5
Data Stewardship, Security, and Policies

Discretion is the perfection of reason, and a guide to us in all the duties of life.
—Sir Walter Scott

Key Findings
- Data responsibilities tended to follow an organizational division of labor within institutions. Business and academic units most often had primary responsibility for data accuracy, access decisions, and privacy, whereas central IT units tended to have primary responsibility for data formats, metadata management, and notification of affected parties following data breaches.
- Only one-third of institutions reported formal data stewardship policies. Institutions with such policies were more likely to have formal processes for classifying data according to risk, and they had slightly higher mean enterprise data quality scores, but they did not differ from other institutions in respondent agreement about the security of data from unauthorized access.
- The data-related policy portfolios at respondent institutions were modest. Although majorities reported policies on acceptable use of data and employee responsibilities regarding data security, and half had policies regarding notification of parties following data breaches, fewer than half had policies regarding sharing, storage, or transmission of institutional data; data encryption; or cloud-related data practices.
- About 30% of institutions reported at least one data breach in the 24 months prior to the survey that required notification of affected parties. Incidents were more common at larger institutions and more common at doctoral and master’s institutions than at other Carnegie types.

One of the ruder shocks of the Internet era has been the discovery that reducing data friction—the difficulty of finding, copying, and transporting information—often creates a corresponding need to enhance and formalize data security. Not long ago, the greatest danger to institutional information security (barring the odd spy with a pocket microphone) was the departmental Xerox machine, and a few locks and keys, a logbook, and an electromechanical copy counter constituted reasonable protection from data breaches. Today, data practically begs to be copied, and a thumbnail-sized key disk can collect more information in a few seconds than a fleet of photocopiers could reproduce in a year.
Small wonder that data security has been among the top three items named as critical for strategic success in EDUCAUSE’s annual Current Issues Survey every year since 2003.\textsuperscript{1} But taking good care of institutional data means much more than just keeping it out of the wrong hands. It also means ensuring that data is accurate and timely, that it’s formatted in ways that make sense, and that when it does leak, the institution responds appropriately. This range of responsibilities, sometimes conceptualized under the name stewardship, raises many questions about who should shoulder which tasks and how they should be carried out. In this chapter, we investigate how institutions approach these challenges.

**Business/Academic Unit and Central IT Data Responsibilities**

“Data,” writes Burton Group analyst Noreen Kendle, “is a representation of a business.”\textsuperscript{2} This simple fact sums up one of the key dilemmas of data management: When should responsibilities for managing data fall on those whose competencies lie in representing and processing information—typically the IT unit—and when should they fall on those who best know the processes that generate, consume, and are managed by that information—the business and academic units? IT administrators rarely wish to second-guess business managers’ judgments about their own data needs, yet IT professionals can often improve the accuracy, usability, and security of data if they have a hand in managing it.

To find out how institutions divide up these tasks, we asked respondents whether business/academic units or central IT had primary responsibility for each of 10 data management activities. We knew ahead of time that our question was somewhat artificial, because it seemed likely—and results reported in Chapter 3 confirmed—that central IT and business/academic units often work together on data issues. Still, we thought it would be revealing to see how respondents, given the choice, would divide IT and line unit activities.

We found a fairly clear division of labor, with aspects regarding the content and use of data generally assigned to the business/academic units and those concerned with data formatting or classification to central IT. This was especially evident in activities related to data quality and data specification. As Figure 5-1 shows, the great majority of respondents told us that business and academic units had primary responsibility for data accuracy, timeliness, and fitness for purpose, whereas only one in five or fewer said central IT was primarily responsible. (A small number of respondents in each case said that neither was involved, or that the item was not applicable at their institution.) Two items involving the need for technical skills—specifying data formats and metadata management—followed an almost exactly opposite pattern: The great majority of institutions said central IT was primarily responsible (though the “neither/not applicable” responses were more numerous for these activities).

