CHAPTER 14

Information Technology Operations

This chapter includes contributions by Meenu Agarwal (IBM).

According to the IBM 2009 Global CIO Study, today's CIOs spend an impressive 55 percent of their time on activities that spur innovation. These activities include generating buy-in for innovative plans, implementing new technologies, and managing nontechnological business issues. CIOs spend the remaining 45 percent of their time on essential, more traditional tasks related to managing the ongoing technology environment. This includes reducing IT costs, mitigating enterprise risks, and leveraging automation to lower costs elsewhere in the business.

IT is a key stakeholder in the entire information governance program because it is normally the custodian of the data repositories. In addition, IT is the business sponsor for information governance in situations that are generally associated with IT cost savings. Across the entire sample in the IBM 2009 Global CIO Study, CIOs spend about 14 percent of their time removing costs from the technology environment. One of the top management priorities cited by a banking CIO was to "position the IT organization to handle increased activity with minimal additional cost." Simply put, CIOs aspire to do more with less. A retail CIO in the United States described the dual challenge, "The balance between new projects and cost control is the dichotomy of my life."

Here are the best practices to sell information governance internally, within IT operations:

1. Engage with the vice president of applications to retire legacy applications while retaining access to the underlying data.

2. Work with the vice president of applications to reduce storage costs through an archiving strategy.
3. Ensure trusted data when consolidating applications, data warehouses, data marts, and operational data stores.
4. Support the vice president of testing by automating the creation of test data sets.
5. Work with the enterprise architecture team to enforce consistent information-architecture standards.

These best practices are discussed in detail in this chapter.

1. Engage with the VP of Applications to Retire Legacy Applications but Retain Access to Underlying Data

In a 2010 study, IT industry analyst IDC predicted that information volumes would increase by a factor of 44 over the next decade. Most large IT shops have legacy applications that are no longer in use but consume a significant portion of the IT budget. For a variety of legal and regulatory reasons, the business might need access to the underlying data, but not the application itself. For example, there might be "legal holds," or regulations that require the retention of certain data sets.

IT can establish significant cost savings by retiring the legacy applications while retaining access to the underlying data. The key technical challenge is to be able to provide the data on demand, which requires the ability to maintain the referential integrity of the data.

2. Work with the VP of Applications to Reduce Storage Costs Through an Archiving Strategy[1]

Gaining a complete understanding of which areas are accumulating the most information allows an organization to apply the most effective information lifecycle governance strategy. Data duplication has significantly contributed to growth statistics. Organizations frequently clone or copy production databases to support other functions, or for application development and testing. They also maintain several backup copies of critical data, or implement mirrored databases to protect against data loss. Finally, disaster recovery plans require data duplication, to store critical data in an alternate location. All of this duplication has created what is known as the "data multiplier effect."

As data is duplicated, storage and maintenance costs increase proportionally. Figure 14.1 provides an example of a production database that contains one terabyte of data. When that database is copied for backup, disaster recovery, development, testing, and user acceptance, the total data burden increases to six terabytes.

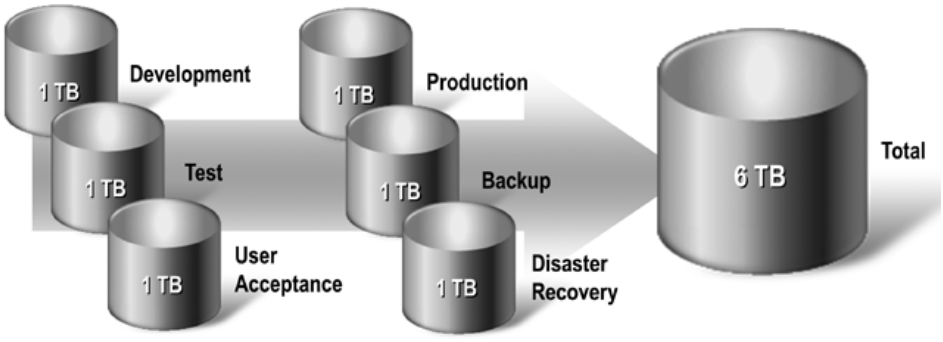

Figure 14.1: The actual data burden equals the size of the production database and all replicated clones.

