Situated Information Flow Theory

Sebastian Benthall
New York University
spb413@nyu.edu

ABSTRACT

A key component of recent privacy rules is restriction on the flows of personal information or data based on information categories. This tendency conflicts with the fact that data's meaning is not stable but depends on how it was formed and with what other information it is combined. These properties of information challenge naive intuitions that information 'flows' like a fluid, such as water or oil. Rather, we build on on Dretske, Pearl, and Nissenbaum to develop situated information flow theory (SIFT): a view of information flows as causal flows with nomic associations due to a larger context of causal relations. The semantics of situated information flow are precise within the statistical framework of Bayesian networks. We argue this understanding of information flow has three policy implications. (1) Restrictions on data transfers are more precise and enforceable than restrictions on information flow. (2) Information 'categories' or meanings must be defined relative to a particular class of observers and take into account their reasonable background information. (3) The semantics of data are ambiguous when there is uncertainty about causal structure, and this structure is learned from data aggregation. Hence, the information asymmetry between data aggregators and individual data subjects are one reason why data processors are 'opaque' and difficult to regulate.

KEYWORDS

information flow, Bayesian networks, contextual integrity

1 INTRODUCTION

Information technologies get their regulatory relevance from operating in the world at large. They are embedded in large social systems, and the information processed by these systems gets their meaning from this context. This is evident in contemporary breakdowns in the management of privacy online. Recent examples of privacy violation in a "big data" context are due to the fact that the semantics of data are not clear or stable, but rather that due to data mining techniques

and other innovations [51] [56], "data's meaning has become a moving target" [43] because the inferences data enable depend on its sources and what other data it is combined with.

We argue that a conceptual problem is at the heart of these privacy violations. There is a disconnect between how "information flows" are discussed in natural language in the discussion of data protection and privacy regulations, and how the term is used in mathematics and engineering practice. On one hand, almost by definition information technology operates on information, and the specific details of this processing in any case can be the result of many hours of engineering labor. On the other, regulators have imposed conditions on the collection, transfer, and use of information and data in terms that they understand. At issue is that again and again, intuitive understandings of what information *means* are violated by technical systems that discover new meanings through data mining and machine learning. The lack of a shared theory of what information is and how it gets its meaning disconnects laws and public expectations from engineering practice.

A scientific theory of how information flows and gets its meanings is needed to clarify the causes and potential solutions of these challenges to regulating privacy. Ideally, such a theory would bridge between law, social theory, statistics, and computer science, bringing each field into alignment. We build on prior work on Contextual Integrity [58], a theory of privacy inspired by law and social theory with many applications in computer science [4, 6, 9, 16, 22, 44, 46, 47, 62, 71, 74, 84], which identifies privacy as "appropriate information flow". These motivations for this paper are elaborated in Section 2.

Section 3 reviews data protection and privacy laws for how they refer to information and data flows. This review of regulations demonstrates the pervasiveness of regulatory language about "information flow", showing the practical implication of understanding what the term means for the purpose of engineering compliant systems. Special attention is paid to how recent laws, such as the California Consumer Privacy Act and Europe's General Data Protection Regulation distinguish between *categories* of information. These regulations show a general recognition that data moves and has meaning, but are imprecise and sometimes in implicit disagreement about how and why. This paper argues "information flows" in a way that is disanalogous to water or oil or electricity. Rather, we propose that "information flow" is actually two things: a flow of causality, whereby one event influences the outcome of another, and regular associations between events. The associations give meaning to an observed event (which can include data from a user interface, for example), but those associations depend on a broader context of causal relations. We name this account of information flow *situated information flow*, to emphasize that the meaning of information depends on the entire situation in which it takes part. Situated information flow theory (SIFT) follows from well established formal

representations of causal models used in statistics [65]. We develop SIFT in Section 4.

Section 5 concludes with an outline of the implications of SIFT for data protection and other regulation.

2 MOTIVATION

We motivate the following inquiry with three examples of privacy violation that depend on data having unexpected meaning. We ground the idea of a privacy violation in Contextual Integrity (CI) [57] as a field of privacy scholarship. The motivating examples challenge CI's definition of information flow. A new theory of information flow is needed to address these cases.

2.1 Examples of privacy violations from unanticipated semantics

Consider three examples of privacy violations in which an unexpected "information flow" happens because the full semantics of information shared by the data subject is not known to them.

Example 1. Target, the retailer, has learned that purchases of certain products are correlated with pregnancy and childbirth, but not obviously so. In one story, Target sends coupons for baby related products to a woman's home after she buys scent-free hand lotion, which upsets her father. Target has learned that the coupons are less effective if the detection of pregnancy is too obvious, and has learned to surround these coupons with other coupons for "random" products so that they are not so revealing.[42]

Example 2. Research has shown that a number of sensitive and private attributes–including sexual orientation, political views, ethnicity, and psychographic profiles–can be inferred from social media posts even when these topics are not explicit. [49] Allegedly, social media data has been used to develop psychographically targeted advertising, exploiting personality traits that were never intended to be revealed by users. [13, 82]

Example 3. In 2017, Quartz discovered that Android smartphones were collecting the identifiers of the cell phone towers the phone was collecting and sending them to Google even when location services were disabled [14]. Because cell phone tower identifiers are excellent proxies for the smartphone location, this was deemed by Quartz to be a violation of consumer expectation of privacy. After it was discovered, Google changed the phone software so that it would no longer collect this information.[3]

In these examples, information revealed by the data subject is ostensibly "about" one attribute, but upon investigation the data reveals other attributes as well. We will argue that the true "topics" of any data set in question are not inherent to data, but rather depend on the processes that generate it. That is why there may be hidden private variables (such as pregnancy or personality type) that are unknown to the data subject but known to the analyst who can see behavioral patterns in aggregate.

2.2 Contextual integrity

Contextual Integrity (CI) [57, 58] is a field of privacy scholarship that defines privacy as "appropriate information flow". In the core theory of CI, information flow is "appropriate" based on whether it

adheres to social norms that hold within particular social contexts or spheres. These norms can be expressed in terms of five parameters:

(1) The data *subject*, whom the data is about.
(2) The *sender* of the data.
(3) The *recipient* of the data.
(4) The information *type* or *attribute*, what aspect of the subject it is about.
(5) The *transmission principle*, a condition on flow such as obligation, confidentiality, or reciprocity.

