

Data Governance: A Necessity in an Integrated Information World

By Danette McGilvray

December 01, 2006, 5:00 a.m. EST



You live in a typical neighborhood. You know most of the people who live there. Every household maintains its home according to its own preferences - some people mow the lawn once a week others work in their yards on a daily basis. Now imagine that everyone on the street is picking their bags and leaving their homes. All the occupants of each house are moving in together!

Each household brings its own ways of living, preferences and attitudes. We can immediately see the potential for conflict. Certainly a different level of coordination and cooperation is required to live together productively and peacefully in the same house than was needed as neighbors in separate dwellings.

Any time your company integrates information, such as in an enterprise resource planning (ERP) application or data warehouse, it is as though each of the source systems - with its associated people, business processes and data - is picking up and moving in together. Companies are living in a world that is much more integrated than in the past. Figure 1 shows the evolution of information integration within a company and approaches to managing information.

eBook MDM for customer data

Understanding your customer helps you deliver relevant, personalized content through the right channels at the right time in the purchasing lifecycle...



Figure 1: Evolution of Information Integration

Our applications and business needs for information are integrated, but our behavior has not changed to work effectively in this world. For example, your company may need information to support end-to-end processes and enterprise decision-making, but the information is being created by an individual contributor from the business who has no visibility to other needs for the same information. The application manager is accountable for system speed and performance, not content accuracy. Both focus on and are rewarded by executing the immediate transaction only. Once again, we see conflict between needs and behavior.

We have different business uses of the information; different platforms, systems, databases and applications; different types of data (customer, vendor, manufacturing, employee and finance); different data structures, definitions and standards; and different data, processes and technology customized to fit a specific business, geography or application. All these differences are supposed to be resolved so everyone can live happily ever after in their new integrated world.

How are decisions to be made in this integrated world? In my house example, each family has its own room in one large home. Occupants of a particular room have the right to put down new flooring and decorate the room the way they want. However, none of the occupants can change the plumbing or redecorate the living room (a common area for all) without the agreement of those who live in the building. In some cases, the occupants could bestow authority upon someone to make the plumbing and common-area decisions. They would entrust that person to make decisions for the benefit of everyone who lives in the building, expect to be informed of changes, and be able to raise any issues that need attention. There needs to be roles, responsibilities, rules and processes in place for managing the house. In other words, governance or formalized accountability is required.

How do we handle the corresponding issues in our information world? The answer is also data governance or formalized accounting. Govern means to "control, regulate or direct something; to have or exercise an influence over and to be the defining rule for something." I consider data governance as a process and structure for formally managing information as a resource. Data governance ensures the appropriate people representing business processes, data and technology are involved in the decisions that affect them. Data governance supplies the structure, roles and processes that provide venues for interaction and communication paths for gathering appropriate input, making decisions, identifying and resolving issues, escalating when necessary, implementing changes and communicating actions.

Data stewardship is often mentioned along with data governance. Data stewardship is an approach to data governance that formalizes accountability for managing information resources on behalf of others and in the best interests of the organization.

Stewardship is an important concept within data governance. I promote stewardship over ownership when it comes to data. A steward is someone who manages something on behalf of someone else. I do not promote the use of the word "ownership" when it comes to data and information. Why? Too often people act as though the data belongs to them and not anyone else. This is a nonproductive attitude. Managing that information for one specific use when it actually impacts many other uses across the company can be dangerous.

I promote the use of the word ownership when it comes to business processes. Why? Because in this case it means to acknowledge full personal responsibility, and those with the authority own the business processes in that sense. But even though the business may own a process, anyone who touches the data in the course of carrying out that process is a steward of the data. That is, they have to manage the data to meet not just their own immediate needs, but manage it on behalf of others in the company who also use that data or information. Whatever word you use in your company - steward, owner, custodian, trustee, etc. - be sure it means both acknowledging personal responsibility and the responsibility of managing that data on behalf of others and in the best interests of the organization.

As a guiding principle for any data governance program, information is owned by the enterprise and is maintained in various company systems. In addition, customers, vendors and employees retain certain rights to their own information. Various organizations, teams and individuals in the company are stewards of the information. They have the responsibility to manage the enterprise's data effectively on behalf of others and in a fair, lawful and honest manner.

In addition to the need for data governance due to integrated information, the biggest motivation for companies to institute data governance comes from regulatory and legal requirements. A few of the U.S. requirements stem from The Sarbanes-Oxley Act of 2002, the Data Quality Act, The National Data Privacy Law and more.