You cannot govern data if you do not understand it, so it is critical that you first document your existing data landscape, using data discovery. Data discovery analyzes data values and patterns to identify the relationships that link disparate data elements into logical units of information, or *business objects*, such as customer, patient, and invoice. These business objects provide essential input for archiving. Without an automated process to identify data relationships and define business objects, organizations can spend months performing manual analysis, with no assurance of completeness or accuracy.

From an information governance perspective, an organization needs to archive both structured and unstructured content to reduce storage costs, improve system performance, and ensure compliance with regulatory requirements. In particular, unstructured content in the form of emails and other documents makes up more than 80 percent of the content in a typical enterprise. This content needs to be archived to reduce storage costs.

Storing archived data according to its business value is a logical component of an integrated data management strategy. A three-tier classification strategy is a useful way to approach the problem. Current transactions are maintained in high-speed, primary storage. Reporting data is relocated to mid-tier storage. Reference data is retained on a secure Write Once, Read Many (WORM) device, keeping it available in case an audit request should arise. This approach to a tiered storage and archiving strategy is a good way to reduce costs and maximize business value, as discussed in Case Study 14.1.

Case Study 14.1: Managing data growth at a large telecommunications operator

A large telecommunications operator was struggling with the storage costs associated with large amounts of data, including Call Detail Records (CDRs). The telecommunications operator had nearly 16 terabytes of data in one production environment, and six database clones for disaster recovery, backup, and pre-production environments. The operator needed to fund additional mainframe computing capacity every year to upgrade performance.

The information governance team established a business case that articulated significant cost savings by archiving data to lower-cost storage environments. Important questions to the business included "How many months of data do we need to maintain in production systems?" Because of the initiative, the operator was able to reduce storage costs dramatically. The business benefits are highlighted in Table 14.1.

Table 14.1: The Sanitized Business Benefits from an Archiving Solution at a Large Telecommunications Operator

A.	Size of the production database in gigabytes	16,000
B.	Number of database copies for disaster recovery, backup, and performance testing	6
C.	Annual cost of storage per gigabyte	$50
D.	Current cost of storage ((B + 1) × A × C)	$5,600,000
E.	Percentage of data that can be archived	70%
F.	Post-archiving storage cost savings (D × E)	$3,920,000
G.	Three-year storage cost savings (F × 3)	$11,760,000

3. Ensure Trusted Data when Consolidating Applications, Data Warehouses, Data Marts, and Data Stores

Most large IT departments have multiple, overlapping applications, data warehouses, data marts, and operational data stores that are enormously expensive to maintain. One large bank saved hundreds of millions of dollars by consolidating eight customer information files (CIFs) into one. Another bank had 10 CIFs, each with five or more copies, several of which were built in COBOL. It cost the bank tens of millions of dollars to maintain this unwieldy system. The information governance program was driven by a strong desire to cut costs, especially when a seemingly straightforward project to implement "powers of attorney" cost several millions of dollars because the same functionality had to be re-implemented within each CIF.

When organizations embark on multi-year, multi-million dollar, enterprise resource planning (ERP) implementations, they believe that the business benefits delivered by ERP solutions will far outweigh the costs of implementation. Empirical

evidence shows that approximately 40 percent of the cost of an ERP implementation is around data integration. Sound information governance will ensure that not only is "the data loaded correctly," but that the "correct data is loaded correctly" into the ERP application. When an organization migrates to new ERP applications, the information governance program needs to ensure that data within the source systems is understood, cleansed, transformed, and delivered to the ERP system.[2]

Table 14.2 provides a sanitized version of the business benefits associated with improving the quality of data as part of an SAP rollout. The SAP project will achieve limited savings during the initial rollout due to the upfront cost associated with adopting a new approach. However, the project will produce substantial benefits in subsequent rollouts, with the adoption of a data integration center of excellence. In addition, the business benefits shown do not reflect soft savings associated with improved decision-making based on better data quality.