These norms are legitimized by the ends, purpose, and values of the social sphere they are recognized to be a part of.

CI has had broad uptake as a theory of privacy within the critical, legal, and technical research communities [4, 16, 22, 39, 40, 44, 46, 47, 62, 70, 71, 73, 74, 84]. Recent work [9] has shed light on how computer science researchers in particular have opened up new lines of inquiry within the field, identifying several 'theoretical gaps' that are prompts for new research. An example is the difficulty that Contextual Integrity, as currently resourced, has with conceptualizing how information flows between or within multiple contexts, taking on multiple meanings. This is an important omission to address when so many privacy violations come from cases of context collapse [20, 52].

Information norms, according to CI, treat information's meaning as a function of the social context of its collection and use. As information flows evidently have meaning besides those prescribed by its social context, we must look outside the theory to explain why and how information avails more insight than is expected from social convention. Legal regulation sometimes shares with social convention this naivete about unpredicted uses of information.

An alternative vision of privacy, dubbed Origin Privacy, [61] has considered the value of replacing the "attribute" parameter with a reference to the origin or source of the information. While the work on Origin Privacy has offered mixed results, it has shown that references to information attribute are ambiguous. Any particular information flow may be about many attributes or topics. In some cases it is easy to establish some of these attributes. But in general it is difficult to determine what attributes an information flow is *not* about.

3 DATA PROTECTION REGULATIONS

Data protection statutes and agreements respond to political interest in personal data protection. A challenge in information law is the design and application of regulations to increasingly complex and innovative uses of personal data with information technology. The difficulty of this challenge comes in part from the fact that lawyers and technologists often speak literally different languages [79]. The law is written in what computer scientists call "natural language", the normal ways that human beings learn to communicate. Specialized legal language is taught as part of legal training, but it is largely articulable within normal natural language. Natural language contains terms with necessary ambiguity [55]; Hildebrandt [41] argues that contests over the ambiguity of human language is a productive part of legal practice.

Information technology, on the other hand, is often designed and implemented according to specifications that are described mathematically or in software code. Mathematical and programming

languages are generally not "natural" languages, but rather have different syntax and semantic rules that guarantee strict and specific interpretations. In the case of programming languages, the semantics of the language are guaranteed by the program compiler, which translates the code into machine instructions.

A response to this challenge is the translation of data protection laws into formal logical statements. These logical statements can then be analyzed for the possibility of complete or incomplete automated enforcement. Prior work [5, 17, 18, 24] addresses how particular clauses of data protection regulation the put conditions on the flow of information can be translated into formal logic and implemented. This article addresses what is arguably a more fundamental ontological question of how information gets its semantic content, how laws refer to that content, and how can meaningful information flows be formally modeled. This section compares the wording of several regulations around their definitions of what is being protected: how they refer to it (data or information), what sorts of movements or transfers they are sensitive to, and how they are sensitive to different types of data. Table 1 condenses the findings, which are summarized in Section 3.5.

3.1 International Guidelines

While not laws per se, international guidelines such as those published by the OECD are "soft-law" regulations that reflect international norms.

3.1.1 OECD. The OECD council's "Guidelines Governing The Protection of Privacy and Transborder Flows of Personal Data", initially proposed in 1980 and amended in 2013, recognize the importance of interoperable privacy frameworks given continuous flows of personal data across global networks. It explicitly defines "transborder flows of personal data" as "movements of personal data across national borders".

3.1.2 APEC. The Asia-Pacific Economic Cooperation (APEC) is a regional economic forum whose initiatives synchronize regulatory systems among its 21 members states. Its 2015 Privacy Framework aims to promote electronic commerce while reaffirming the value of privacy. It is modeled on the OECD guidelines, defining personal information as:

> Personal information means any information about an identified or identifiable individual. (II.9)

In its clarifying comment, the guideline is explicit about the possibility of identifying information through aggregation with other information.

> The Framework is intended to apply to information about natural living persons, not legal persons. The Framework applies to personal information, which is information that can be used to identify an individual. It also includes information that would not meet this criteria alone, but when put together with other information would identify an individual. For example, certain types of metadata, when aggregated, can reveal personal information and can give an insight into an individual's behavior, social relationships, private preferences and identity. (II.9)

This regulation makes no special mention of particular categories of personal information.

3.2 US Sectoral Law

In United States federal law, data protection is mainly regulated by sectoral law, that is, laws that protect personal information in particular economic sectors. HIPAA covers the health sector; GLBA covers financial services. COPPA is designed to protect children's information from online data collection. Consumer privacy in sectors unregulated by specific laws are regulated by the Federal Trade Commission, which rules on unfair and deceptive practices. Such deceptive practices can include the violation of terms of service or privacy agreements to which users of online services consent.

3.2.1 COPPA (US). The Children's Online Privacy Protection Act (COPPA) aims to restrict the collection of information from any child under the age of 13 without first parental consent and the subsequent ability for parents to access the information collected. It does this by forbidding data collection without parental consent by web service operators that target children or have actual knowledge that they are being used by children. This restriction has been called a temporal restriction [5], as opposed to a restriction based on attribute. Many significant breaches of COPPA have been discovered in apps that are directed at children and collect data from their users prior to acquiring parental consent [68].

The law refers to a number of examples of kinds of personal information, include social security numbers, screen or user name, customer number held in cookie, and device identifier. It does not treat any one category of personal information differently from the others.

3.2.2 HIPAA. HIPAA is the US law governing personal information in health care. It refers to a sectorally specific category of information: protected health information (PHI). The definition of PHI depends less on the semantics of the information and more on the fact that it has been disclosed to a Covered Entity. An interpretation is that, for example, personal information becomes health information when it is part of a transaction between a patient and a health care provider. An exception in HIPAA is made for psychotherapy notes, which get special protections and are defined as those notes that are derived from a psychotherapeutic counseling session. This policy specific to the origin of the data has inspired Origin Privacy [60].

3.2.3 GLBA. The Graham-Leach-Bliley Act (GLBA) governs the treatment of personal information by financial services institutions. It defines "nonpublic personal information" as:

> (A)The term "nonpublic personal information" means personally identifiable financial information– (i) provided by a consumer to a financial institution; (ii) resulting from any transaction with the consumer or any service performed for the consumer; or (iii) otherwise obtained by the financial institution.

In sum, the regulation is less concerned with the content of the information than the fact that a Covered Entity has obtained it, that it is nonpublic, and that it is personally identifiable.

