Legal Concepts Meet Technology: A 50-State Survey of Privacy Laws

Miriam B. Russom
University of Illinois at Chicago
Dept. of Computer Science (MC 152)
Chicago, IL 60607-7053
mberha2@uic.edu

Robert H. Sloan
University of Illinois at Chicago
Dept. of Computer Science (MC 152)
Chicago, IL 60607-7053
sloan@uic.edu

Richard Warner
Chicago-Kent College of Law
565 W. Adams St
Chicago, IL 60661-3652
rwarner@kentlaw.edu

ABSTRACT

Informational privacy consists of the ability to control how others use our personal information. Over the past several decades, we have lost a significant degree of control as a result of advances in information processing technology and the rise of the Internet. We pose two questions: (1) What legal concepts of privacy and privacy protection have developed in response to changing technology? (2) Are these responses adequate?

To answer these questions, we provide a detailed survey and analysis of current US state privacy and security laws, which illustrate the limits of current practical legal concepts of privacy. Our analysis reveals that when evaluated against the background of relevant technological developments, the answer is that the response has clearly been inadequate. We offer the analysis here as a starting point toward a more adequate response.

Categories and Subject Descriptors

K.4.1 [**Computers and Society**] Public Policy Issues – *privacy, regulation.*

General Terms

Legal Aspects, Security, Management.

Keywords

Privacy, personally identifying information, state law, data security, data disposal.

1. INTRODUCTION

Westin, in his seminal 1967 book *Privacy and Freedom* defined informational privacy as the "claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [14]. Many other similar definitions have been given since, but we still really like Westin's definition.

Since 1967, advances in information processing technology and the Web have deprived us of the control we once had. This raises three key questions: *To what extent* should we be able to control our information? To ensure that degree of control, *what types* of information should we protect? And, *how* should we protect it? Westin's general definition does not answer these questions, but legal regulation attempts to do so. The requirements and the

concepts that laws articulate reveal society's current answers, at least as realized in the laws so far adopted.

Our analysis reveals that current legal regulation is a radically incomplete patchwork of rules that is insufficiently grounded in technological reality. This is a common criticism; our contribution is to identify in detail what is lacking with special emphasis on the technological developments that present critical challenges to privacy regulation. We provide our picture through a survey of state privacy laws. Many have already thoroughly canvassed the relevant federal regulation. We feel no need to retread that well-covered ground. Furthermore the 51 different regulatory regimes (the 50 states and the District of Columbia) offer a rich field of privacy regulation in which commonalities and differences reveal an underlying conception of information privacy. We include state laws regulating online informational security in our survey. Privacy and security go hand in hand. Privacy laws identify one type of information that requires protection, and security laws tell us how we must protect it. We classify a law as a security law if its primary aim is to address the problem of unauthorized access to information. By "privacy laws," we mean laws primarily addressing limitations on the use of personally identifying information (PII). Section 2 examines how state laws specify what types of information count as PII. The specifications vary greatly, a variation that we will argue reveals the lack of a satisfying answer to the question of what types of private information merit legal protection. In Section 3 we turn to what our survey of state laws reveals about security laws. The laws concern data security, encryption, data disposal, and prohibitions on spyware. These laws tell us what we should do to prevent unauthorized access and hence comprise a key part of the state law answer to the question of how we are to protect private data. Our assessment is that the laws provide very little guidance. In Section 4, we examine privacy laws, in particular restrictions on data sharing, access and correction requirements, and breach notification requirements. These laws are privacy laws in our sense, because they are laws addressing limitations on the use of PII; however, like the data security laws, they also comprise part of the answer to the question of how we should protect private data by limiting data dispersal (which increases the likelihood of unauthorized access), by allowing access and correction to protect against errors created by both authorized and unauthorized access, and by requiring a remedial response of notification when unauthorized access does occur. We arrive at the same conclusion as we did for security laws: The laws of the 50 states provide little guidance about what we are actually supposed to do.

We will not discuss how much control should we have over our information. Indeed, we argue in Section 5 that US state laws have largely failed to address this critical question. Despite our critique, our goal is not negative but positive. Our critique is a

step toward an adequate response to the impact of technology on informational privacy.

1.1 Related Work

To the best of our knowledge, this work is the first extensive and integrated survey of the modern electronic privacy and security laws of all 50 states from a technological perspective, so there is no previous work just like ours. Here we mention briefly two works connected to our work here.

Smedinghoff [13] surveys the "take reasonable steps to protect data" requirements found in many federal and state laws that address information security. The survey is not focused on PII as ours is, and, since it aims are different, is not as detailed and complete as ours. Schwartz and Solve [12] critique the use of PII in federal laws as a way to define a type of protected information. Their insightful discussion does not address state laws and is not concerned, as we are, with providing a detailed picture of the patterns of protection.

