FIU Law Review

Volume 17
Number 2 *Symposium: Privacy Discussion Forum – University of Stockholm (Stockholm, Sweden, 2022)*

Article 9

Spring 2023

Calculated Privacy: Tech Meets Law & Law Meets Tech

Tobias J. Oechtering
KTH Royal Institute of Technology, oech@kth.se

Sara Saeidian
KTH Royal Institute of Technology, saeidian@kth.se

Cecilia Magnusson Sjöberg
Stockholm University, cecilia.magnussonsjoberg@juridicum.su.se

Online ISSN: 2643-7759

Recommended Citation
Tobias J. Oechtering, Sara Saeidian & Cecilia Magnusson Sjöberg, *Calculated Privacy: Tech Meets Law & Law Meets Tech*, 17 FIU L. Rev. 383 (2023).
DOI: https://dx.doi.org/10.25148/lawrev.17.2.9

CALCULATED PRIVACY: TECH MEETS LAW & LAW MEETS TECH

Tobias J. Oechtering, Sara Saeidian, and Cecilia Magnusson Sjöberg*

ABSTRACT

The article explores the relationship between technical privacy and legal privacy, specifically within the context of EU data protection law. The interdisciplinary approach taken aims to bridge the gap between law and technology by linking legal data protection principles with statistical concepts. The article argues that the data minimization principle can be related to the concept of a sufficient statistic that cannot be transformed further without losing utility, and that deviations from this request require careful justification. Additionally, the article discusses the importance of using technical privacy measures to rigorously assess privacy risks. Differential privacy and pointwise maximal leakage are briefly reflected upon as promising approaches for achieving legal compliance with data protection legislation. Finally, the article concludes that connecting legal principles with rigorous mathematical concepts can help address the gap between technology and law in privacy and provide lasting design guidelines for technology.

* Tobias J. Oechtering and Sara Saeidian are with Information Science and Engineering Division at KTH Royal Institute of Technology, and Cecilia Magnusson Sjöberg is with Swedish Law & Informatics at Stockholm University, both in Stockholm, Sweden. The authors wish to thank KTH Digital Futures center that supported the research within the collaborative project DataLEASH.

I. INTRODUCTION

The topic addressed here and further analyzed below easily evolves into a kind of mission impossible considering the broad scope surrounding modern technologies.[1] Nevertheless, the conditions for privacy, briefly expressed as the human right to a private sphere and sometimes to be left alone, seem to be a worthwhile task to investigate into. Today´s information society and the kind of specific legal issues that digitalization gives rise to is also a voice in this development.[2]

Over the years, so much has been said and written about privacy that added value can be hard to imagine. However, the quest for a better understanding of important societal issues calls for attention to those issues in the current digital environments. Therefore, we are convinced that even rather general reflections about what we refer to as "Calculated Privacy"—where tech meets law and law meets tech—could contribute to ongoing discussions. To put it simply, the focus here is on when the law is handled by means of figures and forms, rather than by words and semantics. For instance, this means that the primary interest is not Privacy Enhancing Technologies (PETs) as such, but rather, the application of general Information Communications Technologies (ICT) Law. A more specific background is to be found in EU data protection law (GDPR), which we will use as an example below.[3]

In GDPR, the General Data Protection Regulation (EU) 2016/679, there are several provisions that in particular are oriented towards calculated privacy.[4] First mention should be made to Article 5 that contains the fundamental data protection principles relating to the processing of personal data. Furthermore, Article 25 calls for attention when it comes to data protection by design and by default. This should be provided for within a risk-based approach, as laid down in Articles 32 and 33 regarding security of

---

1 *See generally* Russell L. Weaver, *Privacy Discussion Forum: Introduction*, 17 FIU L. REV. 263 (2023).

2 For a current overview of specifically legal implications, *see generally* KATJA DE VRIES ET AL., DE LEGE: LAW, AI, DIGITALISATION (Katja de Vries & Mattias Dahlberg eds., 2022); ROYAL ACADEM. ENG'R SCI. [IVA], DIGITALISERING I VÄLFÄRDEN – DAGSLÄGE OCH FRAMTID: RAPPORT FRÅN IVAS PROJEKT *DIGITALISERING – MÖJLIGGÖRARE I FRAMTIDENS VÄLFÄRD* [DIGITIZATION IN WELFARE – CURRENT SITUATION AND FUTURE: REPORT FROM IVAS PROJECT DIGITIZATION – ENABLER IN THE WELFARE OF THE FUTURE] (2022).