Regulation	Date	Nouns	Movement words	Category verbs
OECD	1980/2013	personal data	"transborder flows", access, disclosure	none
APEC	2015	personal information	collection, holding, processing, use, transfer	none
COPPA	1998	personal information	collection, disclosure	none
HIPAA	1996	personal health information	disclosing	none
GLBA	1999	nonpublic personal information	disclosing, access, security	disclosing
CCPA	2020*	personal information	collecting, sharing, access, selling, transfer	disclose
NY Financial Services	2017	nonpublic information	access, use	none
GDPR	2018	personal data	transfer	reveal, provide

Table 1: Summary of language use in data protection regulations.

At the time of establishing a customer relationship and annually thereafter, a financial institution must disclose the categories of information that it collects about the customer.

3.3 US State Law

Some states in the United States have passed their own data protection laws. These are too numerous and varied to survey in this article. This section summarizes two recent and notable laws that reflect different approaches to data protection. The California Consumer Privacy Act (CCPA) addresses consumer privacy and the sale of personal information specifically. New York State's cybersecurity law focuses on cybersecurity, not privacy, and so protects nonpublic data in general.

3.3.1 CCPA (US, California).

The California Consumer Privacy Act contains special provisions about the sale of personal data, though it also covers a more general scope including the collection and disclosure of personal information.

> "Sell," "selling," "sale," or "sold," means selling, renting, releasing, disclosing, disseminating, making available, transferring, or otherwise communicating orally, in writing, or by electronic or other means, a consumer's personal information by the business to another business or a third party for monetary or other valuable consideration. (CCPA, 1798.140. (9) (t) (1))

The law also specifies that businesses must inform consumers about the categories of personal information they have collected, sold, and disclosed. Some, but not all, of the categories of information that the law considers personal information are listed in the law. They include topics as diverse as a person's protected classification (sex, race, and age), inferred behavioral profiles, and network device identifiers that are linked to a person.

3.3.2 Cybersecurity Requirements for Financial Services Companies (US, New York).

The *Cybersecurity Requirements for Financial Services Companies* of New York apply to any legal persons (individuals or non-governmental entities) operating under banking, insurance, or financial services laws. It refers to "nonpublic information", which includes both business related information that would cause a material adverse impact to a Covered Entity, but also any information concerning an individual based on an identifier. The regulation is an expansive cybersecurity policy, not only a data protection policy, and so mandates the use of risk assessments and the regular disposal of data to reduce exposure. It is not concerned

with 'collection', 'transfer', or 'flow' of data, but rather refers to 'access' and 'use'.

3.4 EU Omnibus Law

The European Union has a history of omnibus data protection regulation, meaning its regulation provides a baseline for all personal data use for all sectors of industry and government. Most recently, in 2018 the EU's General Data Protection Regulation (GDPR) came into effect. Due to strong enforcement clauses in the new law, GDPR has created a visible response from the commercial sector.

3.4.1 GDPR.

The GDPR does not use the term 'information flow', but rather refers to 'flows of personal data', or 'data transfers'.

It also refers 'special categories' of personal data:

> Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited. (Article 1, Section 9)

The conceptualization of data falling into categories has implications for the obligations of data controllers. In particular, data controllers must provide data subjects with information about which categories of information are concerned in data processing and transfer (Article 14.1).

3.5 Summary

The regulations surveyed here variously used the terms "data" and "information". Most also stipulated restrictions on how the data or information moves, as if a substance, especially by changing hands or ownership. This confirms CI's conception of privacy as appropriate information flow between different parties.

The regulations rarely address the specific contents of information. Individually identifiability is the key criterion for protection of information. Some references to information 'attributes', such as health information and financial information, devolve into conditions on the kinds of institutions involved in data transfer. This indicates that these sectoral privacy laws may fail to address how information that is meaningfully relevant to health and finances passes through entities that are not covered.

Some data protections rules (GLBA, GDPR, CCPA) have attempted to address the potential privacy violation of data reuse by

stipulating that data processors must reveal to their data subjects the categories of information that they have collected. It is not clear how enforceable this requirement is in cases where inferable attributes of the information are unknown. Some rules refer specifically to new kinds of information resulting from processing, such as behavioral profiles drawn from other personal information.

COPPA is most specific about how information falls under its scope. It applies to operators that direct a web service at children or have actual knowledge that a child is a user. These conditions, which refer to the total situation of an operator and its users, structure the probability that a particular user is a child.

4 SITUATED INFORMATION FLOW THEORY (SIFT)

Privacy violations due to the unexpected semantics of data and the need to regulate information flows based on the processes that generate them motivate new thinking about how technical and social platforms can be designed (Section 2). It seems that many of our socially comfortable expectations of information flow, for example those expressed by CI, are ontologically mismatched to with how information works in the world. Legal regulations have attempted to address privacy by restricting flows of personal information, often without making it clear either what "information" means or what information means (Section 3).

It is surprising that such a foundational question has not yet met with a scientific, transdisciplinary answer, as the meaning of data is of great pragmatic concern to science and industry. This section briefly reviews interdisciplinary theory (Section 4.1) and proposes a theory of information flow that synthesizing several literatures (Section 4.2). This theory, which we call situated information flow theory (SIFT), builds primarily on the philosophy of Dretske [27] and the statistics of Pearl [67] to argue that information flows are best understood as causal flows situated in the context of other causal relations. This theory is formalized in terms of Bayesian graphical models in Section 4.3. This formalization reveals why many uses of the term "information flow" are ambiguous (Section 4.4). SIFT is then used to shed light on the motivating examples.

4.1 Information across the disciplines

Different academic disciplines variously define "information" with more or less rigor and consistency. A consensus definition of semantic information has eluded philosophers [28]. Linguistic analysis of the modern use of the word "information" has concluded that it is a confused creole of distinct and incompatible meanings [63]. However, in the natural sciences most definitions of information build on Shannon foundational work on of information theory [72], which has had broad application in many fields [33], including physics [45, 85] and biology [1, 21]. Among philosophers, Dretske [27] perhaps most closely followed mathematical information theory for inspiration.

With respect to its application in systems of humans and documents, Library and Information Sciences (LIS) has fruitfully analyzed the term 'information' and discovered that information can be both a process and a thing [11]. In LIS, Brier [10] provides a comprehensive account of "cybersemiotics" that traces the relationship between hierarchical layers of semiotics ranging from the basic

information theoretic sense developed by Shannon [72] to social and linguistic meaning based on the social theory of Luhmann [50]. While a compelling theory, this view currently suffers from a lack of mathematical formulation.