Table 14.2: The Sanitized Business Benefits from Improved Data Quality as Part of an SAP Rollout

Rollout	As-Is Approach	To-Be Approach with Focus on Data Quality	Hard Savings
A. Number of hours for rollout 1	30,000	27,000	3,000
B. Number of hours for rollout 2	25,000	15,000	10,000
C. Number of hours for rollout 3	20,000	8,000	12,000
D. Total number of hours (A + B + C)	75,000	50,000	25,000
E. Labor cost per hour	$75	$75	$75
F. Total labor costs (D × E)	$5,625,000	$3,750,000	$1,875,000

4. Support the VP of Testing by Automating the Creation of Test Data Sets

The creation of realistic test data sets can add significant extra time to projects. For example, a large organization added several months to its project timeline to create thousands of realistic customer data sets for testing. According to the white paper "Enterprise Strategies to Improve Application Testing" (IBM, April 2008), it is typically impractical to clone an entire production database, made up of hundreds of interrelated tables, just for testing purposes. First, there are the capacity, cost, and time issues with provisioning an entirely new database environment just for testing. Second, there is a quality issue: when working with large test databases, developers might find it difficult to track and validate specific test cases.

Here are some of the requirements for effective test data management:

- *Create realistic data.* It is important to create a smaller, realistic subset of data that accurately reflects application production data.

- *Preserve the referential integrity of the test data.* The data subsets need to respect the referential integrity rules enforced within the database and the applications. Typically, application-enforced referential integrity is more complex. For example, the application might include relationships that use compatible but not identical data types, composite and partial columns, and data-driven relationships.
- *Force error and boundary conditions.* Creating realistic subsets of related test data from a production database is a reasonable start. However, it is sometimes necessary to edit the data to force specific error conditions, or to validate specific processing functions.
- *Mask and transform test data.* With the increased focus on data privacy, the ability to transform and de-identify sensitive data in the development and testing environments is critical to preventing data breaches and severe penalties.
- *Compare before and after test data.* The ability to compare test data before and after successive tests is essential to the overall quality of the application. This process involves the comparison of each test iteration against baseline test data to identify problems that otherwise could go undetected—especially when tests potentially affect hundreds or thousands of tables.

Case Study 14.2 provides an example of a large information services company that dealt with several issues around test data. Table 14.3 highlights the potential business benefits from the case study's solution.

Case Study 14.2: Managing test data at a large information services company

A large information services company was focused on improving productivity, increasing efficiency, and providing high levels of customer service for its corporate clients. The business had a strong focus on cost cutting while protecting confidential client data. However, the IT department found that the testing department was making copies of sensitive data for use in test environments. As a result, IT embarked on a journey to quantify the business benefits associated with a more streamlined test environment. Due to the large volumes of data, IT found that:

- Right-sizing the test environments would have a significant impact on storage cost savings.
- Masking the data within pre-production, performance, and test environments would significantly reduce the risk of exposing confidential client information.
- Automating scripts to mask and populate test data would significantly improve the productivity of the testing team, due to reduced downtime.

In addition to the hard-dollar benefits in Table 14.3, the solution also yielded soft-dollar benefits from improved security and privacy of client data.

