2010

Rules, Standards and Geeks

Derek E. Bambauer

Follow this and additional works at: https://brooklynworks.brooklaw.edu/bjcfcl

Recommended Citation
Available at: https://brooklynworks.brooklaw.edu/bjcfcl/vol5/iss1/2

This Article is brought to you for free and open access by the Law Journals at BrooklynWorks. It has been accepted for inclusion in Brooklyn Journal of Corporate, Financial & Commercial Law by an authorized editor of BrooklynWorks.
INTRODUCTION

When it comes to regulating technology, the age-old debate between rules and standards tilts heavily towards standards. Rules, for all their clarity, are seen as slow-changing tools in industries characterized by dynamism. They are also viewed as being both under- and over-inclusive, and in prizing form—one means of achieving a desired result—over substance—the result itself. Moreover, setting legal rules for technology risks creating lock-in, which may cement a given technology in place. In short, standards—particularly standards that look to industry best practices—are lauded as the best means for governing code through law.

This Article, though, argues that rules are preferable for regulating data security, at least under certain conditions. In part, this is so because data security typically focuses on controlling the wrong set of events. Security is often preoccupied with regulating access to data—in particular, with preventing unauthorized access. Yet, strangely, unauthorized access is ubiquitous. Employees lose laptops, hackers breach corporate databases, and information is inadvertently e-mailed or posted to the public Internet.

* Associate Professor of Law, Brooklyn Law School. A.B., Harvard College; J.D., Harvard Law School. The author thanks Lia Sheena, Lia Smith, and Carolyn Wall for expert research assistance. Thanks for helpful suggestions and discussion are owed to Miriam Baer, Ted Janger, Thinh Nguyen, and Jane Yakowitz. The author welcomes comments at derek.bambauer@brooklaw.edu.


2. See, e.g., Daniel Gervais, The Regulation of Inchoate Technologies, 47 HOUS. L. REV. 665, 702 (2010) (stating that “an inchoate technology may provide a better solution than regulation—perhaps industry-based standards will emerge making legal regulation unnecessary at best and potentially counterproductive”).


This Article argues that preventing data breaches is not only the wrong goal for regulators, it is an impossible one. Complex systems design theory shows that accidents are inevitable. Thus, instead of seeking to prevent crashes, policymakers should concentrate on enabling us to walk away from them. The focus should be on airbags, not anti-lock brakes. Regulation should seek to allow data to “degrade gracefully,” mitigating the harm that occurs when a breach (inevitably) happens.

Such regulatory methods are optimally framed as rules under three conditions. First, minimal compliance—meeting only the letter of the law—is sufficient to avoid most harm. Second, rules should be relatively impervious to decay in efficacy over time; technological change, such as increased CPU speeds, should not immediately undermine a rule’s preventive impact. Furthermore, compliance with a rule should be easy and inexpensive to evaluate. In addition, rules are likely to be helpful where error costs from standards are high; where if an entity’s judgment about data security is wrong, there is significant risk of harm or risk of significant harm. Finally, this argument has implications for how compliance should be assessed. When regulation is clear and low-cost, it creates an excellent case for a per se negligence rule, or, in other words, a regime of strict liability for failure to comply with the rule. This Article thus addresses not the desirability of regulation—when data security should be mandated—but rather how to structure that regulation once it is deemed worthwhile.

The debate about framing legal commands as rules or as standards is a venerable one. Scholars have addressed the dichotomy in contexts from real property rights to patent law to antitrust. The merits and shortcomings of each approach have been analyzed from a variety of theoretical perspectives.


12. See Duffy, supra note 1, at 611.

13. See Crane, supra note 1, at 52.

constrained; they help both regulated and regulators assess compliance more cheaply and easily. In addition, they may prevent abuse by conferring less discretion on regulators. However, rules are often under-inclusive—failing to cover behavior that should fall within their ambit, or failing to prevent risks they are designed to address—or over-inclusive—imposing burdens on unrelated actors or activities. Standards, by contrast, are more readily adapted to complex or changing situations, but often at the price of predictability and cost.