There is a long history of literature on information flow in computer security and privacy research [7, 38, 53, 69, 76] This article draws especially on Tschantz et al. [80], who demonstrate that the basic security property of noninterference [36] can be aptly understood in terms of Pearl's [67] theory of probabilistic causal relations.

Pearl's theory of causation is itself attractive for its interdisciplinary acceptance. Cognitive scientists have argued that graphical models such as those used by Pearl are robust models for how humans learn and understand causal relations [34, 35, 37, 75]. Because of these results, we see causal graphical models as a promising bridge between rigorous probability theory based views of information and qualitative legal logic. Within computer science's study of privacy, Pearlian causation has elucidated differential privacy in work by Tschantz et al. [81].

4.2 Situated Information Flow

We posit *situated information flow* as a scientific, statistical view of information that addresses the practical challenges of unintended data meaning. Specifically, this account of information flow builds on the philosopher Dretske [27]. According to Dretske's theory, a message carries information about some phenomenon if a suitably equipped observer could learn about the phenomenon from the message. In other words, a message carries information about any topic that can be learned from it. For an observer to learn from it, the message must have a *nomic* connection with its subject, where here "nomic" means "law-like" or "regular" [27].

To make this more precise, we can model these regularities formally. As has been known since Shannon [72], information only flows when a signal has many different potential values. As famously said by Bateson [8], information is "a difference which makes a difference". The mathematics of probability and statistics, which provide formal tools for understanding the relationships between variables whose values are uncertain, are intimately connected to the mathematics of information for precisely this reason. When two random variables are related in such a way that one can learn about the state of one through an observation of the other, then they have *mutual information*.

Definition 4.1 (Mutual information [15]). For two random variables X, Y with joint probability distribution $P_{XY}(x, y)$, the mutual information $I(X; Y)$ is given by

$$I(X; Y) = E_{P_{XY}} log \frac{P_{XY}}{P_X P_Y}$$

Two random variables X, Y that have no mutual information, $I(X; Y)$, are conditionally independent. We will denote conditional independence with $X \perp\!\!\!\perp Y$, or $X \perp\!\!\!\perp Y|C$ if the variables are conditionally independent given another variable C.

For two variables to have positive mutual information, it is not sufficent for observed events to be empirically correlated; this correlation may be spurious unless it is supported by robust probabilistic relationships between events. Pearl [67] provides a robust and widely used formal account of structural flows of probabilistic influence
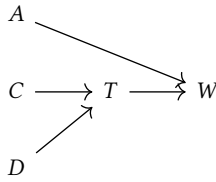
Figure 1: Alice's commute to work.

Figure 2: The electrical grid G controls power to all the buildings B_i.

through causal relationships. Pearlian causation and Bayesian networks can provide a useful and tractable formalism for understanding the meaning and value of information flows. The advantage of this formalism is that it can model the relationships between both technical components and social practices in an apples-to-apples way, because all aspects of a sociotechnical system can be modeled as probabilistic events.

4.3 Causal probabilistic graphical models

In our approach, privacy depends on appropriate information flow [59], where information is defined as that which allows somebody to learn about something else based on its regular associations with it [27]. We find a formalization of this idea in causal Bayesian networks [65], a common formalism in statistics, which represent the relationships between random variables with a directed acyclic graph. Bayesian networks have the attractive property that it is easy to derive some independence relations between variables from the graph structure of a Bayesian network. This formalism rigorously clarifies an ambiguity in the term 'information flow', which refers to both causal flow and nomic (regular) associations between variables. We adopt the term *situated information flow* for this sense of information flow in causal context.

4.3.1 Bayesian networks. A Bayesian network represents the joint probability distribution of a set of random variables with a graph. Consider variables $X_1, ..., X_n$ where each X_i takes on values in some set $dom(X_i)$. We use $\mathcal{X}$ to refer to the set $X_1, ..., X_n$ and $dom(\mathcal{X})$ to refer to their joint domain. A Bayesian network (BN) represents the distribution on $\mathcal{X}$ using a graph whose nodes represent the random variables and whose edges represent direct influence of one variable on another.

Definition 4.2 (Bayesian network). A Bayesian network over variables $\mathcal{X} = X_1, ..., X_n$ is a pair (G, Pr). G is a directed acyclic graph with n nodes, each labeled for one of the variables in $\mathcal{X}$. We use $Pa(X)$ to denote the parents of X in the graph. Pr is a mapping of each node X to a conditional probability distribution (CPD), $Pr(X|Pa(X))$. (G, Pr) specifies the joint probability distribution P as the product of the individual nodes' probabilities $P(X|Pa(X))$. Formally, $P(\mathcal{X}) = \prod_{X \in \mathcal{X}} P(X|Pa(X))$.

Example 4. (Figure 1) Alice will be on time for work W if she sets her alarm A early enough and traffic T allows. Bad traffic can be caused by construction C or an accident on the road D.

Bayesian networks give formal clarity to Dretske's theory of information flow. The conditional dependence functions between random variables are the *nomic relations* between events and messages. If
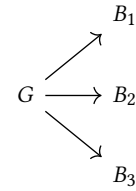
two variables are conditionally dependent on each other, and this conditional dependence is known to the observer of one of the variables, then the observer can infer something (have knowledge of) the other variables. Hence, by our definitions, the variables carry information about each other. If privacy is appropriate information flow, then the privacy of a system will depend on the causal relationships between its components and the environment.

A directed edge between one variable and another indicates a possible conditional dependence between them. Strictly speaking, it does not necessitate that there is a conditional dependence between them. For example, if A is the only member of $Pa(B)$, then there is an edge from A to B, and there is a conditional probability distribution $P(B|A)$. In a degenerate case, this distribution may be such that $A \perp\!\!\!\perp B$; in this case, the network is said to be not *faithful* to the distribution [77]. This has been shown, under reasonable conditions, to be rare in a measure-theoretic sense given the range of all *possible* conditional probability distributions [54].

4.3.2 D-separatedness. A useful property of probabilistic graphical models is that some aspects of the joint probability distribution of all variables represented in the graph can be read easily from the graph's structure. Of particular interest in the analysis of the joint probability distribution is when and under what conditions two random variables are independent.

