CREATING AN
ENTERPRISE DATA STRATEGY

Managing Data as a Corporate Asset

BY WAYNE ECKERSON

Director of Research, Business Applications and Architecture Group, TechTarget

SPONSORED BY: **embarcadero**

Executive Summary

CREATING AN ENTERPRISE data strategy is not for the faint of heart. It first requires a commitment from the top and an acknowledgement that data is a corporate asset that must be managed and protected like any other asset. Given the difficulty of getting executive buy-in, it's not surprising that only one in 10 organizations have an enterprise data strategy.

To harness data as a corporate asset requires a mix of "soft" skills, required to build sustainable strategies and manage change, and "hard" stuff, which applies a portfolio of data management tools and techniques to ensure the delivery of consistent, high-quality data that's aligned with business strategies and initiatives.

One of the key elements in managing data is reconciling enterprise and local requirements for data. Most organizations whipsaw between these two extremes, but astute managers foster a dynamic interplay between the two polarities, embracing both at the same time without getting stuck at either end (or somewhere in the middle.) Organizations with highly centralized approaches to data management need to distribute development back to the departments (as scary as that might be) while maintaining high standards for data integrity and quality; organizations with fragmented data and lots of analytic silos need to implement processes and controls to standardize shared data across functional areas.

A key element in any data strategy is to design and implement a data governance program. The fundamental premise of a data governance program is that the business designs and runs the program and the IT department (or data management team) executes the policies and procedures defined by the data governance team. In addition, a good data governance program incorporates change management practices that accelerate user adoption and ensure long-term sustainability.

Finally, the data management team needs to implement a broad portfolio of data integration tools to accelerate and automate data management tasks. As the type and volume of data that users want to analyze expands, data management teams also need to expand on the tools they can use to manage and manipulate this data. ∎

Data Strategy in Context

**The promise**. It's been said, "The guy with the best data wins."

During my almost 20 years in the business intelligence (BI) industry, I've researched many top-notch organizations that harness data as a corporate asset and reap the rewards. Here are three:

1. **1-800 CONTACTS.com**, an online provider of contact lenses, delivered near-real-time data about sales to its call center staff and executives through a series of dashboards, resulting in $55,000 additional revenue a month and increasing the company's agility to respond to fluctuations in market activity.

2. **Kelley Blue Book Co.**, an online provider of automobile valuations, sources, consolidates, normalizes and analyzes millions of automobile transactions every week to offer the most accurate and timely car valuations in the marketplace.

3. **GE Capital Rail Services**, a provider of railcar and intermodal fleet management services, leverages historical and near-real-time data to optimize railcar repairs, improving customer service and satisfaction and dramatically reducing operational costs.

In these and hundreds of other examples, both private and public sector organizations are using data to increase revenues, reduce costs, improve customer service and loyalty, and improve employee satisfaction. For these organizations, treating data as a corporate asset is a no-brainer. They invest in data—including the people, processes and technology required to manage data—proportionally to any other corporate asset, including cash, securities, people, equipment, facilities and intellectual property.

The reality. I talked recently with Peter Aiken, a well-known expert and consultant in the data management field, who claimed that only one in 10 organi-

FIGURE 1: *What is the status of your data quality plan?*
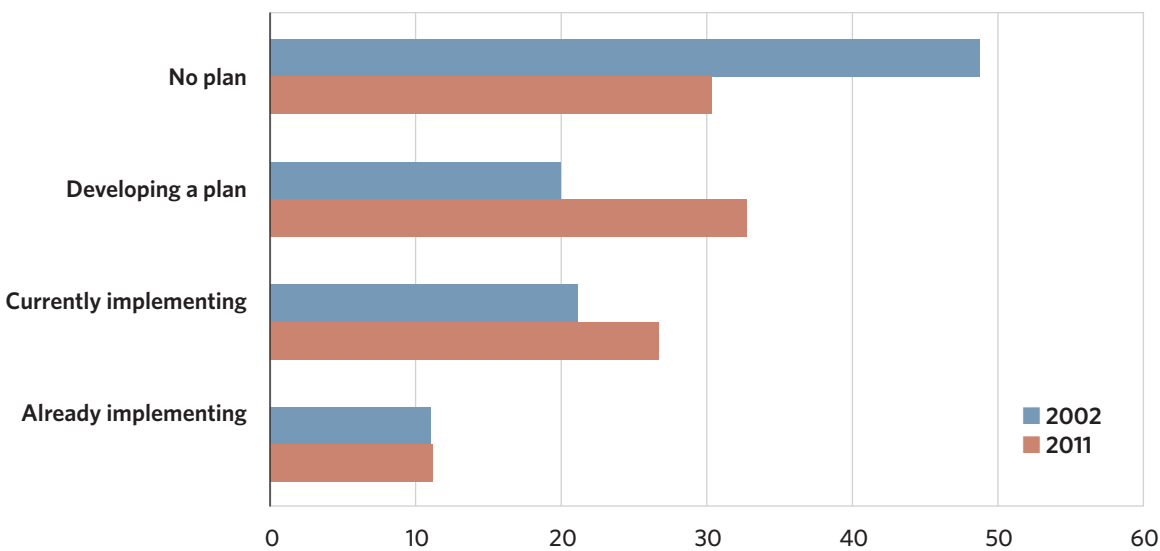
SOURCE: 2011 DATA FROM BI LEADERSHIP FORUM SURVEY (WWW.BILEADERSHIP.COM) OF 96 BI DIRECTORS; 2002 DATA FROM A SURVEY OF 650 BI PROFESSIONALS BY THE DATA WAREHOUSING INSTITUTE.

zations has a data strategy. I was shocked. I pushed back on the figure, but Peter was insistent. It was only until I conducted my own survey as part of this report that I realized Peter was right. As a consultant, Peter usually deals with organizations in data distress; as an industry analyst and thought leader, I engage mostly with innovators and early adopters.

My survey shows that only 11% of organizations in 2011 have implemented a data quality plan (see **FIGURE 1**). Startling. Even more astonishing is that this figure has not changed since I asked the question almost a decade ago![1] The only good news is that the percentage of companies without a data quality plan has dropped from 48% to 30%, and higher percentages are planning or implementing. But still …

To triangulate this sorry state of data affairs, Tony Fisher, a longtime executive at a leading

> *"Only one in 10 organizations has a data strategy."*
> —*PETER AIKEN*
> *founder, Data Blueprint*

---

[1] See Wayne Eckerson, "Data Quality and the Bottom Line: Achieving Business Success through a Commitment to High Quality Data," The Data Warehousing Institute, 2002. www.tdwi.org/research.

data quality software provider, claims that only 15% of organization have reached the two highest levels in his data governance maturity model, where-as 35% are stuck in Level 1 and 50% are in Level 2.

Tony Fisher's Data Governance Maturity Model[2]

LEVEL 1: Undisciplined (35% of organizations).
Reactive, IT-driven projects usually carried out by an individual; duplicate, inconsistent data; inability to adapt to business changes; executives unaware of the cost of poor data; no standards for cleaning or sharing data.

LEVEL 2: Reactive (50% of organizations).
Line-of-business influences IT projects; little cross-functional collaboration; high cost to maintain multiple applications; short-range projects; little executive oversight.

LEVEL 3: Proactive (10% of organizations).
IT and business collaborate; data viewed as a strategic asset; a team of data stewards maintain corporate data definitions and business rules; ongoing monitoring; emphasis on preventing rather than fixing data quality problems.

LEVEL 4: Governed (5% of organizations).
Business requirements drive IT projects; repeatable, automated business processes; personalized customer relationships and optimized operations; data projects funded appropriately; executives and users trust in all data-based decisions.

[2]From Tony Fisher, "The Data Asset: How Smart Companies Govern Their Data for Business Success," Wiley & Sons, 2009.

DATA CHALLENGES
Although the promise of data is high, the reality is abysmal. So, what is the explanation?

■ **Humpty Dumpty.** To put it simply, data is hard. Managing data is a long-term employment opportunity. (And many of us in the industry who make our living educating and consulting organizations on how to leverage data are thankful for that!) Piecing together data scattered across an enterprise is like trying to put Humpty Dumpty back together again. It takes a lot of time to find and fit the pieces together and glue them tight. Except each time the work is nearly done, Humpty climbs back on the wall and soon takes another nose-dive, splitting again into a million pieces. Mergers and acquisitions, new executives and strategies, new technologies, and new regulations and competitors all conspire to fragment organizational data, making it a Herculean task to keep Humpty Dumpty whole and vital.

