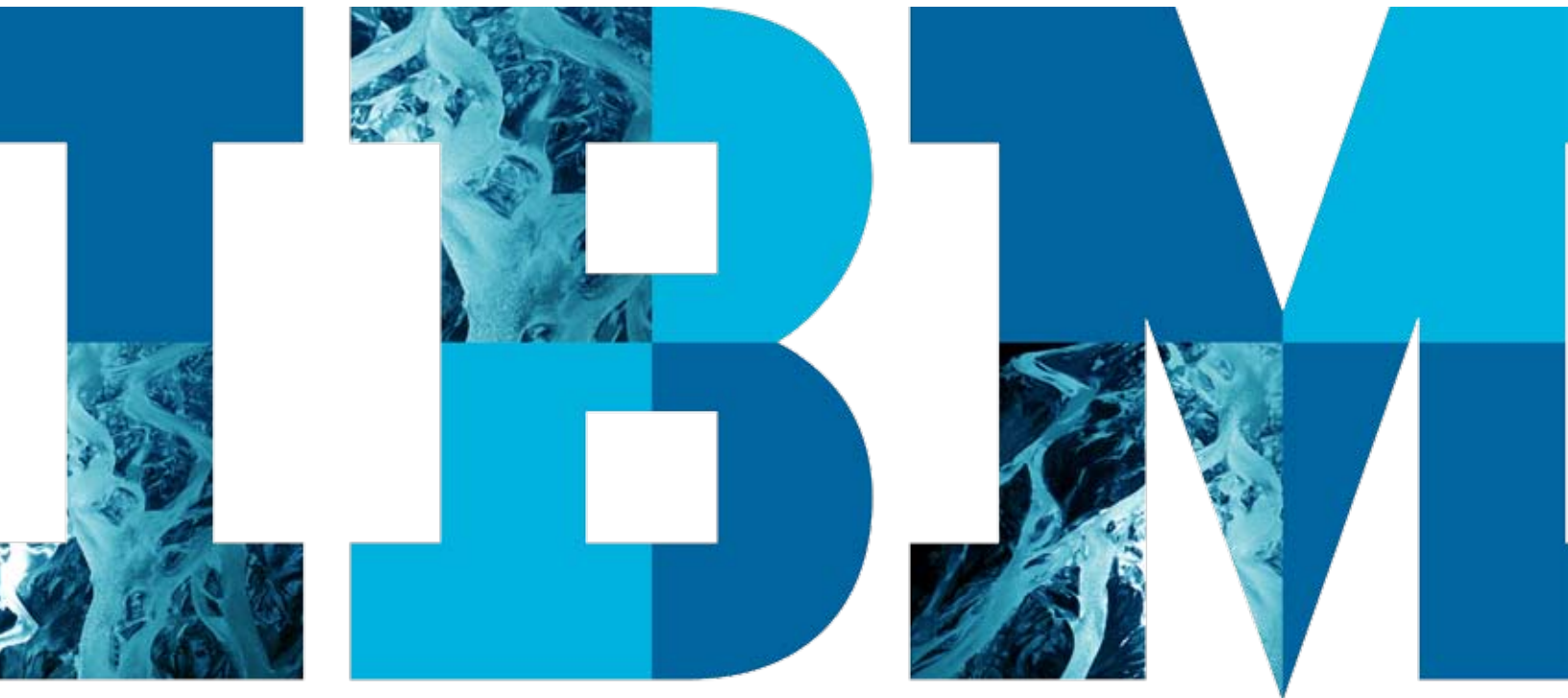


## Six easy steps for smart governance



## Introduction

Consider the modern enterprise and its many sources of internal and external information. Data enters the organization, often with few or no high-level quality controls. From there it is typically spread across multiple silos, with little cross-organizational collaboration. Critical business processes rely on this data, but with few defined governance and stewardship structures the information is often unverified, redundant, incomplete or dangerously out of date. Even small mistakes in upstream data entry can flow downstream through endless data processing rapids before becoming major sources of toxic content and data pollution. The end result: poorly informed decisions that can cripple a business.