Definition 4.3 (Path). A *path* between two nodes X_1 and X_2 in a graph to be a sequence of nodes starting with X_1 and ending with X_2 such that successive nodes are connected by an edge (traversing in either direction).

Definition 4.4 (Head-to-tail, tail-to-tail, head-to-head). For any three nodes (A, B, C) in succession on a path, they may be *head-to-tail* ($A \rightarrow B \rightarrow C$ or $A \leftarrow B \leftarrow C$), *tail-to-tail* ($A \leftarrow B \rightarrow C$), or *head-to-head* ($A \rightarrow B \leftarrow C$).

There are two ways in which a variable A can be conditionally dependent on another variable B without one of them being a descendant of the other. The variables may share an unobserved common cause (a tail-to-tail succession) or they may share an observed common effect (a head-to-head succession).

Example 5. (Figure 2) One building in a neighborhood loses power, B_1. One can guess that other buildings B_i around nearby lost power, because power in each building is dependent on the electric grid G. All the buildings may be affected by the common cause of a grid failure.

Example 6. (Figure 1) Suppose we observe that Alice is late for work W, as per our earlier example. This could be due to many

reasons, including traffic T and missing her alarm A. Traffic may be due to construction C or an accident D. The probability of any particular cause is conditionally dependent on the others, because if any one cause is ruled out, the others are more likely.

The existence of a path between two nodes is necessary for their probabilistic dependence on each other. It is not sufficient, particularly when considering their dependence *conditional on other variables*. For this reason, paths in a Bayesian network can be blocked or unblocked based on a set of variables that is otherwise known or observed, the *conditioning set*.

Definition 4.5 (Blocked path). A path is considered to be *blocked* if either:

- it includes a node that is in the conditioning set C where the arrows point to it do not meet head-to-head, or
- it includes a node where arrows do meet head to head, but neither this node nor any of its descendants is in the conditioning set

Intuitively, association "flows" through a Bayesian network through direct causal connections (head-to-tail), through common causes (tail-to-tail), or through observed common effects (head-to-head).

Definition 4.6 (D-separation). If every path from X_1 to X_2 given conditioning set C is blocked, then X_1 and X_2 are d-separated.

THEOREM 4.7. *If X_1 and X_2 are d-separated conditioned on set C, then $X_1 \perp\!\!\!\perp X_2|C$.*

PROOF. In [83]. □

The converse (that independence implies d-separatedness) is not true in general because specific conditional distribution functions can imply independence. Similarly, it is not generally true that the absence of d-separatedness implies conditional dependence. However, is has been shown that conditional distribution functions implying conditional independence are rare in a measure-theoretic sense [32, 48, 54, 77].

4.3.3 Intervention and causality. Bayesian networks support a causal interpretation (as opposed to a merely probabilistic one) through one additional construct, *intervention* [66]. An intervention on a Bayesian network sets the values of one or more of its values. Unlike an observation of a variable, an intervention effectively creates a new graphical model that cuts off the influence of a set variable on its parents and vice versa. Descendants of the set variable are affected by the intervention according to the probability distribution of the original model.

Definition 4.8 (Intervention). An *atomic intervention* setting variable X_i to x_i' on a Bayesian network G creates a new network G' with post-intervention probability distribution $Pr_{x_i'}$

$$Pr_{x_i'}(X_1, X_2, ..., X_n) = \begin{cases} \frac{Pr(X_1, X_2, ..., X_n)}{Pr(X_i = x_i'|Pa(X_i))} & \text{if} X_i = x_i' \\ 0, & \text{otherwise} \end{cases}$$

The intervention construct gives meaning to the directionality of the edges of a Bayesian network. Without it, multiple Markov equivalent graphs can represent the same probability distribution $P(X)$.
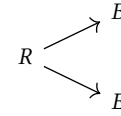

Figure 3: Test score ranks (R) distributed to Bob (B) and Eve (E).

4.4 Ambiguity of information flow

We have drawn a connection between information flow in the Dretske's philosophical sense and Bayesian networks. A Bayesian network is a way of representing the nomic dependencies between phenomena. They are "nomic" because they describe probability distributions that generalize over particular instances of a system's functioning. These nomic relations are factored out as an explicit structure of causal relationships.

This reveals an ambiguity in the very concept of *information flow*, illustrated in the following example.

Example 7. Alice, a teacher tells every student privately their test score's rank R (first in class, second in class, etc.) after every test, with class participation used as a tie-breaker. Alice sends a message B to Bob with the information that he has the second highest rank in the class. Alice also sends a message E to Eve that she has the highest rank in the class. From her message and knowledge of the test environment, Eve learns from her message that Bob was told that he was, at best, second in class. Did information about Bob flow to Eve?

A formal representation of this example makes the root of the ambiguity clear. Consider a three node Bayesian network where R is the test results, B is the message sent to Bob, and E is the message sent to Eve (Figure 3).

There is causal flow along the edges from R to B and from R to E. But an observer of a single variable aware of the system's laws (nomic connections, graphical structure) can learn nomic associations of a message that inform about variables that are not in the message's causal history. Despite E neither causing nor being caused by B, E reveals information about B.

The phrase "information flow" is ambiguous because the word "information" is ambiguous [63]: it can refer to both a message and the contents of a message. We do not favor either sense. Nor do we favor either a view of "information flow" as merely a causal link (a "data transfer"), or as merely an uncaused correlation between observed phenomen. Rather, we propose that to resolve this ambiguity, one has to recognize how the systematic creation and transfer of messages–represented in general by a graph of causal flows–gives each message its meaningful contents. We therefore refer to situated information flows: causal flows that, by virtue of their place in a larger causal structure, have nomic associations. Insisting on this unambiguous formulation of "information flow" is the heart of SIFT.

4.5 Applying SIFT to privacy cases

SIFT can clarify each of the examples of privacy violations raised in Section 2. In the first example, the purchase of scent-free hand lotion gives away to Target the pregnancy of the purchaser. Viewed

$$P \longrightarrow B$$

Figure 4: Pregnancy (P) causes the purchases (B), so B carries information about P. The store, knowing B and this relation, can guess that the buyer is pregnant. B is associated with P though a direct causal link. However, if the causal relation is not known, the association will be opaque.
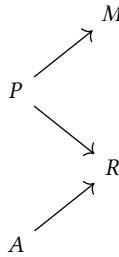
Figure 5: The personality of the data subject (P) causes social media activity (M). It also causes the responses (R) to different kinds of advertisements (A). The analyst knowing these relations can learn about R by observing M because these variables share a common cause. This allows them to choose A to maximize their advantage.

through SIFT, this is due to the fact that pregnancy is a cause of the purchase. This causal relation means that the purchase carries information about the pregnancy (see Figure 4).