Business and academic units also predominantly held primary responsibility for data access; at about three-quarters of institutions, they were responsible for deciding both who has routine access to data and who could have access on an as-needed or exception basis (see Figure 5-2). This makes sense, since the business and academic units are the ones most likely to know both the nature of the data and the parties asking to use it. Still, given the identity management implications of access, it seems likely that close coordination with central IT is common. At Georgia State University, for example, “access is automated based on being in a role,” said J. L. Albert, associate provost and CIO. “Each of the major function areas has a data steward who applies granularity of access. They set the rules,
Institutional Data Management

ECAR Research Study 8, 2009

Metadata management (N = 306)

Specifying data formats (N = 307)

Data timeliness (N = 305)

Data fitness for purpose (N = 306)

Data accuracy (N = 307)

0%             10%          20%           30%           40%           50%          60%           70%          80%           90%         100%
Percentage of Institutions

Figure 5-1. Primary Responsibility for Data Quality and Specification Activities

Notification of affected parties following data breach (N = 306)

Classifying data for risk/sensitivity (N = 305)

Regulatory compliance regarding privacy/disclosure of information (N = 308)

Deciding who has access to data on an as-needed or exception basis (N = 307)

Deciding who has routine access to data (N = 304)

0%          10%         20%        30%        40%         50%        60%        70%         80%        90%       100%
Percentage of Institutions

Figure 5-2. Primary Responsibility for Data-Privacy-Related Activities

Business/academic units
Central IT
Neither/not applicable
and IT implements them.” A similar process applies at the University of North Texas, according to Maurice Leatherbury, acting vice president for information technology and CIO. “Authorization is primarily a business unit concern,” said Leatherbury, “though IT sometimes brokers approvals where people on campus don’t understand the distributed nature of data stewardship.” Another indication that access decisions can be a multiparty process lies in the fact that 46.4% of institutions told us they have a governance structure for setting data access policy.

Classification of data for risk or sensitivity was most often reported as a central IT responsibility. However, with 38.4% of institutions saying business/academic units were primarily responsible and a relatively high “neither/not applicable” response, this was the most evenly divided of the activities we asked about.

Somewhat surprisingly, about two-thirds of institutions said central IT was primarily responsible for notifying affected parties following a data breach, whereas only about a quarter selected business/academic units. This was something of a contrast with our results’ overall picture of business/academic units as the main data owners. However, the virtues of a centralized notification process are easy to see: It can help ensure that regulatory protocols are followed to the letter, and it can enhance the enterprise visibility of data breaches that might otherwise be known only to local units. Also, where breaches result from technical vulnerabilities, central IT is in the best position to carry out security forensics and apply solutions.

Although a pattern of divided labor was fairly evident in these results, few other patterns were. We didn’t find significant differences in the way institutions assigned responsibilities on the basis of Carnegie class, FTE enrollment size, or public/private control, nor any relationships with enterprise data quality scores, the existence of data stewardship policies, or better or worse performance in data management outcomes.

**Data Stewardship Policies**

If, as we noted above, data represents the business of an organization, it is very much a traveling representative. Optimizing business processes and coordinating them at multiple levels of business often requires that data move through a lot of hands. With the advent of real-time integrated administrative suites and sprawling institutional websites, movement of both structured and unstructured data has expanded astronomically. But these enormous increases in the availability of data have not been met by any corresponding decrease in the need for security, privacy, and accuracy—quite the contrary.

To combat this danger, modern data management has introduced a concept of data stewardship—the formal definition and assignment of responsibility for the management of data resources. Acknowledging that data may flow through few or many channels, data stewardship explicitly describes who is primarily responsible for seeing to it that data represent what it’s supposed to represent, what rules and regulations apply to it, and who may see or manipulate it.

For data management expert Robert Seiner, formality rather than assignment of responsibilities is the real key to successful data stewardship. Seiner notes that de facto stewardship often arises out of the logic of the business and its organizational structure. “Data stewards are already there,” Seiner writes. “They may not know they are stewards and the organization may not recognize them as such, but there are already people in your organization that have accountability for the management of data. The challenge is getting them to participate and operate more efficiently and effectively in these data management roles. The data will not govern itself.”

"3
Our survey examined data stewardship by asking whether institutions had documented policies for it, and where they did, by looking at what kinds of data were covered and to whom data steward responsibilities were assigned. In addition, in our qualitative interviews we talked to interviewees about their experience with data stewardship and their approaches to it. We examine the results in this section.

