DETERRING DISCOVERY-DRIVEN DATA DELETION

Daniel Long Sockwell*

Litigation costs frequently frighten companies into deleting valuable data—and this problem has gotten worse with new technology. In the past, litigation costs were dwarfed by the physical costs of storage: keeping letters in filing cabinets was so expensive that companies deleted data without even considering litigation. Now, however, technology has caused physical storage costs to fall dramatically; hard drives are vastly less expensive than filing cabinets. As these costs fell, the litigation costs became more important. Litigation costs now cause a large fraction of data deletion. This Note identifies the rise in discovery-driven data deletion and explains how this data deletion benefits no one and harms society. To solve this problem, this Note proposes reforms to the Federal Rules of Civil Procedure that would prevent discovery costs from incentivizing so much data deletion. This Note closes with practical thoughts on how judges and attorneys operating under the current Rules can minimize costs.

* J.D. Candidate 2015, Columbia Law School; B.A. 2011, Pomona College. The author would like to thank Daniel Richman, Shreya Fadia, Debbie Long, W.D. Sockwell, Bert Huang, Maura Grossman, and the staff of the *Columbia Business Law Review*.

I. INTRODUCTION

Imagine an executive struggling to decide if her company should delete old emails.[1] The executive considers the bene-

[1] For example, many companies' document retention policies require them to delete emails after five years or a similar period. *See infra* notes

fits of preserving the emails[2]—perhaps they could be used to identify the best salespeople[3] or to uncover waste—against the costs of preservation.[4] She then weighs those benefits against the costs of storage, including the long-term costs of hardware,[5] data migration costs,[6] and the cost of searching through larger email archives.[7] After balancing the benefits against this first set of costs, the executive decides that the emails are worth preserving.[8] Before the executive finalizes her decision, however, she remembers that she neglected another set of costs: any preserved documents would be available through discovery to anyone who sues the company, potentially resulting in larger judgments or more expensive litigation.[9] After considering these "discovery costs," the ex-

78–82 and accompanying text (listing corporate and institutional document retention policies).

2 *See infra* Part II.A (describing the benefits of preserving data).

3 *See* Nora Gardner et al., *Question for Your HR Chief: Are We Using Our 'People Data' to Create Value?*, MCKINSEY Q., Mar. 2011, at 1, 2 ("The Bon-Ton chain of more than 280 department stores in the United States, for example, leveraged its data to identify attributes that made cosmetics sales reps successful.").

4 *See infra* Part II.B (describing the costs of data preservation).

5 *See infra* Part II.B.1 (detailing data storage costs, including those from hardware).

6 *See infra* Part II.B.2 (describing information management costs, including data migration costs).

7 *See id.*

8 Companies would generally employ such a cost-benefit analysis when deciding whether to preserve a document. *See infra* Part II (discussing the cost-benefit analysis for determining whether to retain a document).

9 Current discovery rules grant plaintiffs broad access to defendants' data. *See* FED. R. CIV. P. 26(b) ("Unless otherwise limited by court order, the scope of discovery is as follows: Parties may obtain discovery regarding any nonprivileged matter that is relevant to any party's claim or defense . . . ."). The choice about what data to delete is not constrained by current rules: defendants are free to delete nearly any data when not anticipating litigation. *See* FED. R. CIV. P. 37(e) ("Absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information lost as a result of the routine, good-faith operation of an electronic information system."); *infra* notes 60–63 and accompanying text (discussing the case law that applies Rule 37(e)

ecutive decides to delete the emails.

When the executive considered the first set of costs, her company's incentives aligned perfectly with society's incentives. Retaining emails whose value exceeds the cost of storage is good for the company and—since it helps the company without hurting anyone else—is good for society as well. Conversely, storing emails that are worth so little that the physical costs of storage exceed the benefits is a waste for both the company and society—the resources devoted to storing those emails could be used more productively elsewhere.

However, when the executive turns to the second set of costs—discovery costs—the company's interests diverge from society's interests. When she deletes an otherwise valuable email because of discovery costs, her company saves the discovery costs, but some other party loses matching benefits.[10] If her company avoids paying a million dollar settlement, some plaintiff loses that same million dollars; from a net social perspective, the loss of a million dollars is not avoided, but just redistributed.[11] Indeed, after the email is deleted, a potential plaintiff is no better off than if the email were not accessible through discovery at all.[12] In this case, a discovery regime designed to grant plaintiffs access to that email

to allow companies to delete virtually any document when litigation is not anticipated). But defendants cannot delete any documents once litigation is reasonably anticipated.