In the second example, the social media user is assumed to have a stable personality that is expressed through their online activity. This personality is also a cause of how they will respond to different kinds of targeted advertisements. For the analyst aware of these connections, the data subject's social media behavior carries information about the subject's susceptibility to ads, because these variables share a common cause (see Figure 5).

In the third example, the data subject attempts to conceal their location by turning off the GPS sensor of their smartphone, which they know is caused by their location. However, their location also causes the phone to connect with specific nearby cell phone towers. The cell phone tower locations carry information about the location of the smartphone user when the connected tower IDs are known, because the connect tower IDs are an observed common effect (see Figure 6).

SIFT shows how each of these privacy violation cases is the result of probabilistic and causal relationships. The narrative of each story can be illustrated in a simple graph, and the logic of how the violation occurred is apparent through qualitative yet rigorous analysis of the simple graph.

5 POLICY IMPLICATIONS

SIFT is intended to be a scientifically valid theory of information flow based on probability theory. It is also intended to have simple qualitative consequences that can clarify the reasons for privacy violations and clarify information policy choices. SIFT indicates that restrictions of data transfers will be more clear than restrictions on
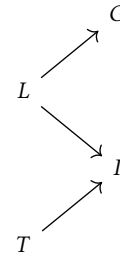
Figure 6: The location of the data subject (L) causes the GPS sensor reading (G). Seeking to hide her location information, the data subject turns off the GPS sensor. However, the location, combined with the locations of the cell phone towers (T) causes a specific sequence of tower IDs I. The smartphone provider is able to use the tower locations T to learn about the location L of the user if they can access the tower ID data I because I is a common cause of L and T.

data flow, because data transfers more specifically refers to causal flows, not flows of association. Where flows of information are specified, these must be framed in terms of observer knowledge or causal structure. And finally, SIFT reveals that the opacity of many data processing systems is due to the complexity of the external causal structure rather than the internal processes of the system. When the generative processes of data can only be learned from data aggregation, this creates a key information asymmetry between data subjects and data processors.

5.1 Focus on data transfers

It can be difficult to determine what topics an information flow is *not* about [60]. SIFT explains this by showing that the meaning of information depends on the structure of its causal context, which may be unknown. However, an information flow requires a causal flow to convey any meaning at all.

This provides some motivation for practical policies to focus on *data transfers*, as opposed to *information flows*. "Data transfer" can refer to a causal flow without any assumptions about the data's semantics. The OECD and APEC guidelines are examples of policies that refer only to data transfers without stipulating any rules about specific categories of information. In contrast, privacy policies and social expectations that restrict information naively based on its contents may be unenforceable or ambiguous. This includes those policies, like the CCPA and GDPR, that require data processors to inform data subjects about the categories of information that they have collected. SIFT predicts that the lists of categories provided to data subjects will almost always be incomplete, because additional meanings can be found in the data with more knowledge of causal context. SIFT confirms that technical solutions that provide secrecy or privacy regardless of data semantics, such as cryptography, are robust and valid means of preserving privacy.

5.2 Observer capabilities and generative conditions

A second implication of situated information flow is that when policies and norms must refer to information 'categories' or meanings, they ought to be defined in terms of either (a) a particular class of observers and take into account their reasonable background information, or (b) a known set of generative conditions that give the information flow its nomic associations.

An example of a law that does this is COPPA. COPPA is very clear about the grounds for associating information with children. It covers service operators that either (a) have actual knowledge that children use the system, or (b) are directed at children. The second condition is a stipulation about the generative structure of the information, or in other words the causal context in which the flow is situated.

A serious challenge for implementing a system that's automatically compliant with such a policy is determining whether and how a system can have any requisite causal knowledge. It has been argued that policy clauses that refer to the knowledge states of human actors cannot be automatically enforced [23]. On the other hand, learning probabilistic graphical structure from data is a well-studied machine-learning problem [29–31]. It may be possible to hold a machine learnt system accountable for inferences that it has been able to make about its data. A direction for future work is developing a theory of information flow security and privacy under conditions where observer knowledge of nomic associations is itself a function of system inputs. In simple cases this may reduce to single program analyses studied in work on Use Privacy [19].

5.3 The opacity of information flows

Scholarly concerns about the opacity of data-driven systems–sometimes referred to problematically as "algorithms" [26]– have raised many questions about the interpetability of machine learning systems. [25, 64] Burrell [12] has identified three sources of opacity: deliberate secrecy of system operators, technical illiteracy of system users, and the properties of machine learning systems that allow them to scale usefully.

SIFT indicates a fourth way that the results of these systems may be "opaque": the causal context of data used and produced by them may be unknown. The meaning of data depends on the network of causal relations that it is situated in. That includes the computational system and the wider environment the system is in. This causal context is rarely represented directly in the data itself. It must be inferred through a separate process.

If the causal context is discovered through the aggregation of many data points, then an individual data subject may be unable to see the meaning of data they give to a web service. This is illustrated by Example 2, about psychographic profiling, in Section 2. A system's opacity can therefore be due to information asymmetry between the user and the system operator. As information assymetries are a well-known source of market failure [2], this raises questions about whether the market for data-driven commercial services is economically efficient [78].

