Digital Hoarders: Breaking the Habit

Technology, Privacy, and eCommerce

Defaults drive behavior. We take the path of least resistance not because we are lazy but because we are overwhelmed. In our information economy, attention is the scarcest commodity. And the volume of information is only growing. When an organization like ACC takes on information governance — the Information Governance Committee launched in July — the questions concern not just what policies are preferable, but how we can address defaults that currently drive behavior at odds with best practices. In particular, data retention by default is unsustainable.

Data has a useful life. It expires. But lifespans differ. Business value, regulatory requirements and legal holds affect how long data should be kept. Most organizations have data retention schedules outlining when to dispose of digital debris. These schedules are largely ignored. Purging expired data requires effort from individuals who see no direct benefit from the exercise. We therefore retain data by default. The resulting data accretion is accelerated by ever-growing data volumes. Data generation is doubling every two years. To stay ahead of this build up, enterprises double their storage capacity every 18 months.

The direct economic implications of these compounding data storage demands are easy to misread. Hard drive space, especially in the form of cloud storage, is getting appreciably less expensive every year. But disk space is a minor element in the cost equation. Security, labor, real estate, power, commercial-grade hardware, software, migration, disaster preparedness and host of other factors drive costs up from somewhere around $0.03 per gigabyte for raw, consumer-grade disk space to $4 per gigabyte for the annual total cost of ownership of managed enterprise storage. Extrapolations based on increasing data generation and the compounding effect of data accumulation can get silly in short order — e.g., annual data storage costs growing from $32/yr to $9,000/yr in the course of a

decade for a single employee hired today.

Organizations may choose less expensive data storage options. When you read about the massive OPM breach, you encounter sentences like this from *The New York Times*: "Much of the personnel data had been stored in the lightly protected systems of the Department of the Interior, because it had cheap, available space for digital data storage." Our companies spend so much on data retention beyond raw storage costs precisely because they want to avoid being a *New York Times* story about losing control of proprietary or personal identifying information. Better to get rid of expired data. That which does not exist cannot be hacked. Nor does it contribute to mounting costs during overzealous and overbroad e-discovery.

There is not just value buried under our mountains of superfluous data but also latent liabilities. Not only do governments tell us what we must keep, they are passing more legislation and creating more regulations to tell us what we must not keep or, if we keep it, for how long and at what level of security. From HIPAA to European privacy laws, our companies are sitting on millions of dollars in potential fines, penalties, civil lawsuit exposure and reputational damage. As their counsel, we need to be cognizant of these risks. The impulse to keep everything so you will have it if the regulator or the court ever asks is understandable but unhealthy.

The response to these mounting pressures has been the development of frameworks like defensible deletion, defensible disposition and defensible destruction. The ideas are excellent. Purging ROT (redundant, obsolete, trivial). Regulatory mapping. Rationalization of legal holds. All worthwhile. But I am no fan of the framing.

The implication that all data deletion must be defended is inconsistent with our actual obligations. Our process for retaining data subject to regulatory requirements and litigation holds must be reasonable. That's six percent of our data. From a legal perspective, we have no dominion over the other 94 percent. Yet, we are setting up a framework that seems to make lawyers the final arbiters of deletion decisions. Instead, deletion should be the default. Because it costs money, we should require reasonable justification to retain data. The onus is on the lawyers to identify the six percent of data to be retained rather than to bless the deletion of whatever portion of the remainder is past its useful business life.

Per usual, technology is essential to solve the problems created by technology. We can, and should, build our retention schedules into our data storage. Retention then requires effort as deletion becomes the default. Defaults drive behavior.

[D. Casey Flaherty](#)

Corporate Counsel

Kia Motors America