the computation outcome.

To illustrate, consider a game in which an adversary's goal is to guess a person's gender. With two possible values for gender (male and female), it would be possible for the adversary to guess the gender correctly with the probability of 50%, even without obtaining any information about that person. This sets, as the baseline, a probability of 50% for guessing the target individual's gender, as it would not be reasonable to establish a requirement that all adversaries must guess correctly with a success rate that is lower than this probability. Additionally, we typically allow for adversaries to win with a probability that is slightly higher than the baseline value (because any system that provides utility necessarily leaks at least a tiny bit of information). How much the adversary is allowed to win beyond the baseline probability while the computation is still considered to satisfy the privacy definition is often quantified as a parameter that can be tuned to provide less or more privacy protection.[45]

For a more detailed explanation, consider the following privacy game from cryptography. Alice wants to send a private message to Bob, but she is worried that a third party, Eve, might be eavesdropping on their communications. Therefore, Alice decides to encrypt the message before sending it to Bob. She wants to be confident that the computation she uses to encrypt the message

[42] When we use the game framework to prove that a class of computations provides privacy, the proof assumes that an arbitrary computation of this class is selected as the uniquely specified computation before the execution of the game begins.

[43]

[44] Note that the game mechanics does not necessarily correspond to any specific "real world" entity. A privacy game is a thought experiment and in general there is not a direct correspondence between the components of a privacy game and the parties in a real-world privacy scenario.

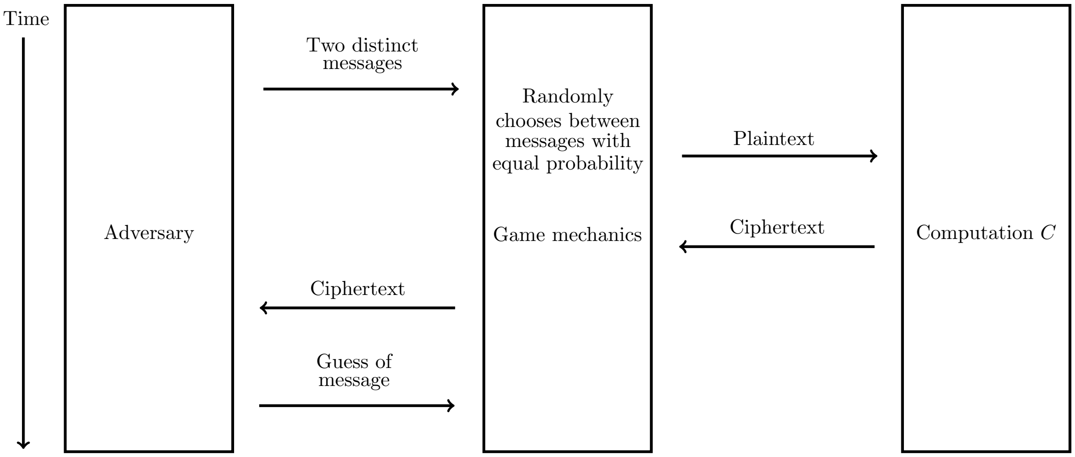
[45] See Section 4 below for further discussion.

will ensure that Eve cannot learn much about the content of the original message (i.e., the plaintext) from seeing the encrypted version of the message (i.e., the ciphertext). To gain confidence in the security of the system, Alice can formalize her privacy desiderata as a game and then use an encryption computation that is proven to meet the game's definition of privacy. Here is one possible mechanics for the game, which are also represented visually in Figure 2:[46]

1. An adversary Eve chooses two distinct plaintext messages and passes them to the game mechanics. Intuitively, she is asserting that she can distinguish between the encryptions of these messages and hence the encryption is insecure.

2. The game mechanics tosses a fair coin to choose between the messages with equal probability, encrypts the chosen message (denoted "plaintext" in Figure 2) with a computation C, and gives the resulting ciphertext to the adversary.

3. The adversary wins if she is able to guess from seeing the ciphertext which of the two original messages was encrypted.



Adversary wins if the guess equals the plaintext message used during the encryption computation.

Figure 2: Example cryptographic privacy game.

Notice that an adversary that ignores the ciphertext she is given in Step 2 of the game and simply outputs one of the messages she selected in Step 1 already has a 50% chance of winning. This means that any reasonable adversary would have a winning probability of at least 50%, and we take the 50% success rate as a baseline for comparison. Therefore, a computation C for which no adversary can win the game with a probability greater than 50%, equal to the baseline, can be considered as providing perfect privacy protection. We typically relax this requirement to allow adversaries to obtain a small advantage over the 50% baseline. That is, a computation C would be considered to satisfy the privacy definition if no adversary can win this game with a probability significantly higher than 50%. The cryptographic standard for encryption is to only

[46] This game is modeled after the notion of indistinguishability of ciphertexts introduced in Shafi Goldwasser & Silvio Micali, *Probabilistic Encryption and How to Play Mental Poker Keeping Secret All Partial Information*, Proceedings of ACM Symposium on Theory of Computing 36 (1982).

allow adversaries a negligible advantage over the 50% probability, say a 50.00000000001% chance of winning the game.[47] Now that her privacy desiderata have been formalized, Alice can use any encryption computation that is mathematically proven to meet this definition of privacy to send an encrypted message to Bob with confidence that Eve cannot learn much from eavesdropping on their communication.

Although privacy laws are not written with an explicit game-based definition of privacy, we argue that it is possible to extract a suitable privacy game from a law, its legislative history, and corresponding commentaries and guidance. Furthermore, the privacy game that is extracted can be used to establish that particular computations meet the privacy requirements of the law. In Section 4 below, we extract such a privacy game from FERPA, and based on this game we sketch in Section 5 a proof showing that all differentially private computations provide sufficient privacy protection to satisfy the requirements of FERPA.

2.3 An introduction to legal approaches to privacy

Information privacy laws around the world vary substantially with respect to their scope of coverage and the protections they afford.[48] Hence, this discussion focuses on a subset of laws that restrict the release of statistical information about individuals or groups of individuals, whether released as raw data, de-identified data, or statistical summaries. The applicability of such laws typically turns on the definition of terminology such as personal information, personal data, personally identifiable information, or a similar term.[49] If the information to be released falls within a particular law's definition of personal information, then the law typically applies and restricts the disclosure of the information.[50] If it does not meet the particular law's definition of personal information, then the information is often afforded no or minimal protection under that law.[51] In addition, some laws expressly exclude a category of *de-identified information*. Information that has been transformed such that it satisfies the law's de-identification standard can be shared under relaxed conditions.

[47] The difference of 0.00000000001% between the baseline of 50% and the set threshold of 50.00000000001% determines the adversary's benefit-cost tradeoff. For example, if Eve's goal is to differentiate between the two messages ATTACK AT DAWN and ATTACK AT SUNSET, then she would have to accumulate a number of encrypted messages that is inversely proportional to this difference (i.e., in the order of magnitude of 10^{12} messages). The allowed difference may also affect the efficiency of the cryptographic scheme, for instance it can affect Alice's computational costs. Exactly how much higher the adversary is permitted to go above the baseline probability is often captured as a parameter that can be tuned to different privacy (and computational cost) levels. For instance, if Alice feels that the message she is sending is not especially sensitive, she might choose that it is acceptable for an adversary to win the game with a probability of 50.000001%.

[48] For a broad survey and classification of privacy laws across many jurisdictions, see Bert-Jaap Koops, Bryce Clayton Newell, Tjerk Timan, Ivan Škorvánek, Tomislav Chokrevski, & Maša Galič, *A Typology of Privacy*, 38 U. PA. J. INT'L L. (forthcoming 2016) (surveying constitutional protections and privacy scholarship from nine North American and European countries).

[49] For an overview of various definitions of personal information, see *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814 (2011).

[50] *See, e.g.,* Children's Online Privacy Protection Act, 15 U.S.C. § 6502(a)(1) (providing that "[i]t is unlawful for an operator of a website or online service directed to children, or any operator that has actual knowledge that it is collecting personal information from a child, to collect personal information from a child in a manner that violates the regulations . . ."

[51] *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1816 (2011) (finding that many laws "share the same basic assumption—that in the absence of PII, there is no privacy harm. Thus, privacy regulation focuses on the collection, use, and disclosure of PII, and leaves non-PII unregulated.").

In some cases, de-identified information can even be released publicly without further restriction on use or redistribution.[52]

Definitions of personal information vary considerably, and the inconsistency of the often narrow definitions is widely cited as a weakness of the legal framework for privacy protection.[53] It is beyond the scope of this Article to detail all legal approaches to privacy and all of the regulatory definitions of personal information currently in place around the world. Instead, we provide an overview of selected approaches in order to illustrate a range of different approaches and definitions, and some of the challenges that have arisen in developing, interpreting, and complying with various regulatory definitions and standards.

2.3.1 Selected approaches from the United States

Privacy law in the United States takes a sectoral approach, in which many privacy laws are in place, each of which is rather narrowly defined.[54] For illustration, consider the Video Privacy Protection Act, which protects the privacy of individuals in records of video sales and rentals, defines *personally identifiable information* (PII) as "information which identifies a person as having requested or obtained specific video materials or services from a video tape service provider."[55] In contrast, the California Confidentiality of Medical Information Act defines *medical information* as "individually identifiable health information about a patient's medical history, mental or physical condition, or treatment" which "must include an element that identifies a person, such as name, address, email address, telephone number, or Social Security number, or that can be combined with other publicly available information to reveal a person's identity."[56] Many laws in the US adopt some variation of a binary definition that depends on whether the information "identifies a person."[57] In practice, there has been substantial uncertainty regarding the application of such a standard.[58]

In contrast to laws whose scope is defined by the presence of information that "identifies a person," some laws aim to provide a bright-line standard, by setting forth an exhaustive list of the types of information that the law protects. An example is the Massachusetts data protection regulation, which defines *personal information* with an exhaustive list: "a Massachusetts resident's first name and last name or first initial and last name in combination with any one or more of the following data elements that relate to such resident: (a) Social Security number; (b) driver's license

[52] *See, e.g.,* HIPAA Privacy Rule, 45 C.F.R. § 164.502(d)(2) (providing that"[h]ealth information that meets the standard and implementation specifications for de-identification . . . is considered not to be individually identifiable health information, i.e., de-identified. The requirements of this subpart do not apply to information that has been de-identified in accordance with the applicable requirements . . .").

[53] *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. Rev. 1814, 1893 (2011) (concluding that "there is no uniform definition of PII in information privacy law. Moreover, the definitions that do exist are unsatisfactory.").

[54] For a discussion of the evolution and nature of the US sectoral approach to privacy, see Paul M. Schwartz, *Preemption and Privacy*, 118 Yale L.J. 902 (2008).

[55] 18 U.S.C. § 2710(a)(3) (emphasis added).

[56] Cal. Civ. Code §§ 56–56.37.

[57] *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. Rev. 1814 (2011).

[58] *See, e.g.,* Pineda v. Williams-Sonoma Stores, 246 P.3d 612, 612 (Cal. 2011) (reversing the lower courts and determining that a "cardholder's ZIP code, without more, constitutes personal identification information" within the meaning of the California Song-Beverly Credit Card Act of 1971 "in light of the statutory language, as well as the legislative history and evident purpose of the statute").

number or state-issued identification card number; or (c) financial account number, or credit or debit card number, with or without any required security code, access code, personal identification number or password, that would permit access to a resident's financial account" but notes that it does not include "information that is lawfully obtained from publicly available information, or from federal, state or local government records lawfully made available to the general public."[59] As another example, the Health Insurance Portability and Accountability Act Privacy Rule[60] provides a safe harbor, by which data can be shared widely once all information from a list of eighteen categories of information have been removed.[61]

Beyond a small subset of such laws that attempt to take a bright-line approach to defining *personally identifiable information*, privacy laws in the US generally employ standards that require case-by-case determinations that rely on some degree of interpretation. These determinations are complicated by advances in analytical capabilities, the increased availability of data about individuals, and developments in the scientific understanding of privacy. These developments, in combination with limited guidance on interpreting and applying regulatory standards for privacy, have led individual actors who manage personal data to incorporate a wide range of different standards and practices for privacy protection.[62] Recognizing the need for case-specific determinations, the HIPAA Privacy Rule provides an alternative approach that allows data to be shared pursuant to an expert's determination that "generally accepted statistical and scientific principles and methods for rendering information not individually identifiable" have been applied and provision of documentation 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."[63] Practitioners frequently comment on the ambiguity of these standards and the lack of clarity around the interpretation of definitions such as personally identifiable information.[64]

[59] 201 C.M.R. 17.02.

[60] 45 C.F.R. Part 160 and Subparts A and E of Part 164.

[61] 45 C.F.R. § 164.514. Note that in addition to the removal of the eighteen identifiers, the safe harbor also requires that the entity releasing the data "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." *Id.*

[62] *See, e.g.,* Benjamin C.M. Fung, Ke Wang, Rui Chen & Philip S. Yu, *Privacy-Preserving Data Publishing: A Survey of Recent Developments*, 42 ACM Computing Surveys (2010).

[63] 45 C.F.R. § 164.514(b). The Department of Health & Human Services has declined to provide specific instructions for carrying out an expert determination. *See* HHS Office of the Secretary, Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule HHS.gov, http://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf (2012) ("No single universal solution addresses all privacy and identifiability issues. Rather, a combination of technical and policy procedures are often applied to the de-identification task. [The Office for Civil Rights (OCR)] does not require a particular process for an expert to use to reach a determination that the risk of identification is very small. However, the Rule does require that the methods and results of the analysis that justify the determination be documented and made available to OCR upon request. The following information is meant to provide covered entities with a general understanding of the de-identification process applied by an expert. It does not provide sufficient detail in statistical or scientific methods to serve as a substitute for working with an expert in de-identification.").

[64] For example, the 2008 rulemaking to update FERPA acknowledged the confusion expressed by commentators regarding the potential applicability of the law's definition of personally identifiable information. *See* 73 Fed. Reg. at 74,830–31 (noting comments from the public that "the standard . . . about whether the information requested is 'linked or linkable' to a specific student was too vague and overly broad and could be logically extended to cover almost any information about a student," "a comprehensive list of indirect identifiers would be helpful," a definition of "the concept of indirect identifiers" is needed, and clarification of "which personally identifiable data elements may

In a 2012 survey of commentary on the U.S. legal framework for privacy protection, the Federal Trade Commission (FTC) concluded that "the traditional distinction between PII and non-PII has blurred and that it is appropriate to more comprehensively examine data to determine the data's privacy implications."[65] The FTC, which has authority to bring enforcement actions against companies that engage in unfair and deceptive trade practices including such practices that involve data privacy and security, takes a different approach. The FTC has developed a privacy framework that applies to commercial entities that collect or use consumer data that "can be reasonably linked to a specific consumer, computer, or other device."[66] FTC guidance has set forth a three-part test for determining whether data are "reasonably linkable" under this standard. To demonstrate that data are not reasonably linkable to an individual identity, a company must (1) take reasonable measures to ensure the data are de-identified, (2) publicly commit not to try to re-identify the data, and (3) contractually prohibit downstream recipients from attempting to re-identify the data.[67] The first prong is satisfied when there is "a reasonable level of justified confidence that the data cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer, computer, or other device."[68] Noting that it will follow the flexible standard that it follows in data security cases,[69] the FTC clarifies that "what qualifies as a reasonable level of justified confidence depends upon the particular circumstances, including the available methods and technologies," as well as "the nature of the data at issue and the purposes for which it will be used."[70] The FTC notes that various technical approaches can be used to satisfy this standard, and that it "encourages companies and researchers to continue innovating in the development and evaluation of new and better approaches to deidentification. FTC staff will continue to monitor and assess the state of the art in de-identification."[71] As such, the FTC's approach is likely to evolve over time in response to the development of new technologies for privacy protection.