■ **Data flood.** And the job of managing data is getting harder. Data managers today are beset by a flood of data and new data types. Organizations want more detailed transaction and interaction data to optimize processes, such as supply chains, Web marketing and operations. They want to generate insights from new data sources, such as social networking sites (e.g., Twitter, RSS, Facebook, LinkedIn, blogs), machine-generated data (e.g., sensors, GPS data, radio-frequency identification), geospatial data, and email and text messages. And they want to monitor dozens of new data-driven applications in the cloud, on the Web and on mobile devices. In turn, organizations want all this data faster, as the velocity of business increases, so they can act more quickly to prevent problems and exploit opportunities.

> *The root of the problem is more intrinsic. It has to do with basic human and organizational behaviors.*

■ **Human and organizational behaviors.** The relentless pace of business and technology innovation certainly contributes to our data ills. Also, a lack of funding and sponsorship, suboptimized data processes and missing, outdated or inappropriate technology magnify the issue. Yet the root of the problem is more intrinsic. It has to do with basic human and organizational behaviors.

› **Awareness.** The first stage in leveraging data as a corporate asset is to get business executives to acknowledge that there is a problem. Executives

may view data management as an IT issue, not a business one. They may not see the true cost of poor-quality data. And they may not see that they are both the source of and solution to the problem.

› **Fear of change.** Once the problem is acknowledged, summoning the will power to fix it can be daunting. Individuals fear change and will resist it. They may fear losing their status, privilege or position of power, which are often rooted in managing the flow of data. They may claim they don't have the time or prefer the current methods of accessing and analyzing information.

› **'Stuckness.'** Humans are blind to what they cannot see and have not experienced. They are not open to new ways of doing or thinking about things because they are stuck in the old and blind to the new. Without education and, often, some coercion, most businesspeople won't adopt a new strategy for handling or using data.

› **Lack of trust.** Executives might not trust IT to execute needed changes. And workers (including IT) might not trust executives to create and execute a plan instead of just talking about one. ("Here we go again. This strategy soon will pass.") Also, each department has its own goals, jargon and incentives, creating a virulent "us versus them" mentality. Often, data becomes a victim of organizational turf wars.

› **Organizational dynamics.** Wise leaders know that no single approach to data management (or anything else, for that matter) works for long. To optimize performance, they need to balance opposites rather than eradicate them: short term and long term; enterprise and department; top down and bottom up; speed and standards; creativity and structure; and control and openness. They know success comes from living in the tension created by bringing opposite forces together.

KEYS TO MANAGING ENTERPRISE DATA

■ **Change management.** Enterprise data strategies need to address the above human and organizational dynamics to succeed. In fact, most data management experts agree that the most critical aspect of any enterprise data strategy is managing change and overcoming individual and organizational resist-

ance. To succeed with change management, you almost need dual degrees in sociology and marketing! More realistically, you need a streetwise understanding of what makes people tick, a plan of communication and education, and a lot of persistence and patience.

■ **Data governance program.** A key ingredient of any change management plan is baking new processes into the fabric of the organization; so any change you implement sticks. Once change is embedded into processes, it becomes sustainable. This is the ultimate goal. Data governance is a business-driven program that creates sustainable processes around an enterprise data strategy. A data governance framework defines guiding principles, goals, policies, roles, decision rights, accountability, processes, measurements and oversight committees. It is wise to recruit outside experts to mentor this effort, because it's not a technology initiative; it's a change management endeavor.

Once change is embedded into processes, it becomes sustainable.

■ **Data management portfolio.** Moreover, organizations need to upgrade their portfolio of data management tools to address new challenges in the era of "big data" and "deep analytics." They need data quality and integration tools that can exchange rules and support collaborative development. They also need tools to archive and secure data as well as mask data to ensure privacy. They also need data services that can capture and filter changes from high-volume source applications and trickle feed them into low-latency analytical systems. In addition, they need tools that can federate, replicate and integrate data between corporate data centers and the cloud, and interoperate with new data management platforms, such as data warehousing appliances, Hadoop Distributed File System, NoSQL databases, search-based indexing systems and visual analysis tools.

It is a tall order to address both the "soft stuff" and "hard stuff" required to deliver an enterprise data strategy and treat data as a corporate asset. Entire books have been written on this topic. This report will attempt to highlight the most salient points. ■

Data Problems

SYMPTOMS. Given that the first step toward an enterprise data strategy is building awareness, it might be helpful to list a few symptoms of poorly executed or nonexistent data strategy. If your organization exhibits any of the following characteristics, it is wasting valuable time and money, jeopardizing the success of major initiatives and forfeiting valuable opportunities to gain a competitive edge.

You know you have a data problem when

› Executives don't know how many customers the company has or how many it has lost.

› Meetings degenerate into arguments about whose spreadsheet is right.

› Your company treats existing customers like it doesn't know them very well.

› Business users must submit a request to IT to get a custom view of a report.

› Your company tries to retain customers who cost more money than they generate.

› Businesspeople don't trust IT to deliver applications and data in a timely manner.

› A merger fails to deliver promised synergies because the two companies share more customers than thought.

› There are insufficient controls to ensure employees can't see confidential customer data.

› Software vendors sell to business, bypassing the IT department until the deal closes.

› Changing a data warehouse subject area or corporate report takes weeks because metadata is not centrally managed.

› Executives say data is a corporate asset but don't assign the right business executives and analysts to oversee the program or allocate enough funds to make it work efficiently or effectively.

› Businesspeople think buying tools will address data and systems problems.

› Most departments create and maintain their own data marts, BI tools and reports.

› There is no process for archiving data that must be retained or disposing data that's no longer needed.

› There are thousands of reports with contradictory rules and definitions.

› No one was ever fired for misusing corporate data or violating established data policies.

> *The difference between organizations that are crippled by data problems and those that harness data for competitive gain is awareness.*

These are just a few symptoms of poor data management practices. Every organization can cite a litany of data problems, even those that have implemented processes to treat data as a corporate asset. The difference between organizations that are crippled by data problems and those that harness data for competitive gain is largely one of awareness. The latter have acknowledged they have a data problem and have begun to take action.

■ **Perceptions.** Unfortunately, our research shows that many organizations have misplaced perceptions about the quality of their data, and this behavior
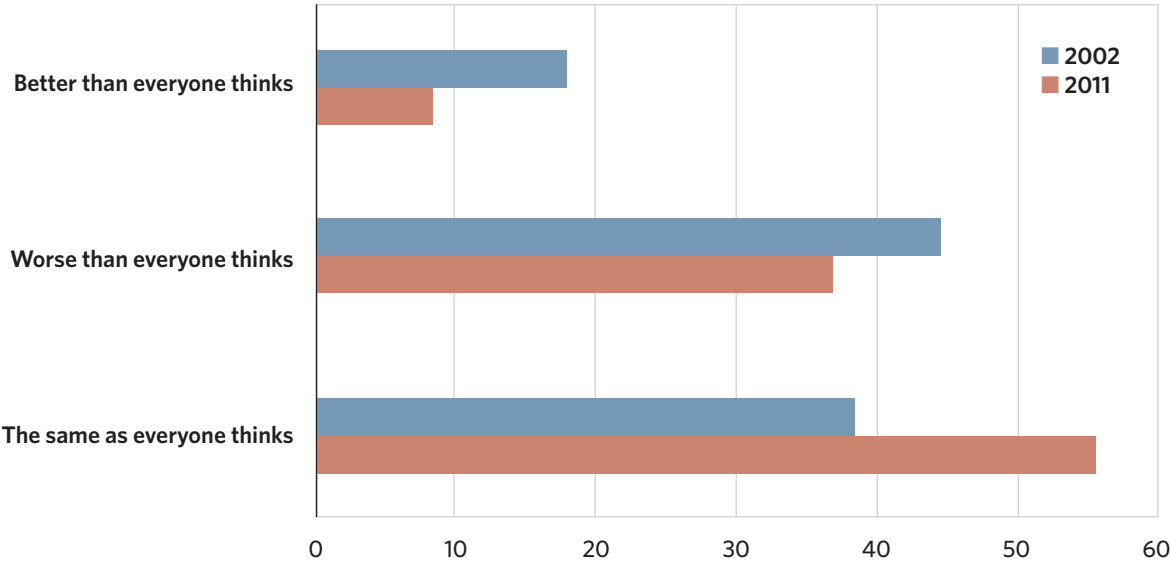
hasn't changed much in the past decade. When BI managers were asked to assess organizational perceptions of data quality, more than a third (37%) said it's "worse than everyone thinks," down only slightly from almost a decade ago. This is still a startlingly high percentage, although now a majority (55%) has an accurate perception (see **FIGURE 2**).

When asked to rate key data challenges today, BI directors cited business and leadership issues above technical ones. The top vote getter was "getting the business to take data quality and integrity seriously" (59%), followed by "reconciling and integrating reference data across applications" (55%), "gaining consensus on key terms and definitions" (55%), and "managing user perceptions of data quality" (52%) (see **FIGURE 3**, page 12). Each of these challenges requires strong executive leadership to acknowledge and evangelize the issue and overcome political resistance from departmental heads and factions.