REFERENCES

[1] Christoph Adami. 2004. Information theory in molecular biology. *Physics of Life Reviews* 1, 1 (2004), 3–22.
[2] George A Akerlof. 1970. The market for "lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics* (1970), 488–500.
[3] Edward C. Baig. 2017. Google stops secretly tracking cellular location info. https://www.usatoday.com/story/tech/talkingtech/2017/11/22/google-stops-collecting-cellular-location-info-after-news-probe/886345001/
[4] A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum. 2006. Privacy and Contextual Integrity: Framework and Applications. In *IEEE Symposium on Security and Privacy*. https://doi.org/10.1109/SP.2006.32
[5] Adam Barth, Anupam Datta, John C Mitchell, and Helen Nissenbaum. 2006. Privacy and contextual integrity: Framework and applications. In *Security and Privacy, 2006 IEEE Symposium on*. IEEE, 15–pp.
[6] Adam Barth, John Mitchell, Anupam Datta, and Sharada Sundaram. 2007. Privacy and Utility in Business Processes. In *Proceedings of the 20th IEEE Computer Security Foundations Symposium (CSF '07)*. IEEE Computer Society, 279–294. https://doi.org/10.1109/CSF.2007.26
[7] Gilles Barthe, Pedro R. D'Argenio, and Tamara Rezk. 2004. Secure Information Flow by Self-Composition. In *CSFW '04: Proceedings of the 17th IEEE Computer Security Foundations Workshop*. 100. https://doi.org/10.1109/CSFW.2004.17
[8] Gregory Bateson. 1972. *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. University of Chicago Press.
[9] Sebastian Benthall, Seda Gürses, Helen Nissenbaum, et al. 2017. Contextual Integrity through the Lens of Computer Science. *Foundations and Trends® in Privacy and Security* 2, 1 (2017), 1–69.
[10] Søren Brier. 2008. *Cybersemiotics: Why information is not enough!* University of Toronto Press.
[11] Michael K Buckland. 1991. Information as thing. *Journal of the American Society for Information Science (1986-1998)* 42, 5 (1991), 351.
[12] Jenna Burrell. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* 3, 1 (2016), 2053951715622512.
[13] Carole Cadwalladr and E Graham-Harrison. 2018. The Cambridge Analytica Files. *'I made Steve Bannon's psyhological warfare tool': meet the data war whistleblower* (2018).
[14] Keith Collins. 2017. Google collects Android users' locations even when location services are disabled. *Quartz Media* 21 (2017), 2017.
[15] Thomas M Cover and Joy A Thomas. 2012. *Elements of information theory*. John Wiley & Sons.
[16] Natalia Criado and Jose M Such. 2015. Implicit contextual integrity in online social networks. *Information Sciences* 325 (2015), 48–69.
[17] Anupam Datta. 2014. Privacy through accountability: A computer science perspective. In *International Conference on Distributed Computing and Internet Technology*. Springer, 43–49.
[18] Anupam Datta, Jeremiah Blocki, Nicolas Christin, Henry DeYoung, Deepak Garg, Limin Jia, Dilsun Kaynar, and Arunesh Sinha. 2011. Understanding and protecting privacy: Formal semantics and principled audit mechanisms. In *International Conference on Information Systems Security*. Springer, 1–27.
[19] Anupam Datta, Matthew Fredrikson, Gihyuk Ko, Piotr Mardziel, and Shayak Sen. 2017. Use Privacy in Data-Driven Systems: Theory and Experiments with Machine Learnt Programs. *arXiv preprint arXiv:1705.07807* (2017).
[20] Jenny L Davis and Nathan Jurgenson. 2014. Context collapse: theorizing context collusions and collisions. *Information, Communication & Society* 17, 4 (2014), 476–485.
[21] Terrence W Deacon. 2015. Steps to a science of biosemiotics. *Green Letters* 19, 3 (2015), 293–311.
[22] Francien Dechesne, Martijn Warnier, and Jeroen van den Hoven. 2013. Ethical requirements for reconfigurable sensor technology: A challenge for value sensitive design. *Ethics and Information Technology* 15, 3 (2013), 173–181.
[23] Henry DeYoung, Deepak Garg, Limin Jia, Dilsun Kaynar, and Anupam Datta. 2010. Experiences in the logical specification of the HIPAA and GLBA privacy laws. In *Proceedings of the 9th annual ACM workshop on Privacy in the electronic society*. ACM, 73–82.
[24] Henry DeYoung, Deepak Garg, Limin Jia, Dilsun Kirli Kaynar, and Anupam Datta. 2010. Experiences in the logical specification of the HIPAA and GLBA privacy laws. In *WPES*. 73–82.
[25] Finale Doshi-Velez and Been Kim. 2017. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608* (2017).
[26] Paul Dourish. 2016. Algorithms and their others: Algorithmic culture in context. *Big Data & Society* 3, 2 (2016), 2053951716665128.

[27] Fred Dretske. 1981. *Knowledge and the Flow of Information*. MIT Press.
[28] Luciano Floridi. 2017. Semantic Conceptions of Information. In *The Stanford Encyclopedia of Philosophy* (spring 2017 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
[29] Nir Friedman, Moises Goldszmidt, and Abraham Wyner. 1999. Data analysis with Bayesian networks: A bootstrap approach. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 196–205.
[30] Nir Friedman and Daphne Koller. 2000. Being Bayesian about network structure. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 201–210.
[31] Nir Friedman, Kevin Murphy, and Stuart Russell. 1998. Learning the structure of dynamic probabilistic networks. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 139–147.
[32] Dan Geiger and Judea Pearl. 1990. Logical and algorithmic properties of independence and their application to Bayesian networks. *Annals of Mathematics and Artificial Intelligence* 2, 1-4 (1990), 165–178.
[33] James Gleick and Rob Shapiro. 2011. *The information*. Books on Tape.
[34] Clark Glymour. 2003. Learning, prediction and causal Bayes nets. *Trends in cognitive sciences* 7, 1 (2003), 43–48.
[35] Clark N Glymour. 2001. *The mind's arrows: Bayes nets and graphical causal models in psychology*. MIT press.
[36] Joseph A. Goguen and Jose Meseguer. 1982. Security policies and security models. In *Proceedings of the IEEE Symposium on Security and Privacy*. 11–20.
[37] Alison Gopnik and Henry M Wellman. 2012. Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological bulletin* 138, 6 (2012), 1085.
[38] James W. Gray, III. 1991. Toward a Mathematical Foundation for Information Flow Security. In *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*. 21–34.
[39] Frances Grodzinsky and Herman T Tavani. 2010. Applying the "Contextual Integrity" Model of Privacy to Personal Blogs in the Blogoshere. (2010).
[40] Mireille Hildebrandt. 2014. Location Data, Purpose Binding and Contextual Integrity: What's the Message? In *Protection of Information and the Right to Privacy-A New Equilibrium?* Springer, 31–62.
[41] Mireille Hildebrandt. 2015. *Smart technologies and the end (s) of law: novel entanglements of law and technology*. Edward Elgar Publishing.
[42] K. Hill. 2012. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes* (2012).
[43] Eric Horvitz and Deirdre Mulligan. 2015. Data, privacy, and the greater good. *Science* 349, 6245 (2015), 253–255.
[44] Hsiao-Ying Huang and Masooda Bashir. 2015. Direct-to-consumer genetic testing: contextual privacy predicament. In *Proceedings of the 78th ASIS&T Annual Meeting: information science with impact: research in and for the community*. American Society for Information Science, 50.
[45] Edwin T Jaynes. 1957. Information theory and statistical mechanics. *Physical review* 106, 4 (1957), 620.
[46] Yunhan Jack Jia, Qi Alfred Chen, Shiqi Wang, Amir Rahmati, Earlence Fernandes, Z Morley Mao, and Atul Prakash. 2017. ContexIoT: Towards providing contextual integrity to appified IoT platforms. In *Proceedings of The Network and Distributed System Security Symposium*, Vol. 2017.
[47] Imrul Kayes and Adriana Iamnitchi. 2013. Aegis: A semantic implementation of privacy as contextual integrity in social ecosystems. In *Privacy, security and trust (pst), 2013 eleventh annual international conference on*. IEEE, 88–97.
[48] Daphne Koller and Brian Milch. 2003. Multi-agent influence diagrams for representing and solving games. *Games and economic behavior* 45, 1 (2003), 181–221.
[49] Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110, 15 (2013), 5802–5805.
[50] Niklas Luhmann. 1995. *Social systems*. Stanford University Press.
[51] Bradley Malin and Latanya Sweeney. 2001. Re-identification of DNA through an automated linkage process.. In *Proceedings of the AMIA Symposium*. American Medical Informatics Association, 423.
[52] Alice E Marwick and Danah Boyd. 2011. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society* 13, 1 (2011), 114–133.
[53] John McLean. 1990. Security models and information flow. In *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*. 180–187.
[54] Christopher Meek. 1995. Strong completeness and faithfulness in Bayesian networks. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 411–418.
[55] Deirdre K Mulligan, Colin Koopman, and Nick Doty. 2016. Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy. *Phil. Trans. R. Soc. A* 374, 2083 (2016), 20160118.
[56] Arvind Narayanan and Vitaly Shmatikov. 2006. How to break anonymity of the netflix prize dataset. *arXiv preprint cs/0610105* (2006).