3 *See generally Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act)*, COM (2022) 68 final (Feb. 23, 2022); *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final (Apr. 21, 2021).

4 *See* the brief conceptual introduction above about the right to be let alone, etc.

processing. Mention should also be made to Article 35 and the regulation about data protection impact of assessment (DPIA).

The theoretical framework of this study is the *interdisciplinary* field of law and informatics. More precisely, the illuminated issues concern the interplay between substantive ICT law and legal system management. There are of course other ways to formulate this research paradigm. Such a situation evolves when law in a digitalized context becomes a particular object of study that primarily emphasizes and acknowledges the impact of statistics, mathematics, computational technologies, etc. In any case, the scientific approach boils down, as already mentioned, to what may be referred to as *calculated privacy*, i.e., where tech meets law and vice versa. The phenomenon especially observed here will be privacy/*integrity* and *personal data processing*. This implies that we will not engage in any in-depth, traditional, dogmatic legal analyses of what the law says according to different law-making bodies, decided court cases, legal scholars, etc., but rather, how the introduction of multi-faceted digital factors could look like from a societal perspective. More precisely, this could concern the impact that information retrieval (IR) search algorithms have on recall and precision within the legal domain, how information standards influence legal document management, and so on. In practice, it is no doubt impossible to once and for all interpret and apply (all) of the legal rules and regulations that are applicable in various digital settings. With this in mind, a set of *critical features* are introduced in our discussion searching for what may be referred to as calculated components of technical privacy on the one hand and traditional legal privacy on the other. However, this distinction is by no means crystal clear.

Adding to the picture is the way in which (a) applications of ICT have consequences for *legal infrastructures* governing particular data processing,[5] specifically, core concepts mirroring legitimate personal data processing, e.g. controllers, processors, consultants, personal data,[6] and consent; (b) in practice, ICT, in many sectors of society, has grown into a legally oriented steering mechanism based on algorithms targeting traditional automation and developing towards artificial intelligence (AI) and machine learning (ML) and furthermore, ICT also functions as a lever for internationalization surrounded by jurisdictional issues; and (c) ICT functions as a conventional tool, although digitalized and materialized in various digitalized platforms for client management, which is also noteworthy

---

5 A similar process might be applicable to this cluster of activities: anonymization, identification (deidentification, reidentification), pseudonymization, and synthetization.

6 In a holistic approach, we will also illuminate a few clusters of concepts (variables) ready for analyses and application. To begin with, there is personal data, sensitive data, certain criminal data, data about children, private data, and synthetic data.

The introduction of this work concludes in a number of *hypotheses for continued research.* Keeping it short, it still seems worthwhile to illuminate and investigate into the interplay between technical privacy and legal privacy. Continuing on, we will therefore once again briefly observe a set of privacy issues. What comes next is a fraction of reflections emanating from generic privacy protection. These major findings may be captured in terms of *calculated privacy*.

There are a few recent technically motivated attempts to bridge the gap between the legal and computer science domains,[7] including even attempts to derive formal claims, called *legal theorem*, to analyze whether technical concepts can satisfy the legal requirements of privacy.[8] A statistical view on GDPR's singling out aspect has been done by introducing the idea of *predictive singling out.*[9] Recently, focusing on anonymization and deidentification, it has been concluded that with the rapid advancements in machine learning (ML), regulations are quickly becoming outdated, such that data protection concepts need to incorporate scientific principles based on mathematical rigor.[10]

This article is inspired by those pioneering works, and it attempts to bridge the gap between law and tech by providing discussions that connect some legal and statistical concepts in the privacy domain.

II. TECHNICAL PRIVACY

In this section, we first discuss several basic principles of technical privacy that explain the general approach. After that, building on those basic principles, we provide a technical interpretation of the data minimization principle.