1.2 Methodology

The information in Tables 1 and 2 below were mainly collected from the National Council of State Legislatures (http://www.ncsl.org/) and the direct search of online repositories of state laws. The compilation is not an exhaustive list of all privacy laws (and does not include statutory recognition of the traditional privacy torts such as appropriation). Rather, we choose the statutes surveyed for their relevance to contemporary informational privacy. We first collected the type of statutes that are relevant to our study and examined each of the 50 states (and DC) by analyzing the commonalities and differences in the definitions of PII and the key provisions of the privacy and security statutes. The statutes considered are mostly found in states' trade and commercial codes, business codes, customer record laws, or internet privacy codes.

2. PERSONALLY IDENTIFYING INFORMATION (PII)

The statutes we examine limit access to, and use of, PII, so naturally we want to know what PII is, beyond the general fact that it is information that identifies a person. Incidentally, several of the statutes we examined use the term "personal information," rather than PII but it is clear from the context that all the uses of "personal information" also mean information that identifies a person.

Many of the surveyed statutes specify PII by a list of specific elements that count as PII. The lists vary greatly, a variation that we argue shows a serious gap in our current conception of informational privacy.

2.1 Specification of PII by List

PII is frequently defined as the first name or first initial and last name in combination with any one or more of the following data elements, when the name and the data elements are not encrypted:

1. Social Security Number.

2. Driver's license number or identification card number.

3. Account number, credit card number or debit card number, in combination with any required security code, access code or password that would permit access to the person's financial account.

There is, however, considerable variation by state and within a state among different statutes. We summarize the definitions in Table 1, which we believe is the first systematic catalogue of the treatment of PII in state laws. Table 1 reveals the patterns, and lack of patterns, in regard to PII in state laws.

Virtually every state identifies (1)–(3) above as PII. Table 1 otherwise shows significant state-to-state variation as well as variation within a state among the different types of statutes. Texas, for example, identifies biometric data, fingerprints, mother's maiden name, routing codes, and voice prints as PII in its data disposal law, but not elsewhere. Similarly, Connecticut and Tennessee treat biometric data in as PII in their data disposal but not in their breach notification statute. Minnesota and North Carolina, on the other hand, define PII consistently across all statute types. Indiana and Utah illustrate state-to-state variation; their spyware laws classify as PII types of information not so classified in other states' spyware laws. Both states include biometric data and employment information; Utah (only) includes digital or electronic signatures; Indiana (only) includes fingerprints, mother's maiden names, routing codes, and voice prints. Neither state, however, identifies records of purchases and web histories as PII although other states spyware laws do.

2.2 Other Approaches to Specifying PII

We note two other approaches to specifying PII: one by reference to subject matter; one by reference to public information.

2.2.1 Subject matter specification

Subject matter approaches typically define PII as certain types of health, genetic, and financial information. In the case of health, 15 states have enacted laws similar to the Health Insurance Portability and Accountability Act (HIPAA) of 1996. Thirty-seven states have genetic privacy laws; 16 have laws protecting the privacy of medical and financial information collected for insurance policy purposes. 19 states protect financial information privacy. There is considerable variation, however, in how the statutes define the relevant type of information.

2.2.2 Specification by Reference to Public Information

Utah's Notice of Intent to Sell Non-public personal information is typical of attempts to delimit PII by reference to public information. Under the statute (Utah Code 1953 § 13-102),

(5)(a) "Nonpublic personal information" means information that:
(i) is not public information; and
(ii) either alone or in conjunction with public information, identifies a person in distinction from other persons.
(b) "Nonpublic personal information" includes:
(i) a person's Social Security number;
(ii) information used to determine a person's credit worthiness including a person's:
(A) income; or
(B) employment history;
(iii) the purchasing patterns of a person; or
(iv) the personal preferences of a person.
(6) "Public information" means a person's:
(a) name;
(b) telephone number; or
(c) street address.

Compare Utah's statute to the federal Gramm-Leach-Bliley (GLB) Act. GLB appeals to "public information" to define the type of financial information it protects, but does not define the term. (15 United States Code § 6809(4)(A)). Utah tries to improve

Table 1. Summary of the elements included in the definition of PII for each major type of statute for each state (and District of Columbia). Rows give data elements, columns give statute type.