10 As discussed below, the total social cost of preservation might be either higher or lower than the private costs due to externalities. However, positive externalities of data preservation likely exceed negative externalities, and thus the social benefits of data preservation are typically at least as great as the private benefits. *See infra* Part II (discussing the potential positive and negative externalities of data preservation).

11 Litigation represents bargaining in bilateral monopoly: because the plaintiff and defendant can only settle with each other, each has an incentive to expend socially-wasteful resources trying to gain at the other's expense in a zero-sum game. *See* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 763–67 (8th ed., 2011) [hereinafter POSNER, ECONOMIC ANALYSIS]. Here, the deleting of valuable data impacts the distribution of wealth between plaintiffs and defendants but does not create additional value; thus, that deletion harms society.

12 *See infra* Part III.C (arguing that plaintiffs do not receive any additional documents through this discovery regime).

did not provide any benefit to plaintiffs—the rule's only effect was to encourage wasteful deletion.

The discovery costs are thus analogous to a tax that, although designed to raise revenue, is set so high that it discourages companies from engaging in the taxed activity.[13] When companies avoid the taxed activity, society loses whatever value the activity would have created and no one benefits—the government does not even collect tax revenue, since the company avoided the activity. Similarly, the discovery cost "tax" on data retention benefits no one when companies delete emails: society loses the value in those emails and plaintiffs do not gain any benefits.

To evaluate the total burden of the discovery "tax," we must know how frequently companies delete data due to discovery costs. If they do so frequently, then the discovery system imposes significant social waste; conversely, if discovery-driven deletion is rare, the social waste from discovery costs is low. Before recent technological changes, discovery costs were small relative to the high cost of storing data—filing cabinets were expensive. As a result, discovery costs were rarely the deciding factor in whether documents were worth preserving, and few documents were deleted solely due to discovery costs.

This Note argues that discovery costs now represent a much higher fraction of total data retention costs and, as a direct result, companies are much more likely to delete data solely to avoid discovery costs.[14] Discovery costs are now so dominant a consideration because they have stayed constant[15] even as all the other costs of data storage have fallen dramatically due to technological change.[16] As a result, the discovery rules increasingly drive companies to delete data

13 In fact, it represents a deadweight loss. *See infra* Part III.A (elaborating on the analogy between costs imposed by the discovery system and the deadweight loss arising from taxes).

14 *See infra* Part III (arguing that litigation costs are now a significant driver of data deletion).

15 Litigation costs are a function of the legal rules, not of technology, and thus have not fallen. *See infra* Part II.B.4 (discussing litigation costs).

16 *See infra* Part II.B.1–3 (discussing falling non-litigation costs).

they could inexpensively preserve. This increased deletion of valuable data creates social waste and justifies reform of the discovery rules.[17] This Note suggests three reforms[18] that could reduce the social loss from document deletion and argues that at least one of these reforms is justified.

Part II of this Note details the current state of data preservation, first describing the significant benefits of data preservation and then examining the practical and legal costs of this preservation. Part III explains how technological changes have created a legal problem: as the practical costs of data preservation have fallen but the legal costs have remained high, legal costs became the driving factor in the decision to delete valuable data. Part IV considers three possible reforms to solve the problem of discovery-driven deletion and concludes that expanding the Rule 26(b)(2)(B) safe harbor would be the most effective reform. Part IV also addresses several serious counterarguments to this proposal. Part V concludes with thoughts for judges and lawyers practicing under the current Rules.[19]

II. COSTS AND BENEFITS OF DATA PRESERVATION

The first step in analyzing changes to data preservation is to develop a model for how companies currently make data retention decisions. In addition to the private costs and benefits to the company, preserving data may also have external social costs and benefits. Understanding current private social costs and benefits of data retention is critical to evaluat-

17 *See infra* Part III.C (arguing that increased social losses should motivate reform).

18 *See infra* Part IV.A (presenting a plaintiff-friendly reform); *infra* Part IV.B (presenting a defendant-friendly reform); *infra* Part IV.C (presenting a reform balanced between benefits to plaintiffs and defendants).

19 This Note uses the term "plaintiff" to refer to the party seeking documents thorough discovery and the term "defendant" to refer to parties seeking to prevent disclosure of documents. Thus a "plaintiff-friendly rule change" would be a rule change that allows for greater discovery of documents. This terminology is not correct in all cases—sometimes, defendants seek broad discovery while plaintiffs oppose it.