The discussion becomes more complex when we recognize that the distinction is continuous rather than binary. Standards can be rule-like, and rules standards-like. Consider two security mandates: “encrypt,” and “follow industry best practice for securing data.” The former looks like a rule, and the latter like a standard. However, “encrypt” could be seen as a standard: the command specifies a method, but leaves the implementation entirely up to the regulated entity. Encryption has been used since the days of Mary, Queen of Scots; its modes range from simple (and simply cracked) transposition ciphers to elliptic curve cryptography. Even a more specific command like “encrypt using asymmetric key cryptography” can be met with a variety of responses. The RSA, ElGamal, and DSS key techniques all meet the criterion, but have important differences among them. Thus, a rule can be transformed into a standard by altering the level of specificity.

Similarly, “follow industry best practice for securing data” could be a rule. If, for example, the industry has standardized on the use of SSL (Secure Sockets Layer) to safeguard sensitive data while it is being communicated over a network, that best practice standard effectively becomes a rule: “use SSL.” Thus, even if an alternative technique were demonstrated to be functionally equivalent, it would not comply with the standard, even though standards are typically viewed as ends-driven and not

20. Id. at 7–8.
means-driven. In short, the line between rules and standards blurs, particularly as a rule’s command becomes more general.

The Article next assesses the conventional wisdom for technological regulation, which holds that standards are the preferred modality. It then turns to arguments in favor of using rules instead, under certain defined conditions. Finally, it closes with observations about the larger role of technology regulation in the context of data security in the payment system.

I. THE VIRTUES OF STANDARDS

Technology changes quickly; law, slowly. Most commentators favor standards when dealing with technological regulation of issues such as security, for at least five reasons.

First, standards allow regulated entities to comply in a more cost-efficient fashion than rules. Requiring a particular technology or approach may be unnecessarily expensive, especially where infrastructures differ significantly, where there are a range of alternatives, or where the endpoint can be achieved without applying technology in some situations.24 Rules can limit creativity in achieving regulators’ goals.25

Second, standards can be less vulnerable to obsolescence. Rule-based specifications may decay quickly when technology changes rapidly. This either undercuts the efficacy of regulation, or forces frequent updates to it. The Clipper Chip controversy of the mid-1990s provides a potent example; regulation that mandated use of one particular encryption technique might well have undercut the deployment of e-commerce and other advances dependent on data security.26

Third, standards can minimize the ill-effects of information asymmetry regarding technology.27 Regulators may not know what technologies are cutting-edge or appropriate or unnecessarily costly. Standards can wrap in expertise from regulated entities while meeting regulatory goals.

Fourth, standards may deal better with interoperability concerns. Most organizations have heterogeneous information technology environments for a variety of reasons: mergers, legacy systems, customer demands, and so forth. Regulations that specify a particular technology, or method of compliance, may make demands that are impossible or inapposite. For example, Deutsche Bank used the IBM operating system OS/2 long after

most other customers had migrated to Microsoft Windows or a UNIX platform. Thus, requirements tied to Windows (for example, using the NTFS file system) or to software only available for that operating system would have forced Deutsche Bank into a costly migration, or to fall out of compliance. In contrast, a standard that specifies its goal, but is technology-agnostic, allows entities with a range of infrastructures to comply adequately.

Finally, selecting one technology for regulatory compliance risks producing market-making effects. Regulation may confer success, or at least widespread adoption, on a single product or company—a problem that worsens if the technology is sub-optimal. For example, the memory chip manufacturer Rambus was able to influence the industry group JEDEC (Joint Electron Device Engineering Counsel) to adopt, as part of its standard for SDRAM (Synchronous Dynamic Random Access Memory), technology over which Rambus held patent rights. (Indeed, Rambus actually amended its pending patent applications to conform better to the JEDEC technology.) This led to lawsuits against Rambus for fraud, and to an initial Federal Trade Commission (FTC, or Commission) finding that the company had engaged in antitrust violations (under Section 2 of the Sherman Act). However, Rambus emerged unscathed from both the suits and the FTC investigation. Similarly, a legal mandate to incorporate a particular technology could create market power for that technology’s owner, particularly if the technology were protected by intellectual property rights such as a patent. Thus, a rule may entrench a single technology into a powerful if not unassailable market position.