**Data Stewardship Policies**

Institutions with documented policies defining data steward responsibilities were the exception rather than the rule among our respondents. Overall, only about one in three reported them. As Figure 5-3 shows, data stewardship policies were most common at institutions with FTE enrollments greater than 15,000, though even among these they constituted only a modest 56.3% majority. At the smallest institutions, those with enrollments of 2,000 or fewer, 15.7% reported data stewardship policies, and roughly one-third of institutions in the size categories in between reported them.

Does having a documented data stewardship policy make a difference in desirable outcomes? We found some limited evidence that it does. Institutions reporting such policies averaged slightly higher enterprise data quality scores than those without them, and they were almost three times as likely to say that they had a formal process for classifying institutional data by risk or sensitivity level (an item we discuss in more detail below). Furthermore, data stewardship policies do seem to help match data with responsible parties: Although 70.7% of respondents at institutions with data stewardship policies agreed or strongly agreed that all institutional data had a designated data steward, fewer than half as many (31.2%) did so where there was no policy. These findings all lend some weight to the idea that good data management practices are mutually reinforcing, an idea that Neil McElroy, dean of Libraries and Information Technology Services at Lafayette College, expressed in one of our interviews. “I suppose we might have understood this better 10 years ago,”

![Figure 5-3. Institution Has Documented Institutional Policy Defining Data Steward Responsibilities](image-url)
he said of changes he witnessed after a “reimplementation” of the institutional ERP system led to systemic improvements to the institution’s data environment. “But it’s all coming together now. Data stewardship depends on shared definitions and data integrity, and it all contributes to more productive work for the college.”

At the same time, we did not find a significant association between having a data stewardship policy and any of the key data management outcomes we asked about, such as getting maximum business or academic value from data, securing sensitive data from unauthorized access, or employee understanding of her or his responsibilities in the use of data. Formalizing data stewardship policies may have practical benefits such as documenting an institutional commitment to responsible data use or permitting better communication about data rules, especially in large institutions. Furthermore, as our finding on data quality scores suggests, data stewardship policies may foster other improvements that in turn ultimately contribute to better data management outcomes. But as far as our evidence permits, we cannot say that documented data stewardship policies are related to either better or worse major outcomes.

**Data Stewardship Scope and Assignments**

Most of the advisory literature on data stewardship focuses on structured business data, though in principle the idea might apply to many other kinds as well. When we asked about the kinds of data covered in documented data stewardship policies, we found that although administrative systems data seemed to be the chief concern—every single institution with a data stewardship policy said that this kind of data was covered by it—coverage was often broadly inclusive (see Figure 5-4).

Although administrative systems data ruled the roost, learning management systems data and content, e-mail, and institutional web content were all reported as covered by more than 6 in 10 institutions. Research data was covered less frequently, perhaps due to its distributed nature or because protocols separate from those in the institutional data stewardship policy govern it; even so, a slight majority (52.8%) covered it. Surprisingly, we found no association between research data coverage and Carnegie class or research/teaching mission emphasis. E-mail and learning management systems data and content were most likely to be covered at doctoral institutions and least likely at BA institutions, but on
the whole we found few differences based on institutional demographics.

As for who the data stewards are, we found assignments concentrated among administrative officers and staff and an overall pattern suggesting the mutual involvement of business/academic units and central IT (see Figure 5-5). Almost 9 out of 10 institutions reported business/academic unit managers as being explicitly assigned data steward responsibilities in institutional data stewardship policies. Central IT leadership was the next most often reported (78.1%), and business/academic executives and staff and other central IT participants all were assigned responsibilities by between 68.4% and 72.0% of institutions. There was a clear break between these parties and other institutional officers such as researchers and teaching staff, auditors, and institutional counsel; none of these parties were assigned responsibilities by even half of institutions, and although higher “don’t know” responses for these groups may indicate that our results somewhat understate the actual figures, they also suggest that these groups are less visibly associated with the data stewardship process.