One of the unique characteristics of data is that it's an enterprise, departmental and individual resource. However, without strong governance, data always defaults to the lowest common denominator, which is, first and foremost, an analyst armed with a spreadsheet, and second, a department head with his own IT staff and data management systems.

FIGURE 2: *In reality, the quality of our data is …*

Legend: 2002, 2011

| Category	
Better than everyone thinks	
Worse than everyone thinks	
The same as everyone thinks	

(horizontal axis: 0, 10, 20, 30, 40, 50, 60)

SOURCE: 2011 DATA FROM BI LEADERSHIP FORUM SURVEY (WWW.BILEADERSHIP.COM) OF 96 BI DIRECTORS; 2002 DATA FROM A SURVEY OF 650 BI PROFESSIONALS BY THE DATA WAREHOUSING INSTITUTE.

■ **Spreadmarts and independent data marts.** These spreadsheets and departmental data warehouses and data marts work fine within their own domain of influence. The problem comes when the organization (i.e., the CEO or chief financial officer) wants to aggregate these localized data sets to obtain an enterprise view of activity or performance. Since every analyst and departmental IT manager defines key entities using his own definitions and rules, it becomes impossible to obtain in a timely manner an enterprise view of key business metrics, such as sales, customers and returns. Given a kaleido-scope of disjointed data, beleaguered CEOs and CFOs run their organizations "blind"—like a pilot whose instruments have gone haywire—and must make decisions on intuition without valid supporting facts.

■ **Painful insights.** At some point, such an organization "crashes"—it experi-ences a major problem because of a lack of integrated, accurate data. It may fail to send bills to customers because of incorrectly coded accounts, losing millions of dollars. Or a faulty spreadsheet formula may cause it to overcharge customers for monthly subscriptions. Or an employee's laptop with sensitive

FIGURE 3: *How challenging are the following data quality activities?*
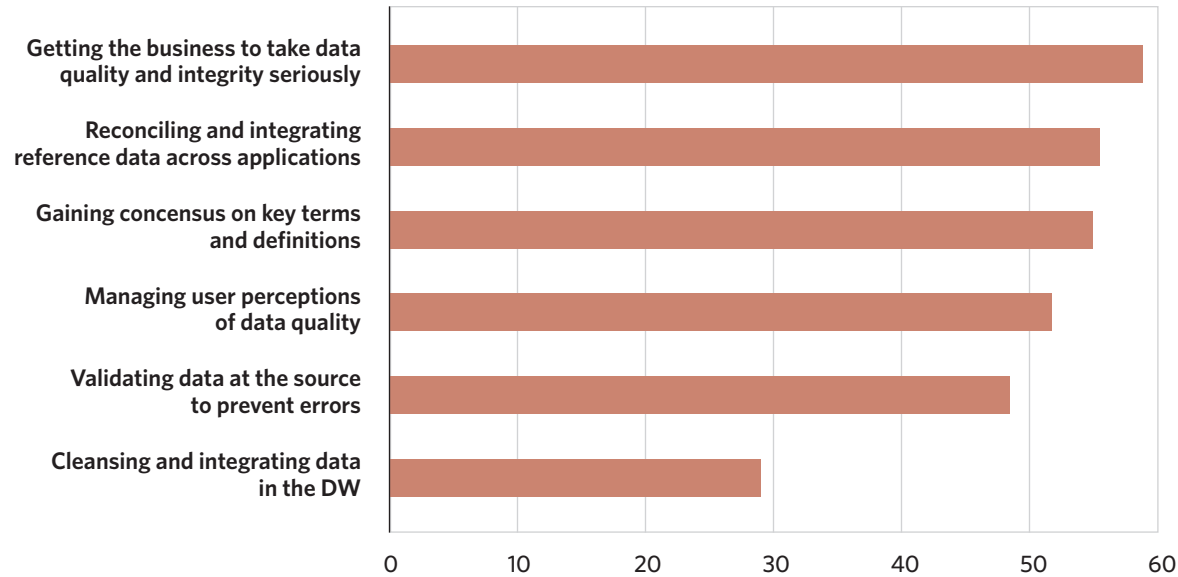
SOURCE: BI LEADERSHIP FORUM SURVEY (WWW.BILEADERSHIP.COM) OF 96 BI DIRECTORS. PERCENTAGES REFER TO RESPONDENTS WHO RATED THE CHALLENGE "HIGH."

customer information gets stolen.

My all-time favorite data quality nightmare happened at a regional bank in the southwestern United States. The bank sent out a mass mailing that addressed one of its major customers, the Mexican Consulate, as "Dear Mexican." The customer was livid, and the CEO had to personally apologize for the incident. The embarrassment caused the CEO to launch a high-profile data quality initiative and assign a senior vice president to oversee it. Kicking off the program, the CEO wrote in a company newsletter:

> *More than 98 percent of our company's assets and those of our customers are managed by data and information—and less than 2 percent are in the form of cold, hard cash. Just as we are careful and meticulous in managing cash and negotiables, we have a duty and obligation to exercise a high degree of care with the data that is the basis for customer relations and decision making.*

Here is a banker saying that data is as important as cash! It is too bad that most business executives need to crash and burn before they recognize that data is an important corporate asset. But given a second chance, these prodigal executives often become stalwart advocates for data governance initiatives.

> *It is too bad that most business executives must crash and burn before they recognize that data is an important corporate asset.*

BALANCING LOCAL AND ENTERPRISE REQUIREMENTS

One of the key challenges in managing data as a corporate asset is to balance local and enterprise data requirements. Most organizations volley between these extremes, suffering all the downsides and reaping none of the benefits of either approach. They either apply too much enterprise control, which becomes repressive and demoralizes employees, or they allow too much freedom and change and slip into chaos, which stresses out workers (see **FIGURE 4**, page 14).

The secret is to allow a dynamic interplay between the two extremes,

FIGURE 4: *Balancing opposites*
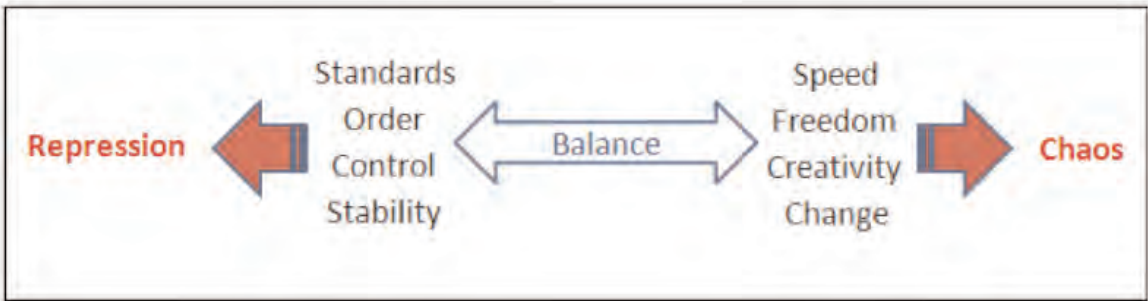
embracing both at the same time without getting stuck at either end (or somewhere in the middle.) Balancing polar opposites creates tension, but it maximizes benefits and minimizes downsides.

■ **BI Maturity Model.** This dynamic between local and enterprise requirements is depicted in TDWI's BI Maturity Model, which I created in 2004. **FIGURE 5**, page 15, shows a view of the five-stage model, which depicts the tension between local and enterprise requirements. As organizations begin their BI journey in the Prenatal and Infant stages, local interests (i.e., analysts and departments) trump enterprise standards. Here, analysts and departments create their own reports and data marts using their own tools, systems, rules and definitions. The mantra is "Think local, resist global." Enterprise standards do not exist or are not heeded or implemented. There is no information alignment or consistency, and there is a lot of reporting overhead from redundant systems and staff.

[At the early stages of BI] analysts and departments create their own reports and data marts. ... The mantra is "Think local, resist global."