In the technical domain, privacy is strongly related to the concept of disclosure control, i.e., how to prevent, limit, and assess the risk of unauthorized access to information. From a technical perspective, one furthermore often distinguishes between data and information. The amount of information regarding something in a data set can be described by the reduction of the uncertainty of knowing the data set. Take, for instance, a certain disease that a person either has or does not have. The disease is an

---

7 Kobbi Nissim et al., *Bridging the Gap Between Computer Science and Legal Approaches to Privacy*, 31 HARV. J. L. & TECH. 687, 690–91 (2018).

8 Kobbi Nissim, *Privacy: From Database Reconstruction to Legal Theorems*, PROC. OF THE 40TH ACM SIGMOD-SIGACT-SIGAI SYMP. ON PRINCIPLES OF DATABASE SYS., 33, 36 (2021).

9 Aloni Cohen & Kobbi Nissim, *Towards Formalizing the GDPR's Notion to Legal Theorems,* 117 PROC. NAT'L ACAD. SCI. 8344, 8345 (2020).

10 Micah Altman et al., A Principled Approach to Defining Anonymization as Applied to EU Data Protection Law 26 (May 10, 2022) (unpublished discussion draft) (on file with the SSRN).

attribute of the person and is described by a binary variable. Knowledge about the value of this variable denotes information. The information might be explicitly stated in a medical record if it has been diagnosed, but it might also be provided only implicitly by some data, e.g., some biomarker records. Then, the attribute is known with only some certainty and the information in the dataset is the reduction of uncertainty regarding that attribute. Technically, it is therefore the information that needs to be protected. However, a dataset might provide information about many more attributes—more conclusions on other diseases or health conditions that can be made from biomarkers. Since the information is in the data and one might not know exactly what an adversary is interested in, protection of the data as a whole might be reasonable as well.

Next, technical privacy is usually developed with the consideration of a specific *adversarial attack* in mind.[11] This means that it has been specified what an adversary aims to do and what resources are available for the attack. Based on this, the *privacy risk* of an adversary's success is assessed (privacy risk assessment). As a next step, technical means to limit or prevent the information leakage might be provided (privacy-by-design).

For instance, if we wish to hide the information of whether a certain person was part of a study, then we would consider an adversary who does a membership inference attack. Differential privacy has been proposed to measure the risk of such attacks and many privacy mechanisms have been proposed in the research literature to reduce the privacy risk. For instance, the well-known Laplacian mechanism adds Laplacian distributed noise on the data to reduce the leakage on the membership information.

The goal of *privacy-by-design* approaches is therefore to take privacy aspects into account from the very beginning of the design and perhaps, choose a system setting that inherently leads to a lower privacy risk. For instance, we might face a situation where data is distributed, and we are interested in processing the data. We might have the option to either collect all the data and then process it centrally or to do a distributed processing where the processing is done locally and only the processing result is shared (assuming the task decomposes so that distributed processing is possible). The latter is always better from the privacy perspective since we never share **more** information. This follows from the fact that if you centrally have all of the data, then the distributed processing can be done by the central processing, among other ways of processing the data.

Because technical privacy is usually designed with a specific adversarial attack in mind, the guarantee is limited to that specific attack scenario. For

---

11 Isabel Wagner & David Eckhoff, *Technical Privacy Metrics: A Systematic Survey*, 51 ACM COMPUTING SURVS. 1, 1, 5 (2018); Maria Rigaki & Sebastian García, *A Survey of Privacy Attacks in Machine Learning*, ARXIV 1, 1–2 (2020).

instance, a statistical disclosure measure such as Bayesian inference risk might be used as a guarantee when considering an attribute inference attack, but it might not give a guarantee for a membership inference attack.

Next, the famous ϵ-*differential privacy* measure (discussed further in Section 3.1) provides a guarantee for membership inference attacks, but the technically relaxed (ϵ,δ)-*differential privacy* measure does not provide any protection with the probability equaling δ, which is an important detail. This means that understanding the value of technical privacy always requires a technical understanding of the assumptions.

Lastly, it is important to note that a technical privacy guarantee is an abstract concept that requires a mathematical proof, meaning that we need an analytical argument that proves the privacy guarantee claim. Theorems that show that one privacy guarantee implies another privacy guarantee are very interesting because they make the privacy guarantee stronger. At the same time, it might be that the former privacy guarantee is too demanding for privacy-by-design if the latter privacy guarantee is considered to be sufficient.