In response, many organizations are turning to data governance, which is the activity of coordinating people and processes to ensure that the data they collect, share and act on is accurate and trustworthy. Data governance filters bad data out of the business, creating trusted information for better decision making and lowering risk across the enterprise.

## Building a strong data governance system

A strong data governance program assigns people to “own” business data and ensure that it is accurate and valid. These “data stewards” continually assess data quality based on policies created by data governance councils composed of business and technical team members. They have defined

action plans for dealing with issues that arise during ongoing monitoring (and tools like IBM® InfoSphere™ Information Analyzer can be used to profile the data).

Data governance programs can take a bottom-up or top-down approach. Some have charters and formal structures; others are ad hoc and informal. Many have cross-organizational mandates, while some are departmental and IT-oriented. In other words, there is no “one size fits all” approach to a data governance model and structure.

In fact, the precise shape of the program—a council, a workgroup or a decentralized data stewardship—matters less than simply having a system. Without a systemic approach to governance, your results will be haphazard and inconsistent. A repeatable system is critical, and should offer consistent processes, rules of engagement, supporting technology and documented results.

This paper describes six key steps for a successful data governance program. These steps will help ensure that your data governance decisions and policies are aligned with goals, grounded in facts, communicated effectively, delivered with measurable results and subjected to continuous auditing. You can apply these steps to any kind of governance program, from IT and Service Oriented Architecture (SOA) to security and more.

## Step 1: Set and communicate goals

Data governance goals should be specific, measurable and directly tied to either the success of the business, or to the processes and initiatives that are most likely to help the business succeed. Broadly speaking, organizations should look at setting two types of goals: situational and sustainable.

**Situational goals** are policy-specific goals based on a measurable deficiency in a program as reported by a key performance indicator (KPI). For example, you might have a business process failure that is impacting data quality. Your sustainable goal is to achieve 90 percent data quality across your entire information supply chain, but the KPI is reporting only 80 percent. Your situational goal, then, would be to improve data quality to 90 percent.

**Sustainable goals** are the goals the program expects to achieve. Do you want to clean up metadata? Drive more revenue? Sustainable goals should be directly tied to the business: cut costs by 10 percent, increase sales by 25 percent, reduce the number of customer complaints by 20 percent.

Sustainable goals should be based on a scientific assessment of where your organization is today and where you want it to be tomorrow. A maturity model assessment is a good way to assess your current state and build organizational support for your sustainable goals. The IBM Information Governance Council, for example, created the IBM Data Governance Maturity Model. It identifies five levels of maturity from which any organization can benchmark its current status and define program goals for future attainment (see Figure 1). In a maturity model workshop, each sustainable goal has an associated roadmap for attainment.

Creating situational and sustainable goals is only the first step. Like any leader, the person driving a data governance initiative can't take it for granted that everyone knows and understands these goals. Communicate them often, reminding everyone—even yourself—why your governance program exists and its charter for change.