The use of standards in technology regulation is a familiar aspect of the data payment system in the United States. For example, the FTC imposed standards-based requirements for the security of non-public information, known as the Safeguards Rule, as part of its rulemaking authority under the Gramm-Leach-Bliley (GLB) Act. The Commission mandates a “comprehensive information security program that is written in one or more

---

30. Id.
readily accessible parts and contains administrative, technical, and physical safeguards that are appropriate to [an organization’s] size and complexity, the nature and scope of its activities, and the sensitivity of any customer information at issue.”\(^\text{34}\) Regulated entities must perform a risk assessment, and then “[d]esign and implement information safeguards to control the risks [it] identifies through risk assessment, and regularly test or otherwise monitor the effectiveness of the safeguards’ key controls, systems, and procedures.”\(^\text{35}\) Thus, the GLB Act is a purposive regulatory standard: it sets goals, and identifies key areas and targets, but is method-agnostic. Financial institutions can implement its requirements using the technology they think best fits their infrastructures and businesses. The Commission’s final rulemaking emphasized that the “standard is highly flexible,” and the notice repeatedly reassured regulated institutions that its approach was fact-specific and contextual.\(^\text{36}\)

Indeed, there are zones of regulatory concern regarding payment data security where standards appear superior. One example is application design. As I have written elsewhere, both custom-designed and off-the-shelf applications in the payment system suffer from security flaws.\(^\text{37}\) Some of these bugs result from coding errors; others, from the inherent complexity of data processing and from interactions between systems and data stores.\(^\text{38}\) As Microsoft’s Patch Tuesday ritual reminds us, bugs are inevitable.\(^\text{39}\) They can be minimized, but not eliminated.\(^\text{40}\) Thus, as with data losses and security breaches themselves, the best regulatory goal for application design is to minimize bugs.\(^\text{41}\) Software design involves the familiar trade-off between time and cost versus greater security, with a minimum optimal bugginess greater than zero.

For application design, then, the critical regulatory issue is methodology: setting parameters for the design, testing, and deployment of the software.\(^\text{42}\) Again, this approach is familiar to the payment industry. The Payment Card Industry Data Security Standard (PCI DSS) Requirements and Security Assessment Standards, promulgated by an industry association founded by payment card networks such as American Express, create

\(^{34}\) \textit{Id.} at 36,494.

\(^{35}\) \textit{Id.}

\(^{36}\) \textit{Id.} at 36,488.

\(^{37}\) Derek E. Bambauer & Oliver Day, \textit{The Hacker’s Aegis}, 60 \textit{Emory L.J.} (forthcoming 2010) (manuscript at 8).

\(^{38}\) See \textit{generally id.} at 8–10.


\(^{40}\) See Bambauer & Day, supra note 37 (manuscript at 8–14).

\(^{41}\) See \textit{generally Frederick P. Brooks, Jr., THE MYTHICAL MAN-MONTH: ESSAYS ON SOFTWARE ENGINEERING} (3rd prtg. 1979).

private law regulation of customer account data. To comply with PCI DSS, an organization must develop its software applications in accordance with the DSS standards, and with industry best practices. Requirements include validating application input to prevent buffer overflow and cross-site scripting (CSS) attacks, checking error handling, validating encrypted storage, validating communications security, and checking role-based access controls. Organizations must implement code review for custom software before deploying applications. Public Web applications are subject to additional standards, such as developing based on the Open Web Application Security Project Guide, and protecting against newly discovered vulnerabilities by using a firewall or vulnerability assessment tools. The goal of these requirements is to prevent breaches from common attacks, such as the SQL injection attack that caused the data spill at Heartland Payment Systems.

PCI DSS, as its moniker suggests, is framed as a standard and not as a rule. This is clear from its focus on process, such as engaging in code review, and on goals, such as protecting against new attacks or vulnerabilities. Thus, for example, PCI DSS requires validating secure communications, not using a particular secure communications technology such as SSL. Application design is a sensible target for standards-based regulation, for at least three reasons. First, history matters. Most financial institutions maintain legacy systems, such as mainframe-based applications, due to the cost and difficulty of upgrading. It may be impossible for them to employ a given technology to achieve security without expensive wholesale changes to their infrastructure. Second, systems heterogeneity means that even applications with a common goal, such as connecting to


45. Id. at 32.

46. Id. at 33.


48. PCI SECURITY PROCEDURES, supra note 44, at 31.

payment networks, likely must be custom-coded. Forcing financial institutions to use one technology or method to gain security ends would drive up their costs unnecessarily. Finally, here rule-based specifications seem more vulnerable to decay. New attacks and vulnerabilities appear constantly. Having a single approach to security across the financial industry may, like monoculture agriculture, leave institutions vulnerable to a single new pathogen. In short, security may well degrade rapidly, rather than slowly. For these three reasons—legacy systems, customized code, and rapid degradation—a standards-based regime is preferable to a rule-based one for application design.