In principle, we can develop a technical privacy measure for every specific adversarial attack scenario, which makes the area very broad. Therefore, one is often interested in either attack scenarios that precisely provide the required privacy protection or attack scenarios that lead to a privacy protection that implies several guarantees against several other attacks.

Finally, besides operational approaches that provide guarantees for specific adversarial attacks, approaches based on privacy definitions that formally do not have an operational meaning but are reasonable and often easier to handle with regards to the design have also been developed. Whether those privacy measures are legally suitable or whether additional arguments are needed to make them suitable remains an open question.

A. *Data Minimization Principle*

Several articles in GDPR can be interpreted from a technical perspective. Such discussions reveal the appropriateness of certain technical principles that a system designer should strive to follow in order to achieve privacy-by-design. Following is a statistical interpretation of the data minimization principle stipulated in Article 5 of GDPR (principles relating to processing of personal data):

> 1. Personal data shall be:
> (c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimization').

The reasonability of the data minimization principle is currently debated as it is argued that it might prevent the exploration of a dataset, i.e., the search for dependencies between attributes in the data (data science). Such studies are interesting in the health domain where a statistical relation between a disease and other attributes might provide new insights on how to prevent or treat the disease if the causality question, i.e., that a certain attribute causes the disease, can be positively affirmed. In this case, the exploration of potential correlations is the actual purpose of the processing where an unreasonable data minimization would indeed limit the value of the dataset and as such, should not be done. Thus, the data minimization principle is not a limitation if the purpose of the processing is correctly identified.

There are, however, cases where not following the data minimization principle results in significant privacy risks. For instance, the privacy risk from sharing a machine learning (ML) model trained with sensitive data is often underestimated. Because the model is the outcome of a complicated learning algorithm that is not fully understood technically, one might wrongly conclude that sharing the ML model is also safe and thus, not expect that a shared ML model can leak information about the training data. Recently, however, it has been shown in an *attribute inference attack* that ML models memorize many details about the training data, which can be also extracted.[12] The attacks are based on unauthorized statistical inferences using the shared ML model. In other words, information is extracted from the ML models by using data analysis tools to draw unauthorized conclusions on attributes available in the training data and memorized during the learning in the shared ML model.

It should therefore be clear that sharing a ML model has a privacy risk and it has a bigger risk if the input data has not been minimized. Unfortunately, sharing a ML model carries an even larger risk. In *model inversion attacks,* a trained model is used to make predictions on unintended attributes,[13] with the reason being that a trained model of, for instance, a supervised learning[14] algorithm should capture all dependencies between all attributes in the input data and the target value (the value that the model should predict). Given some incomplete input data and the target value, the model can be misused for the prediction of some remaining input attributes (model inversion). Thus, the privacy risk from sharing such a model increases

---

12 Giuseppe Ateniese et al., *Hacking Smart Machines with Smarter Ones: How to Extract Meaningful Data from Machine Learning Classifiers*, 10 INT'L J. SEC. & NETWORKS 137, 138 (2015).

13 Matthew Fredrikson et al., *Privacy in Pharmacogenetics: An End-to-End Case Study to Personalized Warfarin Dosing*, PROC. 23 USENIX SEC. SYMP. 17, 18 (2014).

14 In supervised learning, the model is generated using some labeled training data, i.e., when one has the input data and knows the correct target value (label). The general principle is to train a model that has learned the dependencies between the input and the target values.

when the input data has not been minimized and as a result, includes correlation to unnecessary attributes which have been memorized. The following statistical interpretation of the data minimization principle will help to suppress the success of such attacks.

In a statistical context, the principle of data minimization can be related to the concept of a *sufficient statistic*. A sufficient statistic describes a statistic which is a function of an original statistic that is equally good for the inference task.[15] For instance, if we have tabular data and we are interested in the average of the first column, then we can remove all other columns in the table and still compute the average. It is important to note that a sufficient statistic, i.e., the function that is used to produce the sufficient statistic, depends on the data *and* the inference task, i.e., the purpose of the processing that specifies the utility.