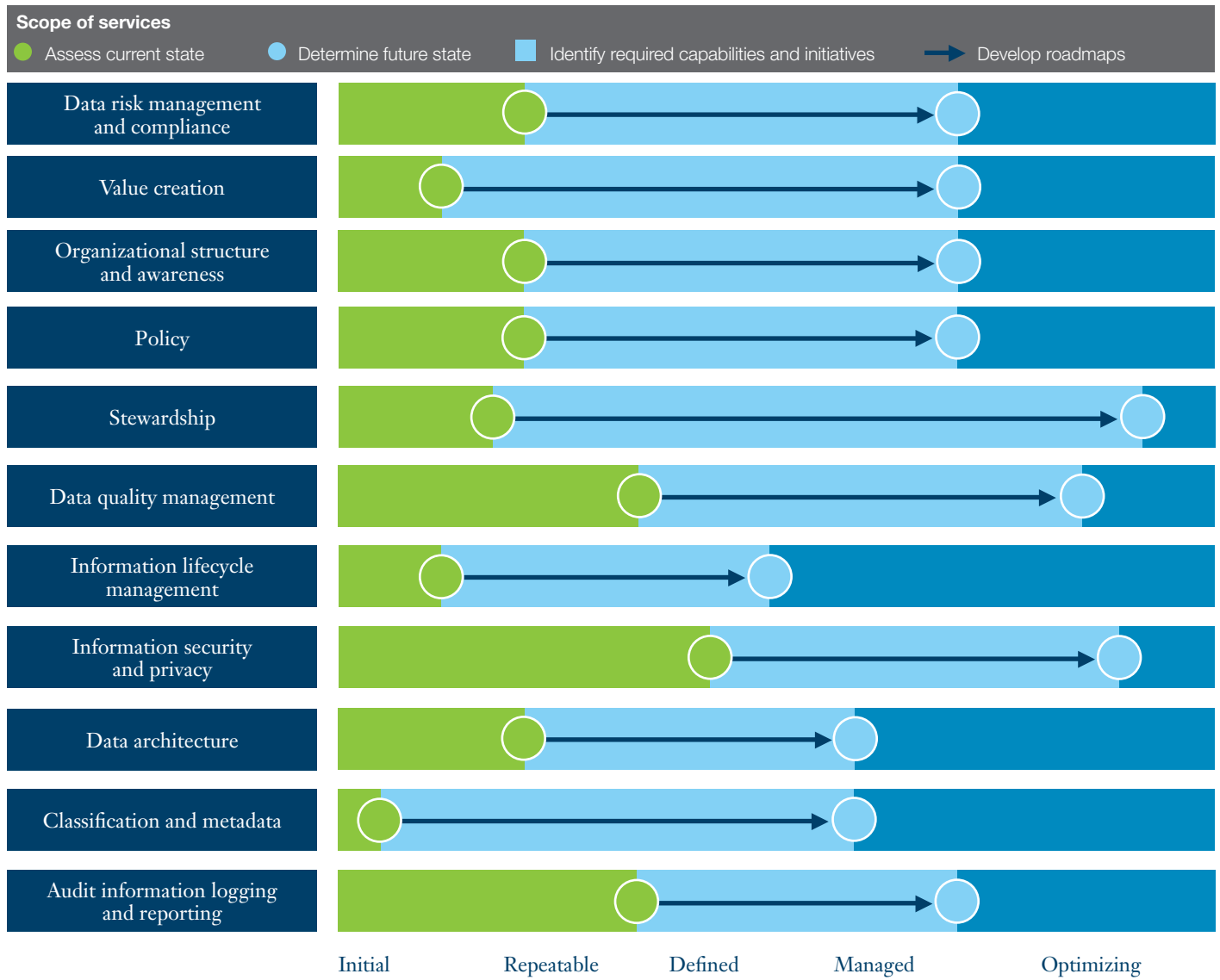


Figure 1: A maturity model assessment can reveal gaps in your data governance program.

---

## Step 2: Define your metrics

How you define situational goals will be based in large part on how you define your metrics and measure success, as well as the kind of information supply chains you are monitoring. Without careful metrics, it is very hard to know if your program is achieving its goals.

A metric should be a real, objective measurement of something, like the impact on the business of better data quality in a particular business process. Not all metrics are created equal—they exist in a distinct hierarchy, with each level building on the one below. At the top of the hierarchy are KPIs, which are synthesized from performance indicators. Performance indicators themselves are built from indicators, and each level is a sum of the correlation of the level below. Those correlations yield meaning, making sense of disparate data points, and that marks the difference between many indicators, several performance indicators and just a handful of KPIs.

**Indicator:** This is the raw data collected from many inflection points in your information supply chain, whether it's a field in a log or the output of a sensor. But indicators are data points without context: like words in a sentence, each has a dictionary definition, but you don't get a complete thought unless you concatenate them into a sentence (see sidebar, "A sentence of indicators").

---

### A sentence of indicators

#### Indicators

INDICATOR 1: Each

INDICATOR 2: part

INDICATOR 3: of

INDICATOR 4: this

INDICATOR 5: sentence

INDICATOR 6: is

INDICATOR 7: an

INDICATOR 8: indicator

INDICATOR 9: .

#### Performance indicator

Each+part+of+this+sentence+is+an+indicator+.