The situation and the need for data governance have been presented. Let's assume you are convinced and want to know how to begin.

Getting Started

Philosophy and Principles. You may be surprised, but the first question to answer is, "What is your philosophy?" You have a philosophy whether you know it or not. Your philosophy is that set of precepts, beliefs, principles or aims underlying your conduct. You will be more effective if you consciously understand your philosophy because it is driving your behavior and what you put into practice. The philosophy of your data governance program should be reflected in your principles - sometimes called guiding principles because these principles guide behavior and determine what is put into practice. The policies, plans, programs, projects, procedures and processes that are put into place will reflect those principles.

Write down your guiding principles as they relate to data governance. Include your philosophy about information and information quality, as this is one of the important reasons for instituting data governance. Don't spend too much time on this, but spend enough time to provide direction to your data governance efforts.

Here are a few examples of guiding principles:

- Information is essential to achieving our company's business objectives. It is a business asset and resource and will be managed as a resource with value - just as our company's products, inventory, facilities, and human and financial resources have value.
- As with other company resources, information requires formalized accountability for managing information resources on behalf of others and for the best interests of the organization. Formalized accountability will be accomplished through the data governance and stewardship program.
- Information quality is the result of applying appropriate business management principles and information management processes, methods, tools and best practices. Data governance and stewardship is a best practice essential to achieving high quality information.
- Information quality is essential to meeting regulatory requirements and lowering the risk of incurring penalties for noncompliance.

If the people you need involved in data governance do not place importance on the guiding principles, they will not take action or be involved in your efforts.

Motivation - The Why. Let me add a few more comments to the earlier statements on why you need data governance. This is where you need to get specific for your situation. Are you considering data governance and stewardship just because it is getting a lot of press lately and it sounds like the right thing? Well, that is just barely a start and won't get you support and commitment. Figure out what is happening in your company and why data governance can help solve the problem. Here are two examples:

You are migrating from several source systems to an ERP system. The ERP project team is working to identify requirements, look at source data and discover discrepancies between the sources. Developers are writing transformation programs, data is being cleansed in the source systems where appropriate, and other required data is being created that does not currently exist in any source system. Everything is being done that can be done to cleanse, transform, prepare and migrate data to meet business and ERP system requirements. What is going to happen after the system goes into production? What happens after go-live? The project team eventually dissolves. Who is responsible for the ongoing quality of the data? The company has invested millions of dollars in implementation. Doesn't it make sense to protect that investment by ensuring the right people representing the business processes, data expertise and system/application knowledge have input to decisions that impact their areas of responsibility in the ERP? Remember, everyone is living in a different world. Different ways of thinking, coordination and cooperation are required.

Another situation. Your data warehouse has been in production for about a year. Unfortunately, it is not being used as much as planned. The data is there to support many business needs, but the users do not trust the data and are not willing to depend on it for their business processes. Here is another example where data governance - with the appropriate roles, responsibilities, communication and escalation paths between the source systems and the data warehouse - is important to improving the quality of the data and increasing the trust in, use of and return on investment from the data warehouse.

Even with the examples stated previously, it is not yet enough to provide reasons for supporting a data governance effort. You need to be able to take it another step forward and answer the question, "Why is data quality important to our company?" You need to answer the questions - "What impact does poor data quality have on my business?" and "Why does it matter?" You need to be able to show specific examples in your company of where changes to data in one area affect another and impact the business.

For example, one ERP team made a change to the product line data element in the product hierarchy. They did not know this data was used by downstream applications outside of ERP. The entire product hierarchy had to be deleted and reloaded. If the issue had not been identified and resolved, the company's sales force would not be correctly compensated. It took at least 30 person days of just category rework to fix the problem. Having a governance structure in place would have prevented this issue.

What are your circumstances? Why does your company need data governance? Write the reasons down, talk to others and keep them available as you gain support to start, design and implement data governance and stewardship.

Scope and Roadmap. While I have been talking about governance as an enterprise-wide effort, it is often more practical to implement governance first within a company initiative such as an ERP migration or data warehouse effort. Then, leverage from what you have developed, and scale governance across the enterprise. It is also possible that upon further investigation, you find multiple governance projects underway in your company that need to be coordinated. In all cases, it is important to understand the scope of your effort and have a long-term plan and timeline, or roadmap, for achieving those results.

Our discussion of data governance and stewardship will continue in my January 2007 DM Review article where I focus on structure, roles and responsibilities, finding appropriate representation, and key considerations when filling roles.

Danette McGilvray