Table 14.3: The Sanitized Hard-Dollar Business Benefits from a Test Data Management Solution at a Large Information Services Provider		
A.	Size of production database in gigabytes	5,000
B.	Number of production environments	1
C.	Number of test environments with cloned copies of production data	8
D.	Annual cost of storage per gigabyte	$50
E.	Total annual cost of storage (A × (B + C) × D)	$2,250,000
F.	Percentage of production data to be cloned to each test environment	20%
G.	Size of each test data environment in gigabytes (A × F)	1,000
H.	Annual storage in gigabytes after right-sizing the test data environments (A + (C × G))	13,000
I.	Future-state cost of storage (H × D)	$650,000
J.	Potential storage cost savings (E – I)	$1,600,000
K.	Number of hours saved on refreshing the test environment per year (4 hours per refresh × 12 refreshes per year × C)	384
L.	Total number of testers affected by refresh cycles	90
M.	Potential productivity impact on testing team (K × L)	34,560
N.	Cost per hour of a tester	$50
O.	Total impact on testing productivity (M × N)	$1,728,000
P.	Annual business benefits from test data management solution (J + O)	$3,328,000
Q.	Business benefits over three years from test data management solution (P × 3)	$9,984,000

5. Work with the Enterprise Architecture Team to Enforce Consistent Information Architecture Standards

In some cases, the information governance team might assume responsibility to ensure that the organization observes standards for information architecture. In many organizations, an IT enterprise architecture review board that has sign-off authority for new IT projects might handle IT architecture governance. This board has an important role in driving overall IT efficiency through the enforcement of standards. For example, the standardization of tools is critical as organizations look to reduce license, software maintenance, and support costs. Keep in mind, however, that there is always a natural tension between IT architecture teams that tend to prefer standardization and business units that prefer tactical approaches to address project needs.

Summary

IT has a critical role as the custodian of data to serve the needs of the business. In addition, IT can benefit from sound information governance to drive operational efficiencies and cost savings.

Notes:

[1] "Control Application Data Growth Before It Controls Your Business," IBM, September 2009.

[2] "IBM Information Server: Easing SAP implementations, migrations and instance consolidation," IBM Corporation, 2007.

CHAPTER 18

OPERATIONS

This chapter includes contributions by Mary Bunzel and Pete Karns and (both of IBM).

Operations is a general category that encompasses a class of recurring activities that occur within the enterprise. For the purposes of this book, we include the following general categories of activities within the term "operations":

- Customer service, including the call center
- Billing
- Mailing of goods, statements, invoices, and marketing materials
- Managing inventories
- Scheduling of fleet, aircraft, and railroad cars
- Scheduling of crew and workforce
- Processing of documents, such as bank mortgages and insurance claims
- Maintenance of assets

Because the scope of functions within operations varies by industry and by organization, it is difficult to establish a uniform set of information governance best practices for this category. Nevertheless, the following best practices, in our experience, have the broadest applicability to sell the value of information governance to the operations department:

1. Standardize nomenclature and improve the quality of asset data to reduce capital expenditure and operating costs.
2. Improve the quality of data for crew scheduling.

3. Enhance the privacy of sensitive customer data within customer service environments.
4. Improve operational efficiencies with better customer data.

These best practices are discussed in detail in the rest of this chapter.

1. Standardize Nomenclature and Improve the Quality of Asset Data to Reduce Expenditure and Costs

Several companies have fragmented asset management systems that do not represent a single view of their overall assets. By implementing governance around asset data, companies can reduce capital expenditures, plant downtime, property tax assessments, and the associated labor costs.

Figure 18.1 shows how industry leaders are using asset information governance as a competitive advantage. The figure highlights key differentiators between laggards and leaders in terms of the management of operational assets. Every operation and maintenance budget includes major categories such as maintenance, repair, and operations (MRO) inventory, labor costs, and unplanned downtime, which result in reactive spending, under-utilization of assets, and capital expenditure on equipment that needs to be replaced because of wear and tear.

As MRO inventory turns increase from less than once per year to two to three times per year, market leaders can reduce annual operating costs by an average of 10 to 25 percent. In addition, market leaders can increase production by 8 to 15 percent by improving asset utilization and reducing unplanned downtime. Finally, market leaders can reduce capital expenditure by three to seven percent, which frees up funding for transformational investment or direct savings.

Case Study 18.1 illustrates information governance around equipment run time to reduce preventive maintenance costs at a hypothetical manufacturer.

In addition, Case Study 18.2 illustrates information governance around the "property value" attribute of asset data at a pharmaceutical company.