---

Here are a few examples of indicators from the data governance world:

- i. Percentage of paper documents under records management (%PAPERDOCS)
- ii. Percentage of electronic documents under records management (EDOCS)
- iii. Percentage of email under records management (EMAIL)
- iv. Average time in days to turn around e-discovery requests (DISCOVERYTIME)
- v. Total storage (GB) and total cost of storage (USD in thousands) (\$STORCOST)

Are individual indicators interesting? Often they are. But are they relevant? Without more context, there's no way to know.

**Performance indicator:** To continue the grammatical metaphor, a performance indicator is a sentence. It is a series of facts that together yield insight into what's going on. It is the result of mathematical analysis, and the sum is greater than the meaning of the parts. Performance indicators tell you how the things you are measuring are performing.

For example, by combining the example indicators above about records, documents and storage costs, an organization could arrive at a performance indicator:

RECORD RETENTION COSTS = ((%PAPERDOCS x \$STORCOST) + (%EDOCS x \$STORCOST) + (%EMAIL x \$STORCOST))

Performance indicators are almost always interesting. Are they relevant? They can be, but organizations need to go even farther.

**Key performance indicator:** This is often abbreviated as KPI, and it's appropriate that the acronym starts with "K", as KPIs represent *knowledge*. KPIs are the über-analysis. They show the trends of change, as well as the ups and downs of performance indicators as they oscillate over time. An organization should have relatively few KPIs, but everyone in the organization should know what they are and what they're called—and should hotly debate what they mean.

Value at risk (VAR) is an example of a KPI from the banking industry. VAR is a ratio of risk, calculated from tens of thousands of indicators and hundreds of performance indicators. Before the credit crisis, almost all bankers saw regular reports on VAR, knew VAR was very important and made decisions based on VAR. Only one or two people in a particular bank (neither of whom would have been the CEO) understood how VAR is calculated and where the data came from, but that's not important. What's important is that monitoring VAR was critical to understanding the risk performance of a bank.

KPIs in any organization should be like VAR: one number that everyone cares about. At our example organization, record retention costs could be combined with other performance indicators to create a KPI:

**KEY PERFORMANCE INDICATOR:** We are wasting money on redundant storage costs

KPIs are interesting *and* they are relevant. If you have more than five KPIs, most of them are either indicators or performance indicators and no one will remember or care about them. Make sure your KPIs really are key, and that you define your indicators, performance indicators and KPIs to demonstrate why and how you are—or are not—achieving your goals.

### Step 3: Define how decisions will be made

Few organizations devote any time to systemically recording their decision-making processes, but it is important nonetheless. Everyone makes mistakes, and you can't possibly learn from them if you don't keep track of how they were made. Data governance is no exception.

Every decision is a policy. It doesn't matter if one person decides or many. To determine why a policy is or is not working, you need metrics about the decision-making process that you can analyze to plan your next move. You may have a simple problem in one department that a data steward can easily address, or you may need to change business processes, deploy new software or sell new data architecture to the organization. In every case, a decision needs to be made. Who participates in the decision, how the metrics were used to justify the decision and how the information was analyzed are all important key decision indicators (KDIs). Record them, and get good at matching the decision-making model to the scope and scale of the decision.

Also, recognize that it doesn't matter if problems are solved by data stewards, governance workgroups, boards or councils. What matters is the decision-making model: who was included in the decision that addressed the problem and how the decision was made.

There are many types of decision-making models, and each has its own strengths. For example, many organizations start data governance programs with a council. Each council member represents different parts of the organization, and each has a vote. If council decisions are made by majority, unanimity or super-majority, this is a representative decision-making model. In contrast, you might simultaneously use a local-empowerment model for some areas where you have data stewards with delegated authority who can make their own decisions without council consultation.

Some decisions require speed and authority, and in those cases you might elect to use an escalation path. That is a hierarchical-type model, in which decisions are sent right to the top, or require consultation with other groups. However, there are situations when having a crowd participate in decisions creates ownership, which is desirable, even if it makes decision times longer. For example, in Cologne, Germany, the government allows citizens to participate in the budget-making process by submitting their own funding

proposals.<sup>1</sup> It not only drives government agencies to be more responsive to citizens' needs, it also includes citizens in decision-making processes by informing them of the options and letting them decide on municipal priorities.