	Data Disposal	Breach Notification	Security Measures	Spyware
Address	CA, KY	DC, MN	MN	AZ, AK, CA, GA, IN, IA, LA, NH, PA, RI, TX, UT, VA, WA
Biometric data	CO, MA, NC, TN, TX	IA, NE, NC, WI	NC	IN, UT
Digital or electronic signature	NC	NC, ND	NC	UT
DNA		WI		
Driver license # or other state identification #	AL, AZ, AK, CA, CO, CT, GA, HI, IL, IN, KS, KY, MD, MA, MI, MT, NV, NJ, NY, NC, OR, RI, SC, TN, TX, UT, VT, WA	AL, AZ, AK, CA, CO, CT, DE, DC, FL, GA, HI, ID, IL, IN, IA, KS, LA, ME, MD, MA, MI, MN, MS, MO, MT, NE, NV, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, UT, VT, VA, WA, WV, WI, WY	AK, CA, MA, MD, NV, NC, OR, RI, TX, UT	AZ, CA, IN, NH, PA, RI, UT, VA, WA
Social Security #	All above excluding AL	All above	All above	All above plus AK, GA, IA, LA, TX
Financial account/ Credit/Debit Card # / access codes	All above including WI	All above excluding CO	All above	Same as immediately above
Employment information	AL, CA, CO	ND		IN, UT
E-mail address	KY, NC	MN, NC	MN, NC	AZ, IN, NH, RI, TX, WA
Fingerprints	KY, NC, TX	NC, WI	NC	IN
Health insurance information	AL, CA, CT, RI, TN	CA, MO		
Medical information	AL, AK, CA, KY, RI, WA, WI	AK, CA, MO, TX	AK, CA, TX	
Mother's maiden name	NY, TX	ND		IN
Passwords	CO, NC, TN	GA, IA, MO, NE, NC	NC	NH, UT
Payment History / Overdraft history				AZ, AK, CA, GA, IN, LA, PA, RI, TX, WA
Physical description	CA, RI	WI		
Phone number	CA, KY	DC, MN	MN	IN, NH, UT
Record of purchases				AK, CA, GA, LA, PA
Routing code	TX	IA, MO, NE		IN
Voice print	TX	WI		IN
Web history		MN	MN	AK, CA, GA, LA, PA

on GLB by offering lists, but the lists we find in statutes vary greatly.

2.3 What the Variation Shows

The variation in state law definitions of PII shows that there is little or no general agreement on what information should be legally protected. This is a serious failing; the mainstay of privacy legislation has been the assumption that if we adequately protect PII, then we adequately protect privacy. One response is to try to improve the definition of PII. Any such attempt faces a serious technological challenge from the power of current de-anonymization algorithms.

2.4 De-Anonymization: Is Non-PII PII?

The difficulties for privacy may be *worse* than the great variation in state laws' definitions of PII suggest, because recent advances in de-anonymization ensure that, in very many cases, non-PII may in fact identify individuals [6]. Using de-anonymization algorithms, it takes surprisingly little information that is *not*—by *any* of the above lists—PII to uniquely identify a person. So if PII is information that identifies a person, *almost all* information turns out to be PII; however, in an economy and culture that depend on a rich transfer of information, no one seriously proposes significantly restricting access to almost all information.

2.5   Lack of Clarity about what to Protect

Particularly in light of the de-anonymization challenge, there is considerable lack of clarity about what types of information merit legal protection. We urgently need to remedy this situation. We might hope to get some guidance by looking in detail at *how* we try to protect PII. How we do things can often be instructive about what we are trying to do, and indeed that is true in this case, but only in the sense of revealing more problems to be solved. A catalogue of problems is, however, better than no guidance at all.

3.   SECURITY LAWS

Our survey of state privacy and security laws reveals how we try to protect PII. We examine the following types of laws: data security, encryption, data disposal, and anti-spyware. In Table 2 we combine the data on the data security state laws discussed in this section, and the privacy laws discussed in Section 4.

3.1   Data Security Laws

As Table 2 shows, 11 states require businesses to adopt reasonable security measures to protect the confidentiality and integrity of collected PII. Of those 11 states, only 2, Nevada and Massachusetts impose specific measures or standards of reasonableness. California is typical of states requiring reasonable security measures. It requires that "a business that owns or licenses personal information about a California resident shall implement and maintain reasonable security procedures and practices appropriate to the nature of the information, to protect the personal information from unauthorized access, destruction, use, modification, or disclosure" (Cal. Civil Code § 1798.81.5 (b)). In addition, California's law requires businesses to take all reasonable steps to destroy customer records containing personal information once there is no longer reason to retained them.

3.1.1   Costs and Benefits

Losses from unauthorized access to online information run in the billions per year [2,10]. We assume it would be better to avoid this loss; that is, that the gains from allowing unauthorized access (e.g., saving the time, effort, and money otherwise spent in prevention) are not sufficient to offset the losses. This assumption is certainly plausible. According to a study by the Ponemon Institute [9], in 2010 the average cost to organizations of a data breach was $214 per compromised record. The analyzed companies spent, on average, $7.2 million per data breach incident. A United Kingdom government study estimates the yearly cost of data breaches to be £21 billion to businesses, £2.2 billion to government, and £3.1 billion to citizens [2].