Gartner analyst Ted Friedman (as well as other data management advisers) recommends that “data stewards should reside in the business, not the IT organization.” Do our assignment results, in which central IT officers figure so prominently, suggest that our respondent institutions are violating this best

Figure 5-5. Parties Assigned Data Steward Responsibilities (Institutions with Data Stewardship Policies)
practice? We did not ask fine-grained questions about the nature of the assigned responsibilities, but on the whole we are inclined to interpret the results as suggesting cooperation and division of duties rather than undue domination by IT over data stewardship.

The data responsibilities results reported in Figures 5-1 and 5-2 suggest a reasonable division of labor in which business and academic units are more often responsible for data content and use, whereas IT is more often responsible for specification and classification. Data stewardship policies have room for both, since they commonly define a variety of data responsibility profiles and a data sensitivity or risk classification scheme that might involve both business unit knowledge of regulatory and use issues and IT knowledge of data formatting and system integration issues. In addition, IT may play a broader role educating users about data security and privacy issues beyond those applying to a particular functional set of data elements.

Our qualitative interviews suggest that institutions implementing data steward policies “get it” when it comes to giving business/academic unit officers the primary role in owning and stewarding data while also giving IT an important supporting role. At Bates College, said Gene Wiemers, vice president for Information and Library Services, “data stewardship is decentralized and the responsibility of the individual departments. But in central IT we see ourselves as educators for creating awareness. We try to teach everyone from data entry clerks to vice presidents about their data-related responsibilities.” Andrew Lawlor, associate vice president for Technology and Communications at Edinboro University of Pennsylvania, recounts IT’s part in getting data stewardship moving and also its emphasis on business unit ownership. “Our information security policy for administrative information covers data stewardship and defines data custodian roles and other requirements. IT developed it based on practices already in place, and the business managers reviewed it. The data custodians call the shots—for example, if somebody wants access, we require them to get approval from the data custodian, then we enable the appropriate process.” And at the University of San Francisco, CIO Steve Gallagher reported a similar sensibility. “We have designated data stewards who have been identified by the appropriate division lead in each of the business units, and they control access to the various administrative functions and the associated data,” Gallagher said. “IT is the facilitator, not the decision maker, and we’re pleased with that arrangement.”

More worrisome than the promiscuous mixing of business/academic and IT officers in our assignment results is the relatively lower level of involvement by faculty, in their roles as researchers and teachers, and by other officers whose understanding of institutional risk could be vital to successful data stewardship. Not only are academic officers now involved in the digital management of large amounts of regulated data, but also their options for collecting, storing, and sharing those data are growing as consumer-oriented cloud tools increasingly compete with institutionally delivered tools. It may be that the data responsibilities of these positions are being addressed in other ways, of course, but incorporation into a broad-based notion of data stewardship could help send the message that everyone who handles institutional data—not just administrators—has a role in minimizing institutional exposure.

Data Stewards’ Councils

Besides their responsibilities in the operational oversight of data, data stewards are sometimes conceived of as subject matter or policy advisers with a role in data governance (or IT governance generally). A council of data stewards is one mechanism for encouraging communication between data stewards and uncovering common practices or problems
that cross departmental boundaries. It also provides a convenient way to draw on the data stewards’ advice and can be used for addressing data policy issues.

The institutional data stewards’ council is, however, a relatively rare body among our respondent institutions. Only 30 of 100 institutions with documented data stewardship policies said that they have such a council. Among those that did, however, most seemed to give it an influential role. Although the most common of the five roles we asked about was to provide advice, 22 institutions also said their council set policy and 21 said it set priorities, and councils at 18 institutions adjudicated conflicts (see Figure 5-6). The power of the purse was largely absent, however: only one institution said its council could authorize funding.

### Data Use, Security, and Access Policies

Although only about a third of our respondent institutions reported data stewardship policies, we found more widespread incidence of other policies relating to data use, security, and access. As Figure 5-7 shows, nearly three-quarters of respondent institutions (73.4%) said they have a documented policy on the acceptable use of institutional data, and a solid majority (59.8%) have one defining individual employee responsibilities for data security practices. It is possible that some institutions see these as alternatives to a more formal data stewardship policy, though institutions reporting a data stewardship policy were also more likely to have these policies.