As time goes on, the corporate BI team begins to establish data standards and policies and implement systems and tools to manage shared data elements, bringing some measure of order to information chaos. In the Teenager phase, the two lines representing local flex-
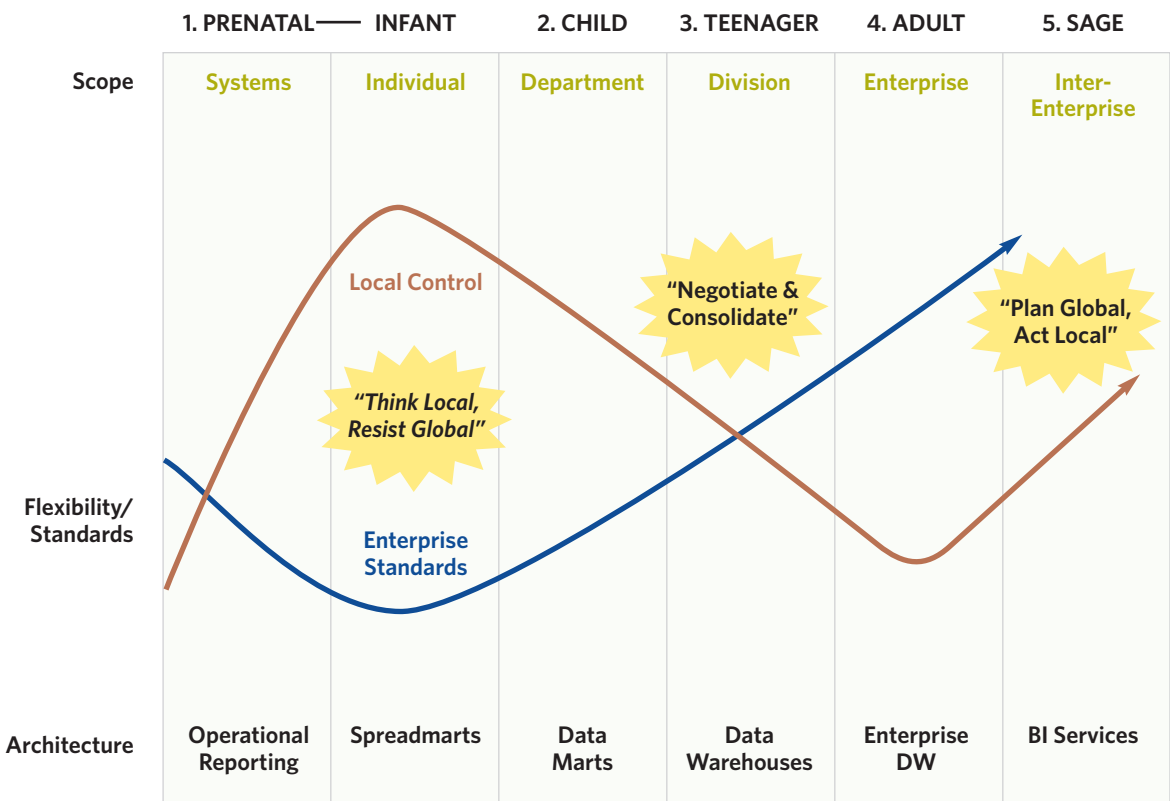
ibility and enterprise standards cross paths, heading in opposite directions.

As the organization moves into the Adult stage, the corporate BI team gains the upper hand as it begins to enforce standards for definitions, rules, tools, architecture, development and projects on an enterprise basis. Unfortunately, as the corporate BI consolidates people, processes and technology, local groups lose their ability to respond quickly to local needs since they no longer have control over their own data and tools.

Unless BI teams can turn the bottom line around (as in **FIGURE 5**) and align local and enterprise requirements, local users will feel disempowered and demoralized. They will either leave the company or circumvent the corporate BI team by creating spreadmarts and independent data marts. If this happens, the organization's BI initiative has fallen into the "chasm" and will revert to earlier stages of development. The organization will struggle for many years to regain its information footing. ■

FIGURE 5: *TDWI's BI Maturity—local versus enterprise requirements*

	1. PRENATAL —— INFANT		2. CHILD	3. TEENAGER	4. ADULT	5. SAGE
Scope	Systems	Individual	Department	Division	Enterprise	Inter-Enterprise
Architecture	Operational Reporting	Spreadmarts	Data Marts	Data Warehouses	Enterprise DW	BI Services

Flexibility/Standards

Local Control

Enterprise Standards

"Think Local, Resist Global"

"Negotiate & Consolidate"

"Plan Global, Act Local"

Reconciling Opposites

SO HOW DO you keep both lines in the BI Maturity Model moving upward in tandem? How do you balance local and enterprise data requirements? How do you marry speed and standards? Freedom and order? Creativity and control? Change and stability? Business and IT?

There are many strategies and tactics. But the ones that work for your organization align with your corporate culture and organizational structure and provide counterweight to the current dynamics of your BI initiative.

In other words, if you have a top-down, centralized approach to data management, you'll need to distribute some development back out to departments. If you have a decentralized BI environment, you'll need to inject standards and processes to improve alignment and economies of scale. If you are just starting out, you'll need to sell management on the value of liberating data from its current jail (e.g., spreadsheets or legacy systems) and quickly deliver new, information-rich applications. Let me give you a few examples.

> *[The data strategies] that work for your organization align with your corporate culture and organizational structure and provide counterweight to the current dynamics of your BI initiative.*

ESTABLISHING ORDER AND CONTROL

Cisco: Cisco Systems Inc. is a classic case of an organization with a decentralized organizational structure that now recognizes that it needs to tighten its data belt and establish uniform, enterprise data management practices that cut across major functional areas and divisions. Cisco is an amalgam of more than 140 companies acquired over the past 25 years. Yet CEO John Chambers has established a corporate culture whose chief aim is to delight customers

and achieve high levels of customer satisfaction and loyalty. To do this, Cisco must consolidate hundreds of data silos, standardize thousands of shared data elements and establish uniform processes to ensure it delivers accurate and relevant data for decision making.

"We are in the process of offering data and reporting as a common, shared service across the organization. Our team is chartered with creating a Cisco-wide business intelligence service that will accelerate and improve operational and financial decisions based on verifiable, trusted data sources," says Naznin Shroff, senior manager in the newly created data foundation, reporting and analytics team, whose scope of work initially spans finance, field operations and channel operations at Cisco. "People are now rallying around the fact that data is critical and we can't do much without it."

> *"People are now rallying around the fact that data is critical and we can't do much without it."*
>
> **—NANZIN SHROFF**
> *Manager of data quality systems at Cisco Systems Inc.*

With strong backing and funding from top management, Shroff is leading the effort to build a unified data foundation at Cisco and has created a multi-year roadmap that will reduce silos of redundant data, standardize key metrics, dimensions and hierarchies, improve data accuracy and relevancy, and align reports with business needs (see **TABLE 1**). From a business perspective, the effort will help streamline customer touchpoints and make it easier for Cisco to integrate new acquisitions.

Table 1. Cisco's 2011 high-level data roadmap

YEAR 0: Establish the vision, charter, goals, benefits and roadmap.

YEAR 1: Establish oversight and execution teams; implement sustainable processes; identify success measures aligned with the business.

YEAR 2+: Execute and evolve operations; drive continuous improvement; achieve best-in-class recognition for data management practices.

■ **Sustainable processes.** Besides strong sponsorship, Shroff said a key to the success of Cisco's data management program is to bake data quality into the DNA of the business. "Creating sustainable processes [around data] is the biggest test. We need to embed data quality into every data project and life cycle at Cisco."

■ **Measurements.** It's also important to define data quality measurements aligned with business value and monitor the ongoing effectiveness of data quality processes. "Without measures, you can't monitor the health of your data, and if you can't monitor, the quality of your data will quickly degenerate," Shroff said. "The whole point is sustainability, and measurements are critical to that."

■ **Relevant data.** Another key is to focus the team's efforts on data that aligns with business insights and not try to manage data for data's sake. That means spending time up-front to understand business strategies and goals and business questions that emanate from those things. "We need to focus on creating business insights more than creating data," Shroff said. "This is a paradigm shift in thinking."

INTUIT: Architecturally, managing enterprise and local requirements is a high-wire balancing act. Intuit Inc., the software maker of Quicken, TurboTax and QuickBooks, among other things, has a decentralized organizational structure, reflecting its numerous acquisitions. Each newly acquired business brings its own systems and information to Intuit, making the delivery of unified and consistent information across the enterprise a major challenge.

In the late 1990s, Intuit created a corporate BI team whose goal was to build an enterprise data warehouse (EDW) to support the reporting and analysis of all the divisions. The EDW would eliminate information silos, reduce costs and deliver a "single version of truth." While the EDW largely succeeded in achieving these aims, it also created a huge project backlog, frustrating the divisions that previously managed their own data resources but now were blocked from getting the information they needed in a timely manner.