Regulation by standards rather than rules is the established norm in the data payment system. Indeed, as the discussion of application design demonstrates, this preference may be sensible in some areas. However, standards are not always superior. The next section explores the virtues of regulation by rules for security.

II. THE VIRTUES OF RULES

Arguing for rules in technological regulation is an uphill climb: they can become obsolete rapidly, may increase costs by forcing entities to comply in a highly specific fashion, and may be both over- and under-inclusive. Yet, this Article argues that rules are preferable to standards when at least three conditions hold: sufficient minima, slow or low decay, and inexpensive verification.

First, a rule is helpful when the specified level of data security—effectively, a minimum—suffices in most or all circumstances. One example would be to mandate that transmission of data take place over a connection protected by 128-bit SSL. SSL certificates are widely and cheaply available, and root certificates are built into all major browsers. Currently, 128-bit SSL traffic is proof against brute-force decryption attacks even when adversaries use clusters or supercomputers. Thus, 128-bit encryption is strong enough to protect data in communication, even if

52. See generally Daniel D. Chiras, Environmental Science 116 (8th ed. 2010).
53. See Pci Security Procedures, supra note 44.
55. Id.
institutions do not take additional measures, such as protecting against eavesdropping.\textsuperscript{57}

A corollary is that rules may be helpful where the impact of a data breach is high, and where the specified technology raises the cost to an attacker or discoverer of captured information. One example here is hard drive encryption. Stories of lost laptops, backup tapes, and USB drives are legion. Here, rules serve not to prevent loss—indeed, hard drive encryption is only useful after the loss has taken place—but to reduce its effects.\textsuperscript{58} Similarly, a rule mandating logging of access to sensitive data cannot prevent an employee from copying down customer account information displayed on a computer monitor, but can aid an institution to detect what has been revealed in the breach, and perhaps to minimize its spread.\textsuperscript{59} This condition requires that the rule specify protection that is good enough in most or all cases.

Second, rules work well when they need not be frequently updated—in other words, when they decay slowly. This reduces the administrative cost of the rule, and allows it to retain effectiveness over time.\textsuperscript{60} 128-bit encryption, for example, will likely suffice against brute-force attacks for at least ten years, given current rates of advance in CPU clock cycles and parallelization.\textsuperscript{61} To take another encryption case study, DES (Data Encryption Standard) was adopted as a Federal Information Processing Standard in 1976.\textsuperscript{62} It remained impervious to commercial-level decryption (as opposed to governmental attacks) until the late 1990s.\textsuperscript{63} A technology-

\textsuperscript{57} “Man in the middle” attacks against SSL are still theoretically possible, but financial institutions (unlike end users) should be sophisticated enough to take steps such as verifying certificate signatures to safeguard against such hacks. See, e.g., Larry Seltzer, \textit{SSL Man-in-the-Middle Attack Exposed}, PCMAG.COM (Nov. 5, 2009), http://www.pcmag.com/article2/0,2817,2355432,00.asp; Ben Laurie, \textit{Another Protocol Bites the Dust}, LINKS (Nov. 5, 2009, 8:03 AM), http://www.links.org/?p=780; Dan Goodin, \textit{Hacker Pokes New Hole in Secure Sockets Layer}, REGISTER (London) (Feb. 19, 2009, 5:38 GMT), http://www.theregister.co.uk/2009/02/19/ssl_busting_demo.


\textsuperscript{60} Sunstein, supra note 14, at 1012–16.