2.3.2 Selected approaches from the European Union

Privacy law in the European Union relies on a definition of *personal data* that is broader than the corresponding definitions in the United States. The Data Protection Directive, for instance, defines personal data as "any information relating to an identified or identifiable natural person," where an "an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity."[72] In addition, the proposed General Data Protection Regulation defines *personal data* as "any information relating to a data subject," where a data subject is defined as a person who "can be identified, directly or indirectly, by means reasonably

be released without consent" should be provided.)

[65] FEDERAL TRADE COMMISSION, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE 2 (2012).

[66] *Id.* at vii. Note also that the framework does not apply to "companies that collect only non-sensitive data from fewer than 5,000 consumers a year, provided they do not share the data with third parties." *Id.* at iv.

[67] *Id.* at 21.

[68] *Id.* at 2.

[69] "The [Federal Trade] Commission's approach in data security cases is a flexible one. Where a company has offered assurances to consumers that it has implemented reasonable security measures, the Commission assesses the reasonableness based, among other things, on the sensitivity of the information collected, the measures the company has implemented to protect such information, and whether the company has taken action to address and prevent well-known and easily addressable security vulnerabilities." *Id.* at 21 n.108.

[70] *Id.* at 21.

[71] *Id.* at 21.

[72] Council Directive 95/46/EC, art. 2, 1995 O.J. (L. 281) 31.

likely to be used."[73] The Directive's provisions do not apply to "data rendered anonymous in such a way that the data subject is no longer identifiable."[74]

The Article 29 Working Party which provides advisory guidelines on the Data Protection Directive has clarified the distinction between an *identified* and *identifiable* person as follows: "a natural person can be considered as 'identified' when, within a group of persons, he or she is 'distinguished' from all other members of the group. Accordingly, the natural person is 'identifiable' when, although the person has not been identified yet, it is possible to do it."[75] It is also clear that "the question of whether the individual to whom the information relates is identified or not depends on the circumstances of the case."[76] For instance, "[a] very common family name will not be sufficient to identify someone - i.e. to single someone out - from the whole of a country's population, while it is likely to achieve identification of a pupil in a classroom. Even ancillary information, such as 'the man wearing a black suit' may identify someone out of the passers-by standing at a traffic light."[77]

The Directive also explains that "whereas to determine whether a person is identifiable account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person."[78] The Article 29 Working Party writes that in interpreting this standard, "the cost of conducting identification," "[t]he intended purpose, the way the processing is structured, the advantage expected by the controller, the interests at stake for the individuals, as well as the risk of organisational dysfunctions (e.g. breaches of confidentiality duties) and technical failures should all be taken into account."[79] It also notes that "this test is a dynamic one and should consider the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed. . . . The system should be able to adapt to these developments as they happen, and to incorporate then the appropriate technical and organisational measures in due course."[80] To provide specific guidance on applying de-identification techniques, the Article 29 Working Party has released an opinion that assesses the strengths and weaknesses of various technical approaches to de-identification.[81] In implementing the Data Protection Directive and applying the standard for determining whether data subjects can be considered "identifiable" or whether the data have been rendered "anonymous," the EU Member States have adopted divergent interpretations.[82] In 2012, the European Council concluded that some Member States, e.g., Denmark, Finland, France, Italy, Spain, and Sweden, "are generally less demanding [than other Member States] with regard to the processing of data that are not immediately identifiable, taking into account the likelihood of the data subject being identified as well as the nature of the data."[83]

Given the wide variation among these definitions and how they are interpreted, the gaps created

[73] Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, at 41, COM (2012) 11 final (Jan. 25, 2012).

[74] Data Protection Directive, at recital 26.

[75] Article 29 Data Protection Working Party, Opinion 4/2007 on the Concept of Personal Data, 12251/03/EN 12 (June 20, 2007).

[76] *Id.* at 13.

[77] *Id.*

[78] Data Protection Directive, at recital 26.

[79] Opinion 4/2007 at 15.

[80] *Id.*

[81] Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques (Apr. 10, 2014).

[82] European Council, Evaluation of the Implementation of the Data Protection Directive, Annex 2, at 15 (2012).

[83] *Id.*

by the narrowness of their scope (particularly within the US framework), the ambiguity regarding the context-specific applicability along the boundaries, and the dynamic nature of the definitions and their interpretation in response to technological developments over time, we argue that an approach to modeling the requirements of these laws more formally is needed so that it can be demonstrated that a privacy-preserving technology satisfies them. In the language of the selected laws introduced in this section, this Article proposes a methodology that can potentially be used to formalize regulatory definitions such as "information which identifies a person;"[84] "a reasonable level of justified confidence that the data cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer, computer, or other device;"[85] and "a person who can be identified, directly or indirectly, by means reasonably likely to be used"[86] so that a privacy technology's compliance with the regulations can be rigorously demonstrated. The approach we propose could also potentially be used by regulators and advisory bodies in future assessments regarding whether an emerging privacy technology satisfies regulatory requirements.

3 Introduction to two privacy concepts: differential privacy and FERPA

In this Article, we demonstrate a methodology for bridging between a technical and a regulatory privacy concept. Although we believe this methodology is generally applicable, in order to illustrate its use, we focus on two specific privacy concepts—one technical (differential privacy) and the other regulatory (FERPA). In this Section, we introduce the two concepts and set forth the definitions that will form the basis of the analysis that will follow in later sections.

We choose to rely on differential privacy in this Article because its rich and developed theory can serve as an initial subject of an examination of how formal notions of privacy can be compared to a legal standard. In addition, demonstrating that differential privacy is in accordance with privacy regulations may be essential for some practical uses of differential privacy.

3.1 Differential privacy

Differential privacy is a recent privacy concept that has emerged in the theoretical computer science literature, in response to accumulated evidence of the weaknesses of traditional techniques for privacy protection such as de-identification.[87] Differential privacy, first presented in 2006, is the result of ongoing research to develop a privacy technology that provides robust protection even against unforeseen attacks. Differential privacy by itself is not a single technological solution but a definition (sometimes referred to as a standard) that states a concrete requirement, and technological solutions are said to satisfy differential privacy if they adhere to the definition. As a strong, quantitative notion of privacy, differential privacy is provably resilient to a very large class of potential misuses. Differential privacy therefore represents a solution that moves beyond

[84] Video Privacy Protection Act, 18 U.S.C. § 2710(a)(3)

[85] Federal Trade Commission, Protecting Consumer Privacy in an Era of Rapid Change 2 (2012).

[86] Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, at 41, COM (2012) 11 final (Jan. 25, 2012).

[87] This discussion of differential privacy is adapted from Kobbi Nissim, Thomas Steinke, Alexandra Wood, Mark Bun, Marco Gaboardi, David O'Brien & Salil Vadhan, *Differential Privacy: An Introduction for Social Scientists*, Working Paper (forthcoming 2016). For the literature on differential privacy, see sources cited *supra* note 1.

the penetrate-and-patch approach that is characteristic of many traditional approaches to privacy, which must be continually updated as new vulnerabilities are discovered.

3.1.1 The privacy definition and its guarantee

In this section, we offer an intuitive view of the privacy guarantee provided by differentially private computations.[88]

Consider a hypothetical individual John, who has the opportunity to participate in a study exploring the relationship between socioeconomic status and medical outcomes. All participants in the study must complete a questionnaire covering where they live, their health, and their finances. John is aware of re-identification attacks that have been performed on de-identified data and is worried that some sensitive information about him, such as his HIV status or annual income, might be revealed by a future analysis based in part on his responses, should he participate in the study. If leaked, this personal information could, for example, lead to a change in his life insurance premium or affect the outcome of a future bank loan application.

If an analysis on the data from this study is differentially private, then John is guaranteed that even though his information is used in the analysis, the outcome of the analysis will not disclose anything that is *specific to him*. To understand what this means, consider a thought experiment, which we refer to as *John's privacy-ideal scenario* and illustrate in Figure 3. John's privacy-ideal scenario is one in which his personal information is omitted but the information of all other individuals is provided as input as usual. Because John's information is omitted, the outcome of the computation *cannot* depend on John's specific information.

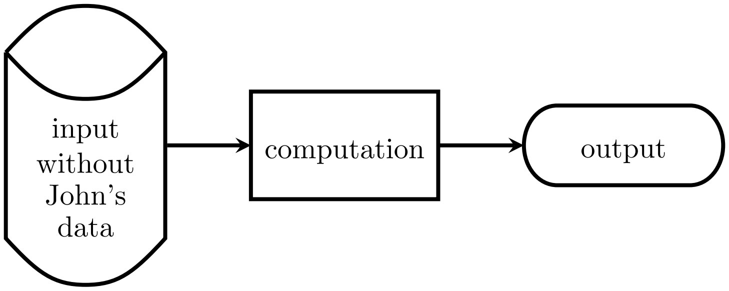


Figure 3: John's privacy-ideal scenario.

Differential privacy aims to provide John with privacy protection in the real-world scenario that approximates his privacy-ideal scenario. Hence, what can be learned about John from a differentially private computation is (essentially) limited to what could be learned about him from everyone else's data *without him being included in the computation.* Crucially, this very same guarantee is made not only with respect to John, but also to every other individual contributing his or her information to the analysis!

A parameter quantifies and limits the extent of the deviation between the privacy-ideal and real world scenarios. As shown in Figure 4 below, this parameter is usually denoted by the Greek letter ϵ (epsilon) and is referred to as the "privacy parameter," or, more accurately, the "privacy loss parameter." The parameter ϵ measures the effect of each individual's information on the output of

[88] For the mathematical definition, see the Appendix.

the analysis. It can also be viewed as a measure of the additional privacy risk an individual could incur beyond the risk incurred in the privacy-ideal scenario. Note that in Figure 4 we replaced John with a prototypical individual X to emphasize that the differential privacy guarantee is made simultaneously to *all* individuals in the sample.

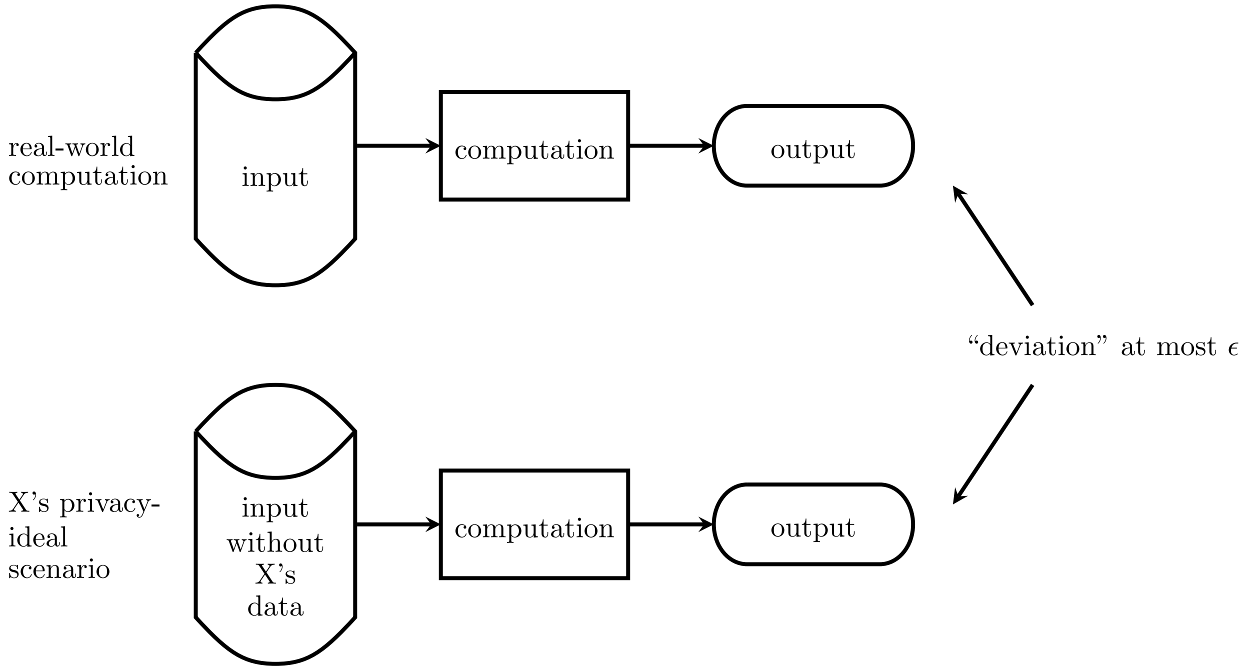


Figure 4: Differential Privacy. The maximum deviation between the privacy-ideal scenario and real world computation should hold simultaneously for each individual X whose information is included in the input.

It is important to note that differential privacy does not guarantee that an observer will not be able to learn anything about John from the outcome of the survey. Consider an observer, Alice, who possesses prior knowledge of some information about John, such as that he regularly consumes a lot of red wine. If the study reports a correlation between drinking red wine and the occurrence of a certain type of liver disease, Alice might conclude that John has a heightened liver disease risk. However, notice that even if information about John is not used in the study, Alice would be able to draw the conclusion that he has a heightened liver disease risk just like other red wine drinkers. In other words, this type of risk is present in both John's privacy-ideal scenario and his real-world scenario.

John may be adversely affected by the discovery of the results of a differentially private computation (for example, if sales of red wine were made illegal as a result of the discovery). The guarantee is that such harm is not due to the presence of John's data; that is, it would occur also in his privacy-ideal scenario.

3.1.2 Differential privacy in the real world

Despite being a relatively new concept, differential privacy has already found use in several real-world applications, and more applications are currently under development. The U.S. Census Bureau makes available an online interface for exploring the commuting patterns of workers across the United States, using confidential data collected through the Longitudinal Employer-Household Dynamics program over the period of 2002–2014.[89] Users of this interface interact with synthetic datasets that have been carefully generated from confidential agency records. The computations used to synthesize the data provide formal privacy guarantees and meet a variant of differential privacy.[90]

Google is also experimenting with differentially private applications. The RAPPOR system developed by Google employs differentially private computations to collect information from users of the company's Chrome web browser, in order to gather statistics used to monitor how unwanted software hijacks the browser settings of their users.[91] This application allows analysts at Google to study trends present in the extensive Chrome user base, with strong guarantees that not much can be learned that is specific to any individual user.