Another example could be where we have a table listing the location of people, e.g., in a cellular network, and we wish to know the distribution of the distance to a certain point, e.g., because we wish to find a good location for our shop. We can then transform the two-dimensional location data into a one-dimensional distance data and still compute the distance distribution from the new data set. The distance data set provides us with a new statistic, where the purpose is preserved but the data set is made smaller. The exact position of a person is not available anymore, but rather, only the distance is kept. Accordingly, a sufficient statistic describes a transformation of the dataset that is equally good for the purpose, i.e., has the same utility, but it might not be reversible.[16] A non-reversible transformation might map several data points in the original dataset on the same data point in the new dataset. For instance, in a tabular database, we might have several entries where the first column has the same number (e.g., age) while all other columns have different entries. If we remove all columns but the first, then all rows with the same number will result in the same entry in the new dataset.

Following this discussion then, the principle of data minimization can then be seen as the request for a *sufficient statistic that cannot be further reduced*. This means that we are looking for the non-reversible transformation that produces a sufficient statistic which cannot be transformed further with non-reversible transformation without losing its utility. Such transformed dataset would be statistically per definition "limited to what is necessary to the purposes."

One might wonder if such a sufficient statistic that cannot be further reduced is unique. In statistics, there exists the concept of a *minimal sufficient*

---

15 *See* GEORGE CASELLA & ROGER L. BERGER, STATISTICAL INFERENCE 272–73 (Carolyn Crockett ed., 2nd ed. 2002).

16 A transformation (function) is reversible if a one-to-one correspondence exists.

statistic, which describes a sufficient statistic that can be produced from any other sufficient statistic. A minimal sufficient statistic would obviously satisfy the data minimization principle.[17] Unfortunately, the existence of a minimal sufficient statistic for any inference task and dataset is not guaranteed because the request that the minimal sufficient statistic be produced by any other sufficient statistic is too demanding. There might be different non-reversible transformations that lead to different sufficient statistics that cannot be further reduced and cannot be produced from each other. Each of those sufficient statistics would satisfy the data minimization principle.

As mentioned above, the data minimization principle becomes important when one wants to share (processed) data while it is unclear what information is included in the data to be shared. This may happen if one wants to share a ML model which has been trained using personal data. For instance, in personalized medicine, the input can be patient data spanning from basic information, such as gender or bio markers, to genetic data, but also, societal background information with information on one's lifestyle might be used. The target value might be a prediction on a certain disease or the best dosage of a medicine or treatment. The power of ML is the capability to learn the dependencies between the training data and the target value. If the training data has not been minimized, then we will face a privacy risk because the model has memorized the dependencies between training data and target values, in which it has been shown how this information can also be extracted. Thus, in such scenarios, it will be helpful if the ML model is trained using a sufficient statistic that cannot be further reduced.

For inference tasks, where the relation between the training data and the inference target is more complex, the function to produce a sufficient statistic that cannot be further reduced can be arbitrarily complex. While the concept of a sufficient statistic that cannot be further reduced provides us with a framework to characterize what is theoretically possible, we might not strictly follow the principle due to complexity reasons. Thus, the request to find a transformation that leads to a sufficient statistic that cannot be further reduced might be too demanding in some cases, such that a strict request might not be always reasonable.

A simplified request can be formulated using the concept of *irrelevant data*, which is closely related to the concept of a sufficient statistic. Considering a dataset where we can identify two parts of data, i.e., two sets of columns in our tabular data, we can call one part of the data irrelevant for an inference task if the remaining data is a sufficient statistic for the inference

---

17 Such is true because if a non-reversible transformation to further reduce the minimal sufficient statistic exists, then we cannot produce the minimal sufficient statistic from this reduced sufficient statistic as the transformation was non-reversible.

task. In our example above, where we are interested in the average age and the first column of the data provides us with a sufficient statistic, the remaining columns are irrelevant data. Thus, a simplified data minimization request could be that we require the removal of all irrelevant data from a given dataset. This is then not necessarily a sufficient statistic that cannot be further reduced as more complex non-reversible transformations might still possible, but the removal of irrelevant data is a simple task.

Another reason to deviate from the request of a sufficient statistic that cannot be further reduced might be data efficiency reasons. For instance, a certain ML model can be more easily learned if the training data is not reduced to a sufficient statistic that cannot be further reduced. However, one should request that any deviation be carefully justified.