In 2008, Intuit's corporate BI team created a hybrid data warehousing (DW) architecture in which the corporate BI team manages the EDW and maintains definitions and rules for data shared across divisions and functional groups. It then allows each division, if desired, to leverage this data infrastructure so it can build its own data marts using corporate-issued extract, transform and
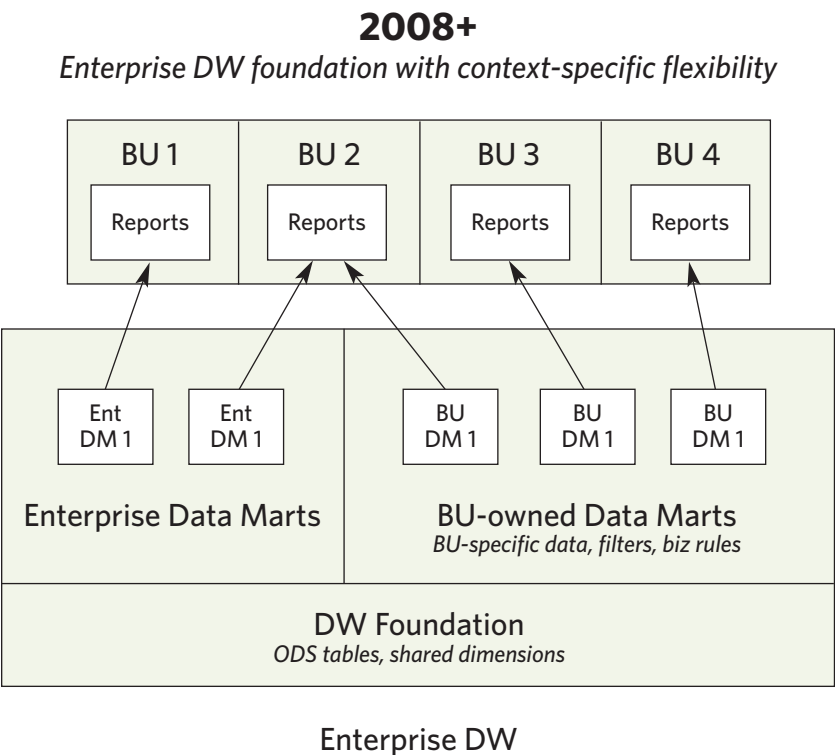
load (ETL) tools. The technical leads in the divisions are allocated space in the EDW and encouraged to use common database naming conventions, error-handling procedures and other corporate standards for managing data in the enterprise data environment. If a division doesn't have the IT resources or time to build its own marts and reports, the corporate BI team fills the gap (see **Figure 6**).

"The path to durable and large-scale success of the BI program involves striking a balance between centralization and decentralization," said Brian Treu, BI director at Intuit. Treu says the hybrid architecture enabled divisions to use a centralized model when a "high degree of rigor" is required and a decentralized model when "high flexibility" is needed.

■ **Center of excellence.** The key to making a hybrid data architecture work is communication and education. First, the corporate BI team needs to identify BI developers and data management practitioners in each business division and then establish channels of communication with these individuals and work-

Figure 6: *Intuit's hybrid DW architecture*

2008+
Enterprise DW foundation with context-specific flexibility

COURTESY INTUIT

groups. It needs to create a BI center of excellence that can enable this extended team of BI and data professionals to share best practices, define operating standards, specify enterprise toolsets, implement controls to secure sensitive data, offer skills training and career development opportunities and standardize BI career paths with avenues for advancement.

"In this kind of environment, it's imperative to overcommunicate," Treu said. "We don't always succeed, but we're heading in the right direction. So far, the results have been encouraging."

EMBRACING SPEED AND AGILITY

Harley-Davidson: Harley-Davidson Motor Co. is an iconic U.S.-based manufacturer of motorcycles. Its data management processes reflect its collaborative business culture. It has an active integration competency center (ICC) and well-defined processes for sourcing, validating and delivering data. Any application or project that requires data must seek the approval of the ICC, which generally also performs the work to ensure the delivery of clean, consistent, integrated data throughout the company.

However, Harley-Davidson's business strategy is changing, and top management is keen to make sure that data and application management operations align with future growth opportunities in global markets and new business models, such as e-commerce and factory customization offerings. Consequently, Jim Keene, systems manager of global information services at Harley-Davidson, is creating a new information strategy to meet emerging business requirements.

■ **Go faster.** The new strategy will assess current capabilities against an ideal future state and recommend strategies for closing the gap. One common theme emerging from the analysis is the need for Harley-Davidson's information management capabilities to be more agile and flexible.

"Harley-Davidson doesn't want access to information to be controlled by a monolithic waterfall development process. They want it to be more iterative and agile. They want to be able to make trade-offs between the completeness and accuracy of data and the ability to access information quickly that they need to run their business," Keene said.

Currently, Harley-Davidson's information management team handles all business requests the same way, regardless of the nature or urgency of the

request. In the future, to go as fast as the business wants, the team will need to break free from its "one size fits all" methodology and take a more customized approach to business requests. "We need to make sure our processes don't get in the way of the business," Keene said.

■ **Proactive modeling.** Keene also envisions working more proactively to meet business requirements. Currently, the team does project-centric logical and physical modeling, but Keene would like to flesh out conceptual models and high-level key performance indicators prior to kicking off a systems project. "We'd like to get ahead of the business, make sure there are no big gaps, by using higher-level models and hierarchies," Keene said.

The new data strategy will likely specify a number of initiatives designed to improve the agility and flexibility of Harley-Davidson's information management team to meet business needs quickly:

> **Agile development.** Speed the creation of new reports and data warehouse subject areas.

> **Self-service BI.** Empower business users to create their own reports and reduce IT backlog.

> **Superuser network.** Recruit tech-savvy business users in each department to help balance ad hoc report creation with standardized report delivery.

> **Packaged analytic applications.** Buy rather than build analytic capabilities to speed time to value.

> **Analytic sandboxes.** Provide business analysts with a centralized resource to explore and merge data without creating analytic silos or waiting for IT to create a custom data set.

> *"We need to make sure our processes don't get in the way of the business."*
> —JIM KEENE
> *systems manager of global information services at Harley-Davidson Motor Co.*

› **BI search.** Implement BI search tools to improve ad hoc query capabilities for executives and managers.

› **Mobile BI.** Deliver BI via mobile devices to supply users with information wherever they go.

› **Analytic platforms.** Implement analytic databases and appliances to speed query performance and support advanced analytic processing.

Each of these initiatives will take time to implement and roll out. But the goal is to balance Harley-Davidson's established data governance processes with the agile delivery of information to support new business opportunities. The goal is to be consistent, accurate and fast.

STAYING A STEP AHEAD

As with Harley-Davison, the problem with shared services and corporate information management teams is that they have a hard time staying ahead of business needs. Often, department and division heads who have their own IT budgets and are incented to optimize performance in their local domains view corporate IT as an impediment to getting things done. They often do end-runs around corporate IT. They negotiate directly with systems, applications and tools vendors and build their own data marts and reports outside of IT's purview. This is how analytic silos get built.

■ **New analytic silos.** In the 1990s, the departmental tools of choice were spreadsheets and Access databases. Today, they are primarily in-memory visual analysis tools, followed by cloud and open source BI tools. But any inexpensive tool that is easy to install and use is a prime candidate for a departmental data delivery system. "If you don't care about data security, control

> *"If you don't care about data security, control and governance, choose an in-memory solution. When used for anything but creating prototypes, they result in complete data anarchy."*
>
> **—PEDRO PERFEITO**
> *BI consultant at Novabase*

and governance, choose an in-memory solution. When used for anything but creating prototypes, they result in complete data anarchy," wrote Pedro Perfeito, a BI consultant at Novabase, a Portuguese BI consultancy, on a BI Leadership Forum discussion thread on LinkedIn, railing against the latest type of spreadmart-inducing toolset.

However, most information management directors understand that you can never completely stamp out these analytic silos, so the best strategy is to coopt or quarantine them. This is easier to do if IT already has an established information infrastructure and data governance processes. Some self-service BI vendors, such as Quest Software Inc., recognize the analytic silo problem and offer tools (such as Toad) that make it easy for business and data analysts to access, filter and join data virtually, without having to download subsets of data that could turn into spreadmarts and analytic silos.

■ **Coopt.** For example, Darrell Piatt, director and chief architect at Corporation Service Company (CSC), believes the new in-memory toolsets offer a nice supplement to the company's existing BI toolset. He plans to select a tool shortly and roll it out to a few analytics users who can benefit from speed-of-thought visual analysis. However, to prevent the analysts from using the tool to create analytic silos, he will link the tool to specific subject areas in the company's EDW so users only use authorized data.

"It's my hope that by getting out in front of this and partnering with our business users, we'll steer things in the right direction," Piatt said.

■ **80/20 rule.** Mike Masciandaro, BI director at The Dow Chemical Co., concurs with this approach and takes it one step further. He won't try to control usage of these types of tools up-front but rather use them as the vanguard for new development in the corporate data environment.

"I don't think it's healthy to think that your central BI solution can do it all. The ratio I'm going for is 80% corporate, 20% niche. I also like the prototyping nature of these tools and when appropriate, I'll take the output and scale it within our central environment. Rather than think of these smaller solutions as competitors, I use them to raise our central capability, which can beat the pants off of every one of them in a 1:1 bakeoff," Masciandaro said.

