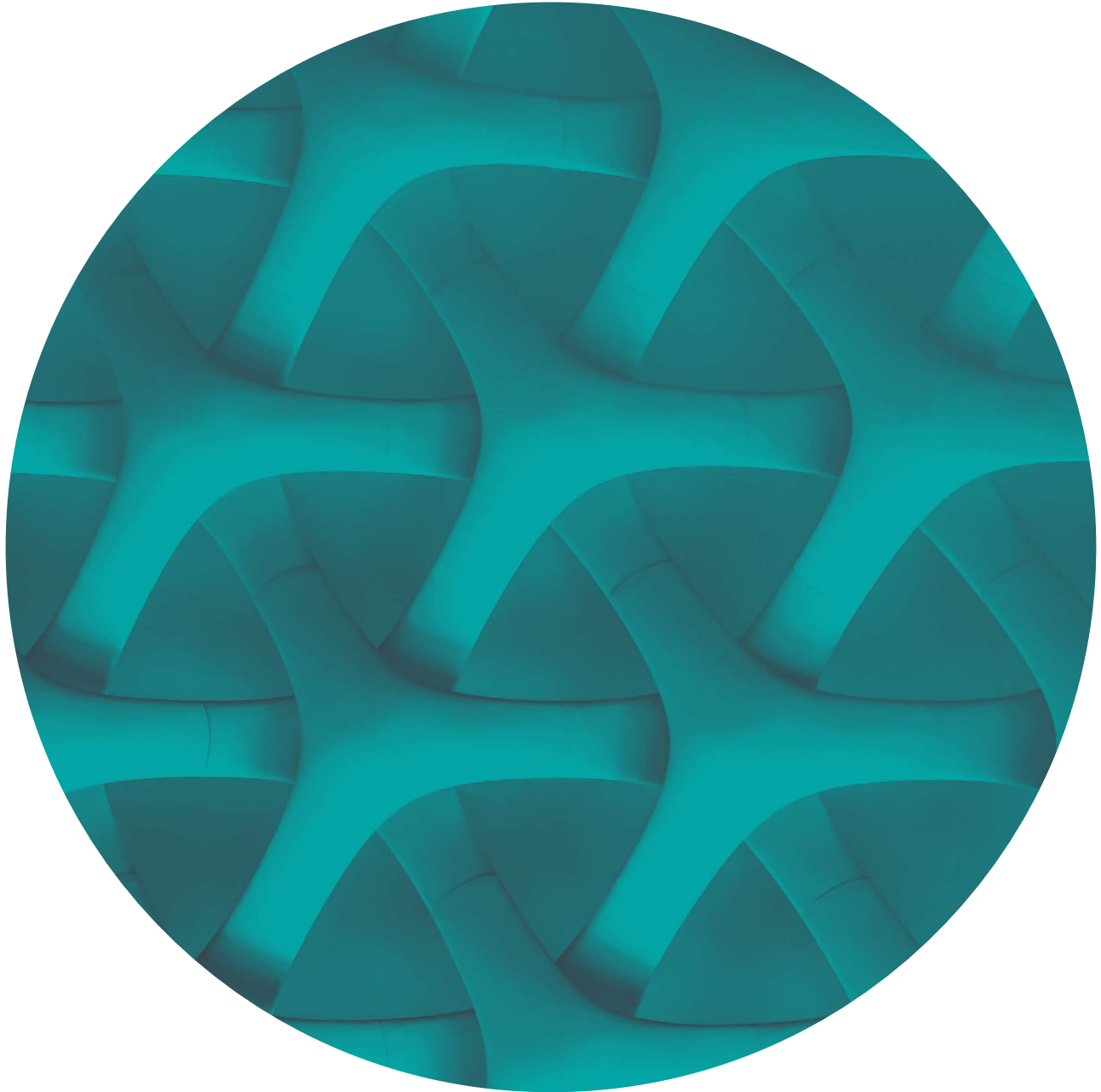


WHITE PAPER

A Strategic Approach to Enterprise Data Management

Step-by-Step



By Kevin Lewis, Data Strategy and Architecture

06.20 / DATA ANALYTICS / WHITE PAPER

teradata.

Table of Contents

- 2 Introduction
- 3 Symptoms
- 4 Causes
- 5 Strategies
- 7 Steps
- 9 Enterprise Data Management—
Core Capabilities
- 11 Conclusion: Intelligent
Strategy Development
- 11 About Teradata

Introduction

Implementing a successful and valuable Enterprise Data Management (EDM) program requires more than picking the right technologies and installing and configuring them correctly. It requires consideration of the people, processes, *and* technologies involved. Data management professionals must apply pragmatic strategies to ensure that the program meets the immediate needs of the business as quickly as possible and adds additional value with each new project. These incremental steps must not only deliver the intended business results, but contribute to an integrated and trusted data resource with each delivery. In addition, just as the solution should be deployed incrementally, in practical steps, the EDM program should be matured incrementally over time. This way, with each new deployment of data, the organization becomes more and more effective at delivery, operations, and analytics.