III. CALCULATED PRIVACY

The result of this brief study shows the relevance in *distinguishing* between technical privacy and legal privacy. In this context, a *conceptual* approach to data protection has potential with regards to privacy regulation in comparison to traditional formal legislation. *Recitals* associated with certain rules and regulations within EU law is one way of clarifying matters without letting the text entities formally adhere to specific provisions. Something else to be aware of concerns the possibility of structuring technical commentaries—recitals—and their contents so that they mirror the previously observed categorization of technical vs legal privacy.

Given legal and societal development, it becomes obvious that the notion of *privacy* somehow needs to be *interpreted and implemented*. How this is to take place in practice is difficult to foresee. However, in the future of legal research, calculated privacy appears to be one promising way forward.

Summing up, several keywords can be extracted in a tentative abstract prior to, for instance, publication. Evidently, technical privacy as well as legal privacy would be important to take into consideration. Yet another dimension can be expressed in terms of interpreted and applied (customized) privacy. The meaning of integrity can, in and itself, be quite a challenge. In the Scandinavian countries, language barriers may, for example, lead to personal integrity being understood as a privacy protection and not related to dignity, for example, which would be more adequate. In this context, calculated privacy is important for at least three reasons: first, because it mirrors *reality*; second, because it helps to strike a *balance* between law and ICT; and third, in order to achieve legal *compliance* with the governing legal framework, e.g., data protection legislation, specifically, Articles 5, 25, 32, 33, and 35 in GDPR. What can then be expected by future research? This is

difficult to predestine if at all feasible. The dynamic research community surrounding privacy indicates, however, that already today, interesting activities are opening up for code in the legal domain.

From a technical point of view, calculated privacy requires a notion of a *privacy measure*, i.e., a mathematical expression with the goal of describing how private a given data processing system is. Roughly speaking, a privacy measure takes as input a *probabilistic* description of the data processing system and outputs a number that describes the system's privacy level. As a simple but concrete example, consider a data processing system which takes one bit of data as input, and produces one bit of data in the output. If the input bit has a value of 0, the system flips an unbiased coin and produces either a 0 or a 1 in the output with equal probability. If the input bit has a value of 1, then the system flips a biased coin which produces a 0 with a probability of 0.3 and a 1 with a probability 0.7. Supplied with such a description, a privacy measure then calculates a number that reflects the privacy level of the system.

In general, privacy measures are defined with specific *adversarial attack*s in mind. An adversarial attack describes a real or hypothetical scenario in which an adversary with pre-defined capabilities (e.g., in terms of computational power) interacts with a data processing system in order to achieve a certain objective. For instance, we may imagine an adversary with unlimited computational power who monitors the output of our single-bit system, described above, to guess the value of its input bit. Adversarial attacks are also used to define *operational meanings* for privacy measures. Heuristically, the operational meaning of a privacy measure describes the specific notion of privacy that is meant to be preserved, or equivalently, the types of adversarial attacks whose risks it measures.

Let us now introduce the two approaches that are commonly used to assign operational meanings to privacy measures: We say that the operational meaning of a privacy measure *succeeds* its definition if we first come up with the mathematical formulation of a privacy measure, and then interpret the formulation in terms of an adversarial attack scenario. Conversely, we say that the operational meaning of a privacy measure *precedes* its definition if we first describe an adversarial attack scenario, and then obtain the definition of a privacy measure by analyzing the attack. Naturally, privacy measures of which operational meanings succeed their definitions rely on intuition and are easier to come up with. It is therefore not surprising that numerous such measures have been proposed in the literature, including differential

privacy,[18] local differential privacy,[19] Bayesian differential privacy,[20] differential identifiability,[21] Pufferfish privacy,[22] and membership privacy,[23] to name a few. Differential privacy is by far the most widely adopted notion of privacy, which will be discussed in further detail below. On the other hand, there have been very few privacy measures of which operational meanings precede their definitions, including maximal leakage[24] and pointwise maximal leakage.[25] Pointwise maximal leakage is briefly discussed below.