[57] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.
[58] Helen Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.
[59] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.
[60] Helen Nissenbaum, Sebastian Benthall, Anupam Datta, Michael C Tschantz, and Piot Mardziel. 2018. *Origin Privacy: Protecting Privacy in the Big-Data Era*. Technical Report. New York University.
[61] Helen Nissenbaum, Anupam Datta, Michael Tschantz, and Sebastian Benthall. 2017. *The Need for Origin Privacy*. Technical Report. New York University.
[62] Alexander Novotny. 2015. Signs of time: Designing social networking site profile interfaces with temporal contextual integrity. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*. Springer, 547–558.
[63] Geoffrey Nunberg. 1996. Farewell to the information age. *The future of the book* (1996), 103–138.
[64] Frank Pasquale. 2015. *The black box society: The secret algorithms that control money and information*. Harvard University Press.
[65] Judea Pearl. 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. (1988).
[66] Judea Pearl. 1993. [Bayesian Analysis in Expert Systems]: Comment: Graphical Models, Causality and Intervention. *Statist. Sci.* 8, 3 (1993), 266–269.
[67] Judea Pearl. 2009. *Causality*. Cambridge university press.
[68] Irwin Reyes, Primal Wiesekera, Abbas Razaghpanah, Joel Reardon, Narseo Vallina-Rodriguez, Serge Egelman, and Christian Kreibich. 2017. "Is Our Children's Apps Learning?" Automatically Detecting COPPA Violations. (2017).
[69] Andrei Sabelfeld and Andrew C Myers. 2003. Language-based information-flow security. *IEEE Journal on selected areas in communications* 21, 1 (2003), 5–19.
[70] Rath Kanha Sar and Yeslam Al-Saggaf. 2014. Contextual integrity's decision heuristic and the tracking by social network sites. *Ethics and Information Technology* 16, 1 (2014), 15–26.
[71] Andrew D Selbst. 2013. Contextual expectations of privacy. *Cardozo L. Rev.* 35 (2013), 643.
[72] CE Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423.
[73] Pan Shi, Heng Xu, and Yunan Chen. 2013. Using contextual integrity to examine interpersonal information boundary on social network sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 35–38.
[74] Katie Shilton and Kirsten E Martin. 2013. Mobile privacy expectations in context. (2013).
[75] Steven Sloman. 2005. *Causal models: How people think about the world and its alternatives*. Oxford University Press.
[76] Geoffrey Smith. 2015. Recent Developments in Quantitative Information Flow (Invited Tutorial). In *Proceedings of the 2015 30th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS) (LICS '15)*. IEEE Computer Society, 23–31. https://doi.org/10.1109/LICS.2015.13
[77] Peter Spirtes, Clark N Glymour, Richard Scheines, David Heckerman, Christopher Meek, Gregory Cooper, and Thomas Richardson. 2000. *Causation, prediction, and search*. MIT press.
[78] Katherine J Strandburg. 2013. Free fall: The online market's consumer preference disconnect. *U. Chi. Legal F.* (2013), 95.
[79] P Swire and A Antón. 2014. Engineers and lawyers in privacy protection: Can we all just get along? *IAPP Privacy Perspectives, Jan* (2014).
[80] Michael Carl Tschantz, Amit Datta, Anupam Datta, and Jeannette M. Wing. 2015. A Methodology for Information Flow Experiments. In *Computer Security Foundations Symposium*. IEEE.
[81] Michael Carl Tschantz, Shayak Sen, and Anupam Datta. 2017. Differential Privacy as a Causal Property. *arXiv preprint arXiv:1710.05899* (2017).
[82] Zeynep Tufekci. 2014. Engineering the public: Big data, surveillance and computational politics. *First Monday* 19, 7 (2014).
[83] Thomas Verma and Judea Pearl. 1990. Causal networks: Semantics and expressiveness. In *Machine Intelligence and Pattern Recognition*. Vol. 9. Elsevier, 69–76.
[84] Primal Wijesekera, Arjun Baokar, Ashkan Hosseini, Serge Egelman, David Wagner, and Konstantin Beznosov. 2015. Android Permissions Remystified: A Field Study on Contextual Integrity.. In *USENIX Security Symposium*. 499–514.
[85] David H Wolpert. 2008. Physical limits of inference. *Physica D: Nonlinear Phenomena* 237, 9 (2008), 1257–1281.