Indeed, the consensus is that the cost of improving protection against unauthorized access to an extent that would significantly reduce losses would be considerably less than the aggregate cost of the losses thereby avoided. For example, both general-purpose software engineering textbooks and articles (e.g., [11]) and those devoted specifically to creating software without security vulnerabilities (e.g., [3]) all repeat the statistics on how errors in general, and security vulnerabilities in particular, are by far less

Table 2. Summary of state (and DC) laws. For each of the listed types of security and privacy law, shows which states have enacted a state statute as of June 2011. Top four rows are data security, bottom six rows privacy.

	Type of law	# states	States
Data Security	Data Security measures requirement law	11	AK, CA, MD, MA, MN, NV, NC, OR, RI, TX, UT
	Data Disposal/ Destruction law	30	AL, AZ, AK, CA, CO, CT, GA, HI, IL, IN, KS, KY, MD, MA, MI, MO, MT, NV, NJ, NY, NC, OR, RI, SC, TN, TX, UT, VT, WA, WI
	Spyware law	16	AL, AZ, AK, CA, GA, IN, IA, LA, NV, NH, PA, RI, TX, UT, VA,WA
Data Privacy	Security Breach Notification law	46 (+ DC)	AL, AZ, AK, CA, CO, CT, DE, DC, FL, GA, HI, ID, IL, IN, IA, KS, LA, ME, MD, MA, MI, MN, MS, MO, MT, NE, NV, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, UT, VT, VA, WA, WV, WI, WY (All except AL, KY, NM, SD)
	Online Privacy law	2	MN, NV
	Data Sharing	2	CA, UT
	Privacy Policy on Web sites requirement law	4	CA, CT, NE, PA
	Health Information Privacy	15	CA, ME, MI, MS, NE, NJ, NM, ND, OR, TN, TX, VA, WA, WI, WY
	Genetic Information Privacy	37	AL, AZ, AK, CA, CO, DE, FL, GA, HI, ID, IL, IA, LA, MD, MA, MI, MN, MO, NE, NV, NH, NJ, NM, NY, OR, RI, SC, SD, TN, TX, UT, VT, VA, WA, WV, WI, WY
	Insurance Information Privacy	16	CA, CT, ME, MA, MI, MN, MT, NE, NM, ND, OH, OK, OR, PA, VA, WA
	Financial Information Privacy	19	AL, CA, CT, FL, IL, ME, MD, MN, MS, MO, NH, NC, ND, OK, OR, SC, TN, UT, WA

costly to fix if caught during design, and progressively more expensive to fix during the implementation, testing, and post-release phases. For example, Graf and Van Wyk give a cost ratio of 60 to 1 for post-release fixes versus design-time fixes [3]. Top management of companies believe that the cost of improving protection against unauthorized access is, in theory, well worth paying: "C-level executives [CEOs, CFOs, CIOs, etc.] believe the cost savings from investing in a data protection program of £11 million is substantially higher than the extrapolated value of data protection spending of £1.9 million. This suggests a very healthy ROI for data protection programs" [10]. Unfortunately, in a wide range of cases, there is little consumer willingness to pay higher prices for improved security; thus, a business that improves security when its competitors do not will likely be at a competitive disadvantage. This creates a race to the bottom.

3.1.2 *The problem with reasonableness requirements*

One way we for US state laws (or federal law) to respond to this economic is to require *reasonable* steps to improve data security. However, the reasonableness requirements are problematic because of the role of custom in establishing reasonableness. Conformity to custom and prevailing industry standards is evidence of reasonableness, and it is extremely difficult to overcome a business's claim that it followed industry practice and hence proceeded reasonably.

On occasion, courts *have* rejected such reasonableness claims. The classic famous example studied by most first-year law students is *The T. J. Hooper* case [15]. In March 1928, two tugboats, the *Montrose* and the *T. J. Hooper*, encountered a gale while towing barges, and the tugs and the barges sank. The tugs did not have shortwave radios. With them, they would have received reports of worsening weather, and put in at the Delaware breakwater to avoid the storm. Shortwave radios, however, were new technology, and the custom and practice was for tugs *not* to have a radio. The court nonetheless held that the tugs were negligently unseaworthy because they lacked shortwave radios. There are two key differences between shortwave radios of *The T. J. Hooper* and information security. First, the cost of shortwave radios was relatively small. Thus the cost did not put a barge owner at a competitive disadvantage; indeed, it arguably conferred one since the owner could offer lower risk transport at the same cost as competitors. This was a critical factor in making it unreasonable not to acquire a radio—*even in the market context at the time*. The second difference between shortwave radios in the time of the *T.J. Hopper* and information technology today is that barge owners could easily make a rough and ready comparison between the radio's cost and the expected losses avoided by its use. The losses, when they do occur, could be huge; and, while the occurrence of violent storms is difficult to predict, their occurrence from time to time is certain. This was a key factor in justifying the holding of negligence.