Organizations also might use social networking solutions to engage the entire organization in crowd-based decision making, using internal stakeholders to build buy-in for enterprise-wide decisions. That is considered a market-based model.

Each of those examples describes a different political decision-making model. Each involves small or large amounts of people and expertise, and each policy decision made has narrow or wide-ranging implications. In a smart data governance system, it's important to be agnostic about your political models and choose the model that best meets your sustainable or situational goals, KPIs and policy needs. By matching your decision-making model to your policy goals, you can take advantage of direct power or indirect expertise, fast or slow institutional support or direct democracy or hybrid approaches.

#### **Step 4: Communicate your policies**

Regardless of which political model you use to set policy, you must communicate the policy effectively to achieve your desired results. Verbal announcements, email and written documents are all examples of policy communication. But

policy also can be communicated by software, using methods such as changes in business glossary definitions, database table structures, encryption or data transformation. Even the acquisition of new software solutions is an example of policy communication (the decision was to buy new software; the communication was in the acquisition and implementation).

The point is that you need to be just as rigorous in your policy communications as you are in documenting the reasons for policy change—as well as your political model for policy decisions. Good policies don't change anything if you communicate them poorly. Communicate your policy decisions in the best way possible so they have an impact. For example, if your policy is to restrict access to sensitive information, which requires changing access control policies in your content repository, then the policy was the decision and the software changes are the communication. Or, if you have a policy to optimize your information supply chain for Sarbanes-Oxley Act compliance reporting, you need to set compliance metrics to measure how your communication tools are succeeding in achieving it.

Remember that to achieve collective goals, all people being governed need to understand your policies, so it's important to measure not only your KPIs and KDIs, but also how well your policies are being communicated.



## Step 5: Measure your outcomes

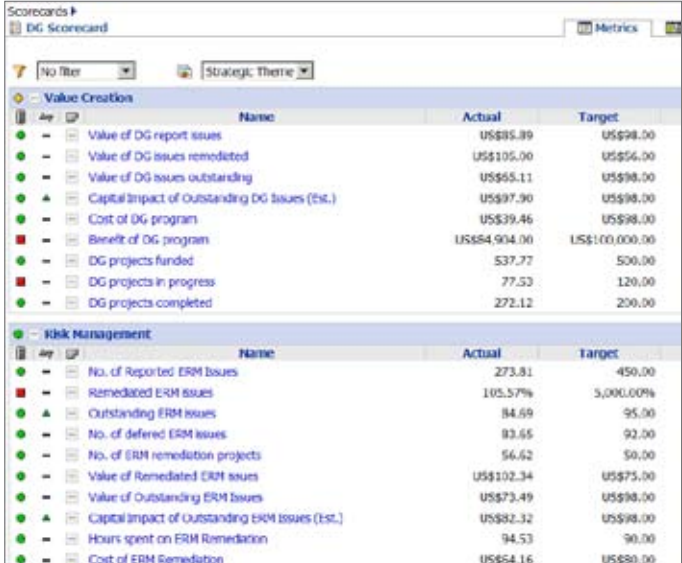
Make no mistake: you govern for outcomes, so you need to measure those, too. Specifically, you must measure how well your policies achieve the sustainable and situational goals of your program.

To measure data quality, you may need to monitor data processing steps, track business and IT process failures and benchmark the impact of systemic failures on data delivery. In fact, just having the right KPIs can change some organizational behavior, once people become aware that their activities are being monitored and their success or failure will be measured.

A recent *New York Times* article about New York Police Department crime statistics illustrates this point. It seems that crime had been falling every year in New York City, even during the last two years when the economy worsened. What's interesting about this is that crime generally rises during a recession. An audit of crime reporting practices revealed that precinct captains were rewarded for reporting lower crime rates. Consequently, they had developed creative practices for reducing crime reporting rates, including persuading victims not to file complaints, or reporting stolen good values based on looking up the values of used goods on eBay instead of reporting victim replacement costs. These practices led to lower reported crime rates, thanks to some new forms of corruption.<sup>2</sup>

A best practice for measuring data quality includes collecting information at defined points in your infrastructure. For example, if you have an information supply chain that delivers financial market data to decision makers, make sure you are measuring data quality at every transformation stage. You don't want business users complaining to you about data quality when it's already too late to change it. You can avoid serious data quality issues by monitoring the critical information supply chains that deliver trusted business information, and then aligning those measurements to KPIs that are reported to your data governance program.