The academic community is also in the process of developing practical platforms for performing differentially private analyses. The *Putting Differential Privacy to Work* project at the University of Pennsylvania strives to build a general system that enables the average programmer to employ differentially private computations in a range of applications.[92] As part of the *Privacy Tools for Sharing Research Data* project at Harvard University,[93] differential privacy will be integrated with TwoRavens,[94] a browser-based software interface for exploring and analyzing data that are hosted on the Dataverse repository platform, which is currently used by institutions throughout the world.[95] This will allow researchers to see the results of statistical analyses created by differentially private computations on sensitive datasets, including datasets that cannot otherwise be shared widely due to privacy concerns.

3.2 The Family Educational Rights and Privacy Act of 1974 (FERPA)

FERPA is a federal law requiring the protection of personal information held in education records in the United States.[96] The law protects education records, which are records that directly relate to a student,[97] and are maintained by an educational agency or institution that receives funding

[89] *See* U.S. Census Bureau, OnTheMap Application for the Longitudinal Employer-Household Dynamics program, http://onthemap.ces.census.gov (last visited Apr. 30, 2016).

[90] *See* Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, & Lars Vilhuber, *Privacy: Theory Meets Practice on the Map*, Proceedings of the IEEE 24th International Conference on Data Engineering 277 (2008).

[91] *See* Úlfar Erlingsson, *Learning Statistics with Privacy, aided by the Flip of a Coin*, Google Research Blog (Oct. 30, 2014), http://googleresearch.blogspot.com/2014/10/learning-statistics-with-privacy-aided.html; Úlfar Erlingsson, Vasyl Pihur & Aleksandra Korolova, *RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response*, Proceedings of the 21st ACM Conference on Computer and Communications Security (2014).

[92] Putting Differential Privacy to Work project, http://privacy.cis.upenn.edu (last visited Apr. 30, 2016).

[93] Privacy Tools for Sharing Research Data, http://privacytools.seas.harvard.edu (last visited Apr. 30, 2016).

[94] Institute for Quantitative Social Science, TwoRavens, http://datascience.iq.harvard.edu/about-tworavens (last visited Apr. 30, 2016).

[95] Institute for Quantitative Social Science, Dataverse, http://datascience.iq.harvard.edu/about-dataverse (last visited Apr. 30, 2016).

[96] Family Educational Rights and Privacy Act of 1974, 20 U.S.C. § 1232g; 34 C.F.R. Part 99 (2013).

[97] 20 U.S.C. § 1232g(a)(4).

under a program administered by the U.S. Department of Education.[98] This includes elementary and secondary schools, school districts, colleges and universities, state educational agencies, and other institutions providing educational services or directing institutions that do.[99]

FERPA provides parents with certain rights associated with their child's education records, rights which transfer to an "eligible student" upon turning 18.[100] The rights provided under FERPA include the right to inspect, request amendment to, and consent to the disclosure of education records.[101] The law distinguishes between two types of information from education records: *directory information* and *non-directory personally identifiable information*.[102] Generally, a parent or eligible student must provide written consent before an educational agency or institution can disclose non-directory personally identifiable information from an education record.[103] Information designated by a school as directory information can be disclosed without consent, as long as parents were provided with notice and an opportunity to opt out of the disclosure of directory information.

An exception to FERPA that permits the disclosure of *de-identified information* is intended to establish "an appropriate balance that facilitates school accountability and educational research while preserving the statutory privacy protections in FERPA."[104] This exception to FERPA permits the disclosure of de-identified information without consent, and indeed without any restriction, "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."[105] This de-identification provision enables the widespread publication and use of statistics on educational programs in the United States.

In the discussion below, we provide the definitions adopted in FERPA, as well as interpretations of and rationale behind the definitions that can be found in Department of Education guidance. These definitions, as interpreted by the Department of Education,[106] form the basis of our formal model of FERPA's privacy requirements in Section 4.

3.2.1 The applicability of FERPA's requirements to differential privacy

FERPA's protections likely apply to the release of statistics produced by methods that satisfy formal privacy models such as differential privacy. Therefore, it is important to understand exactly how information privacy laws such as FERPA would govern the use of new technologies based on formal privacy models, in anticipation of their practical implementation. In this Section, we point to sources that highlight the ways in which FERPA seems likely to apply to the release

[98] 20 U.S.C. § 1232g(a)(1)(A).

[99] 34 C.F.R. § 99.1(a)(1-2).

[100] 34 C.F.R. § 99.3.

[101] §§ 99.10, 99.20, 99.30.

[102] This distinction is discussed in depth in Section 3.2.2 below.

[103] § 99.30. *Disclosure* means "to permit access to or the release, transfer, or other communication of personally identifiable information contained in education records by any means, including oral, written, or electronic means, to any party except the party identified as the party that provided or created the record." 34 C.F.R. § 99.3.

[104] *Id.*

[105] 34 C.F.R. § 99.31(b)(1).

[106] The interpretation of FERPA used in this Article draws from the FERPA regulations, 34 C.F.R. Part 99, and the 2008 rulemaking to update the FERPA regulations, 73 Fed. Reg. 74,806–55 (Dec. 9, 2008), in which an extended discussion of the definition of *personally identifiable information* was provided in justification of the latest revision to the definiton.

of differentially private statistics, and set the stage for later sections which aim to interpret this language more formally.

Generally, FERPA governs the disclosure of *non-directory personally identifiable information* about students in education records maintained by educational agencies and institutions. Here, *disclosure* is defined broadly, meaning "to permit access to or the release, transfer, or other communication of personally identifiable information contained in education records by any means, including oral, written, or electronic means, to any party except the party identified as the party that provided or created the record."[107] *Personally identifiable information* is also defined broadly to include "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."[108] Because as we discuss above in Section 2.1.3 it has been demonstrated that releases of aggregate statistics can leak some information about individuals, these definitions, which cover any communication of any information linkable to a specific student with reasonable certainty, are arguably written broadly enough to encompass privacy risks associated with statistical releases.

This interpretation of FERPA is supported by the 2008 rulemaking to update the FERPA regulations,[109] which explicitly addresses privacy issues in releases of statistics. For example, the 2008 rulemaking refers to the capability of "re-identifying statistical information or redacted records."[110] Guidance on interpreting FERPA, by explictly referring to the privacy risks associated with both "statistical information" and "redacted records" instructs educational agencies and institutions to consider the privacy risks in releases of data in both aggregate and individual-level formats. The rulemaking discusses specific examples to illustrate some of the privacy risks associated with the release of statistical information.[111] The Department of Education notes, for example, that "a school may not release statistics on penalties imposed on students for cheating on a test where the local media have published identifiable information about the only student (or students) who received that penalty; that statistical information or redacted record is now personally identifiable to the student or students because of the local publicity."[112] In addition, the rulemaking explains how the publication of a series of tables about the same set of students, with the data broken down in different ways, can in combination reveal personally identifiable information about individual students.[113] It also notes that educational institutions are prohibited from reporting that 100 percent of students achieved specified performance levels, as a measure to prevent the leakage of personally identifiable information.[114] These references from the 2008 rulemaking serve as evidence that the Department of Education recognizes privacy risks associated with the release of aggregate statistics.

Educational agencies and institutions share statistics from education records with other agencies, researchers, and the public, for the purposes of research and evaluation. The law mandates the release of certain education statistics,[115] unless such statistics would reveal personally identifi-

[107] 34 C.F.R. § 99.3.

[108] *Id.* The definition of personally identifiable information is discussed in more detail below in Section 3.2.3.

[109] 73 Fed. Reg. 73 Fed. Reg. 74,806–55 (Dec. 9, 2008)

[110] 73 Fed. Reg. at 74,832.

[111] *See id.*

[112] *Id.*

[113] *See id.* at 74,835.

[114] *See id.*

[115] *See* 20 U.S.C. § 6311(h); 20 U.S.C. § 9607.

able information as defined by FERPA.[116] For instance, educational agencies and institutions are prohibited from releasing tables containing statistics on groups of individuals falling below some minimum size, where the minimum size varies by state.[117] Each state has established additional procedures for protecting privacy in the release of statistics, including "various forms of suppression, top and bottom coding of values at the ends of a distribution, and limiting the amount of detail reported for the underlying counts."[118] The Department of Education's National Center for Education Statistics has released implementation guidance devoted to helping such institutions apply privacy safeguards in accordance with FERPA when releasing aggregate statistics.[119] This guidance clarifies the goal of FERPA in the context of aggregate data releases:

> Protecting student privacy means publishing data only in a manner that does not reveal individual students' personally identifiable information, either directly or in combination with other available information. Another way of putting this is that the goal is to publish summary results that do not allow someone to learn information about a specific student.[120]

The guidance further "demonstrates how disclosures occur even in summary statistics," discusses how common approaches to privacy may fall short of FERPA's standard for privacy protection, and provides some best practices for applying disclosure limitation techniques in releases of aggregate data.[121] This practical guidance developed by the Department of Education is further evidence that the agency recognizes some leakages of information about individuals from aggregate data releases to be FERPA violations.

Not only does the Department of Education require educational agencies and institutions to address privacy risks in the release of aggregate statistics, but the scientific literature on privacy also suggests that this is a prudent approach. Numerous attacks have demonstrated that it is often possible to link particular individuals to information about them in aggregate data releases.[122] Moreover, the potential leakage of personally identifiable information through releases of aggregate statistics is a concern that is anticipated to evolve and grow over time, as analytical capabilities advance and the availability of large quantities of personal information from various sources increases. It is likely that the approaches identified in current agency guidance will in the future be shown not to provide adequate privacy protection, while also significantly limiting the usefulness of the data released. In response, educational agencies and institutions may turn to formal privacy models,

[116] 34 C.F.R. § 200.7(b) ("A State may not use disaggregated data for one or more subgroups . . . to report achievement results . . . if the results would reveal personally identifiable information about an individual student . . . [under the requirements of FERPA]").

[117] "Individual states have adopted minimum group size reporting rules, with the minimum number of students ranging from 5 to 30 and a modal category of 10 (used by 39 states in the most recent results available on state websites in late winter of 2010)." National Center for Education Statistics, Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting, SLDS Technical Brief: Guidance for Statewide Longitudinal Data Systems (SLDS), https://nces.ed.gov/pubs2011/2011603.pdf (Dec. 2010), at 1.

[118] *Id.*

[119] *Id.*

[120] *Id.* at 4.

[121] *Id.* Specifically, best practices identified in this report include "publishing the percentage distribution across categories of outcome measures with no underlying counts or totals; publishing a collapsed percentage distribution across categories of outcome measures with no underlying counts or totals; publishing counts but using complementary suppression at the subgroup level when a small subgroup is suppressed; limiting the amount of detail published for school background information; recoding the ends of percentage distributions; and recoding high and low rates." *Id.*

[122] Privacy risks associated with de-identified and aggregate data are discussed above in Section 2.1.3.

such as differential privacy, as promising solutions for releasing data while providing stronger and quantifiable privacy protection and, in some cases, improvements in data utility.[123]

In particular, formal privacy models provide a solution to a problem that the Department of Education currently identifies as an open problem. The 2008 rulemaking to update FERPA recognizes the growing understanding that privacy risk from data disclosures is cumulative, and accordingly requires educational agencies and institutions to take into account the cumulative privacy risk from multiple disclosures of information:

> The existing professional literature makes clear that public directories and previously released information, including local publicity and even information that has been de-identified, is sometimes linked or linkable to an otherwise de-identified record or dataset and renders the information personally identifiable. The regulations properly require parties that release information from education records to address these situations.[124]

However, agency resources do not provide guidance on addressing cumulative privacy risks from successive disclosures. Instead, the Department notes that "[i]n the future [it] will provide further information on how to monitor and limit disclosure of personally identifiable information in successive statistical data releases."[125] Research from the computer science literature demonstrates that it is very difficult to account for the cumulative privacy risk from multiple statistical releases.[126] In contrast to all other privacy concepts we are aware of, differential privacy provides guarantees on how risk accumulates from successive data releases. To address this concern as well as the other privacy concerns identified by the Department of Education in the discussion above, differential privacy is an attractive approach for future releases of statistics from education records. This emphasizes the value in demonstrating that differential privacy satisfies the requirements of FERPA.

3.2.2 The distinction between directory and non-directory information

FERPA distinguishes between *directory information* and *non-directory personally identifiable information.* Directory information is defined as "information contained in an education record of a student that would not generally be considered harmful or an invasion of privacy if disclosed."[127] Each educational agency or institution produces a list of categories of information it designates as directory information.[128] FERPA provides some categories of information for illustration of the types of information an educational agency or institution may designate as directory information:

[123] For a discussion of the advantages formal privacy models provide over traditional approaches, see the discussion above in Section 3.3.

[124] 73 Fed. Reg. at 74,831.

[125] *Id.* at 74,835.

[126] *See* Srivatsava Ranjit Ganta, Shiva Prasad Kasiviswanathan & Adam D. Smith *Composition attacks and auxiliary information in data privacy*, Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 285 (2008).

[127] 34 C.F.R. § 99.3.

[128] When releasing data, educational agencies and institutions must assess the privacy-related risks in light of publicly-available information, including directory information. *See* 73 Fed. Reg. at 74,834 (noting that "the risk of reidentification may be greater for student data than other information because of the regular publication of student directories, commercial databases, and de-identified but detailed educational reports by States and researchers that can be manipulated with increasing ease by computer technology. . . . [T]he re-identification risk of any given release is cumulative, i.e., directly related to what has previously been released, and this includes both publicly-available directory information, which is personally identifiable, and de-identified data releases."). Because the scope of directory information varies from school to school, knowing what information may be available to a potential adversary is ambiguous, creating a challenge for an educational agency or institution assessing the privacy risks

> Directory information includes, but is not limited to, the student's name; address; telephone listing; electronic mail address; photograph; date and place of birth; major field of study; grade level; enrollment status (e.g., undergraduate or graduate, full-time or part-time); dates of attendance; participation in officially recognized activities and sports; weight and height of members of athletic teams; degrees, honors, and awards received; and the most recent educational agency or institution attended.[129]

The law does not require consent for the disclosure of directory information as long as the relevant educational agency or institution has provided parents and eligible students with public notice of the types of information it has designated as directory information, and an opportunity to opt out of the disclosure or publication of directory information.[130] An educational agency or institution may also disclose directory information about former students without complying with the notice and opt out conditions.[131]

By contrast, educational agencies and institutions must take steps to protect *non-directory personally identifiable information* from release. This category of informcation can only be disclosed without consent under certain exceptions, such as the sharing of data for the purposes of developing predictive tests, administering student aid programs, improving instruction, or auditing or evaluating a federal- or state-supported education program.[132] In addition, information from education records that would otherwise be considered non-directory personally identifiable information can be released to the public, without consent, if it has been rendered *de-identified*, meaning "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."[133]

The definition of non-directory personally identifiable information and how it has been interpreted by the Department of Education serves as the basis for the formal model of FERPA's requirements we construct in this Article. Hence, we turn next to how the definition of personally identifiable information has been interpreted and applied in guidance and practice.

3.2.3 The definition of personally identifiable information

The analysis in this Article focuses on FERPA's definition of *personally identifiable information* and how it has been interpreted by the Department of Education. FERPA defines personally identifiable information by way of a non-exhaustive list of categories of information included within the definition. The definition is as follows:

associated with a planned data release. In Section 4 below, we propose an approach to formally modeling directory information and privacy risks in the release of education data more generally, which addresses the ambiguity created by differences in classifying directory information across different educational agencies and institutions. Specifically, to address this ambiguity in our model, we make a "worst-case" assumption that a potential privacy attacker is given unrestricted access to information that does not require consent for release, including all directory information.