In this paper, we outline common symptoms leading organizations to consider an enterprise approach to data management, the typical root causes of those symptoms, strategies that underpin a successful enterprise data management approach, and four specific steps any organization can take to initiate and implement a successful EDM program. Finally, we briefly describe the key practices within enterprise data management. The emphasis of this paper is the strategic approach. In our experience, EDM programs typically fail not because of a lack of understanding of the elements of data management—metadata management, data quality management, and so on—they fail because the overall strategy is not adequately aligned to business initiatives or to the broader goals and operating model of the enterprise.



Symptoms

Many large organizations experience similar symptoms leading them toward Enterprise Data Management. Although the primary driver may differ, common symptoms include:

Excessive time spent finding, assembling, and analyzing data

Business analysts often report spending approximately 80 percent of their time finding data and 20 percent making use of the data for business purposes. Also, many end users have developed extraordinary skills in technologies such as Microsoft Excel, SQL, and other tools to collect, assemble, and assess the quality of the data. While these are impressive capabilities, this time would be better spent analyzing the data and looking for business insights that lead to valuable actions, or responding to insights sensed and proactively communicated through the analytic environment itself.

Inability to easily answer cross-domain business questions

Many enterprises have developed analytic systems haphazardly over time. Some of these “data marts” are simply copies of operational databases. Others are formatted more appropriately for analysis, but still hold a limited set of data domains.

For example, assembling a vendor scorecard might require collecting data about sales, on-time delivery rates, damages, inventory balances, transportation costs, invoice match rates, and other information. The data for these needs is stored across many systems and often requires Herculean efforts by end users to combine and curate the data for use. This effort is usually invisible to executives who only see a shiny, well-formatted briefing book after the hard work of collecting the data is done.

Redundant efforts collecting, managing, and distributing similar data

Continuing with the vendor scorecard example, not only is it necessary to assemble the needed data, but assembling this data only for the vendor scorecard misses significant opportunities to use the same data for many other purposes. For example, a supply chain management program might require much of the same data that the vendor scorecard requires. It could be used to understand overall costs of the supply chain, issues related to specific products throughout the distribution network, and so on. Decision makers involved in a customer journey program may wish to offer deals on excess inventory to their best customers. The number of times companies collect the same data from the same sources for different purposes is nothing short of stunning.

Redundant entry and maintenance of data

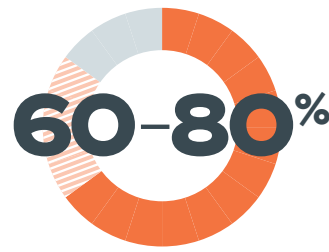
Customers often call a customer service department and at some point in the conversation ask, “Don’t you have that information about me already?” Customers don’t understand why they have to provide the same information multiple times to different people within the same enterprise. Sometimes this is due to acquisitions and mergers and the resulting redundancies, but customers will only accept that excuse for so long. Some companies that have never acquired or merged with another company have the same problem simply because systems are disconnected across department and business units. And of course, the redundancy applies not only to customer data, but also to product, vendor, organization, and other data domains.

Excessive IT costs and complexity, including development and maintenance

Collecting and managing data takes time and cost. In fact, it has been regularly estimated that data integration consumes approximately 60 percent to 80 percent of a data and analytics project. Much of this time and expense are duplicated unnecessarily. In addition, the quality of data is suspect because each point of entry and integration present opportunities to translate the data slightly differently. Each project must develop its own mechanisms for checking data quality and providing descriptive definitions to end users (metadata), along with other services beyond simply collecting the data. And when organizations take a step back and survey the total set of systems—for efforts such as disaster recovery planning or cloud migration for example—they realize that they have built a complicated house of cards that must somehow be replicated in or moved to another location without tumbling down.

Difficulty adhering to regulations and internal policy

Because of the complexity of the overall environment, it becomes difficult to conform to regulations and policies related to certain data domains. For example, personally identifiable information (PII) and protected health information (PHI) are subject to both external regulation and company policy. When this data, such as customer or employee data, is scattered across many systems, it is much more likely to be compromised by internal or external threats. Stories of breaches, fines, and degradation of customer trust are becoming commonplace in the business and mainstream press.



◀ **estimated
data integration
consumption of a
data analytics project**

Causes

There are a number of common root causes of the challenges associated with data proliferation, regardless of the industry or organization:

Numerous independent projects with limited sharing and coordination.

Individual projects are designed to meet a specific set of business objectives. Hence, solutions are designed to meet those objectives as efficiently and quickly as possible. Although corners are often cut and there are challenges in the project, eventually a solution is delivered that meets business needs. However, when this is repeated hundreds of times for different business objectives to serve different business units and departments, the result is a proliferation of solutions connected in a complex web of integration points. Data is one among many casualties of this redundancy and complexity.