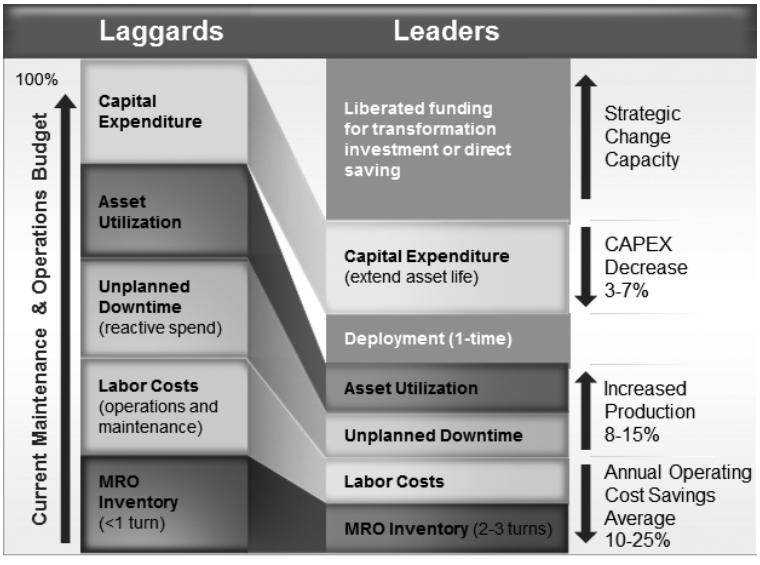

Figure 18.1: Industry leaders are using asset information governance as a competitive advantage.

Case Study 18.1: Information governance around equipment run times

A manufacturer used large stainless steel tubs to mix product at one of its larger plants. Each tub required a large mixer powered by a motor, which in turn required quarterly preventive maintenance at an average cost of $15,000 per occurrence.

The maintenance department implemented governance around the actual run time of each tub. As a result, they found that only a few motors required maintenance because their actual run time exceeded the parameters. In other cases, the motors did not require maintenance because they had not exceeded the maximum threshold of run time during a quarter. As a result, these motors could receive minor preventive maintenance and still attain corporate goals of equipment availability.

By governing the data around equipment run times, the maintenance department was able to reduce maintenance costs at just one plant by more than $1 million per year. It is noteworthy that the maintenance department would not have been able to drive these savings in the absence of standard naming conventions. If the same tub had been called different names, the maintenance department would have found it impossible to standardize maintenance practices across the plant.

Case Study 18.2: Information governance around property values

A major pharmaceutical company paid taxes on assets at each of its locations. The operations department embarked on a major governance program around its asset data.

As the operations department cataloged and validated the asset inventory, it discovered major discrepancies in its ledgers. In some cases, the discrepancies exceeded more than 10 percent of the asset values that were the basis for the company's tax assessments. By improving the accuracy of the asset registries, the operations department was able to reduce the amount paid in property tax assessments.

The essence of effective parts inventory management is to have the right parts available for the right equipment at the right time, and not before. Maintenance departments need to ensure the appropriate information governance around the complete bill of materials for equipment. For example, the correct cross-reference between parts and equipment improves visibility to safety stock requirements, eliminates obsolete inventory, and reduces overstock of inventory. In addition, companies may maintain multiple repositories for their parts inventory data at different locations and even within the same location. This fragmentation may be due to a number of reasons, including acquisitions and the lack of an enterprise asset management system. As a result, the same part may be represented by multiple part numbers within the same company.

Companies can generate substantial cost savings by standardizing the data around their part numbers and by having standard cross-references between parts and equipment. Table 18.1 lays out the hypothetical business benefits associated with standardizing parts inventory at a manufacturer. The manufacturer currently spends $100 million on MRO inventory items across 10 commodity groups. It holds $10 million in parts inventory, and the current inventory carrying cost is 18 percent.