[129] 34 C.F.R. § 99.3.

[130] 34 C.F.R. § 99.37(a).

[131] *Id.*

[132] *See* 34 C.F.R. §§ 99.31(a)(6), (a)(3).

[133] *See* 34 C.F.R. § 99.31(b)(1) ("An educational agency or institution, or a party that has received education records or information from education records under this part, may release the records or information without the consent required . . . 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.").

> “Personally Identifiable Information”
>
> The term includes, but is not limited to–
>
> (a) The student’s name;
>
> (b) The name of the student’s parent or other family members;
>
> (c) The address of the student or student’s family;
>
> (d) A personal identifier, such as the student’s social security number, student number, or biometric record;
>
> (e) Other indirect identifiers, such as the student’s date of birth, place of birth, and mother’s maiden name;
>
> (f) 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
>
> (g) Information requested by a person who the educational agency or institution reasonably believes knows the identity of the student to whom the education record relates.[134]

This analysis focuses in particular on paragraph (f) of the above definition, which forms the basis for the development of a formal model of FERPA’s privacy requirement in Section 4. The rulemaking history[135] and agency guidance[136] provide an aid to its interpretation. The definition of personally identifiable information, by referring to “a reasonable person in the school community”[137] makes use of an objective, reasonableness standard. The reasonable person standard is a common legal standard, referring to a “hypothetical, rational, prudent, average individual.”[138] By referring to “a reasonable person in the school community,” the Department of Education intended to “provide the maximum privacy protection for students” because ”a reasonable person in the school community is also presumed to have at least the knowledge of a reasonable person in the local community, the region or State, the United States, and the world in general.”[139] The agency notes the standard was not intended to refer to the “technological or scientific skill level of a person who would be capable of re-identifying statistical information or redacted record.”[140] Rather, it refers to the knowledge a reasonable person might have “i.e., based on local publicity, communications, and other ordinary conditions.”[141] The regulatory history includes some examples of members of the school community, such as “students, teachers, administrators, parents, coaches, volunteers,” and others at the local school.[142]

Additional clues to interpreting the definition of personally identifiable information can be found in the rulemaking history. In 2008, the Department of Education updated the definition. Previously,

[134] 34 C.F.R. § 99.3.

[135] 73 Fed. Reg. 74,806–55.

[136] National Center for Education Statistics, Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting, SLDS Technical Brief: Guidance for Statewide Longitudinal Data Systems (SLDS), https://nces.ed.gov/pubs2011/2011603.pdf (Dec. 2010).

[137] 34 C.F.R. § 99.3.

[138] 73 Fed. Reg. at 74,832.

[139] *Id.*

[140] *Id.* at 74,831.

[141] *Id.* at 74,832.

[142] *Id.*

it had included "information that would make the student's identity easily traceable" in lieu of clauses (e)–(g).[143] The Department of Education explained that it removed the "easily traceable" language from the definition "because it lacked specificity and clarity" and "suggested that a fairly low standard applied in protecting education records, i.e., that information was considered personally identifiable only if it was easy to identify the student."[144] This rationale indicates that the agency intends the standard to be interpreted to provide strong privacy protection and to take into account some sophisticated approaches to defeating privacy safeguards.

The rulemaking history also explains the adoption of an objective, "reasonableness" standard, in light of this revision to the definition of personally identifiable information:

> The "reasonableness" standards in paragraphs (f) and (g) of the new definition, which replace the "easily traceable" standard, do not require the exercise of subjective judgment or inquiries into a requester's motives. Both provisions require the disclosing party to use legally recognized, objective standards by referring to identification not in the mind of the disclosing party or requester but by a reasonable person and with reasonable certainty, and by requiring the disclosing party to withhold information when it reasonably believes certain facts to be present. These are not subjective standards, and these changes will not diminish the privacy protections in FERPA.[145]

In providing guidance on interpreting the definitions of personally identifiable information and de-identified information, the Department of Education acknowledges that the standard involves some degree of ambiguity. It notes that the removal of "nominal or direct identifiers," such as names and Social Security numbers, "does not necessarily avoid the release of personally identifiable information."[146] Furthermore, the removal of other information such as address, date and place of birth, race, ethnicity, and gender, may not be sufficient to prevent one from "indirectly identify[ing] someone depending on the combination of factors and level of detail released."[147] The regulatory history also emphasizes that de-identified information that is released may be linked or linkable to "public directories and previously released information, including local publicity and even information that has been deidentified," rendering it personally identifiable.[148] The agency concludes that "[t]he regulations properly require parties that release information from education records to address these situations."[149] In other guidance, the agency clarifies that FERPA requires educational agencies and institutions to ensure that the statistics they produce from education records "do not

[143] The previous definition of personally identifiable information, promulgated in 1988, appeared as follows:

> "Personally identifiable information" includes, but is not limited to—
> (a) The student's name;
> (b) The name of the student's parent or other family member;
> (c) The address of the student or student's family;
> (d) A personal identifier, such as the student's social security number or student number;
> (e) A list of personal characteristics that would make the student's identity easily traceable; or
> (f) Other information that would make the student's identity easily traceable.

53 Fed. Reg. 11943 (Apr. 11, 1988).

[144] 73 Fed. Reg. at 74,831.

[145] *Id.*

[146] *Id.* at 74,831.

[147] *Id.*

[148] *Id.*

[149] *Id.*

allow someone to learn information about a specific student."[150] The Department of Education's National Center for Education Statistics provides an overview of statistical disclosure limitation techniques for aggregate data, such as "bottom or top coding the results at the tails of the percentage distribution, or for high and low rates. This is typically done by coding all percentages above 95 percent as greater than 95 percent and coding all percentages below 5 percent as less than 5 percent. This is done to avoid reporting the fact that all, or nearly all, of the students in a reporting subgroup share the same achievement level or the same outcome or that very few or none of the students have a particular outcome."[151]

In some cases, the examples provided in the rulemaking history contribute to a lack of clarity. For instance, the Department of Education provides the following example to illustrate how the language "personal knowledge of the relevant circumstances," found in paragraph (f) of the definition of personally identifiable information, should be interpreted:

> [I]f it is generally known in the school community that a particular student is HIV-positive, or that there is an HIV-positive student in the school, then the school could not reveal that the only HIV-positive student in the school was suspended. However, if it is not generally known or obvious that there is an HIV-positive student in school, then the same information could be released, even though someone with special knowledge of the student's status as HIV-positive would be able to identify the student and learn that he or she had been suspended.[152]

This example seems counterintuitive because it does not address whether members of the school community might know, or might in the future learn, that a particular student was suspended. Possession of such knowledge would enable one to learn that this student is HIV-positive. Enabling this type of disclosure through a release of information—highly sensitive information such as a student's HIV status, no less—is likely not what the agency intended. However, it is not clear what the agency did in fact intend to convey with this example. In another example, the agency notes that "if teachers and other individuals in the school community generally would not be able to identify a specific student based on the student's initials, nickname, or personal characteristics contained in the record, then the information is not considered personally identifiable and may be released without consent."[153] This seems to imply a weak privacy standard, as a student's "initials, nickname, or personal characteristics" are likely to be uniquely identifying in many cases, regardless of whether such characteristics are considered to be generally known within the community.

The ambiguity of the definition of personally identifiable information is also reflected in a possible discrepancy between the regulatory history and an interpretation of the regulation developed by the Privacy Technical Assistance Center (PTAC), a Department of Education contractor that develops guidance on complying with FERPA's requirements. In guidance on interpreting the standard used to evaluate disclosure risk when releasing information from education records, PTAC advises that "[s]chool officials, including teachers, administrators, coaches, and volunteers, are not considered in making the reasonable person determination since they are presumed to have inside

[150] National Center for Education Statistics, Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting, SLDS Technical Brief: Guidance for Statewide Longitudinal Data Systems (SLDS), https://nces.ed.gov/pubs2011/2011603.pdf (Dec. 2010), at 4.

[151] *See id.* at 23.

[152] 73 Fed. Reg. 74,832.

[153] *Id.* at 74,831.

knowledge of the relevant circumstances and of the identity of the students."[154] Such individuals would seem to be "reasonable person[s] in the school community," based on the plain meaning of this language. Indeed, in the 2008 rulemaking, the Department of Education provides an example involving an analysis of whether "students, teachers, administrators, parents, coaches, volunteers, or others at the local high school" are generally aware of a fact, making these categories of individuals part of the case-specific equation.[155] The apparent conflict between the PTAC guidance and the rulemaking history is further evidence that FERPA's privacy standard can lead to various interpretations.

Commentators have likewise expressed uncertainty regarding this standard. For example, the 2008 rulemaking refers to many public comments seeking clarification on the de-identification standard. Comments refer to the standard as being "too vague and overly broad," as the definition of personally identifiable information "could be logically extended to cover almost any information about a student."[156] Other commenters question whether the standard provides privacy protection as strong as the agency intends, in light of concerns about the difficulty of de-identifying data effectively.[157] One commenter noted the ambiguity of the standard, but viewed it in a positive light, by arguing that "ambiguity in the terms 'reasonable person' and 'reasonable certainty' was necessary so that organizations can develop their own standards for addressing the problem of ensuring that information that is released is not personally identifiable."[158]

Given the ambiguity of FERPA's privacy standard, it is difficult to know with certainty whether a particular release of information or a particular disclosure limitation technique satisfies the standard. In Section 4, we aim to overcome ambiguities in the regulatory standard by modeling it formally, based on conservative, "worst-case" assumptions. We believe this approach can be used to demonstrate that a privacy technology satisfies any reasonable interpretation (or at least a large family of reasonable interpretations) of a legal standard for privacy protection, in a way that is both legally and mathematically rigorous.

To construct this formal model of FERPA's privacy standard, we look to the regulatory definitions and guidance on interpreting these definitions from the rulemaking history and other materials prepared by the Department of Education. We rely on many of the guidelines for interpretation discussed in this Section. In partcular, our model of the FERPA standard aims to give force to

[154] Privacy Technical Assistance Center, Frequently Asked Questions—Disclosure Avoidance (2015), http://ptac.ed.gov/sites/default/files/FAQ_Disclosure_Avoidance.pdf.

[155] 73 Fed. Reg. 74,832.

[156] *See* 73 Fed. Reg. at 74,830–34 (including comments such as "the standard . . . about whether the information requested is 'linked or linkable' to a specific student was too vague and overly broad and could be logically extended to cover almost any information about a student," "'relevant circumstances' in paragraph (f) is vague," "a comprehensive list of indirect identifiers would be helpful," a definition of "the concept of indirect identifiers" is needed, clarification of "which personally identifiable data elements may be released without consent" should be provided, "the regulations should provide objective standards for the de-identification of education records," and "examples to help districts determine whether a nontargeted request will reveal personally identifiable information" are needed).

[157] *See id.* at 74,833–834 (referring to comments noting that "complete de-identification of systematic, longitudinal data on every student may not be possible," that "many institutions and individuals have the ability to re-identify seemingly deidentified data and that it is generally much easier to do than most people realize because 87 percent of Americans can be identified uniquely from their date of birth, five-digit zip code, and gender," and that "re-identification is a much greater risk for student data than other kinds of information because FERPA allows for the regular publication of student directories that contain a wealth of personal information, including address and date of birth, that can be used with existing tools and emerging technology to re-identify statistical data, even by non-experts").

[158] *Id.* at 74,830.

the regulators' intent to "provide the maximum privacy protection for students,"[159] by providing protection against a strong adversary, including one that is not of a limited "technological or scientific skill level,"[160] possesses the knowledge about students that "a reasonable person in the school community"[161]—not just a reasonable person—might have, potentially has the capacity to learn something about an individual in the data even if the individual's identity is not "easily traceable,"[162] and has motives that are unknown.[163] The model we outline below does not require a determination of the subjective judgment or inquiries into an attacker's motives and declines to presuppose the attacker's goal. In this way, the model is able to consider attackers with very different goals, different pieces of outside knowledge of students' information, and different ways of using the information. Additionally, there are indications that the regulators intended FERPA adapt to growing privacy risks over time.[164] Accordingly, our model aims to be able to address future privacy risks, including those that are *currently* unrealizable or unknown, by not limiting its assumptions about the scientific or technological skills of the attacker.

3.3 Gaps between FERPA and differential privacy

The emergence of formal privacy models such as differential privacy represents a shift in the conceptualization of data privacy problems and effective solutions to such problems. However, because FERPA and implementation guidance from the Department of Education seem to have been drafted with traditional privacy approaches in mind, the two approaches differ significantly. To illustrate the gap between these two privacy concepts, we outline a number of the key points of departure below. Some of these gaps represent substantial challenges for translating between the two notions.

Overall scope of privacy protection. FERPA does not apply across the board to protect all types of data in all settings. Rather, the law is designed to address certain types of information in specific contexts. FERPA applies only to educational agencies and institutions and protects only certain types of information from education records known as non-directory personally identifiable information. It primarily seems to be aimed at preventing record linkage attacks, or the linkage of a named individual to a record in a release of data. In contrast, differential privacy is designed to be broadly applicable and to provide formal bounds on the leakage of any information about an individual, not just an individual's identity.

Range of attacks. While guidance on interpreting the law expresses the regulators' intent to provide strong privacy protection that addresses a wide range of privacy risks, in effect the provisions of the law seem to address a narrower category of attacks. By permitting the release of de-identified information, or information from which certain pieces of information have been removed, the law seems primarily aimed at addressing record linkage attacks that could enable the linkage of a named individual with a record in a released set of data. In contrast, differential privacy is a strong, quantitative *guarantee* of privacy that is provably resilient to a very large class of potential data misuses. This guarantee holds no matter what inferential strategy, computational resources, or external information the adversary brings to bear. For instance, differential privacy provides

[159] 73 Fed. Reg. at 74,832.

[160] *Id.* at 74,831.

[161] 34 C.F.R. § 99.3.

[162] 73 Fed. Reg. at 74,831.

[163] *Id.*

[164] *See id.* at 74,834 (issuing new guidance to educational agencies and institutions based on the recognition that "since the enactment of FERPA in 1974, the risk of re-identification from [previously released] information has grown as a result of new technologies and methods").

protection against inference attacks and attacks unforeseen by the individual applying the privacy-preserving technique.[165]

Scope of private information. The FERPA privacy standard focuses on protecting personally identifiable information, and draws a binary distinction between personally identifiable and de-identified information. It provides a non-exhaustive list of the categories of information that should be protected from release. Accordingly, a common practice is for data holders to withhold or redact certain pieces of information, such as names, Social Security numbers, and addresses when disclosing information from their education records. The computer science literature recognizes that privacy risks are not limited to certain categories of information. In particular, "[a]ny information that distinguishes one person from another can be used for re-identifying anonymous data."[166] For example, a small number of data points about an individual's characteristics, behavior, or relationships can be sufficient to identify an individual.[167] Differential privacy takes this conception of privacy risks into account by putting puts formal bounds on any leakage of information about an individual from a system.