Many forward-thinking companies use the IBM Cognos® Data Governance Scorecard to define and track KPIs across their organizations (see Figure 2). The scorecard maps KPIs to the IBM Data Governance Maturity Model and provides a convenient interface for tracking a program's sustainable goals, current progress and situational requirements.



Value Creation			
	Name	Actual	Target
●	Value of DG report issues	US\$85.89	US\$98.00
●	Value of DG issues remediated	US\$105.00	US\$56.00
●	Value of DG issues outstanding	US\$65.11	US\$98.00
●▲	Capital Impact of Outstanding DG Issues (Est.)	US\$97.90	US\$98.00
●	Cost of DG program	US\$39.46	US\$98.00
■	Benefit of DG program	US\$84,904.00	US\$100,000.00
●	DG projects funded	537.77	500.00
■	DG projects in progress	77.53	120.00
●	DG projects completed	272.12	200.00
Risk Management			
	Name	Actual	Target
●	No. of Reported ERM Issues	273.81	450.00
■	Remediated ERM issues	105.57%	5,000.00%
●▲	Outstanding ERM issues	84.69	95.00
●	No. of defered ERM issues	83.65	92.00
●	No. of ERM remediation projects	56.62	50.00
●	Value of Remediated ERM issues	US\$102.34	US\$75.00
●	Value of Outstanding ERM Issues	US\$73.49	US\$98.00
●▲	Capital Impact of Outstanding ERM Issues (Est.)	US\$82.32	US\$98.00
●	Hours spent on ERM Remediation	94.53	90.00
●	Cost of ERM Remediation	US\$54.16	US\$80.00

Figure 2: The Cognos Data Governance Scorecard displays goals and measurements for KPIs.

Each reported KPI can become the basis for fact-based policy change requests. When you report facts about the problems in your information supply chains, you gain the operational awareness to prevent problems from becoming future disasters—an absolute prerequisite for smart data governance.

You also can use business and data event monitors to capture critical operational information in your information supply chains to populate KPIs. This should be as automatic a process as possible. IBM has solutions that can help audit your infrastructure and report operational changes to the Cognos Data Governance Scorecard.

No policy will ever achieve 100 percent of its intended goals, because every aspect of the policy-making and communication process is performed by human beings (and we aren't perfect). Be happy with incremental success—and measure progress faithfully. That may mean achieving 50 percent of your goals, figuring out what went wrong and then going back to your KPIs and KDIs to see how you can get to 100 percent.

Remember, governance isn't an end or a state; it's a means, a process. Outcomes will always fall short of goals in an absolute sense. And that brings us to the final and most important tool in your data governance program:

### **Step 6: Audit**

Auditing lies at the heart of every aspect of a smart governance program. Without auditing, smart data governance isn't possible. After all, you can't learn from your mistakes if you don't know what they are. Auditing is the key process and technique underlying many of the measurable steps above. But don't just audit yearly or monthly—set up audits all the time. And for each audit record, do some forensic investigation: find out why it happened, and keep a record of that, too. Over time, you'll have a rich operational history of errors and omissions that will help you avoid past mistakes and make better governing decisions in the future.

Every part of the governance process itself needs to be audited, because compromises and mistakes are part of human activity. Be realistic about incremental change, recognizing that metrics, policies and actions will reflect the political realities of your organization. By documenting each step you can look back later to discover what went wrong, learn from it and improve for the future. Put all the steps together, and you have a complete process. Figure 3 shows an example involving a bank that wants to improve data quality in branch operations.

## Drive bad data out and reap business benefits

Governance is a system with the goal of serving the needs of its customers. It doesn't matter who is doing the governing or the type of governance program in play: if there isn't a systemic approach with auditing at the center, and information flowing evenly to many parties, your data governance program will fail. By following these six simple steps, you will have a data governance program that will be open, transparent, accountable and capable of driving bad data out of your organization. What modern enterprise doesn't want that?