Information security is not like this. Adopting practices that significantly reduce vulnerability costs enough to put businesses at a significant competitive disadvantage, and it is quite difficult for businesses to compare the costs of reducing vulnerabilities to the gains from doing so. The latter point may seem wrong. We argued earlier that we lose billions from unauthorized access, and that improving security would yield a huge net savings. But, we simply know that if we invested more than we now do, we would be better off overall. We do not know exactly *how much* to invest, and that is what businesses need to know. We want them to invest *the right amount*, not too much or too little. The more businesses invest, the less is left for other important business goals. The less they invest, the greater the risk of loss from unauthorized access and hence the greater the risks to society as a whole, however those risks may be divided between businesses and consumers. The investments we want to promote are those that make the optimal tradeoffs. Which ones are those? Without an answer, the law is reluctant to impose requirements that depart from industry practice.

3.2 Encryption Requirements

The intent of encryption is to render information unreadable except for the intended recipient, and its effectiveness depends on the standards and practices adopted. There are currently very few encryption requirements anywhere in American law for data in transit. Only two states, Nevada and Massachusetts, the two states that impose specific standards in their data security laws, specifically impose encryption requirements, and they both have requirements for encrypting data both at rest and in transit. Several other states do have provisions whereby the use of encryption is a "safe harbor" that suffices to show that a company has taken reasonable measures or confers some other legal benefit. For example, in some states, encryption stored data exempts a business from the requirements of breach notification statutes. On the federal level, HIPAA requires the encryption of transmitted health data, and GLB requires the encryption of transmitted financial data. Finally, the Federal Information Security management Act (FISMA) standards impose encryption requirements on federal government entities themselves.

3.2.1 *Industry Standards*

Industry standards play such an important role in this area that we should comment on them at least briefly. Businesses often have a strong market incentive to meet industry standards as the failure to do so may put them at a competitive disadvantage; in addition, it can be a basis for a lawsuit.

The Payment Card Industry Data Security Standards (PCI-DSS), for example, require encryption when financial data is transmitted wirelessly or over a public network [7]. Apart from financial data, encryption during transmission was not, until recently, an industry standard. Web businesses used an HTTPS connection only for the user authentication session. This may be changing. Google recently enabled an HTTPS connection by default for Gmail users, and it offers HTTPS as an option with its search engine and various other services. Facebook has also started allowing an entirely HTTPS connection; however, it is not the default setting.

3.2.2 *Benefits of Encryption*

Encrypting data at rest can provide protection against unauthorized access by insiders (e.g., nosy employees). The biggest potential benefit, however, is probably when a computer is lost or stolen, a common occurrence in this era of laptops, smartphones, and working remotely.

A major benefit of encrypting data in transit is that eavesdroppers cannot capture the data. The greatest vulnerability here is probably to end-users connecting wirelessly to the Internet. However, there may be some benefit to companies too. A 2008 Ponemon Institute study [8] found that organizations with an encryption strategy have a statistically lower rate of data breaches, especially organizations that adopt an enterprise-wise encryption strategy. (Of course, correlation does not imply causation, and there are multiple plausible explanations of that finding.) Certainly many of the large number of data breaches of the past decade would have been prevented by one or both of the encryption of data at rest and data in transit.

3.2.3   Costs of Encryption

Cryptography in general imposes costs in the form of complexity, the performance cost of actually carrying out the encryptions and decryptions, increased data management, and often licensing fees for the encryption tools.

Encryption of data at rest is relatively easy and relatively low cost, except for the nuisance cost to the user. For the case of any given employee's laptop, it is almost painfully easy, as encryption services for files are built into Windows 7, and have been in Mac OS X for many years now.

In general, encryption of data in transmission specifically imposes costs in the form of additional CPU load, network overhead, and latency. Other things being equal, the providers of "free" web services (e.g., search, social networking, and email) would prefer to use the additional CPU cycles for a variety of competing purposes instead of deploying them for encryption and decryption. Equally or more importantly, businesses facing competition do not want to negatively affect users' perceptions of response times and the overall experience, and encryption necessarily adds at least some delay.