The policy groundwork on data issues related to storage and use outside the institution, as in outsourcing, was modest. Slightly fewer than half of respondents reported documented policies for sharing, storing, and transmitting institutional data (for example, with or to ISPs, external networks, or contractors’ systems), despite the common practice of trading data with outside entities. What’s more, given how easy it’s becoming for users to directly use consumer-oriented cloud services (often for free) without any IT assistance, it’s noteworthy that well under one in five institutions (17.6%) reported a documented policy for storage of institutional data on extra-institutional web-delivered (cloud) systems, and fewer still (10.9%) had one for use of extra-institutional web-delivered services to do institutional business. Perhaps some institutions feel that these issues are implicitly covered by their acceptable-use or sharing/storing/transmitting policies. We speculate, however, that most institutions simply have not yet caught up with the marketplace for cloud services.

Given the recent proliferation of laws relating to compromised personal information, we were surprised to find that only half

![Figure 5-6. Role of the Data Stewards’ Council (N = 30)](image-url)
Institutional Data Management

ECAR Research Study 8, 2009

of institutions reported a documented policy for notification of affected parties following data breaches. In addition, though encryption of institutional data on user devices is one way of mitigating the exposure caused by data breaches and may be grounds for exemption from notification requirements, it has not found a lot of traction among our respondents: Only about a quarter said they have documented policies for it.

The most commonly reported policies were sufficiently widespread that we found few distinguishing demographic patterns in their adoption. However, policies for data breach notification; sharing, storing, and transmitting data; and encryption were all more common among larger institutions (see Figure 5–8). The difference was especially notable for data breach notification policies, which were twice as prevalent (74.5%) among institutions with enrollment greater than 15,000 than among those with enrollment of 4,000 or fewer (36.9%).

One can infer from our results that many institutions feel the lack of documented policies relating to data use, security, and access and aspire to have them. In many cases the percentage of institutions reporting a policy under development would, if the policy were completed, dramatically increase the proportion of those with policies. The most striking examples were encryption of institutional data on user devices and the two cloud-related items, for which the percentages reporting a policy under development equaled or exceeded those of institutions that have a policy already.

Although these initiatives may well represent aspiration to a more formal data policy regime, policy development can be informal, low priority, unfunded, and fraught with political obstacles—and can fail. Simply adding the “under development” results to the “have policy” results in Figures 5–7 and 5–8 would probably overstate the likely future incidence of these policies. The percentages
of institutions reporting policies for employee information security responsibilities and for sharing, storing, and transmitting institutional data were actually lower in these results than those saying they had “implemented” similar policies in ECAR’s 2006 IT security study.\(^5\) Differences in the way the questions were posed in the two surveys prevent comparing them conclusively, but our findings certainly do not suggest a substantial increase in data policy formulation.

### Classifying Data by Risk or Sensitivity

Regulatory requirements, access authorization rules, and better organization of data for analytics use are only some of the reasons why institutions might wish to categorize data, and in particular to systematically distinguish data whose inappropriate use creates institutional exposure. Although we recognize that many institutions might handle these risks informally or in a distributed fashion, we wanted to know how many have created a formal process for classifying institutional data by risk or sensitivity level.

Not many, as it turned out: About one in four institutions told us they have such a process. As Figure 5-9 shows, data classification processes were most common among larger institutions, though even among those with more than 15,000 FTE enrollments, fewer than half (45.8%) reported one. Among Carnegie classes, doctoral institutions were the most likely to report a classification process (49.2%) and baccalaureate institutions the least likely (10.8%), with master’s and associate’s institutions in between.

Data classification is often a component of data stewardship programs, so we were not surprised to find that institutions reporting a documented policy defining data steward responsibilities were almost three times as likely to report a formal data classification process.

---

**Figure 5-8. Data-Related Policies, by FTE Enrollment**
process (45.0%) than were those without a data stewardship policy (16.3%).