■ **Quarantine.** In some cases, it makes sense for the corporate group to forestall the creation of analytic silos by creating a temporary solution that can be

easily transitioned to the corporate environment at a later date.

For example, the DW team at Blue Cross and Blue Shield of North Carolina (BCBSNC) was under pressure to provide unified reporting from both old and new claims management systems. Since it would take a long time to integrate the new claims management system with the EDW, the team spent three months creating a temporary repository to house both sets of data and satisfy the immediate business need. Without the repository, the business would have created multiple data marts that would have been hard to shut down, said Celia Fuller, former director of information stewardship and director of data warehousing strategy at BCBSNC.

"We try to develop capacity to do immediate, short-term things to meet business needs while building out our data management infrastructure," Fuller said. In this case, the team used EDW naming conventions and models so it would be easier to migrate users from the temporary repository to the EDW. The team also closely monitored usage of the repository and retired seldom-used tables. It also kept a mailing list of users sorted by category of data so it could inform them of new developments in the EDW and transition them to the new system as quickly as possible.

Part of any cooption strategy is to make sure the final shared service is much more attractive to business users than their temporary system or custom-built data silo. In BCBSNC's case, the EDW offered better query performance across more subject areas than the temporary repository and came with training to help business users learn the new system quickly. In contrast, query performance in the temporary repository was slow and users had to write their own joins. "We didn't want to make the temporary system too easy to use," Fuller admitted.

Although it took years to retire the "temporary" repository, the strategy prevented critical corporate data from fragmenting into business-driven analytic silos. ∎

Data Governance: Embedding Data Quality into Processes

THE ULTIMATE GOAL of an enterprise data strategy is to create sustainable processes that bake data quality into the culture of an organization. A data governance program is the vehicle organizations use to create sustainable processes that ensure appropriate levels of data integrity and quality.

According to Jill Dyché, partner and co-founder of Baseline Consulting Group Inc., data governance does the following things:

> › Establishes business stakeholders as information owners.
> › Positions enterprise data issues as cross-functional.
> › Aligns data quality with business measures and acceptance.
> › Removes IT from business/data conflict ownership.
> › Ensures data is managed separately from applications.
> › Expands change management to include data.
> › Requires monitoring and measurement.

Baseline Consulting has created a data governance framework, based on dozens of client engagements, to help organizations implement data governance programs. The framework consists of five categories of tasks: corporate drivers, business framework, process and policy, data management and data governance execution mechanisms (see **FIGURE 7**, page 26).

CRITICAL SUCCESS FACTORS

Most organizations adhere to a framework to guide their data governance initiatives. Or they hire an outside consultancy to help them step through the process and garner support from top executives. Besides providing a methodology for implementing data governance, consultants emphasize best practices based on first-hand experience. Here are some critical success factors mentioned by data governance experts and practitioners:

1. Align data with business strategy. The first and most important step in a data governance program is to understand the data dependencies of mission-critical business strategies and how a lack of clean, consistent, complete, timely or relevant data hampers the organization's ability to achieve its goals. The challenge here is that many organizations don't have a clearly defined business strategy, or if they do, it's a moving target.

"I always start by focusing on what the business is trying to achieve and relate that to dependency on data," said David Loshin, president of Knowledge Integrity Inc., a data quality and business intelligence consultancy.

2. Start with a cross-functional project. Data governance programs don't start in a vacuum; they need a catalyst, according to Mike Stiffler, data governance consultant for Trillium Software. The best catalysts are cross-functional,

FIGURE 7: *Baseline Consulting's Data Governance Framework*
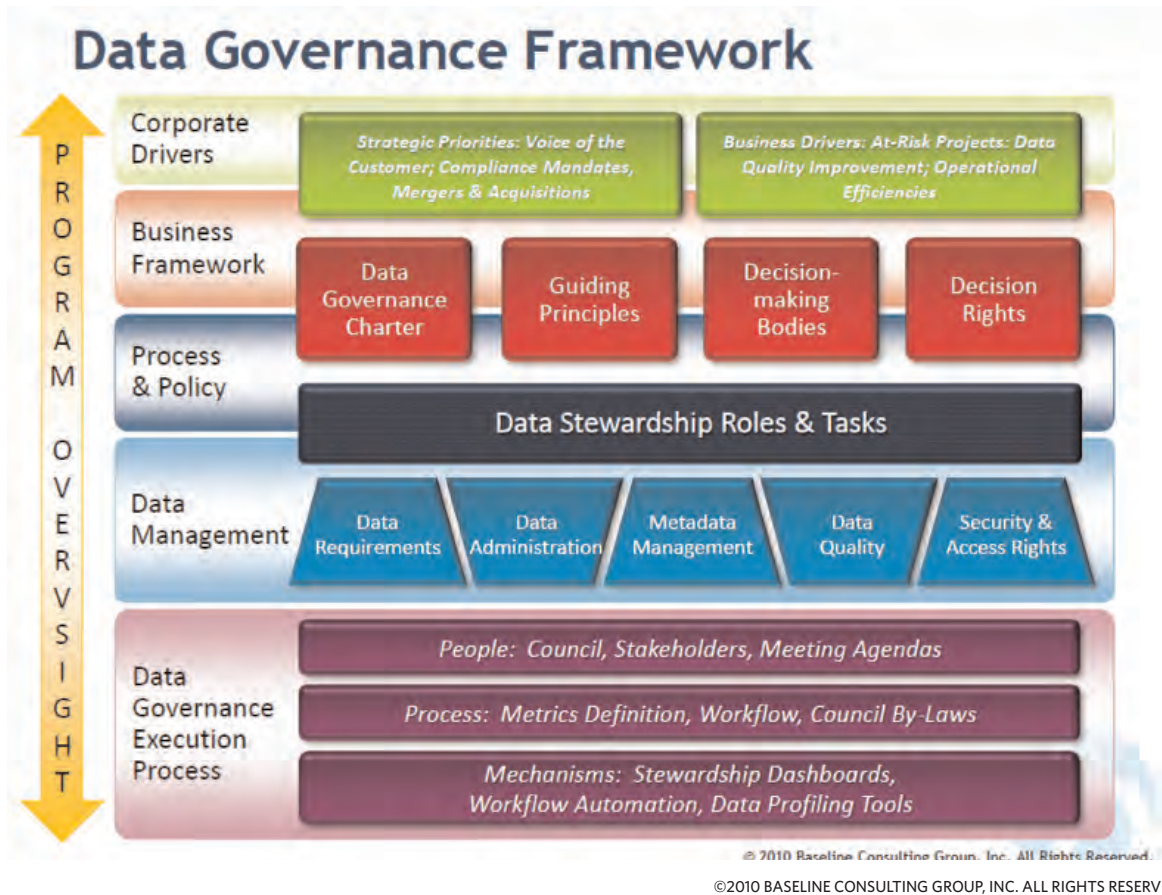
data-oriented projects, such as an enterprise application upgrade or migration, for which there is executive-level buy-in across the company. So the key is to think global, but start local and build on success by expanding the cross-functional project to other initiatives.

"Data governance is an insurance policy—it helps drive the success of any data-enabled business initiative," said Baseline consultant Kimberly Nevala. "It ensures that data is part of the conversation at the outset of a business effort, not at the end, when everyone's pointing fingers about what went wrong."

3. Find the right data stewards. Data stewards manage data on behalf of a business owner who is accountable for quality and outcomes. Data stewards are businesspeople, not IT administrators. They decide which data source to use, define access controls for that data and work collaboratively with other stewards in the business community. They ensure compliance with governance policies and procedures, define business rules and guidelines for managing data elements, ensure the use of standard definitions and rules, assess data quality issues and correct data quality problems if necessary.

"We talk a lot about data stewardship but don't do it in practice," said William McKnight, president of McKnight Consulting Group LLC, a BI and data management consultancy. Data stewards should be good facilitators who also have specialized knowledge of data in a particular area and understand high-level business and data concepts. The role of data steward shouldn't be taken lightly; whether individuals volunteer for or are assigned to the position, their job description should reflect their new responsibilities, and they should be held accountable for the work they're expected to do.

4. Put the fox in the henhouse. There will always be naysayers to a data governance program. These may be mid-level managers whose power base depends on controlling access to data, subject matter experts who are wedded to their data definitions and rules, or IT professionals who don't want businesspeople to dictate data management priorities. Whatever the case, it can sometimes help to put these people on a cross-functional stewardship council to make them part of the solution instead of the problem. If they are open-minded to any extent, this assignment will open their parochial eyes to broader, enterprise concerns, and they will work to serve both the company's local and global interests. It also helps if sponsors provide incentives that shift individuals' focus from their own interests to enterprise needs.