The cost of encryption can be measured. For example, in 2010, Kounavis et al. determined that on a 3 GHz Intel i7 processor encrypting the content of a typical 140KB banking transaction required 2.3–4.8 million clocks, and that performing an RSA 1024 private decrypt operation required 0.9–1.4 million clocks [5]. To make this even more specific to the case at hand, let us consider the performance of TLS (sometimes also referred to as SSL/TLS), since if the typical interaction between a large company and a consumer is to be encrypted, it would be via a secure web connection using TLS. Coarfa et al. [1] reported that a secure TLS web server imposes a factor of 3.4 to 9 overhead (depending on e.g., which cipher suites are used, key size, and size of the transferred file) versus an insecure web server on the same hardware. Still, these costs seem modest compared to the costs of some of the famous data breaches.

3.3. More Encryption Requirements?

Should we have broader legal requirements for encryption in both transmission and storage? Or should we leave the resolution of the extent of encryption to the market? We see no clear answer at present. Further study is in order.

3.3   Data Disposal Laws

Old, residual data typically remains where it is; for sensitive data, this raises concerns about unauthorized access. Thirty states require the disposal of records that contain PII when the holder of the data decides it no longer wants to maintain that data. Arkansas, California, Connecticut, Georgia, Indiana, Kansas, Kentucky, Massachusetts, Michigan, Montana, Nevada, New Jersey, New York, Rhode Island, South Carolina, Texas, Tennessee, Vermont and Wisconsin specifically outline how the data must destroyed: shredding, erasing, or otherwise modifying the personal information to make it unreadable. The remaining 11 states, Alaska, Arizona, Colorado, Hawaii, Illinois, Maryland, Missouri, North Carolina, Oregon, Utah and Washington, require the use of a reasonable disposal standard without specifically enumerating the methods of data destruction. California's data disposal law is typical of the those that enumerate methods of data destruction: "A business shall take all reasonable steps to dispose, or arrange for the disposal, of customer records within its custody or control containing personal information when the records are no longer to be retained by the business by (a) shredding, (b) erasing, or (c) otherwise modifying the personal information in those records to make it unreadable or indecipherable through any means" (Cal. Civ. Code §1798.81).

California's law illustrates a lack of understanding of technology, an aspect of legal regulation that is unfortunately all too common. The law requires businesses to take reasonable steps to dispose of data; given current technology, it is not reasonable to attempt to make data "unreadable or indecipherable through any means." This is like telling a high school basketball player who is incapable of dunking the ball, "Make a reasonable effort to put the ball in the basket, and, by the way, be sure you dunk it." One occasional legal rationale for imposing impossible-to-meet standards is to force technology to develop to enable businesses to comply; however, there is no indication that any such rationale is in play here; lack of understanding is the likely explanation.

In fact, the technical computer security community knows a lot about making disposed data difficult to recover. (See Hughes et al. [4] for a good overview.) A summary of what we know is that one has to consider what level of effort to recover the data one is guarding against. Very few, if any, methods are secure against a major national-government level of effort, which presumably is not what was intended. There are reports of data being recovered from residual pieces of a hard disk that was physically split into pieces, if those pieces were not too small. Since "unreadable by any means" is problematic, consider California's choice of "erasing." (Presumably "shredding," is meant for paper records.) In the realms of electronic records the word "erasing" is open to many interpretations. The simplest form of erasing, simply to delete a record and empty the trash is certainly not sufficient for data to make it unrecoverable by a moderately sophisticated adversary. It seems a big stretch to interpret "erasing" as "erasing and then overwriting," and even the efficacy of overwriting against a sufficiently sophisticated adversary depends on the number of overwriting rounds

3.4   Anti-Spyware Laws

Sixteen states have enacted Anti-Spyware laws that make it illegal to install software without consent on someone else's computer in order to collect PII (as variously defined in the statutes). California's Consumer Protection Against Computer Spyware Act is typical. It prohibits an unauthorized person from knowingly installing or providing software that performs certain functions, such as taking control of the computer or collecting personally identifiable information, on another user's computer located in California. The act requires "willful or intentional deceptive actions" to trigger a violation (Cal. Bus. & Prof. Code § 22947.2).

The problem with spyware laws is that the statutes only apply where consent to the installation of the software is absent, and the "consent" requirement here is very weak. All that is required is that the businesses provide relevant information in a written statement, and that, after being presented with the information, consumers indicate—not necessarily their free and informed choice—but merely their willingness to proceed with the transaction. Different implementations suggest or require different indications of willingness—clicking on an "I agree" button, for example; or, simply using the web site or making a purchase on a page that has a "legal terms and conditions" hyperlink. Call this "in-practice consent." Even though it is well documented that the vast majority of consumers do not read standard form contracts, in-practice consent may be a good proxy for real free and informed consent in the purchase of well-understood products, for example, the purchase of a coffee maker or toaster. Consumers

are familiar with the products. What we have done is extend this treatment to technological sophisticated areas in which the requisite consumer understanding is lacking. Consumers do not have a good understanding of how software works and of the changes it can make in their hard drives when it is installed, and the vast majority of consumers do not read the disclosures anyway. Advertisers exploit this fact to legally distribute tracking software (and tracking cookies) with in-practice consent. This is unacceptable, or at least highly questionable, where in-practice consent is a poor proxy for real consent.