Form of data release. By relying on terminology such as personally identifiable information and de-identification, the FERPA privacy standard seems to be written with microdata—or individual-level data—as its primary use case. Where the guidance on FERPA provides recommendations for releasing aggregate data, the recommendations are limited to the specification of minimum cell size and related approaches from the traditional statistical disclosure limitation literature. In addition, by referring explicitly to de-identification and permitting the disclosure of information from which categories of personally identifiable information have been removed, FERPA appears to endorse heuristic de-identification techniques. These references to microdata and de-identification make it harder to generalize to other techniques. For emerging techniques like differential privacy that provide solutions for non-microdata formats and are categorically different from traditional privacy approaches, it is not clear how the FERPA privacy standard should be applied.

Scope of guidance. FERPA, and how the law has been interpreted in guidance materials, typically emphasizes the most obvious extreme cases of privacy risks. For example, FERPA permits the release of information from which direct and indirect identifiers have been removed. As such, the FERPA privacy standard is primarily directed at preventing the release of full student records with names, student identification numbers, and certain demographic information attached. In addition, FERPA's definition of personally identifiable information is imprecise, as the boundary between personally identifiable information and de-identified information is not clear. Actors in this space typically use heuristics to make decisions regarding the application of de-identification techniques. Where guidance materials refer to aggregate data, they flag the disclosure risks associated with the release of statistics derived from as few as one or two individuals. Most of the examples provided in FERPA are anecdotal and provide little guidance on making difficult decisions along the margins. Outside of the extreme examples, determining whether the steps that have been taken to protect privacy are sufficient is unclear. Because FERPA does not provide a clear privacy goal, it is difficult to determine when a privacy-preserving technology provides protection that is adequate to satisfy

[165] See Arpita Ghosh and Robert Kleinberg, *Inferential Privacy Guarantees for Differentially Private Mechanisms*, Working Paper (2016).

[166] Arvind Narayanan & Vitaly Shmatikov, *Myths and Fallacies of "Personally Identifiable Information,"* 53 Communications of the ACM 24, 26 (2010).

[167] *See* Yves-Alexandre de Montjoye et al., *Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata*, 347 Science 536 (2015); Yves-Alexandre de Montjoye et al., *Unique in the Crowd: The Privacy Bounds of Human Mobility*, 3 Nature Sci. Reps. 1376 (2013).

the standard.

On first impression, it seems likely that the conceptual differences between legal and computer science approaches are too great to be able to make an argument that a standard from one discipline could satisfy one from the other. However, as we argue in the Sections that follow, it is possible to reason formally about privacy standards from law and computer science and bridge between the two. Note also that, while this article focuses on FERPA for an in-depth analysis, other laws such as the HIPAA Privacy Rule share many of the characteristics of FERPA identified in this Section. This suggests that the analysis and methodology presented in this article can be applied beyond the context of FERPA to information privacy laws more generally.[168]

3.4 Value in bridging these gaps

Legal and computer science concepts of privacy are evolving side by side, and it is becoming increasingly important to understand how they can work together. The field of computer science can benefit from an understanding of legal thinking on privacy, and the legal field can similarly be influenced by computer science thinking. The influence of one discipline on another can be very valuable in the future. We argue, however, that to overcome the substantial gaps between the two concepts, a rigorous approach is required.

Bridging the gap between technical and regulatory approaches to privacy will help in bringing formal privacy models such as differential privacy to practice. Uncertainty about compliance with strict regulatory requirements for privacy protection stands as a barrier to adoption and use of emerging techniques for analyzing sensitive information in the real world. If data holders can be assured that the use of formal privacy models will satisfy their legal obligations, they will be more likely to begin making new data sources available for research and commercial use.

At the same time, this approach is also important for the future of robust privacy regulation. Policymakers will need to understand the technology and its guarantees in order to approve and support the use of formal privacy models as a means of satisfying regulatory obligations. In addition, a clear understanding of the principles underlying formal approaches to privacy protection and the ways in which they differ from the existing regulatory approach help illustrate the weaknesses in the regulatory framework and point to a path forward. Taken together, these insights will help pave the way for the adoption of robust privacy practices in the future.

4 Extracting a formal privacy definition from FERPA

To demonstrate that a formal privacy model such as differential privacy satisfies the requirements of a privacy regulation such as FERPA, we propose an approach inspired by the game-based privacy definitions used in computer science. As discussed in Section 2.2, within this approach privacy is defined via a hypothetical game in which an adversary attempts to learn private information from the output of a computation performed on private data. If it can be shown that the adversary cannot win the game “too much,” the computation is considered to provide privacy. Recall the example privacy game from Section 2.2. In this scenario, Alice wants to send an encrypted message to Bob with confidence that an eavesdropper cannot learn much about the content of their communication by seeing the encrypted text. She is able to formalize her desiderata as a privacy game, empowering

[168] Indeed, in Section 6 we provide a general methodology for bridging between regulatory and technical approaches to privacy.

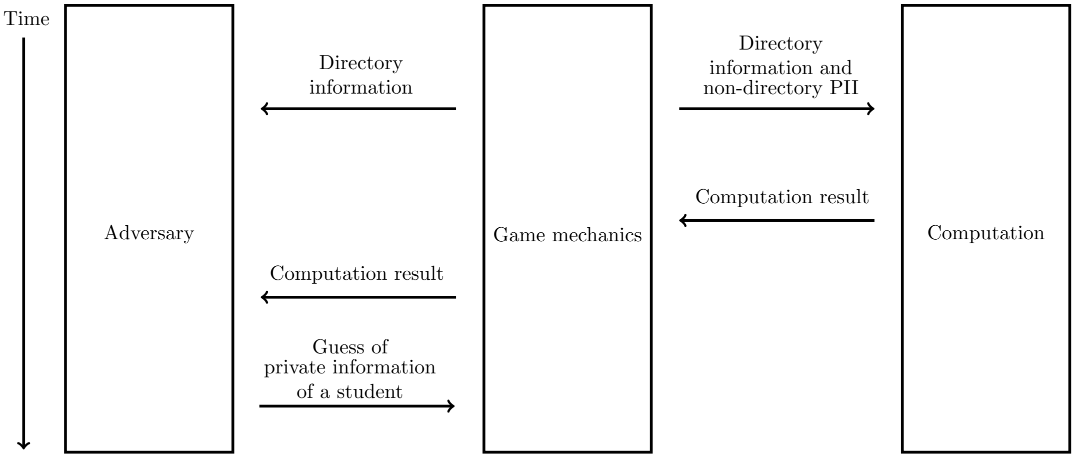
her to use encryption algorithms with confidence if they have been mathematically proven to meet her definition of privacy.

The goal of our approach is to define a game that faithfully encompasses the privacy desiderata that were envisioned by the regulators who drafted FERPA. To do so, we must carefully define the capabilities of the adversary and the mechanics of the game in ways that capture the privacy threat that FERPA was enacted to protect against. As FERPA is written in regulatory language and not as a formal mathematical definition, it necessarily contains ambiguities and open to different interpretations. Therefore, our desideratum is to design a game that conservatively accounts for (ideally) any reasonable interpretation of the regulatory text, or, at least, accounts for a very large class of such interpretations. As we will describe in detail below, this requires us to design games that give the adversary what might be considered to have an unrealistic advantage. However, if we can prove that the system provides privacy even given extremely conservative assumptions, we will have also proven that it provides privacy protection in more realistic scenarios.

We begin with a simple, informal view of a privacy game based on FERPA's requirements. We can imagine a game in which a school classifies the student information maintained in its education records into two distinct categories: *directory information* (which can be made publicly available, and hence in our modeling we make it available to the attacker) and *non-directory personally identifiable information* (which should be protected from disclosure, so we do not make it available to the attacker). A statistical computation is performed by the game mechanics over this information and the result is passed to the adversary. The adversary wins the game if he can successfully guess the sensitive attribute of a student, i.e., link a named student to non-directory personally identifiable information about that student. Figure 5 represents this game visually.

To make this game framework more concrete, consider the following scenario. The National Center for Education Statistics releases some data having applied a certain k-anonymization algorithm to the original private dataset. A data broker attempts to glean non-directory personally identifiable information about students from the k-anonymized data so that it can sell the improved data to a third party. If the data broker (perhaps using some external information) is able to successfully extract a student's non-directory personally identifiable information from the release, then the privacy of that student has been violated and the data broker has won the game. In this scenario, the data broker is playing the role of the adversary and the k-anonymization algorithm is the computation that is supposed to be providing privacy. The data broker wins the game by successfully guessing a student's non-directory personally identifiable information.

Note, however, that the fact that an adversary wins the game does not necessarily mean that a privacy breach has occurred. For instance, consider a scenario where gender is treated as a private attribute. If challenged to guess the gender of an arbitrary student from a typical co-educational school, the adversary can be expected to win the game with a probability of 0.5 even if the guess is based on a random coin toss. The adversary has this chance of success without knowing anything about the student or having access to any statistic computed on that student's private information. If the student's name was publicly available, the adversary might have an even greater chance of success (e.g., if the student's given name was "Sue," the adversary could correctly guess the student's gender with probability close to 1). The example of gender might be contrived, but even in richer domains the adversary always has some probability of successfully guessing private student information, and thus winning the privacy game, without having any access to the results of computations performed on that private information. In these cases a privacy breach has not occurred, since no private information has been leaked; rather, the adversary has won by chance.



Adversary wins if the guess of the private information of a student equals the private information of that student.

Figure 5: Simplified FERPA game. Note that time progresses from top to bottom.

In Section 4.7.2, we further discuss how to account for the success probability an adversary may have even without getting any information from the computation.

The subsections that follow present a methodology for formalizing a game-based privacy definition for FERPA, provide justifications for the assumptions made, and the game-based formalization itself. The modeling attempts to capture a broad scope of potential interpretations of the privacy risks that FERPA is intended to protect against. As a result, any computation that can be proven to satisfy the resulting privacy definition can be used to release statistics regulated by the law with high confidence that it satisfies the requirements of the law.

4.1 A conservative approach to modeling

While the legal definitions found in statutes and regulations are typically more precise than language used in everyday interactions, they often involve some degree of ambiguity. On the one hand, this is an advantageous feature, since it builds some flexibility into legal standards and leaves room for interpretation, judgment, and adaptability, particularly as practices tend to evolve over time. But on the other hand, technological solutions that rely on formal mathematical models require the specification of exact, unambiguous definitions. This presents an issue: how can a legal standard such as the privacy protection required by FERPA be translated into precise mathematical concepts that a technology can be evaluated against?

As an example of ambiguity in FERPA, consider the disclosure of the social security numbers of students. FERPA considers a student's social security number to be personally identifiable information.[169] Assuming that a social security number does not constitute directory information, it would clearly be unacceptable to publicly release all nine digits of a student's social security number. It would also clearly be acceptable to release zero digits of the number, as no information about the social security number is leaked. But would it be acceptable to release five digits of a social security number, or six, or seven? At what point, between disclosing zero and nine digits, is

[169] 34 C.F.R. § 99.3.

the threshold between an acceptable data release and a prohibited data release?[170]

One could approach the problem of ambiguity by analyzing the text of the law and choosing a reasonable interpretation. However, what seems reasonable to one person might not seem reasonable to another. Judges, lawyers, and legal scholars frequently disagree when it comes to interpreting statutory text and applying it in various cases. Therefore, if a model of FERPA's privacy requirements relies on a particular interpretation of the law, this interpretation could be disputed by other legal experts in the future, especially in light of how standards are interpreted to evolve over time and adapt to new contexts. To address this concern, models should try to err on the side of a very conservative interpretation of the law's requirements. That is, whenever there is a choice between different interpretations, the more restrictive interpretation should be chosen. For instance, in our model, because we are uncertain about the potential capabilities of an adversary, we assume that the adversary is well-resourced and fully capable of conducting a sophisticated attack on the security of a system. This approach effectively strengthens the claim we are making that a given technical approach to privacy protection satisfies a particular legal requirement of security. If a system can be proved to be secure in strongly antagonistic circumstances, including in even unrealistically antagonistic circumstances, it is certainly secure given assumptions that are friendlier and more realistic.

In Figure 6 we provide a visual representation of this desidarata. There could be many possible interpretations of the privacy requirements of FERPA. In this figure, we view each interpretation as specifying a set of computations that it considers to provide privacy according to the law, as well as a set of computations that it considers to fail to provide sufficient privacy. If an interpretation is not stated with a mathematical language, then there are most likely also computations that fall into a gray area and are neither unambiguously endorsed nor rejected in that interpretation. A conservative approach to modeling the law attempts to identify those computations that fall unambiguously within the intersection of all (or, more realistically, a large class of) reasonable interpretations of the law. That is, if a conservative model considers a certain computation to provide privacy, then all these reasonable interpretations of the law would also consider it to provide privacy.

In the following sections we describe the model that we developed to capture the privacy desiderata of FERPA's regulators. While at many points in the modeling process we made conservative decisions, we should also note that there are places in our model where we make some assumptions that are not fully conservative, and we plan to extend our analysis to reflect a more fully conservative interpretation in future work. As we describe our model we will note those places where we make assumptions that are not fully conservative.

[170] This example is not a purely abstract one. Real-world regulators are struggling with this type of problem. For instance, guidance provided by the Department of Health and Human Services for de-identifying data in line with HIPAA's safe harbor clause states that in general "parts or derivatives" of identifiers cannot be released. However, the HIPAA regulators have decided that the first three digits of a ZIP code can be safely released without violating privacy if the populations of all ZIP codes that begin with those three digits sum to over 20,000 individuals. *See* HHS Office of the Secretary, Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule HHS.gov, http://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf (2012).

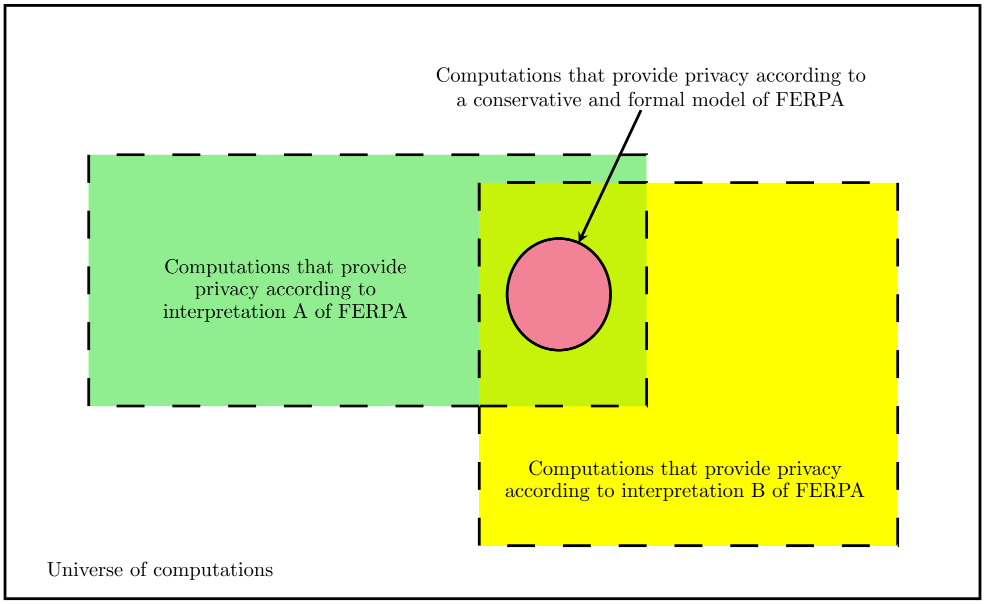

Computations that provide privacy according to a conservative and formal model of FERPA
Computations that provide privacy according to interpretation A of FERPA
Computations that provide privacy according to interpretation B of FERPA
Universe of computations

Figure 6: A conservative approach to modeling. The outer box delineates the universe of all computations. The inner boxes delineate computations that are considered to provide privacy according to particular interpretations of FERPA. We draw the borders of these boxes with dashed lines, since at the "edge" of a (non-formal) interpretation there are often computations that are not clearly accepted or rejected by that interpretation. A conservative approach to modeling attempts to identify those computations that fall safely within the intersection of all reasonable interpretations. In a formal model there is no ambiguity about whether a computation provides privacy by its standards, so we can draw its border with a solid curve.