5. Implement a data quality dashboard. Every data governance program needs to publish and monitor a dashboard that contains metrics about data quality. These metrics should form the basis for service-level agreements between the data governance council and business groups. Each metric should have targets associated with it so data owners and stewards can be measured and held accountable for improving data quality. These dashboards should be published openly so everyone can measure their performance relative to their peers. This engenders friendly competition, communicates the idea that data quality is everyone's job and helps improve data quality quickly.

PITFALLS

Many data governance initiatives fail for a variety of reasons. Here are some of the more common pitfalls mentioned by data governance experts:

1. Fail to enforce. The main output of a data governance program is policies and procedures for the definition and handling of data. But unless the policies are enforced and procedures followed, then all the meetings and agreements are for naught. One way to enforce data governance rules is to gain the backing and cooperation of the company's audit and compliance teams, which presumably already enforce other enterprise policies and procedures.

"You don't want to be a toothless tiger," Stiffler said. "If you don't enforce data policies, organizations will quickly backslide." Added John Ladley, president if IMCue Solutions, a data management consultancy and author of *Making Enterprise Information Management Work for Business* (Morgan Kaufmann, 2010), "If no one has ever gotten fired for violating a data policy, you probably aren't enforcing it."

2. Fail to execute. It's also imperative to follow up words with action. The data governance program needs a data management team within the IT department to implement policies defined by the data governance council. (See "Data Management" below.) The data management team comprises various members of the IT team who manage data, including data analysts, data administrators, application developers and project managers. These folks ensure data policies are adhered to and use various tools to automate data management processes. Without a strong partnership between business and IT, data governance programs grind to a halt. One sign of a strong and respect-

ful partnership is that the business adequately funds data management roles, tasks and projects.

3. IT-driven project. A sure sign that a data governance program will fail is when the CEO assigns the chief information officer or director of IT to head the program. This signals that the business considers data a technical task, not a corporate asset. However, it's OK for IT to initiate the conversation about data governance and evangelize its importance to the business. IT can make a business case and try to win converts among executive ranks. Ideally, however, IT should enlist a supportive business executive who has clout among business executives to sell the program. "It doesn't matter who comes up with the idea; the key is to get buy-in from different organizations," Stiffler said.

4. No change management plan. Every data governance program needs to anticipate backlash. Wresting data from the hands of individuals and departments and making it a corporate asset is bound to threaten and infuriate a good portion of the user community, especially mid-level managers whose power base often comes from managing the flow of data upward and downward in the organization. "Resistance is normal," Ladley said.

In his book, Ladley describes many ways the businesspeople exhibit resistance:

› Foot dragging
› Bargaining for exemption
› Retreating to old ways of doing things
› Poor attendance at meetings or increased absenteeism
› Hostile or cynical comments
› Lack of endorsements
› Acts of political sabotage

And he writes that people resist a data governance initiative (or any other issue for that matter) for a variety of reasons:

› Loss of identity, power or prestige
› Personality traits (introvert vs. extrovert, capacity to embrace change)
› They are overloaded with current responsibilities, with no time to learn new skills

> They feel there's no value in changing for them personally ("What's in it for me?")
> No one sought their input to devise the solution or program
> Past initiatives failed

Change management needs to be baked into any data management or governance initiative. Typically, this involves identifying who is going to be affected by the new program and crafting communications, education and incentive strategies to get those employees to adopt new behaviors. It's often wise to seek counsel from your own marketing department to help craft a detailed plan that segments and describes end-user constituencies and devises a communications strategy tailored to each. The plan should define which messages get delivered to which constituencies via which channels at what times.

OTHER PITFALLS
Here are other common pitfalls:
> The data governance program and strategy don't evolve as the business changes.
> The data steward role is not formally recognized in formal job descriptions.
> The business or IT believes a tool is all that's needed to address data quality problems.
> The initial scope of the data governance program is too big or too fraught with risk.
> The business forms a data governance council in response to a data quality crisis and does not spend the time to create a plan and gain cross-functional buy-in.
> The organization tries to force fit a top-down data governance methodology in a bottom-up culture. No one identifies indirect stakeholders affected by a new program, including customers, suppliers, executives. ∎

Data Management Portfolio

DATA MANAGEMENT ENCOMPASSES the execution of data governance policies, procedures and rules. The purpose of data management is to eliminate data silos and increase data reuse through the use of common tools, technologies, architectures, standards and methodologies.

Organizations embarking on an enterprise data strategy need a broad-based portfolio of data management tools to support various data integration and quality tasks and automate processes where possible. As the types and volumes of data expand, the data management portfolio must expand with it.

A decade ago, data architects and administrators focused on managing largely structured data in legacy systems (mainframes, minicomputers) and relational databases running inside corporate data centers in a tightly controlled environment. Today, corporate data is strewn across the landscape, spanning corporate data centers, the public Web and cloud-based services. And there is more data than ever before, including more detailed transaction data, Web clickstream data, social networking data (e.g., Twitter, Facebook and LinkedIn), email, text messages from mobile devices and machine-generated data from an explosion of sensors in everything from automobiles to shipping pallets.

Ten years ago, BI teams loaded data warehouses in batch, usually daily, weekly or monthly, to support standard types of reporting and analysis applications. During the past decade, users have demanded access to more timely information. In fact, many new applications, such as customer self-service and operational dashboards, require delivering information to users on a near-real-time basis—virtually instantaneously—to optimize decision making and actions.

Consequently, the standard portfolio of data management tools no longer

> *As the types and volumes of data expand, the data management portfolio must expand with it.*

suffices in today's modern age. Modeling, managing and moving large volumes of heterogeneous data requires an expanded portfolio of data management tools. To understand the scope of data management tools, we'll step through the primary data management functions defined in Baseline Consulting's data governance framework.[3]

1. Data requirements. Business requirements analysts define the questions that business users want to ask the data. The primary tools here are interviews, surveys and joint application development sessions. Some specialty tools, such as Balanced Insight's Consensus, enable business users to collaborate around the definition of business dimensions, attributes, hierarchies and metrics and even vote on various options. Consensus even generates a star schema and semantic layers for leading BI tools.

2. Data quality. Data stewards, along with data quality administrators, often manage data quality functions, including evaluating source data, defining data quality rules, defining data latency requirements, identifying, standardizing and consolidating duplicate records, matching records from identical households, and monitoring and reporting data quality metrics.

Today, traditional data quality vendors, such as Trillium Software, now offer a full suite of integrated data profiling, data quality, and data cleansing software that works with batch or real-time data feeds and delivers dashboards for tracking data quality. The identity resolution software built into most data quality tools now sits at the heart of most master data management (MDM) systems that logically or physically consolidate records from multiple applications to standardize information about key business entities, such as customers, products, parts, partners, organizational hierarchies and suppliers.

3. Data architecture. Data architects define how data is designed, conceptually, logically and physically, and build programs to map data from any source to any target. For example, data modeling tools, such as Embarcadero's ER/Studio XE, automate routine data modeling tasks and support collaborative development, enabling designers to more quickly analyze, optimize, manage and share database, data warehouse, metadata and process models.

---

[3] See the Baseline white paper, "A Functional Model for Data Management," which details these functions and the roles and responsibilities that support data management: www.baseline-consulting.com.

There has been a huge expansion in the number and diversity of data integration tools that create source-to-target mappings and physically capture and move data between systems. There are a variety of categories of data integration tools, ranging from ETL tools, change data capture and replication tools, to data movement and virtualization tools. (See "Data Integration Portfolios: A Tool for Every Requirement.")

4. Data administration. Data administrators work with data stewards to ensure data is properly formatted, stored and secured. Data administrators

→ # Data Integration Portfolios: A Tool For Every Requirement

SPOTLIGHT

TODAY, DATA MANAGEMENT teams require an extensive portfolio of data integration tools to keep up with a rapidly changing data landscape. With more data coming faster from more and varied sources, it's imperative that data management professionals have the right tool for every occasion. Software vendors have recognized the need and have greatly expanded their own suites of data integration tools, and many startup companies offer novel approaches to accelerate or automate down-and-dirty data management tasks.

For instance, open source data integration vendor Talend started out in 2005 by delivering extract, transform, and load (ETL) functionality but has since expanded to deliver data profiling, data quality, data cleansing, master data management, data synchronization and Web services application integration. Ideally, data rules created in one data integration tool can be reused in all other tools and stored for reuse in future projects, all of which speeds development and time to value for business users.

New data platforms, such as Web- and cloud-based applications (e.g., Salesforce.com, NetSuite, Twitter, YouTube and Google), mobile devices, search tools, and Hadoop and NoSQL databases require new tools to access, integrate and consolidate the data in these systems along with traditional data sources. Rising to the occasion, cloud-based data integration vendor SnapLogic Inc. offers a simple Web services application programming interface (API) based on representational state transfer (REST) that connects applications with a variety of data sources. The company also offers an online

use data profiling, data quality and data cleansing tools as well as data modeling and metadata management tools. Data administrators often work hand in hand with database administrators, who use tools like those from Embarcadero to set up, administer and tune SQL databases.