\textsuperscript{63} Press Release, Electronic Frontier Foundation, “EFF DES Cracker” Machine Brings Honesty to Crypto Debate (July 17, 1998), http://w2.eff.org/Privacy/Crypto/Crypto_
specifying rule that remains effective for over twenty years is relatively low-cost to update and relatively impervious to decay.64

Finally, rules are particularly effective when monitoring is low-cost and accurate. An ongoing problem with data security breaches is the causation of downstream harm. For example, if a bank suffers a data spill, and its customers later suffer identity theft, is there a causal connection to the spill? Courts have largely interpreted the causation requirements built into tort law to exempt data owners or storehouses from liability.65 This may result in insufficient incentives to take precautions. A rule, for example, that requires data holders to encrypt data usefully serves as a bright-line negligence test—especially when compliance is relatively low-cost. Holding institutions responsible for downstream consequences of harms related to the spilled information provides strong incentives to comply with the rule—including that liability can be avoided entirely (under the current doctrine) simply through encryption.66 Concerns about over-deterrence, or excessive investment in precautions, are minimized (if not eliminated) where the entity can avoid liability relatively simply and cheaply, and where errors in adjudication are unlikely. When a rule is effective, both initially and over time, and where regulators can assess compliance cheaply and with confidence, a rule is likely to be superior to a standard in specifying technological measures for data security. Thus, data security rules can helpfully act as a forcing device that reduces the level of harm from breaches.

One example of a data security rule that appears beneficial (though it is sufficiently new that empirical data is lacking) is the data breach notification scheme added to HIPAA (the federal Health Insurance Portability and Accountability Act of 1996, which set data privacy and security rules for personally-identifiable health information) by the HITECH Act of 2009.67 The HITECH Act regulates information security indirectly: if a covered entity under HIPAA has a breach of “unsecured protected health information,” that entity must inform people whose data was released and, in the case of a breach affecting more than 500 people,
must also inform the news media. Unsecured protected health information (PHI) is PHI that is neither encrypted or destroyed. Thus, a breach of encrypted data does not impose a notification requirement, while a breach of unencrypted PHI does. The HITECH Act is specific about the encryption technologies that meet its mandate, pointing covered entities to a list of methods certified by the National Institute of Standards and Technology (NIST). Examples of NIST-approved encryption methods include the use of Transport Layer Security (TLS), SSL, or IPSec for data communications, and the NTFS file system for data storage. The new HIPAA data security mandate acts like a rule: there is a bright-line test for compliance—either PHI is encrypted with an approved method, or it is treated as unsecured—and the consequences of non-compliance are clear—the entity assumes responsibility for notification in case of a data breach. While the mandate is a soft one—covered entities need not comply if they are willing to notify if a breach occurs—it is nonetheless structured as a rule. The HITECH requirement meets all three conditions specified above. First, encryption is sufficient to mitigate or prevent most harms; second, the NIST-specified standards are relatively slow to decay; and third, compliance is easy to measure—either data is encrypted or it is not.

Even if a rule risks being under-protective, such as where it decays relatively quickly in efficacy (potentially violating the second condition outlined above), it may still be valuable, especially if paired or reinforced by a standard. This is likely to be true where technological changes are not rapid enough to call for a standard, but are faster than, for example, the changes in encryption effectiveness described above. For example, security regulation could employ a rule specifying encryption with a 256-bit symmetric key algorithm, and a standard requiring stronger encryption where industry best practices so indicate. Such a move incorporates both strict liability—failure to utilize 256-bit or greater encryption creates per se liability—and negligence-based analysis—failure to use stronger encryption when one’s industry does so can create liability. This hybrid approach increases compliance costs, as potentially liable entities must engage in additional investigation to determine the standard of care, and also

69. Id. at 42,768.
71. Guidance to Render Unsecured Protected Health, supra note 70.
72. The difference between cleartext and ciphertext is obvious even to a layperson—one is readable text and one appears to be gibberish—although the level of encryption used to encode the ciphertext is not.
monitoring costs, as enforcers must perform the same task.\(^{73}\) However, it can usefully augment a bright-line rule where there are significant concerns that the rule may become under-protective.

This framework suggests, by way of example, three areas where rule-based regulation will be helpful: data storage, data transport, and access logging.