4. PRIVACY LAWS

The state privacy laws we survey divide into the following: data sharing laws; access requirements, correction requirements; and, breach notification requirements.

4.1 Data sharing laws

California is the only state that requires a business to disclose, upon consumer request, the type of PII it shared and the identity of the third parties with whom it shared the information. Utah requires commercial entities with the intention to sell personal information to notify the customer before providing the data. Two states (and only those two)—Minnesota and Nevada— require consent from subject of the information before an ISP may disclose PII as defined in the statutes (ISPs may disclose PII to other ISPs if necessary to enforce the ISPs Terms of Use agreement). Otherwise, the states regulate data sharing only through subject-specific statutes. There are no other laws that regulate private business data sharing generally.

Advances in information processing technology and the Internet have made data sharing common. As evidenced by the minimal legislation in this area, the law has been slow to respond. This is a particularly pressing concern. One danger is that we will share too little information. The Internet's information sharing potential depends on the willingness of its users to share information. One key factor limiting the willingness to share information is the lack of control over its subsequent use and distribution. Individual transfers of information, seemingly innocuous in the context in which they occur, may, when aggregated, constitute an invasion of privacy. Another danger is that we will share too much. Privacy advocates fear a world in which a trail of information, readily accessible by others, follows us everywhere, and they warn of disastrous consequences. There is certainly an argument for greater consumer control over information sharing, either through law, market forces, or custom.

4.2 Access and Correction Requirements

None of the laws surveyed so far give an individual the right to access *and* correct records. The data sharing laws discussed above provide access but do not allow for correction.

Access and correction requirements appear attractive as a way for consumers to regain control over their information. But there are three problems. First, data sharing ensures that information about a consumer is spread over a vast range of databases owned by many different entities, which makes automated checking and correction problematic. Checking all the data is a task few, consumers can, or want, to perform. Second, consumer access to information burdens the business that must provide it; a very large number of requests could impose serious costs. Third, access alone provides little control over a consumer's information unless the consumer has the power to correct errors, and none of the statutes surveyed grants consumers a right to correction.

Despite these difficulties, some employment-related law, medical-records-related law, and state-agency-related laws may include correction requirements.

4.3 Breach notification laws

Forty-six states and the District of Columbia have enacted security breach notification laws that require organizations that own or license personal information to notify individuals when there is reason to think there has been unauthorized access to PPI of which they are the subject. In some cases, the notification requirement applies only to unencrypted data.

The rationale behind breach notification statutes is that the consumer can take protective steps to prevent identity theft, financial fraud, or other undesirable consequences. It is far from clear that this is a good rationale. Consider identity theft. While the losses from an incident may be larger, the likelihood that compromised information will result in identity theft is very small; consequently, the expected loss is very low, and it is likely that the protective steps will actually cost more than the expected loss. This does not mean losses from unauthorized access are not a problem. *Aggregate* losses run in the billions. Table 3 illustrates the extent of the problem. Given the low *individual* expected losses, empowering consumers to defend themselves is not a viable response to the problem of large aggregate losses.

5. CONCLUSION

To conclude, we return to the three questions we posed at the beginning. How much control should we have over our information? To ensure the right degree of control, what types of information should we protect? And, how should we protect it? The laws surveyed do not provide adequate answers.

How much control should we have over our information? The question concerns tradeoffs between protecting informational privacy and permitting its processing to achieve a variety of ends, including improved economic efficiency, security, inventory control, marketing, business planning, and customer relationships. Few want to forego these advantages entirely, so we need to strike a balance between realizing the benefits and promoting informational privacy. Unfortunately, the laws surveyed fail to address the issue; in this way, they have simply failed to keep pace with the possibilities created by technology.

To ensure the right degree of control, what types of information should we protect? The laws surveyed provide a clear but problematic answer: protect PII. The variation in the statutes, however, shows a lack of consensus about what we should protect under the PII rubric, and the power of de-anonymization algorithms is a serious challenge to *any* attempt to delimit PII; even very small combinations of information turn out to be sufficient to uniquely identify a person, so a great deal of "non-PII" is PII. No one, however, proposes restricting access to almost all information in an economy and culture that thrive on a rich transfer of information.