Clinton Smullen, director of Academic and Research Computing Services at the University of Tennessee at Chattanooga, described how the data stewardship and classification systems work at UTC. “The data stewardship program began in tandem with the data classification program,” he said, “and when departments classified their data, they nominated a steward too.” The classification system is based on a state framework that attempts to identify key areas of exposure without getting bogged down in unproductive details. “The state of Tennessee devised a three-tier classification scheme,” Smullen explained. “There’s public data and proprietary data, which we don’t have to report about should it get exposed, and then there’s sensitive data—information about individuals, which requires reporting. The classifications are gross, but if you get any finer, you get into real trouble trying to classify or control it.” Departments take a first cut at classifications, then negotiate them with IT.

As these remarks suggest, data classification is a big and possibly never-ending process, and very few of our respondents (4.0%) were so bold as to tell us that they had classified all their data (see Figure 5-10). On the other hand, no institutions at all said they had classified none of their data, and more than three-quarters (77.3%) said some or most was classified.

Do classification processes for restricted/sensitive data, and for that matter progress in classifying data, make a difference? We found some evidence that they do. Institutions with data classification processes averaged stronger agreement than those without when assessing the statement that restricted or sensitive data at their institution was identified and appropriately classified (see Table 5-1). Looking at the data another way, 80.8% of institutions with data classification processes agreed or strongly agreed with the statement, whereas only 58.4% of those without such processes agreed or strongly agreed. Among institutions with classification processes, those reporting
more progress in classifying data also agreed more strongly with the same statement than did those reporting less progress.

Although these results lend some weight to the hypothesis that formal classification processes make a difference in the important matter of appropriately classifying sensitive data, they also give some hints about why relatively few institutions choose to create them. Nearly 6 in 10 institutions without such processes agreed or strongly agreed that they are accomplishing the task of classifying data appropriately, and most of the rest were neutral rather than disagreeing, for a mean response of 3.47 out of 5.00. What’s more, we found no significant agreement difference between those with and without formal classification processes regarding the key practical end of data classification—“restricted/sensitive data is secure

### Table 5-1. Restricted/Sensitive Data Is Identified and Classified, by Data Classification Characteristics

<table>
<thead>
<tr>
<th>Data Classification Characteristic</th>
<th>Restricted/sensitive data is identified and appropriately classified.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean*</td>
</tr>
<tr>
<td>Does institution have a formal process for classifying institutional data by risk or sensitivity level?</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>3.47</td>
</tr>
<tr>
<td>Yes</td>
<td>3.96</td>
</tr>
<tr>
<td>How much data classified?</td>
<td></td>
</tr>
<tr>
<td>A little/some</td>
<td>3.63</td>
</tr>
<tr>
<td>Most/all</td>
<td>4.27</td>
</tr>
</tbody>
</table>

*Scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

![Figure 5-10. How Much Data Has Been Classified (Institutions with Formal Institutional Data Classification Process, N = 75)](image-url)
from unauthorized access”—or in their agreement that the institution gets maximum business value from data. Even if they think they could do a better job by formalizing the classification process, many institutions may feel that existing informal or decentralized approaches are good enough and that they can get a better return by investing scarce resources elsewhere.

Data Breach Notifications

In recent years, probably nothing has caused more lost CIO sleep than the possibility of an embarrassing data breach exposing sensitive information. Since the passage of California’s landmark Senate Bill 1386 in 2003, more than 40 U.S. states have adopted laws requiring formal notification of individuals when private information about them has been compromised. Similar legislation has become commonplace internationally as well. Despite the highly distributed nature of institutional data and the near impossibility of establishing central IT control over it all, institutions tend to look to the IT organization when breaches occur. As we reported in Figure 5-2, 67.0% of institutions told us that central IT was primarily responsible for notification of affected parties following a data breach.

If actual occurrences of this particular source of CIO insomnia aren’t exactly widespread, they aren’t so rare as to be negligible, either. Asked how many data breaches requiring notification of affected parties had occurred at their institutions in the past 24 months, 30.3% reported one or more. Of these, about half (14.8% of all institutions) reported a single incident, and the remaining institutions reported between 2 and 10, though only a handful (1.8%) said they had had more than 4 (see Figure 5-11).