The manufacturer implements a centralized electronic catalog with standardized part descriptions. In just the first year, the manufacturer experiences a reduction of 30 percent in parts inventories due to the elimination of duplicate part numbers. As a result, the manufacturer can reduce parts inventories by $3 million and inventory carrying costs by $540,000.

Table 18.1: The Hypothetical Business Benefits from Standardizing MRO Parts at a Manufacturer		
A.	Annual MRO spend across 10 commodity groups	$100,000,000
B.	Average MRO inventory	$10,000,000
C.	Inventory carrying cost	18%
D.	Annual carrying cost of the MRO inventory (B × C)	$1,800,000
E.	Reduction in MRO inventory due to the elimination of duplicate part numbers	30%
F.	Savings in carrying cost of the MRO inventory (D × E)	$540,000

Figure 18.2 describes the maintenance steering committee at a large automobile manufacturer. The global maintenance director chaired the maintenance steering committee. Engineering, purchasing, IT, and safety were also represented on the committee. Finally, the committee included maintenance representatives from some of the larger plants that led the entire organization in terms of maintenance best practices.

The maintenance steering committee approved the enterprise asset management template, which defines critical workflows such as the one in Figure 18.3. In addition, the template also defines required maintenance response times for key assets, based on a trade-off between cost and associated benefits. For example, the steering committee approved a maintenance response time of eight to 13 seconds to restore a robot that was critical to the assembly line.

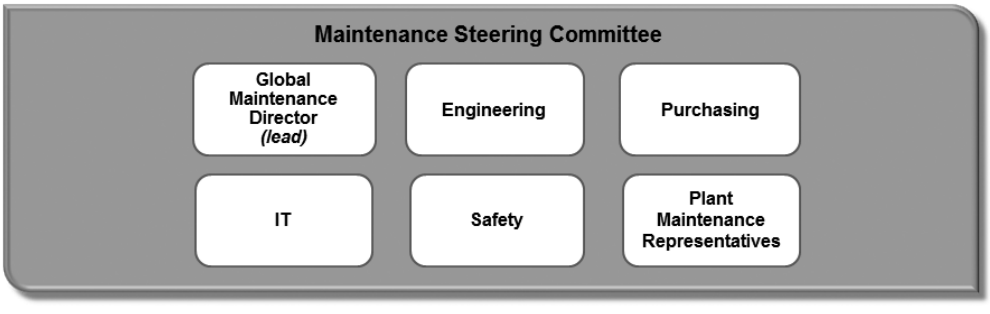

Figure 18.2: The maintenance steering committee at a large automobile manufacturer.

Figure 18.3 describes the process for maintenance change requests at the same automobile manufacturer. The engineering team at a plant might initiate a change request. The request might be to change the configuration of a pump, or to prevent corrosion by switching a bolt from iron to stainless steel. If multiple storerooms use the item, the change request is routed to the plant control group for approval. After approval by the plant control group, the change request is routed to finance for approval. After approval by finance, the change request is routed to purchasing for changes to supplier contracts.

Figure 18.3: The process for maintenance change requests at a large automobile manufacturer.

2. Improve the Quality of Data for Crew Scheduling

Crew scheduling is highly dependent on good data about employee skills and certifications. Accurate scheduling involves matching the right person with the right job. Crew scheduling applications require a detailed understanding of the skill requirements for a given job and the potential people who meet those criteria in terms of availability, skills and certifications, and other constraints such as shift rules.

Organizations need to establish common definitions for skills. Consider a hypothetical example of a large railroad that has repair depots in various cities. If site A defines skills one way, and sites B and C define them another way, then the company will find it hard to move personnel between sites for specialized repair jobs. The company will also find it hard to drive consistency in terms of repair and job plans, and an understanding of how long it takes to complete a given job based on the skills required. This affects business outcomes, such as asset availability and operation and maintenance (O&M) costs.

3. Enhance the Privacy of Sensitive Customer Data Within Customer Service Environments

Many customer service jobs are entry-level positions characterized by low wages and high turnover. By the very nature of their jobs, customer service representatives need to have access to sensitive customer data such as name, address, telephone number, and account balances. As discussed in chapter 7, on travel and transportation, a vacation planner used information from a reservation system to alert an organized crime ring when people would not be in their homes.