Both data storage and data transport can be governed by a simple rule: encrypt. Data encryption technology is ubiquitous, inexpensive, and reliable, yet the wave of data spills suggests that data owners and distributors have insufficient incentives to employ it.\(^{74}\) A rule requiring entities to encrypt data during storage and transport, on pain of facing liability for all harms resulting from breaches or spills, would usefully create incentives for protection and would also drive ineffective or incompetent data handlers from the market. Typical concerns about over-deterrence do not apply where compliance is relatively low-cost and where errors in evaluating it are rare if not absent entirely. Encryption for storage and transport meets the three preconditions this Article posits for rules. First, encrypting data when it is stored or sent should protect against misuse in most circumstances.\(^{75}\) While sophisticated adversaries can decrypt protected information, doing so requires time, technology, and resources. Encryption raises the cost of data misuse, even if it does not affect the likelihood of data spills. Second, a rule requiring encryption is relatively obsolescence-proof. While faster GPUs and CPUs are decreasing the time necessary to decrypt data without authorization, current protocols are likely to be sufficient for at least ten years.\(^{76}\) Finally, detection is cheap and easy. Encryption can be verified through visual inspection. Moreover, given that encryption is strong protection against data misuse, courts might even adopt a presumption that misused data was, in fact, not protected. Res ipsa loquitur is a traditional cost-saving enforcement mechanism that could also helpfully force regulated entities to verify encryption or to enable it by default.\(^{77}\)

Access logging—tracking who has accessed, changed, or deleted data—is also a strong candidate for rule-based regulation.\(^{78}\) Moreover,

---

\(^{73}\) See generally W. KIP VISCUSI, REFORMING PRODUCTS LIABILITY 121–23 (1991) (discussing enforcement and information costs).


\(^{75}\) See supra Part II.

\(^{76}\) Mitchell, supra note 61.

\(^{77}\) See generally THOMAS J. MICELI, THE ECONOMIC APPROACH TO LAW 63–64 (2004).

access monitoring is an example of a mitigation effort rather than a prevention effort; recording who has access to data does not impede copying or misuse directly, but can deter attackers and can also make cleanup efforts easier and more effective. A rule for access logging could be quite specific, mandating that entities capture the user credentials, time of access or alteration, and location of access or alteration in durable form. In addition, the rule could allow some flexibility—become more standard-like—by prescribing what must be captured, when, and how, but not by mandating a particular mode of access control. For example, specifying that a system of electronic medical records must record what records are accessed, what changes are made, by whom (user name, for example), from where (IP address or computer host name, for example), and when, would provide a clear trail that would enable recovery efforts after a data spill. Access logging also meets this Article’s three preconditions. Knowing who—or, at least, whose credentials—accessed the data is helpful to divining downstream data access after a breach; thus, even minimal tracking is quite effective. Second, access logging has changed relatively little since the days of mainframe data storage; users still authenticate via credentials such as names and passwords. Even access controls that employ digital signatures or keys are only variants on this basic technique. Finally, verifying compliance is straightforward: either the entity keeps logs of access, or it does not. Protective techniques such as checksums and hashes can easily test for ex post alteration of access logging, preventing malefactors from obscuring evidence. Thus, not only is access logging usefully regulated by a rule, but it also serves as an example of a necessary shift in regulatory focus: from prevention to mitigation.

As these examples demonstrate, regulation by rule has considerable virtues for technology, at least where the technology has effective minima, slow decay, and easy verification.

CONCLUSION

The default assumption for regulating information technology is that standards are not only the superior choice; they are nearly the only choice. This is because scholars and policymakers have focused on the wrong

80. Id.
81. See, e.g., GARY P. SCHNEIDER, ELECTRONIC COMMERCE 493 (8th ed. 2009); Julie Webber, Software Alone Can’t Protect Your Data, PC Managers Warn, INFOWORLD, Mar. 28, 1988, at S2.
problem: they seek to prevent data spills, rather than to mitigate their impact. Rules can helpfully reduce the effects of a breach. For technology, rules are preferable when they can specify a minimum level of protection that is relatively effective against most risks or attacks; where obsolescence occurs slowly; and where monitoring the rule’s implementation is relatively low-cost and accurate.83 Standards are not always superior, nor are they always inferior—instead, the preferred embodiment of regulation varies with the characteristics of the technological problem at issue. While application design is best governed by standards, due to the critical role of process, the transport and storage of data, along with identification of access to information, are best dealt with via rules.84 This Article questions the prevailing consensus in favor of standards for regulating technology, and also seeks to create testable predictions about when rules will work better. In short, I argue sometimes geeks require rules, not standards.

83. See supra Part II.
84. See supra Part II.