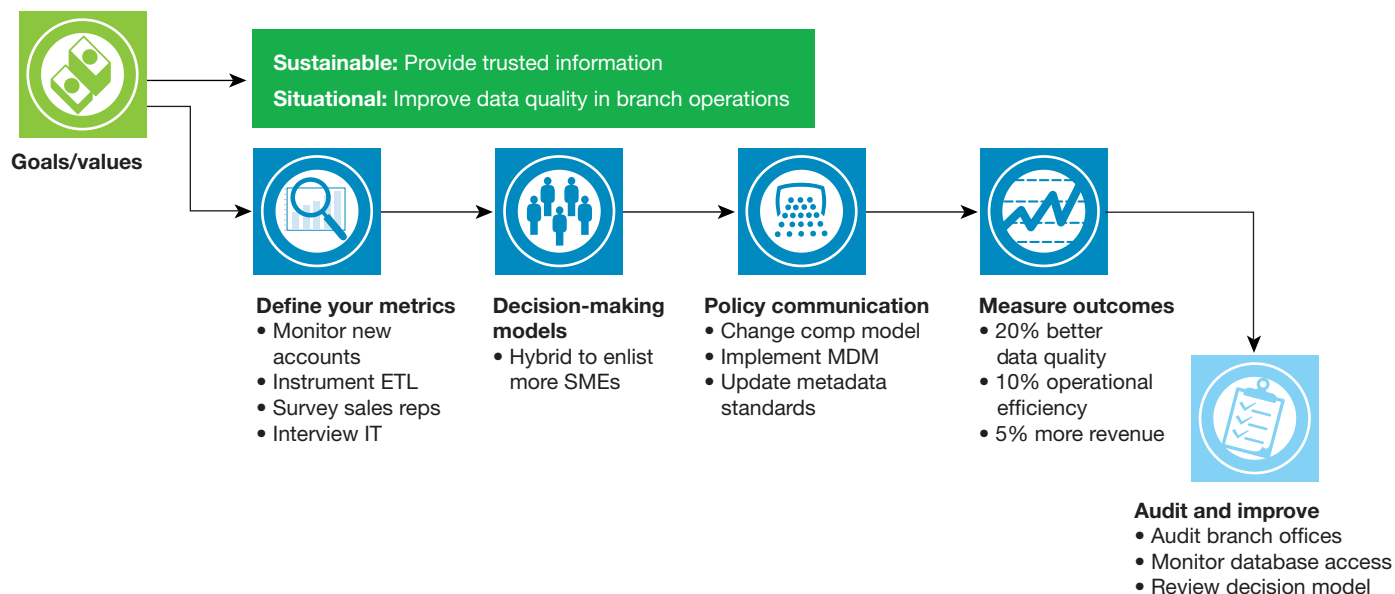


Figure 3: Following the six steps to smart governance, a bank can move from setting its goals through to measuring outcomes and auditing its records.



## For more information

For more information about smart data governance and IBM governance solutions, please visit:

[ibm.com/ibm/servicemanagement/data-governance.html](http://ibm.com/ibm/servicemanagement/data-governance.html)

---

© Copyright IBM Corporation 2010

IBM Software Group  
Route 100  
Somers, NY 10589  
U.S.A.

Produced in the United States of America  
June 2010  
All Rights Reserved

IBM, the IBM logo, ibm.com, Cognos and InfoSphere are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Other company, product or service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. All statements regarding IBM’s future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

- 1 “Cologne—the participatory budget.” [http://s3.amazonaws.com/connected\\_republic/attachments/15/Cologne\\_the\\_participatory\\_budget.pdf](http://s3.amazonaws.com/connected_republic/attachments/15/Cologne_the_participatory_budget.pdf)
- 2 Rashbaum, William K. “Retired Officers Raise Questions on Crime Data.” The New York Times, Feb. 6, 2010. [www.nytimes.com/2010/02/07/nyregion/07crime.html](http://www.nytimes.com/2010/02/07/nyregion/07crime.html)



Please Recycle