A. *Differential Privacy*

Consider a data processing system where data collected from a number of individuals is stored in database D. Our goal is to analyze the database through an algorithm (or function) denoted by A, but also, to assure participants that they will not face negative consequences as a result of contributing their data to the database. For example, Alice's insurance premium may increase if her insurance company finds out that she has contributed her data to a study on diabetes. This is exactly the goal of differential privacy. Essentially, differential privacy, in its original form,[26] guarantees that the outcome of data processing is not too affected by whether or not each individual participates; thus, hiding each person's contribution. To understand how this is achieved, suppose D_1 and D_2 describe any two databases that differ only in the data of a single individual. We say that algorithm A satisfies ϵ-DP (where ϵ is a non-negative number) if for each of the possible outcomes of algorithm A, denoted by o, it holds that:

$$\frac{\Pr[\mathcal{A}(D_1) = o]}{\Pr[\mathcal{A}(D_2) = o]} \leq e^{\epsilon}.$$

---

18 *See generally* CYNTHIA DWORK & AARON ROTH, THE ALGORITHMIC FOUNDATIONS OF DIFFERENTIAL PRIVACY (2014).

19 John C. Duchi et al., *Local Privacy and Statistical Minimax Rates*, 54 ANN. IEEE SYMP. ON FOUNDS. COMPUT. SCI. PROC. 429, 429–38 (2013).

20 *See generally* Bin Yang et al., *Bayesian Differential Privacy on Correlated Data*, 2015 ACM SIGMOD INT'L CONF. ON MGMT. DATA 747 (2015).

21 Jaewoo Lee & Chris Clifton, *Differential Identifiability*, 2012 ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING 1041, 1041–49 (2012).

22 *See generally* Daniel Kifer & Ashwin Machanavajjhala, *Pufferfish: A Framework for Mathematical Privacy Definitions*, 39 ACM TRANSACTIONS ON DATABASE SYS. 1 (2014).

23 *See generally* Ninghui Li et al., *Membership Privacy: A Unifying Framework for Privacy Definitions*, 2013 ACM SIGSAC CONF. ON COMPUT. & COMMC'NS SEC. 889 (2013).

24 *See generally* Ibrahim Issa et al., *An Operational Approach to Information Leakage*, 66 INST. ELEC. & ELECS ENG'RS TRANSACTIONS ON INFO. THEORY 1625 (2020).

25 *See generally* Sara Saeidian et al., *Pointwise Maximal Leakage*, 2022 INST. ELEC. & ELECS. ENG'RS INT'L SYMP. ON INFO. THEORY 1 (2022).

26 DWORK & ROTH, *supra* note 18, at 211, 226.

The above expression should be read as follows: In the numerator, we have $\Pr[\mathcal{A}(D_1) = o]$ which describes the probability of A producing outcome o upon operating on database D_1. Similarly, in the denominator, we have $\Pr[\mathcal{A}(D_2) = o]$ which describes the probability of A producing outcome o upon operating on database D_2. Because D_1 and D_2 differ in the data point of a single individual by restricting the ratio, we ensure that these two probabilities are not too different. In other words, we have restricted the effect of a single data point on every possible outcome of the data processing. This is done by setting the parameter ϵ. Thus, smaller values of ϵ ensure a smaller contribution from each single data point.

It is worth noting that the above description is an interpretation of equation (1), and over the years, a number of works have challenged this interpretation. For instance, some works have claimed that statistical dependencies between the data points in the database may water down the guarantees of differential privacy,[27] while others have claimed that the adversary considered by differential privacy is not necessarily the strongest.[28]

B. Pointwise Maximal Leakage

Pointwise maximal leakage (PML) is a privacy measure that takes a more abstract and general approach to privacy. While differential privacy is designed specifically with individuals contributing their data to a centralized data processing system in mind, PML tries to protect any type of data containing sensitive information, which we refer to as the *secret*. The secret may be an entire database, a single data point collected from an individual, a password typed on a keyboard, and so on. Let the secret be denoted by S.

As mentioned earlier, pointwise maximal leakage is a type of privacy measure of which operational meaning precedes its definition. In fact, it has been shown that PML can be obtained by analyzing two distinct (but mathematically equivalent) adversarial attack scenarios. In what follows, we briefly describe these two scenarios. Note that unlike differential privacy, PML makes all assumptions about the adversary and the secret explicit in the model; thus, making its privacy guarantees more concrete and leaving little room for interpretation.