While the vast majority of personnel abide by organizational policies, vice presidents of customer service need to be vigilant about safeguarding sensitive customer data. The information governance program should institute a number of policies including restricted views, database monitoring, and data masking, as discussed in Case Study 18.3.

Case Study 18.3: Information governance around customer data privacy at a wealth manager

A wealth management institution was very concerned about the leakage of sensitive financial data regarding its high-net-worth customers. These financial details included information such as net income, net worth, and asset allocation. If this information were publicly disclosed, it would cause potential embarrassment to clients, as well as affect the market reputation of the wealth manager itself.

The information governance program instituted a number of policies to safeguard the interests of the organization. For example, the wealth manager restricted access to the records of its extremely high-net-worth clients to only the relationship manager. For all other accounts, the wealth manager masked sensitive customer data such as Social Security numbers. Finally, it implemented a program to monitor database access by privileged users such as customer service agents. Whenever a SELECT statement was issued against a client record, an alert was sent to a third party outside the customer service organization. This prevented unauthorized access to customer records, even if driven by a customer service agent's idle curiosity. Table 18.2 lays out the hypothetical business benefits from this solution.

Table 18.2: The Hypothetical Business Benefits Associated with Customer Data Privacy at a Wealth Manager

A.	Total market capitalization of wealth manager	$5,000,000,000
B.	Probability of a publicly announced breach of customer records (pre-information governance)	2%
C.	Probability of a publicly announced breach of customer records (post-information governance)	0.5%
D.	Expected reduction in market capitalization in the event of a publicly announced breach of customer records	2%
E.	Positive impact on market capitalization as a result of information governance around customer privacy (A × (B – C) × D)	$1,500,000

4. Improve Operational Efficiencies with Better Customer Data

There are a number of examples where improving the quality of customer data can increase operational efficiencies. Here are a few examples:

- *Call center*—A health plan had data for its members fragmented across 10 different silos. As a result, the call center agents had to go to 10 different screens to complete a change of address for a member. The information governance team embarked on an MDM program around member data, as discussed in Case Study 18.4. The customer service department was able to update a member address only once in the MDM hub, which propagated the changes to the 10 legacy systems. This initiative dramatically reduced customer service costs at the health plan.

Case Study 18.4: One-time data stewardship at a European financial institution

A European financial institution needed to implement a one-time data stewardship process to cleanse customer data that was being migrated to the MDM hub. The information governance program calculated that it needed to manually assess more than 100,000 suspected duplicate customer records, to determine if there were any matches. The information governance team had budgeted for only one data steward to manage customer data on an ongoing basis. However, it needed vastly more resources to complete the initial assessment of the customer duplicate suspects.

Senior management identified the customer service department as the key stakeholder in the success of the customer MDM initiative. As a result, each customer service representative was tasked with reviewing 25 customer duplicate suspects each day. With a 200-person team, the customer service department was able to get through the entire data set in about a month (200 customer service representatives × 20 working days × 25 customer records per day = 100,000 customer duplicate suspects).

- *Mailings and fulfillment*—Mailing and fulfillment departments spend huge amounts of money on postage and the actual costs of the items themselves. Operations departments within the United States can reduce postage costs by implementing ZIP+4 postal codes and by improving the quality of mailing addresses to result in fewer returned mailings. Chapter 12, on marketing, includes a case study on the savings in mailing costs by reducing the number of customer duplicates and implementing a householding program.
- *Route planning*—Chapter 7, on travel and transportation, discusses the case study of a logistics provider that reduced fuel costs through better route optimization. The logistics provider reduced the length of delivery trips by developing a better understanding of different business customers who would be at different addresses but under "the same roof."

Summary

Operations can derive significant business benefits from information governance. These benefits touch many facets of the business, including capital budgets, maintenance, crew scheduling, customer service, billing, and logistics.