Data breach notification incidents were reported by institutions in all major Carnegie classes and FTE enrollment size ranges, but they were more common at master’s and doctoral institutions than at associate’s and baccalaureate institutions, and they were more often reported by larger than smaller institutions. As Figure 5-12 shows, exactly half of responding doctoral institutions reported...
Institutional Data Management

one or more data breach notification incidents in the past 24 months, close to five times the rate among baccalaureates. Fewer than one in five institutions with FTE enrollments of 4,000 or below reported one or more incidents, whereas more than half of those with enrollments above 15,000 did so.

Aside from these demographic associations—which may simply reflect the fact that having more data, devices, and people creates more chances for exposure—we found few variables related significantly to data breach notification incidents. One intriguing exception was the existence of a documented policy for notifying affected parties following a data breach; those with such a policy were far more likely to report one or more incidents (46.6%) than those without a policy (8.8%) or those with a policy under development (20.0%). This doesn’t necessarily mean that institutions with policies experienced more data breaches per se, but it may be that institutions with documented policies feel more obliged to make notifications, suggesting that policies may take a more aggressive view of the need to notify than the law strictly requires. Or the existence of a documented policy may simply increase awareness of the need to notify.

Institutions reporting data stewardship policies were not significantly more or less likely than those without them to have experienced incidents, nor did we find that having an institutional data policy office or an IT governance committee dedicated to data made a difference. Institutions that had experienced incidents did not do better or worse in enterprise data quality score or in major data management outcomes, either. Although suggesting that such incidents are merely random would be going much too far, it does appear that they strike all kinds of institutions pursuing many kinds of data management regimes.

Data Stewardship, Formal and Cultural

We confess to being surprised that formal data stewardship policies didn’t draw a
bright line down the middle of the findings presented here. We did, of course, find some evidence that these policies, so prominent in data management best practices advice, make a difference in desirable outcomes. Institutions with documented data stewardship policies do a little better than others on enterprise data quality, and they are more likely to have a process for classifying data by sensitivity, and slightly more likely to agree that data is, in fact, appropriately classified. But data stewardship policies did not prove to be a magic bullet: The one-third of institutions reporting them didn’t do better (or worse) on major measures such as getting business value from data, agreeing that restricted data is secure, or being less likely to have experienced a data breach notification event.

Looking at our data as a whole, however, we noted other results suggesting that some of the spirit of data stewardship may be found in broader practices related to data responsibilities. Concerned that institutions might regard central IT as the default data owner, we found in fact that the great majority assign business and academic units the primary responsibility for such critical matters as data accuracy, access decisions, and regulatory compliance, leaving central IT mainly responsible for appropriate format and metadata management duties. Though we found a rather modest portfolio of data-related policies in place, 60% or more of institutions had policies defining employee data security responsibilities and acceptable use of data (and others reported such policies in development). Many institutions may simply regard data stewardship as part of professional practice and culture and see no extra benefit in formalizing it, regardless of the data management advisory literature.

It’s striking, in fact, that after a decade or more of intense concern over data security and privacy issues, our study found institutions more strongly in agreement that data is secure and appropriately classified for risk than they are about other key measures, such as getting business value from data and effectively managing all the varieties of data and digital content that the institution needs. (We examine such data management outcomes in more detail in Chapter 8.) Although security and privacy concerns remain essential to responsible IT administration, they have also been lavished with resources and attention in recent years. Declaring victory in this arena would be premature (and always will be), but it might be reasonable to consider that progress has been made, that awareness is broad, and that what was once a set of issues primarily understood only by IT professionals has become part of a larger set of professional practices.

Despite some lingering concerns, then, that relatively few institutions have chosen to follow the commonly recommended practice of establishing formal data stewardship policies, we note that this doesn’t necessarily distinguish successful from unsuccessful data management. Although formal data stewardship may be one way for institutions to inculcate security awareness and a culture of responsibility toward data, other initiatives—in data quality, analytics, training, and content management—deserve consideration as well when addressing data-related needs.

Endnotes

7. A single outlier institution reporting more than 30 data breach notification incidents was not included in the results reported here.