5. Metadata management. Metadata describes the properties and attrib-

(CONTINUED FROM PAGE 33)

store, where customers can purchase SnapLogic-based data connectors from SnapLogic and third-party developers that work with a wide variety of data sourcesand applications.

Some data integration vendors focus on technology niches. Datawatch Corp., for example, specializes in extracting data from legacy applications, in particular, mainframe greenbar reports, operational reports, PDF reports, and invoices and statements and converting them into interactive Web-based reports or Excel documents. If all the data users need is contained in the legacy reports, Datawatch obviates the need for a data warehouse or operational data store, especially if re-creating the legacy data in a new system would be prohibitively expensive and risky.

Another vendor, Attunity Ltd., has long been known for its connectors to mainframe and legacy applications, but today, the fastest growing part of its business involves delivering change data capture (CDC) and replication services. As the velocity of business increases, organizations want data delivered faster. CDC captures changes from system log files and feeds them to ETL, messaging buses or file transfer protocol folders where they can be loaded into a data warehouse or target system. CDC is a critical technology to support near-real-time data delivery.

Replication has traditionally been used to create a mirror image of data for disaster recovery purposes but is now increasingly used to support active-active architectures as well as dual data warehouses, one for top-down reporting and another for bottom-up, ad hoc analytics.

Data virtualization (or data federation) has recently become an important component within a data integration portfolio. Instead of expending the time and money involved in creating a data warehouse or data mart, data architects use data virtualization tools to enable users to access and join data on the fly from wherever it lies. Although there are limits on the volume of data and complexity of queries that data virtualization tools support, they've become useful tools for BI SWAT teams that need to build applications quickly and create data services that abstract back-end data resources from users and applications. ∎

utes of data, including its origins, lineage, and upstream and downstream dependencies as well as its owner and steward. Effective metadata management accelerates development and time to value, reducing delays and number of staff needed to maintain and manage data resources.

Ideally, developers source metadata from every component in an analytical environment and maintain them in a single-object-oriented repository, such as ASG's Rochade, so they can quickly find the right data elements to use in a project. Metadata includes technical metadata, which is used by data and IT administrators to track lineage and perform impact analysis, and business metadata, which describes data elements in business terms so businesspeople can understand its definition and origins. "Business metadata is often overlooked," Trillium Software's Stiffler said.

Compliance policies can prevent unauthorized access to data and eliminate data defects.

6. Privacy and compliance. Compliance officers and security administrators oversee data privacy, security and compliance policies. According to Baseline, these policies are often put into practice via a discrete data management function responsible for implementing and monitoring controls on data. These controls can prevent unauthorized access to data and eliminate data defects that undermine confidence in publicly reported data. They also establish policies, procedures and systems to prevent the theft or loss of sensitive company data as well as masking confidential customer data that must be shared with internal or third-party developers, archiving data that must be retained and disposing of data that is no longer needed. ■

Summary

DATA STRATEGY WORKFLOW

Creating an enterprise data strategy requires executives to acknowledge that data is a corporate asset that must be managed and protected like any other asset. With that as a starting point, it's imperative that you understand how your organization currently manages data. Does it manage data centrally or invite individuals and departments to create and manage their own data?

Once you identify your organization's data profile, you can begin plotting an appropriate data strategy. Organizations that tightly control data resources need to adopt new processes, technologies and organizational designs that push development back out to departments and individuals in a controlled manner, while organizations with fragmented data need to establish policies and processes that manage data as a shared service, not as the property of any individual or department.

Once an organization has assessed its current state and readiness to implement an enterprise data strategy, it needs to design a business-driven data governance program that specifies the people, processes and technologies that the enterprise will use to manage data as a corporate asset. The program needs to be designed like a business plan for a new company, with a vision, mission, values, guiding principles, roles and responsibilities, and detailed project plans that show how the program will be designed, implemented and sustained over time. The plan should also specify detailed education and communications initiatives to overcome resistance to change and accelerate adoption.

Finally, the IT department needs to take a back seat to the business if the data governance program is to succeed. If the business doesn't assume its rightful role or follow through on its commitments, then IT should not try to plug the gap. Its job is to implement the policies and rules developed by business stewards, not to create them. It needs to create a broad portfolio of data integration tools that can be used to accelerate and automate data management tasks in support of a data governance program and enterprise data strategy.

DATA STRATEGY TIPS

Of course, you can have the best methodology in the world but still miss the mark if you aren't aware of some of the key trends and pitfalls involved in designing and implementing an enterprise data strategy. Here are general guidelines and recommendations for getting your organization to treat data as a corporate asset and avoiding the pitfalls that beset many data governance and data management initiatives.

■ **Sell based on risk.** Although you can justify a data strategy based on cost savings from consolidating redundant, inconsistent reporting systems and reducing duplicate mailings, the best way is to educate executives about the risks that poor-quality data poses to strategic projects, goals, partnerships and decisions. Showing the dependency of mission-critical business initiatives and processes on clean, consistent and timely data is key to selling an enterprise data strategy.

■ **Learn the hard way.** Unfortunately, risk is virtually invisible until something bad happens. That's why most business executives don't perceive data as a critical asset until they've been burned by poor-quality data. Perhaps a well-publicized merger doesn't deliver promised synergies because of a larger-than-anticipated overlap in customers, or perhaps customer churn is increasing, but they have no idea who is churning or why. If you can't convince business executives to invest in data quality, then just wait for a data catastrophe. It will happen!

■ **Project alignment.** Even with a catastrophic data-induced failure, the only way to cultivate data fastidiousness is to do it one project at a time. Data governance for data governance's sake does not work. Businesspeople must have tangible, self-evident reasons to spend time on infrastructure and service issues rather than immediate business outcomes on which they're being measured. The best projects for which to build an enterprise data strategy are cross-functional in nature and data-intensive, such as a major upgrade to a corporate order management system.

■ **Data defaults.** In the absence of strong data governance, data always defaults to the lowest common denominator, which is first, an analyst armed with a spreadsheet, and second, a department head with his own IT staff and data management systems. This is kind of like the law of entropy: It takes a lot

of energy to maintain order and symmetry but very little for it to devolve into randomness. Only an executive-sponsored, business-run data governance program can offset the tendency of data to dissolve into anarchy.

■ **Business driven.** Although IT can evangelize the importance of having an enterprise data strategy, it cannot allow the business to put it in charge of one. Data strategy and governance is not an IT project or program; it's a business program. Executives must recruit business managers and analysts to formulate and oversee a data strategy, define rules, policies and procedures to maintain the accuracy, completeness and timeliness of critical data elements, and partner with IT to execute the program.

■ **Sustainable processes.** The ultimate objective for managing any shared service is to embed its care and tending into business processes that become part of the corporate culture. When this happens, managing data becomes everyone's business, and no one questions why it's done. This is what creates a sustainable process.

■ **Reconciling extremes.** The key to managing data (or any shared services or strategy) is to balance extremes by maintaining a free interplay between polar opposites. A company in which data is a free-for-all needs to impose standard processes to bring order to chaos. On the other hand, a company with a huge backlog of data projects needs to license certain people and groups to bend or break the rules for the benefit of the business.

■ **A touch of chaos.** Instead of trying to beat back data chaos, BI managers should embrace it. Spreadmarts are instantiating of business requirements, so use them to flesh out the enterprise BI and DW environment. "I don't think it's healthy to think that your central BI solution can do it all. The ratio I'm going for is 80% corporate, 20% niche," Dow's Masciandaro said about the newest incarnation of spreadmarts: in-memory visualization tools.

■ **Safety valves.** Another approach to managing chaos is to coopt it. If users threaten to create independent data marts while they wait for the EDW to meet their needs, create a SWAT team to build a temporary application that meets their needs. If they complain about the fast-and-dirty solution, they know there is a better solution in the offing. ■

embarcadero®

- **ER/Studio vs. CA Erwin product comparison: An independent look at today's top data modeling tools**

- **Data Warehousing & BI Resource Center**

- **Case Study Newmont Mining: An ER/Studio Success Story**

About Embarcadero Technologies, Inc.:
Embarcadero Technologies, Inc. is a leading provider of award-winning software tools for application developers and database professionals so they can design systems right, build them faster and run them better, regardless of their platform or programming language. Ninety of the Fortune 100 and an active community of more than three million users worldwide rely on Embarcadero products to increase productivity, reduce costs, simplify change management and compliance, and accelerate innovation. Additionally, Embarcadero's licensing innovations make software management and deployment easier. Founded in 1993, Embarcadero is headquartered in San Francisco, with offices located around the world. Embarcadero is online at www.embarcadero.com.