How should we protect private information? Data security laws rely on an unexplained reasonableness standard, which does not provide adequate guidance about what we are actually to do. An encryption requirement is arguably a concrete step toward developing a detailed conception of reasonableness; however, the laws surveyed do not go very far in this direction. Anti-spyware statutes are also arguably a concrete step toward specifying reasonableness. The statutes prohibit the unauthorized installation of software that collects PII; however, mere in-practice consent is taken as sufficient for authorization even though such consent is a

poor proxy in these cases for genuine free and informed consent. Limitations on data sharing can arguably be seen as a reasonable step to reduce the likelihood of unauthorized access (by limiting its dispersal), but there is little legal constraint on data sharing even though it would be desirable to give consumers more control over data sharing than they now have. Access and correction requirements are not primarily concerned with limiting or responding to unauthorized access, and, in any case, such requirements are not widespread and face serious practical and technical barriers. Breach notification statutes are a remedial response to unauthorized access, but they are not an effective response to what is the real problem—large aggregate loss.

Our ultimate aim is not critique but progress. We intend our criticisms to define a starting point for working toward an adequate response to the impact of technology on informational privacy requires combining the expertise of computer scientists and policy makers. The necessary foundation for such cooperation is a shared understanding of the problem, and a common language in which to discuss and analyze proposed solutions. We offer this paper as a step toward that goal.

6. ACKNOWLEDGMENTS

Table 3. History of Data Breaches from 2002–2011. Data from Open Security Foundation: Data Loss Database

Year	# of incidents	# of compromised records	Breach notification enactments by year
2002	7	268,998.00	
2003	15	7,061,950.00	California*.
2004	25	3,717,590.00	
2005	142	55,988,256.00	Arkansas+, Delaware*, Florida*, Georgia*, Nevada*+, New Jersey*+, New York, North Carolina+, North Dakota*, Tennessee* and Washington*+.
2006	538	51,251,706.00	Arizona* +, Colorado* +, Connecticut*+, Idaho*+, Illinois, Kansas*+, Louisiana+, Minnesota*, Montana*+, Nebraska*+, Ohio+, Pennsylvania*+, Rhode Island* and Wisconsin+.
2007	551	165,262,288.00	Maine, Massachusetts*, Michigan*, New Hampshire+, Oregon, Utah, Vermont, Wyoming and Washington D. C.
2008	788	86,930,908.00	Hawaii*, Iowa, Maryland, Oklahoma*, Virginia and West Virginia*.
2009	616	221,693,990.00	Alaska*, Indiana*, Missouri, South Carolina* and Texas.
2010	460	28,081,731.00	
2011	NC	NC	Mississippi*.

*: States with safe harbor for encryption .

+: States with risk based notification: breach notification is mandated where there is potential harm to the individuals.

7. REFERENCES

[1] Coarfa, C. et al. 2006. Performance analysis of TLS Web servers. *ACM Trans. Computer Systems*. 24, 1 (Feb. 2006), 39–69.

[2] Detica and Cabinet Office of the United Kingdom 2011. *The Cost of Cyber Crime*.

[3] Graff, M. and Van Wyk, K.R. 2003. *Secure Coding: Principles and Practices*. O'Reilly Media, Inc.

[4] Hughes, G.F. et al. 2009. Disposal of Disk and Tape Data by Secure Sanitization. *IEEE Security & Privacy*. 7, 4 (2009), 29–34.

[5] Kounavis, M.E. et al. 2010. Encrypting the internet. *Proceedings ACM SIGCOMM 2010* (2010), 135–146.

[6] Narayanan, A. and Shmatikov, V. 2010. Myths and fallacies of personally identifiable information. *Commun. ACM*. 53, 6 (2010), 24–26.

[7] PCI DSS Data Security Standards Overview: *https://www.pcisecuritystandards.org/security_standards /*. Accessed: 2011-07-04.

[8] Ponemon Institute 2008. *2008 Annual Report: U.S. Enterprise Encryption Standards*.

[9] Ponemon Institute 2011. *2010 Annual Study: U.S. Cost of a Data Breach*.

[10] Ponemon Institute 2010. *Business Case for Data Protection: A Study of CEOs and other C-level Executives in the United Kingdom*.

[11] Pressman, R. 2009. *Software Engineering: A Practitioner's Approach*. McGraw-Hill.

[12] Schwartz, P.M. and Solove, D.J. 2011. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*. 86, (2011), forthcoming.

[13] Thomas J. Smedinghoff 2008. Defining the Legal Standard for Information Security: What Does

"Reasonable" Security Really Mean? *Securing Privacy In The Internet Age*. Chander et al., eds. Stanford University Press. 19–40.

[14]   Westin, A. 1967. *Privacy and Freedom*. Atheneum Press.

[15]   Circuit Court of Appeals, 2nd Circuit 1932. The TJ Hooper, Vol. 60 *Federal Reporter* 2nd ed. 737–740.