---

27 Graham Cormode, *Personal Privacy vs Population Privacy: Learning to Attack Anonymization*, 2011 ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING 1253, 1256 (2011).

28 Daniel Kifer & Ashwin Machanavajjhala, *No Free Lunch in Data Privacy*, 2011 ACM SIGMOD INT'L CONF. ON MGMT. DATA, 193, 204 (2011).

i. First Approach to Defining PML

First, consider an adversary who is interested in guessing the value of some attribute (or property) of S, denoted by T. This is a very general model where we are assuming that the adversary either wants to know S (as S is an attribute of S) or any other information that can be extracted from S. For example, when S is a database, T may be a set of rows in S. As before, we process the secret S through using algorithm A. Suppose the adversary sees an outcome o of A and constructs a guess of T based on this observation. We also assume that the adversary has full knowledge of the system, i.e., that she knows the probabilistic descriptions of S, T, and also, algorithm A.

The goal of PML is to capture the adversary's gain in knowledge about S due to observing o. It does so by considering two probabilities: (1) the probability of correctly guessing the value of T without any observations, i.e., the *prior probability of success*, which essentially represents the adversary's knowledge before interacting with the system; and (2) the probability of correctly guessing the value of T after observing o, i.e., the *posterior probability of success*, which is a measure of adversarial knowledge after interacting with the system. PML is then defined as the ratio of the posterior probability of success to the prior probability of success:

$$\max_{T} \frac{\Pr[\text{Correctly guessing the value of } T \text{ given } \mathcal{A}(S) = o]}{\Pr[\text{Correctly guessing the value of } T]}.$$

In the above expression, we are maximizing the ratio of probabilities over all possible T's because we may not know a priori which T the adversary is interested in. As a result, maximizing allows us to consider the worst-case scenario. While the above expression may look complicated, it has been shown that it can be considerably simplified to yield a workable mathematical expression.[29]

ii. Second Approach to Defining PML

The second adversarial model that can be used to obtain PML considers adversaries who are interested in maximizing the average value of a *gain function* g. Simply put, gain functions are mathematical descriptions of adversarial objectives. For instance, if the secret S is a database, then there is a gain function that models an adversary who is interested in finding out whether or not Alice is part of the database. Alternatively, if the secret S is a password typed on a keyboard, then there is a gain function that describes an

29 *See* Saeidian et al., *supra* note 25, at 2–3.

adversary who can make a certain number of attempts at correctly guessing the password. In this second adversarial scenario, we define PML by considering the *prior gain* of the adversary and the *posterior gain*. The prior gain describes the average value of g the adversary can achieve before interacting with the system and before observing any outcome of the system. The posterior gain, on the other hand, is the average value of g the adversary can achieve after observing an outcome o. More concretely, PML is defined as:

$$\max_{g} \frac{\text{Average value of } g \text{ given } A(S) = o}{\text{Average value of } g}.$$

Note that the maximization over g renders PML a robust privacy measure in the sense that it considers any adversary whose objective can be described by a gain function. Once again, it has been shown that the above expression can be considerably simplified, and that, in fact, simplification yields the same expression as the one obtained in the first adversarial scenario.[30]

IV. CONCLUSION

The gap between tech and law in privacy causes insecurity among professionals. System designers ask what privacy measures and what levels of such privacy measures are legally adequate. Legal advisors ask what new privacy risks we face due to recent technology developments. Both are prevalent concerns that cause discomfort among the decision makers. We believe that the inter-disciplinary exchange will help mitigate the situation. Connecting legal principles with mathematical concepts will help bridge the gap and provide some timeless design guidelines for technology that system designers can strive to implement and legal advisors can request be implemented. In this essay, we argue that the data minimization principle as applied to a statistical inference problem basically requests a sufficient statistic which cannot be transformed further with non-reversible transformations without losing its utility, and that a deviation from this request should be carefully justified. We also argue that privacy risk assessments call for a formal and mathematical privacy analysis. It should be rigorously quantified how much information is leaked by disclosing some data. In particular, the employed privacy measure needs to be chosen such that it protects against the relevant adversarial attacks that may occur. The

30 *Id.* at 2.

new notion of pointwise maximal leakage is an interesting, robust, and flexible alternative to the celebrated differential privacy measure, which often seems to be blindly adopted.