4.2 Modeling FERPA's implicit adversary

FERPA does not specify an explicit model of the adversary it is intended to thwart. By this we mean neither the statutory nor regulatory text specifies what the regulators considered to be the capabilities of a person who should fail when attempting to defeat measures taken to protect the personally identifiable information held in education records. Despite the lack of an explicit description of the adversary they envisioned, the regulators left some evidence as to the types of attacks and attackers they had considered when drafting the text of the law. In particular, FERPA's definition of *personally identifiable information* provides a description of what or whom the law is designed to protect against. In this section, we argue that this definition and how it has been interpreted by the Department of Education provides useful details that can serve as a basis for modeling the implicit adversary the regulators had in mind when they formulated FERPA's requirements.

FERPA prohibits the disclosure of non-directory personally identifiable information from education records, except with the consent of the student or her parent or guardian, or in accordance with one of the limited exceptions set forth by FERPA.[171] We might naturally ask how the regulators who drafted FERPA originally envisioned such an improper disclosure occurring. In our investigation of this question, we are particularly interested in the case in which a school or educa-

[171] See 34 C.F.R. § 99.30.

tional agency releases information pursuant to the exception to FERPA permitting the release of de-identified data.[172] This exception provides that

> [a]n educational agency or institution, or a party that has received education records or information from education records . . ., may release the records or information without the consent required . . . 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.[173]

To qualify under this exception, the released data must not contain non-directory personally identifiable information. As discussed in Subsection 3.2.3, FERPA defines personally identifiable information to include direct and indirect identifiers, as well as

> 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.[174]

By the explicit inclusion of this quoted language, the drafters emphasized the concern that a member of the school community might be able to learn non-directory information about a student from otherwise de-identified data, and identified this as a category of privacy breach that must be addressed. Because as described in Subsection 3.2.1 this is also the category of privacy breach most relevant to the use of formal privacy methods, we take the "reasonable person in the school community, who does not have personal knowledge of the relevant circumstances" to be the implicit adversary embedded within FERPA's requirements. But who is considered to be a "reasonable person in the school community," and what sort of knowledge does such an individual have?

The rulemaking history for FERPA provides some limited guidance on these questions. As noted in Section 3.2, a "reasonable person" is described in the preamble to the amendment as a "hypothetical, rational, prudent, average individual."[175] Furthermore, beyond enjoying the insider information that comes from being a member of the "school community," this individual is "also presumed to have at least the knowledge of a reasonable person in the local community, the region or State, the United States, and the world in general."[176] The authors of the commentary explain that this assumption is intended to provide "the maximum privacy protection for students."[177] On the other hand, since the adversary does not have "personal knowledge of the relevant circumstances," the adversary must have some uncertainty about the student information.

[172] 34 C.F.R. § 99.31(b)(1).

[173] *Id.*

[174] 34 C.F.R. § 99.3 (emphasis added). Note that FERPA's definition of personally identifiable information also includes "[i]nformation requested by a person who the educational agency or institution reasonably believes knows the identity of the student to whom the education record relates." *Id.* We do not extensively address this part of the definition in our privacy model. We note, however, that the model we derive for FERPA guarantees that, no matter the adversary's *a priori* knowledge about a student, that knowledge does not improve much by seeing the result of a computation that meets our definition of privacy. That is, even if the adversary believes that an aggregate statistic reflects private information of a particular student, the adversary can learn essentially nothing specific to that student from that statistic.

[175] 73 Fed. Reg. 74,806, 74,832 (Dec. 9, 2008).

[176] *Id.*

[177] *Id.*

4.3 Modeling the adversary's knowledge

The regulatory language therefore suggests that a "reasonable person in the school community" brings with her some knowledge. We model this knowledge of the potential adversary of the system in the privacy game we construct. This modeling will affect the adversary's success probabilities in identifying a student both with and without access to the computation.[178]

We choose to model the adversary's knowledge about students as probability distributions over student attributes.[179] We believe that probability distributions provide a sufficiently rich language to model the type of knowledge the regulators envisioned the adversary having about students. For instance, distributions can describe statistics that the adversary knows about the entire student body (e.g., demographic information), as well as beliefs that the adversary might hold about a particular student. For each student, we presume that the adversary has some *a priori* beliefs about that student, which we represent as a probability distribution. That is, in our model each individual student is associated with a probability distribution that represents the adversary's beliefs about the non-directory personally identifiable information of that student.[180] The following examples demonstrate the versatility of probability distributions for modeling the adversary's knowledge.

Probability distributions can be used to describe general knowledge about the demographics of a school district. Say that there is a school district with two schools, A and B. 95% of the students in A are low-income, while only 40% of the students at B fall into this category. Furthermore, say that 35% of the students in A and 15% of the students in B scored below proficient on a statewide assessment. Without knowing anything else about her, if the adversary knows that Alice attends school A, he might believe that there is a 95% chance that Alice is from a low-income family and a 35% chance that she scored below proficient on the exam. On the other hand, the fact that Grace attends B might lead the adversary to believe that there is a 40% chance that she is from a low-income family and a 15% chance that she scored below proficient on the exam. We can model this general knowledge with two distributions. A student sampled randomly from the first distribution has a 95% chance of coming from a low-income family and a 35% chance of scoring below proficient on the assessment; a student sampled randomly from the second distribution will fall into each of these categories with a likelihood of 40% and 15%, respectively. Furthermore, distributions can be tailored to reflect more complex beliefs about trends across demographics. For instance, the two distributions described above could be reworked to reflect the adversary's belief that a student from a low-income family has a much greater chance of scoring below proficient on the examination than a student who is not from a low-income family. For example, consider the distributions given in Table 1. These distributions still reflect the fact that 95% of the students at A are low-income and

[178] Note that while having more knowledge makes it easier for the adversary to win the privacy game, having more knowledge also increases the baseline probability of success without access to the computation. While the former makes violating privacy easier (presumably, having more knowledge makes it easier to identify a student), the latter makes violating privacy harder (the baseline success probability is higher). It follows that an adversary having more knowledge does not necessarily make privacy protection harder. To see why this (somewhat counterintuitive) argument holds, consider the extreme case where an adversary has complete knowledge of all the students' personal information. The adversary's success in identifying a student should not in this case be considered a failure to protect privacy.

[179] Intuitively, a probability distribution describes the properties of random members of a population. Probability distributions are often used to model uncertainty in elements of a population.

[180] We assume that student attributes are drawn independently from the distributions; that is, the fact that one student has certain attributes does not affect the probability of another student having particular attributes. This is a limitation of our model, and will be discussed later in this section and in Section 4.10.

	Distribution describing student from A		Distribution describing student from B	
	Low-income	Not low-income	Low-income	Not low-income
Scored proficient or higher	0.601	0.049	0.26	0.59
Scored below proficient	0.349	0.001	0.14	0.01

Table 1: Distributions describing students at A and B. Numbers are given as probabilities. For instance, if we were to sample randomly from the distribution describing a student from A, the sampled student scored at least proficient on the exam and is from a low-income family with a probability of 0.601. On the other hand, the probability that the sampled student scored below proficient and is not from a low-income family is only 0.001.

40% of the students at B are low-income, and that 35% of students at A scored below proficient on the exam and 15% at B scored at this level. However, the distributions now also reflect the fact if a student (from either school) is from a low-income family, that student scored below proficient on the exam with a likelihood of about 35%, whereas if a student is not from a low-income family, the likelihood that the student scored below proficient is only about 2%.

In addition to describing beliefs based on demographic information, distributions can also reflect more specific beliefs about individuals. This can be used to model the case where the adversary has a substantial amount of outside knowledge about a particular student. For instance, the adversary might have heard that a student, Ashley, has a fiery temper. This knowledge might lead the adversary to believe that there is a greater probability that Ashley has been disciplined at school than the average student. Additionally, say that the adversary is aware that Ashley's mother is a professor of mathematics. Because of her family environment, the adversary might believe that Ashley is more proficient at math than the average student and so has a greater chance of having passed the state math exam. We can model the adversary's beliefs by associating with Ashley a distribution that describes a student that has a higher than average probability of having been disciplined and a higher than average probability of having passed the math exam.

One important limitation of our current modeling is that we assume that the characteristics of a given student are independent of the attributes of all other students. This means that we do not model *correlations* between students. We might expect to see such correlations in the real world. For example, there might be correlations between a small number of students, such as siblings. If one sibling comes from a low-income family, then all the other siblings should also have this characteristic. The best we can currently do in our model is to give each sibling an equal probability of coming from a low-income family. There might also be correlations among larger groups of students. For example, it might be the case that if at least 30% of the students in the school passed the state math examination, then a particular student Alfred must have passed the exam. Alfred's personal attribute (i.e., whether he passed the exam) is dependent on the attributes of all the other students in his school (i.e., what percentage of them passed the exam). While our current model does not account for correlations between students, in Section 4.10 we discuss some ways that we could address this limitation in future work.

4.4 Modeling the adversary's capabilities and incentives

FERPA's rulemaking history also hints at the capabilities of the adversary. Specifically, the Department of Education has explained that the "reasonable person" standard "was not intended to describe the technological or scientific skill level of a person who would be capable of re-identifying statistical information or redacted records."[181] From this, we understand that the regulators make no assumption about the adversary's ability to re-identify students from "de-identified" data. Accordingly, we do not assume anything about the skill level of the adversary in our model—we make no assumptions about the analysis the adversary can perform, and about the computational resources available to her.[182] Furthermore, we make no assumptions about the motivation of the adversary.

Modern cryptography emphasizes the design of cryptographic systems that are secure not only against known attacks, but also against attacks that have not been known at the time the cryptographic system was designed. Predating this emphasis, cryptographic techniques typically only addressed known privacy threats and were often shown to be vulnerable to other, newly developed attacks. The failure of existing cryptographic methods to provide privacy against new attacks would motivate the development of novel cryptographic techniques, but these techniques in turn could be broken by even more sophisticated attacks. To break this cycle, modern cryptographers strive to create cryptographic methods that are guaranteed to be invulnerable to any attack that is feasible under a general computational model. Likewise, by not making assumptions about the capabilities of the adversary in our modeling (beyond adhering to a general computational model), we ensure that our model is robust not only to currently known privacy attacks, but also to attacks developed in the future.

The rulemaking history of FERPA also supports this approach. The Department of Education has noted the increasing capabilities of attackers. For instance, the Department recommends limiting the extent of information publicly released in directories, with the justification that "since the enactment of FERPA in 1974, the risk of re-identification from such information has grown as a result of new technologies and methods."[183] Furthermore, FERPA does not contemplate only one type of attack. While traditional re-identification attacks against microdata are emphasized, implementation guidance also addresses exploits aimed at learning private student information from aggregate data.[184] Because the regulators do not state the full scope of the types of attacks that they are considering, a conservative approach is to assume that our model must account for *any* attack that can be perpetuated by an adversarial observer.

In contrast with our approach, some frameworks for measuring the privacy risk of a data release explicitly require making assumptions about the motives and capabilities of the adversary. For instance, some experts suggest that estimates of the motives and capabilities of a potential adversary should be used as an input when calculating the re-identification risk of a data release.

[181] *Id.* at 74,831-32.

[182] We note, however, that the our modeling does not focus on the security of the particular implementation. I.e., we assume that the adversary only tries to learn private information through observing a data release. We are not concerned in this work with an adversary who might steal a hard drive containing sensitive data, or hack into the system to inspect its memory, or actively manipulate it. We believe that this modeling is consistent with the privacy threats envisaged by the regulators in the context of releasing anonymized data.

[183] 73 Fed. Reg. at 74,834.

[184] National Center for Education Statistics, Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting, SLDS Technical Brief: Guidance for Statewide Longitudinal Data Systems (SLDS), https://nces.ed.gov/pubs2011/2011603.pdf (Dec. 2010).

If it is believed that any potential adversary will have limited resources to apply towards a privacy attack or little incentive to attempt an attack, the risk of re-identification is assumed to be small, so the data can presumably be safely released with fewer protections.[185] We take the more conservative stance that the adversary is fully capable of any privacy attack and fully incentivized to attack.

4.5 Modeling student information

Like the regulations, in our model we draw a distinction between directory information and non-directory personally identifiable information. Because directory information can be made publicly available, our model assumes that the adversary has access to it.[186]

Non-directory personally identifiable information is private information generally not available to the adversary, although the adversary might hold some *a priori* beliefs about it. As mentioned in the previous subsection, we model these beliefs via probability distributions over student attributes.

4.6 Modeling a successful attack

FERPA regulates the non-consensual disclosure of non-directory PII. According to the regulations, disclosure "means to permit access to or the release, transfer, or other communication of personally identifiable information contained in education records by any means, including oral, written, or electronic means, to any party except the party identified as the party that provided or created the record."[187] Therefore, a data release by a school should not communicate any non-directory personally identifiable information to the recipient or user of the released information. For these reasons, we say that the adversary wins the privacy game if she is able to correctly guess non-directory personally identifiable information of a particular student after seeing the output of a computation performed on that private data.

More precisely, the adversary wins the game if she successfully guesses a function of the private information of a student. This is a more conservative approach than only considering the adversary to have won the game if the adversary guesses the exact private information, and models the fact that often learning (or inferring) something about private information could be a breach of privacy, even if the private information itself is not learned directly or completely. For instance, consider a scenario where the private information of a particular student is his numerical test score (in a range from 0 to 100), which happens to be a 54. If the adversary is able to learn from a data release that the student failed the test, we consider the adversary to have won the privacy game, even if the adversary has not learned the student's exact numerical score.

Commentaries on the regulations stress that a data release must preserve uncertainty about private student information. As discussed in Section 3.2.3, many states prohibit reporting that 100% of students achieved certain performance levels, as academic performance is a non-directory attribute that can be linked to specific students through this data release.[188] As a more nuanced example, consider the release of a table that contains information about student test scores. The table distinguishes between native English speakers and English language learners, and gives the number of students from each category who scored at a below basic level, a basic level, a proficient

[185] *See* Khaled El Emam & Luk Arbuckle, Anonymizing Health Data ch. 2 (2013).

[186] Note that, in reality, some directory information may only be available to certain members of the school community, such as teachers and parents, not the general public.

[187] 34 C.F.R. § 99.3.

[188] *See* 73 Fed. Reg. at 74,835.

level, and an advanced level. Say that the table reports that one English language learner achieved a proficient score on the exam and the other nine English language learners scored at either a below basic or basic level. This constitutes a nonconsensual disclosure of private information, as the sole English language learner who achieved a proficient score learns that her peer English language learners scored at one of the lower levels.[189]

For similar reasons, guidance to de-identifying data protected by FERPA suggests that very low or high percentages should be "masked" when reported. For example, consider a statistic reporting what percentage of a class of 41 students scored at the proficient level on an exam, where only one student, Siobhan, scored at the proficient level. Guidance from the National Center for Education Statistics recommends reporting that $\leq 5\%$ scored at a proficient level, instead of the true percentage, which is around 2.5%. If the true percentage were reported, Siobhan would learn that all her peers failed to achieve this level, since she knows that she accounts for the entire 2.5% of students that was reported. However, if the "masked" percentage is reported, Siobhan would no longer be able to make this conclusion with certainty, since another student could have scored at her level (5% of 41 is slightly more than 2). As there is only a 1/40 probability that a given student besides Siobhan scored at a proficient level, Siobhan might be able to form a strong belief about the performance of each of her peers, but she has not learned with certainty the performance level of any particular classmate. Consequently, the data release is considered to preserve uncertainty about private student information.[190]

We need to model the regulators' desire that uncertainty about private information is preserved during a data release. We do so by requiring a stronger property: we say that a system provides privacy only if gaining access to the system does not change the adversary's beliefs about the private information of students very much. In our model, the adversary holds some *a priori* beliefs about the non-directory personally identifiable information of each student, and a system provides privacy if the adversary's beliefs after seeing a data release are about the same as his *a priori* beliefs. Specifically, we allow the adversary's beliefs to change by a multiplicative factor that is constrained to a small range.[191] For instance, we might say that the adversary's beliefs should only change by a factor of between 0.9 and 1.1. This means that if the adversary holds the *a priori* belief that a particular student failed an exam with a likelihood of 40%, he might believe after seeing the output of a privacy-providing computation that the student actually failed with a likelihood of 44%; his belief about the performance of that student has not changed very much. Our approach also preserves uncertainty when the initial uncertainty is small. If the adversary initially believes that a student failed with a likelihood of 99%, we guarantee that the output of a privacy-providing computation would not lead him to believe that the student actually failed with a likelihood of more than 99.1%. Intuitively, this is because not only is the adversary's belief that the student *failed* the exam allowed to only change by a small amount, but the adversary's belief that the student *passed* the exam is also only allowed to change by a small amount. Since his initial belief was that there was a 1% chance that the student passed the exam, this belief can only change by a multiplicative

[189] This example is adapted from Example 1 in NATIONAL CENTER FOR EDUCATION STATISTICS, STATISTICAL METHODS FOR PROTECTING PERSONALLY IDENTIFIABLE INFORMATION IN AGGREGATE REPORTING, SLDS Technical Brief: Guidance for Statewide Longitudinal Data Systems (SLDS), https://nces.ed.gov/pubs2011/2011603.pdf (Dec. 2010), at 5.

[190] *Id.* at 27-29.

[191] More formally, this multiplicative factor is captured by the parameter ε (epsilon), which is typically a small constant. The adversary's belief is allowed to change by a factor of $e^{\pm\varepsilon}$. For small epsilon, $e^{\varepsilon} \approx 1+\varepsilon$ and $e^{-\varepsilon} \approx 1-\varepsilon$. The epsilon parameter can be tuned to provide different levels of privacy.

factor between 0.9 and 1.1, which means that the computation output will not lead the adversary to believe that the student passed the exam with a likelihood of less than 0.9%.

4.7 Towards a FERPA privacy game

In the following subsections we give an informal description of two *privacy games* and corresponding *privacy definitions* that fit the model that we have extracted based on FERPA's requirements for protecting privacy in releases of education records.[192] We first discuss a game that reflects a *targeted* attack, where the adversary is attempting to learn information about a particular student. To capture the targeted character of the attack, in this scenario, the adversary commits to attacking a specific student *before* seeing the output of a computation performed on private student data, and then can only attempt to guess private information about that one student. We then discuss a game that reflects an untargeted, *fishing* attack, where the adversary can try to guess the private information of *any* student. In this scenario, the adversary sees the computation output and is then able to *adaptively* choose which student to target. We distinguish between these two scenarios to provides some insight into how the model we extracted from FERPA can be adapted to address more specific privacy concerns. It can be shown that in our formalization the untargeted scenario provides a strictly stronger privacy guarantee than the targeted scenario. In other words, every computation that is considered to provide privacy against an untargeted attack also provides privacy against a targeted attack, while the opposite does not necessarily hold.[193]

Before we do so, we need to tie up two loose ends in our model. Earlier in this section, we described how student information is split into directory information and non-directory personally identifiable information. However, we did not describe the contents of the student information. Secondly, in the previous subsection we made the argument that the adversary wins the game if he is able to correctly link a student's identity with a function of that student's sensitive information. However, there is always some baseline probability that the adversary is able to do so, even without receiving the output of the computation. How can we account for this fact as we attempt to define whether a computation provides privacy?

4.7.1 Accounting for ambiguity in student information

Before we can use our game, we need to define what constitutes directory information and non-directory personally identifiable information. The adversary will have direct access to the directory information and will have some indirect knowledge of the private information. The computation will use both types of information to produce some output.

There is a degree of ambiguity in the regulatory definitions of directory information and non-directory personally identifiable information. The regulations give a non-exhaustive list of types of

[192] For a more formal treatment of the games and definitions, please see the Appendix.

[193] To illustrate how privacy against a fishing attack can be more restrictive consider a dataset containing the information about 1000 students. In an attempt to produce information about the dataset while providing privacy, a computation picks one of the students at random and exposes all her information. This computation may seem to provide reasonable privacy with respect to a targeted attack, as an attacker committing to a specific student before the computation is performed will have a chance of merely 1 in 1000 to gain sensitive information about that student from the outcome of the computation. On the other hand, the computation fails to provide privacy with respect to an attacker that can choose which student's sensitive information to guess adaptively, i.e., after seeing the outcome of the computation, as such an attacker will succeed in producing a correct guess with probability 1.

directory information,[194] and in fact each educational agency or institution is granted discretion in determining what to classify and release as directory information. Seeking generality, we do not make any assumptions in our model about the content of directory information. Instead, *we allow the adversary to decide what information is published as directory information.* While this may seem to be a highly unintuitive and unrealistic modeling choice—it might be unlikely that any adversarial agent would have this ability in the real world—this modeling decision helps to establish a privacy requirement that will be robust both to potential interpretations of FERPA and to the differences between specific settings under the mandate of FERPA. This modeling also exhibits conceptual clearness that we believe is instrumental in understanding intricate concepts like privacy. To recap this modeling decision, we are effectively allowing the adversary to choose directory information that is the worst-case scenario for the privacy of the system. If we are able to prove that the system provides privacy even in this worst-case scenario, then we will have confidence that it provides privacy no matter what the directory information actually is.

Similarly, there is uncertainty about exactly what information constitutes non-directory personally identifiable information. From the definition of personally identifiable information, we know that the system must protect information that "is linked or linkable to a specific student."[195] However, because we do not want to make any assumptions about the capabilities of the adversary or the methods he might use to identify a student in released records that have been stripped of personally identifiable information, it is impossible to say what information falls into these categories. Furthermore, we do not want to make any assumptions about the auxiliary knowledge that the adversary may have about students. Accordingly, as with the directory information, we allow the adversary to have a say about the non-directory personally identifiable information of the students he is trying to attack. More precisely, for each student in the dataset, we allow the adversary to choose the distribution over student attributes that models that student's non-directory personally identifiable information.

For example, the adversary could choose the student information presented in Table 2. The directory information consists of the name, age, and zip code of three students. The adversary picks concrete values for these attributes. The private attributes are whether a student has a learning disability and whether a student has been suspended in the last year. For each student, the adversary assigns a probability of each combination of these attributes being true. For instance, the adversary chooses that Samuel Strudwick, a seventeen-year-old student from zip code 00034, has a 10% likelihood of both having a learning disability and having been suspended in the last year.

4.7.2 Accounting for the adversary's baseline success

Recall the cryptographic privacy game from Section 2.2, in which one of two plaintext messages is encrypted, the adversary is given the resulting ciphertext, and the adversary must then identify which of the two original plaintext messages is behind the ciphertext. Since each message was encrypted with a probability of 0.5, the adversary can win the game with a probability of 0.5 by just guessing randomly between the two messages, without even examining the ciphertext. This is the adversary's baseline success. Even if a "perfect" cryptographic computation was used to encrypt the message, the adversary can still be expected to win the game 50% of the time.

[194] 34 C.F.R. § 99.3.

[195] 34 C.F.R. § 99.3.

Directory information			*Private information*			
Name	Age	Zip code	Disability & suspended	Disability & not suspended	No disability & suspended	No disability & not suspended
Robi McCabe	18	00034	0.3	0.3	0.2	0.2
Launo Cooney	17	00035	0.2	0.2	0.3	0.3
Samuel Strudwick	17	00034	0.1	0.2	0.3	0.4

Table 2: Example student information chosen by the adversary. Name, age, and zip code are the attributes given as directory information. The private information for each student is represented by a probability distribution describing the probabilities of all combinations of the following two binary attributes: whether that student has a learning disability, and whether that student has been suspended in the last year.

Similarly, it is unreasonable to expect that the adversary will never win the FERPA privacy game that we propose, as there is always the possibility that the adversary guesses correctly by chance. For example, consider that Siobhan attends a school where 50% of the students scored below proficient on the state reading exam. Without any other knowledge, the adversary might guess that Siobhan scored below proficient on the test and would have a 50% chance of being correct, provided that each student scored below proficient with an equal probability.

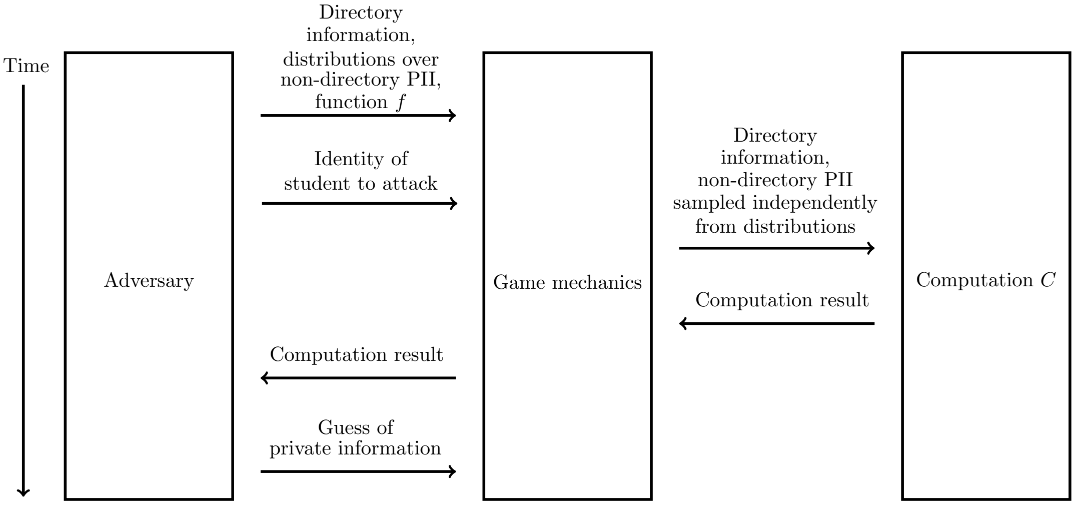
A system can still be considered to provide privacy even if the adversary wins the game with some likelihood (e.g., due to chance). To see why this is the case, consider a computation C that outputs the word "aardvark" on every input. Because the outcome of C does not depend on the input, it provides perfect privacy. Suppose that the adversary knows that 80% of the students in a school have a learning disability. If C is performed on the private student data, the adversary receives only the output "aardvark," which provides her with no useful information for learning the private information about students (i.e., whether that student has a learning disability). However, if the adversary guesses that a given student has a learning disability, she will win the game with a probability of 0.8. The fact that the adversary wins with a high probability should not be considered in contradiction to C providing perfect privacy, as the adversary could win the game with the same probability without access to the outcome of C.

To account for the adversary's baseline chance of success, we say that a computation provides privacy if the probability of the adversary winning the game when she has access to the system is not much greater than the probability of her successfully guessing sensitive information without having access to the system. This approach closely mirrors the approach taken in Section 3.1. By this standard, the "aardvark"-producing computation C from the previous example is considered to provide privacy, since the adversary's probability of winning the game has improved not one iota by having access to the output of this computation.

4.8 The targeted attack scenario

This game represents a scenario in which the adversary is committed to trying to learn non-directory personally identifiable information about a particular student. For instance, consider the case of a local journalist trying to learn the private information of a school's star basketball player from a data release about school disciplinary actions. The journalist only cares about learning the information of this one student.[196] We illustrate the privacy game for the targeted attack scenario in Figure 7.

4.8.1 Mechanics


Time
Adversary
Directory information, distributions over non-directory PII, function f
Identity of student to attack
Computation result
Guess of private information
Game mechanics
Directory information, non-directory PII sampled independently from distributions
Computation result
Computation C

Adversary wins if guess equals result of applying function f to targeted student's private information.

Figure 7: Targeted FERPA privacy game.

In line with a conservative approach to modeling as described in Section 4.1, we allow the adversary to choose a directory of public student information. The adversary assigns to each student a probability distribution that describes the adversary's *a priori* beliefs about the non-directory personally identifiable information for that student, and also chooses a function f whose domain is private student information. Intuitively, f represents the aspect of private student information that the adversary is interested in learning about. For example, consider the situation where the private information of a student is his numerical test score. The adversary might want to learn whether some student Bill passed or failed an exam, but might not care about the exact score that Bill earned. In this case, the adversary can choose a function *processScore* that takes Bill's numerical score and outputs "passed" if he earned a passing score and "failed" otherwise.

The adversary passes the directory, the probability distributions over student information, and the function f to the game mechanics. Additionally, the adversary chooses a particular student to attack (i.e., the adversary commits to trying to learn the non-directory PII of that student, and that student alone), and passes the identity of this distinguished student to the game mechanics.

The game mechanics instantiates a table of non-directory student information by making a random draw from each of the probability distributions chosen by the adversary. This table, along

[196] This scenario is related to subsection (g) of FERPA's definition of personally identifiable information, which includes "[i]nformation requested by a person who the educational agency or institution reasonably believes knows the identity of the student to whom the education record relates." 34 C.F.R. § 99.3. In both cases, the contemplated adversary is trying to learn information about one particular student. However, in our model the adversary does not request education records; instead, the adversary observes the result of a statistical computation performed on education records.

with the directory information, is given to some computation C, resulting in an output statistic. Note that C is not supplied the identity of the student that the adversary is attacking or the function that the adversary has chosen.

The game mechanics passes the result of C to the adversary. Based on this result, the adversary reports a guess about some aspect of the student's non-directory personally identifiable information. The game mechanics declares that the adversary has won if the guess matches the result of applying f to the distinguished student's non-directory personally identifiable information; otherwise, the game mechanics declares that the adversary has lost. To continue the example from above, the adversary will guess either "passed" or "failed", and will win if her guess matches the result of applying $processScore$ to Bill's test score.[197]

4.8.2 Privacy definition

In this scenario, we say that a computation C provides privacy if the probability that the adversary correctly guesses a function of the student's non-directory personally identifiable information after seeing the output of C is roughly equal to the probability that the adversary would have correctly guessed based only on knowing the probability distribution from which the student's non-directory personally identifiable information is sampled, and not having seen the output of C. That is, the adversary has some *a priori* belief about the non-directory personally identifiable information of the distinguished student. Based on this belief alone, the adversary can make a guess about the student's non-directory personally identifiable information. A computation C provides privacy if the probability that the adversary's guess is correct after seeing the output of C is about the same as the probability the adversary's guess is correct without seeing the output of C.

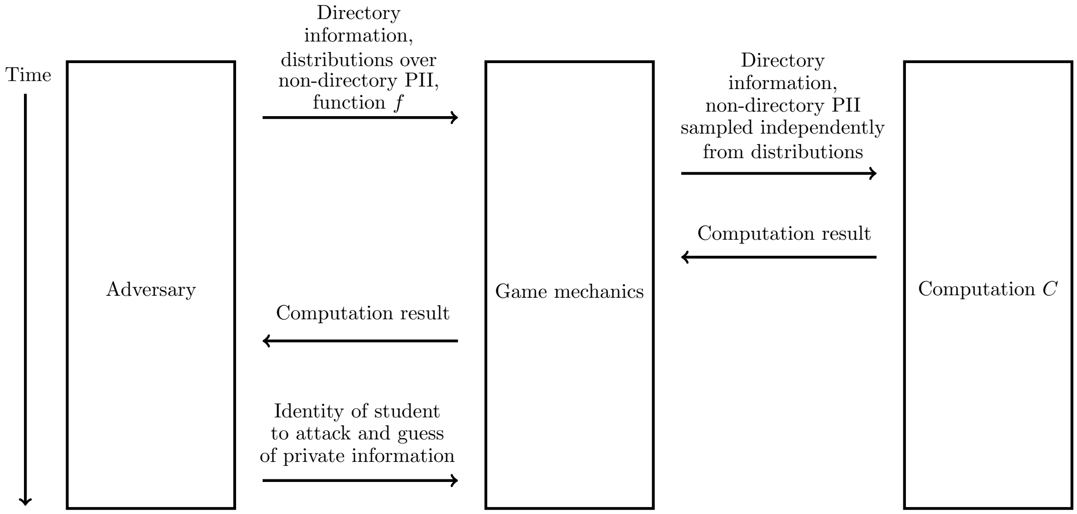
4.9 The untargeted attack scenario

In this scenario, the adversary is interested in learning the non-directory personally identifiable information of any student in the dataset, as shown graphically in Figure 8. Here the attacker does not need to decide which student to attack before seeing the output of the computation. This scenario could correspond to a data broker mining a dataset for private student information. The data broker is not committed to targeting a particular student; rather, it is hoping to learn private information about any student in the dataset.

4.9.1 Mechanics

As before, the adversary chooses a directory of public student information and associates with each student in that directory a probability distribution that describes his beliefs about the non-directory personally identifiable information of that student, and also chooses a function f whose domain is private student information. Unlike the previous game, at this point the adversary does not commit to trying to guess the non-directory PII of any particular student. Instead, the adversary passes just the directory and the probability distributions to the game mechanics.

[197] The game mechanics' declaration of whether the adversary won or lost is only for purposes of the privacy definition (as a proxy for measuring the adversary's success rate) and does not correspond to an actual "real world" declaration. In fact, such a declaration would itself constitute a privacy breach (e.g., if the adversary has multiple chances to guess).


Time
Adversary
Game mechanics
Computation C
Directory information, distributions over non-directory PII, function f
Directory information, non-directory PII sampled independently from distributions
Computation result
Computation result
Identity of student to attack and guess of private information

Adversary wins if guess equals result of applying function f to chosen student's private information.

Figure 8: Untargeted FERPA game.

The game mechanics builds a table of non-directory student information by sampling randomly from each of the probability distributions given by the adversary and passes this table, with the directory, to a computation C, which outputs some result.

The game mechanics then gives this result to the adversary. After seeing this result, the adversary chooses a student to try to attack, and makes a guess about an aspect of that student's non-directory personally identifiable information. The adversary gives the identity of this student and the guess to the game mechanics, who declares that the adversary has won if his guess matches the result of applying the function f to that chosen student's non-directory personally identifiable information. Otherwise the adversary is considered to have lost the game.[198]

4.9.2 Privacy definition

We say that a computation C provides privacy if the probability that the adversary correctly guesses any student's non-directory personally identifiable information after seeing the output of C is roughly equal to the probability that the adversary would have correctly guessed some student's private information based only on knowing the probability distributions from which the students' non-directory personally identifiable information is sampled, and not having seen the output of C.

4.10 Discussion of games

As explained previously, a privacy game is a device for formalizing privacy desiderata so that we can prove that a computation provides a sufficient level of privacy. Some of the modeling choices we have made might not seem realistic, such as allowing the adversary to choose what constitutes directory information. While it is unlikely that a real-world adversary would have this type of ability, we emphasize that we make this choice to account for the fact that there is uncertainty in the regulations. To avoid making an assumption that might not fully reflect the regulators'

[198] *See* explanation *supra* note 197.

privacy desiderata, we assume a powerful adversary. This means that if a computation is proven to provide privacy in the extremely antagonistic environment of our formalization, it should also provide privacy in a more realistic scenario.

In a sense, we have attempted to fill in gaps in the guidance on complying with FERPA, where the regulators have not provided explicit details regarding whether certain types of attacks or attackers fall within the scope of the regulations. For instance, the regulators refer to attacks beyond re-identification attacks, but do not explicitly state the extent of the other types of attacks they are contemplating. We account for this uncertainty in our model by not limiting the type of attack the adversary might attempt. Similarly, the regulations do not clearly state the nature of directory information and non-directory personally identifiable information, so we allow the adversary to choose what constitutes this information and how student attributes are distributed. In addition, the guidance is not clear regarding how much an observer may permissibly learn about a specific individual from a data release, while in our model we can quantify exactly how much an observer's knowledge can improve given access to the system in a worst-case scenario.

Because we have chosen to formalize FERPA primarily as a demonstration of our methodology, we have made some simplifications in our model that would need to be addressed in a more complete formalization of the regulations. As noted in Section 4.3, we treat the attributes of a student as being independent of the attributes of any other student. This does not fully capture reality, as sometimes the fact that one student has some attribute makes it more likely that another student has a certain attribute. For instance, if a student has a learning disability, it is highly likely that his identical twin also has that disability. Our analysis could be modified to protect privacy when there is dependence between members of small groups of students. In addition to correlations between a few individuals such as family members, there could be global correlations among the students. Our model does not account for these correlations. For example, the adversary might have the auxiliary knowledge that a particular student Greg's exam score is 5 points higher than the average exam score. If Greg's school releases the average grade on the exam (or even a noisy version of the average), the adversary is able to infer Greg's performance. In general, it is not possible to provide non-trivial inferential privacy guarantees when the adversary has arbitrary auxiliary information about correlations concerning members of a dataset.[199] However, guarantees can be given when the adversary's knowledge about these correlations meets certain conditions.[200] Therefore, one possible direction for addressing this issue in our model would be to allow the adversary to construct limited types of correlations between students. In this case, we would still be able to give privacy guarantees when students are not independent, provided that the correlations between the students fit the ones allowed in the model. We intend to explore the suitability of this direction in future work.

The regulations also require that, before de-identified data can be released, a determination must be made that "a student's identity is not personally identifiable, whether through single or *multiple* releases."[201] Our formalization currently only accounts for a single release: that is, we assume that the adversary only has access to the output of a single computation performed on the private data. [We will address solutions to this issue more fully in the final version.]

In our privacy definitions, we have stated that the adversary should not be able to guess private student information after seeing the output of the computation performed on the private data with

[199] *See* Cynthia Dwork & Moni Naor, *On the Difficulties of Disclosure Prevention in Statistical Databases or The Case for Differential Privacy*, 2(1) Journal of Privacy and Computation 93 (2010).

[200] *See* Arpita Ghosh & Robert Kleinberg, *Inferential Privacy Guarantees for Differentially Private Mechanisms*, Working Paper (2016).

[201] 34 C.F.R. § 99.31(b)(1) (emphasis added).

much more success than if she had not seen the computation result. A parameter captures exactly how much better the adversary is allowed to do. How to set the parameter to a level that represents an acceptable privacy loss is the subject of ongoing research.

5 Proving that differential privacy satisfies FERPA

[Recall the guarantee made by differential privacy: the output of a computation over a dataset is roughly the same as the output of the same computation performed on a neighboring dataset (i.e., one like the original except with at most one row of data modified). Intuitively, this means that essentially nothing can be learned about an individual in a dataset from the output of a differentially private computation performed on that dataset that could not be learned from the output of the same computation performed on the dataset with that individual's information arbitrarily changed. In this section we will give some intuition for why every differentially private computation meets the definitions for privacy we have extracted from FERPA. We will include formal proofs of this claim in the appendix.]

5.1 The targeted scenario

[It can be shown that, when the attributes of members of a dataset are drawn independently from known distributions, differential privacy provides inferential privacy: that is, given that the attributes of individuals in a dataset are drawn independently according to prior distributions known to the adversary, the adversary's beliefs about every individual in the dataset after seeing the output of a differentially private computation performed on that dataset are about the same as his *a priori* beliefs.[202] Therefore, because in our model we assume that student attributes are drawn independently from distributions known to the adversary, we can conclude that the output of a differentially private computation performed on student data will not allow the adversary to update his beliefs about any student very much. In the case of the targeted scenario, this means that the adversary's beliefs about the targeted student do not change very much by seeing the computation output. In turn, when it comes to guessing that student's private information, the adversary can essentially do no better than guessing based on his *a priori* beliefs. This meets the definition of privacy given in our formalization of the targeted scenario.]

5.2 The untargeted scenario

[As explained in our discussion of the targeted scenario, differential privacy provides inferential privacy in the context of our model. Since the inferential privacy guarantee extends to all members of a dataset, the adversary's belief about *any* student in the dataset after seeing the output of a differentially private computation are essentially the same as his prior beliefs. Therefore, the probability of him successfully guessing private student information is roughly the same as the probability of him doing so based only on his prior knowledge about the students. This meets the definition of privacy given in our formalization of the untargeted scenario.]

[202] For the more general case, see Arpita Ghosh & Robert Kleinberg, *Inferential Privacy Guarantees for Differentially Private Mechanisms*, Working Paper (2016).

6 A general methodology

[In this section, we will summarize the methodology we have used to formalize FERPA and prove that differential privacy meets the privacy requirements of the law. We will also claim that this methodology generalizes and can be applied to other technologies besides differential privacy and other regulations besides FERPA.]

6.1 A methodology for bridging legal and technical approaches

[In brief, our methodology involves conservatively constructing a formal privacy game that captures the privacy desiderata of the regulators. With such a game in hand, we can prove that certain computations (or classes of computations) meet the privacy requirements of the law.]

6.2 Justification for generalization

[Many regulations are implicitly concerned with privacy in statistical computation (the setting of this paper). We believe that by focusing on this particular context our methodology will generalize to other regulations. The computational focus allows us to work at a level of abstraction at which many of the differences between particular regulations become less important. Additionally, given the formalization of a regulation's privacy desiderata as a mathematical definition, we can test other computations against that definition besides differentially private ones.]

7 Discussion

7.1 Key findings

On first impression, the gaps between legal standards for privacy protection and formal mathematical models of privacy may seem vast and insurmountable. However, this article presents a methodology for demonstrating that it is indeed possible to bridge between these diverging privacy concepts, using arguments that are rigorous from both a legal and a mathematical standpoint.

In addition, many information privacy laws seem to adopt a framing of privacy risks that is based on an understanding of traditional approaches to privacy protection like de-identification. This methodology potentially helps to lower the barrier for adoption for new technologies based on formal privacy models, by addressing the substantial conceptual gap between formal and traditional approaches.

Further, the level of abstraction demonstrated by the analysis in this Article has many advantages. Although there is a great number of privacy regulations and each entails very different requirements, at the high level of abstraction adopted in this analysis, many of them ought to share general features. This perspective brings valuable simplicity to the analysis of privacy regulations.

7.2 Potential applications of the methodology

Many government agencies, most notably statistical agencies, are interested in adopting public-facing tools for differentially private analysis. However, before sharing sensitive data with the public, agencies typically must demonstrate that the data release meets relevant regulatory requirements and satisfies various institutional policies, such as those related to any applicable internal disclosure limitation review. The proposed methodology could be used in this case to demonstrate that an agency's use of a formal privacy model satisfies its obligations to protect the privacy of data subjects pursuant to applicable laws such as the Privacy Act of 1974[203] or the Confidential Information Protection and Statistical Efficiency Act.[204]

Organizations such as research universities and corporations currently manage large amounts of personal data that hold tremendous research potential. These organizations are reluctant to share data that may contain sensitive information about individuals due to recent high-profile privacy breaches and the specter of legal liability. This methodology could enable them to begin sharing data with confidence that the risk of a lawsuit under the relevant information privacy laws is low.

In addition, federal regulators such as the Department of Education and corresponding state agencies often develop specific guidance on implementing measures to comply with federal privacy standards. Such agencies could employ the proposed methodology in the future to evaluate and potentially approve the use of new technologies, based on a rigorous determination regarding whether they satisfy existing regulatory and statutory requirements for privacy protection.

7.3 Policy implications

The proposed methodology can contribute to efforts to bring new privacy technologies to practice. This Article makes the argument that differential privacy likely satisfies the requirements of FERPA and future extensions may show that it likely satisfies the requirements of similar laws as well. This demonstration can serve as part of the foundation supporting data holders' use of tools for differentially private analysis in the real world.

This Article focuses on the requirements of a current regulation because it is important for practical reasons to bridge gaps with respect to current standards. However, there are many known gaps in the current legal framework for privacy, and reforming information privacy laws based on a modern scientific approach to privacy could bring greater certainty and stronger privacy measures to practice. Such reforms could be informed by the understanding of the gaps between formal privacy models and traditional approaches that are uncovered by an analysis using the proposed methodology, as well as potential extensions to this work. In particular, the type of analysis set forth in this Article could form the basis for additional tools for making policy and guidance decisions based on rigorous conceptions of privacy. Privacy is a very complex problem with both legal and technical dimensions. Solutions based on combined legal-technical concepts such those proposed in this Article are well-suited to enable stronger privacy protection and increased certainty about data management decisions in the future.

8 Conclusion

[In the final version of the paper, the conclusion will identify future research directions for the proposed methodology, covering topics such as setting the privacy parameter (*epsilon*) based on regulatory requirements, addressing composition (cumulative privacy risk from multiple analyses), and addressing correlations between students.]

[203] 5 U.S.C. § 552a.

[204] 44 U.S.C. § 3501 note.