Lack of strategic planning and foresight

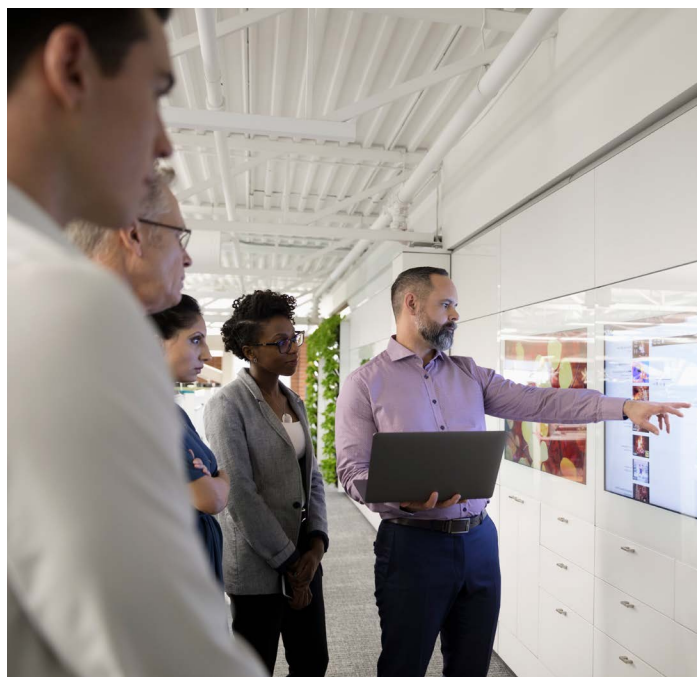
Consider the following complaint from a database administrator: “I’m tired of application teams bringing me a data model that I’m supposed to just implement, with no time at all for optimization or rationalization. How do we fix this?” Although the DBA typically expects an answer related to the development process, handoffs, or enforcement mechanisms (all important topics) the real root of the problem is planning and foresight. The data management team should be aware of the data needs of various projects well in advance of implementation. Most companies have implemented strategic planning processes to develop roadmaps for business unit and departmental solutions, but are often missing the chance to proactively find the shared data opportunities. This also leads to difficulty planning capacity, staffing, and stewardship, and ensuring that data management functions are in place when needed.

Parochial thinking (and funding)

Departmental funding leads to departmental solutions. It is not reasonable to expect a customer journey project funded by marketing to also meet the needs of finance/accounting and loss prevention for their initiatives. It is sometimes possible to appeal to departmental executives and their concern for the “greater good,” but relying on force of personality to sell the idea is hit-and-miss at best.

Skepticism that an enterprise approach is possible

Many business leaders believe in the idea of shared information and a rational approach to delivering solutions, but they believe execution is impractical. Questions are raised such as, “How is it possible to build a shared data structure to serve the needs of many projects when you don’t know all the detailed requirements of even the known projects, much less the unknown projects?” This skepticism is compounded by previous failed attempts to build structures that allow data sharing across applications and projects. The organization is often unaware of principles that can be applied so these objections can be overcome. These challenges have been dealt with many times in real life.



Strategies

As a result of experience with numerous large enterprises in every industry, Teradata has found strategies that ensure the EDM program focuses on delivering real business value, gets buy-in from executives, and that shared data structures stay integrated as new data domains and capabilities are introduced. The strategies are to:

Leverage work that will happen with or without an EDM program.

Executives often see a proposal of shared data strategies, such as an enterprise data warehouse (EDW), an enterprise service bus (ESB), a data lake, and other proposals, as just more work and expense on top of many important initiatives important initiatives on their desks waiting for approval. They are not only concerned with the funding required, but also concerned that these efforts could be a distraction from successfully delivering the most important business initiatives of the company. When data management professionals present the inherent value in rationalizing data and data management practices across the organization, they are frustrated when this garners little excitement. Even a careful and accurate return on investment (ROI) calculation often does not get the expected response. However, if the proposal begins by pointing out how EDM will help deliver all of those other initiatives that are already expected by the business, the rest of the message will be accepted more readily. For example, if an insurance company has projects on the books related to risk management, cross-selling, fraud detection, and customer care improvements—all requiring similar data—then the message should state how the EDM program will help make those efforts more successful while also providing significant cost sharing opportunities. It can also show that while helping make projects successful, the resulting integrated data will be available for analysis across more and more data domains, and the data can be used in new and sometimes unpredictable ways to add value.

Link the EDM program to corporate strategies and philosophies

Walking through corporate offices, it is usually easy to identify the philosophies that are expected to permeate all aspects of the organization. You only have to look for the posters. Different companies use different names, but many are based on common industry trends and practices that sweep through the business world every few years. For example, manufacturers and other organizations have implemented Six Sigma or Lean Management programs for improvements not only in their manufacturing processes, but business processes as well. Many enterprises want to leverage artificial intelligence and machine learning to automate and augment organizational capabilities wherever possible. Because most corporate initiatives and philosophies have a strong information component, data management leaders should make sure that any message promoting the EDM clearly shows how it is an example and an enabler of the corporate philosophy.

Link data management capabilities to the business-oriented projects and programs that will benefit

We often hear some variation of the statement: “We are planning on implementing data quality management by June of this year. We have heard many complaints regarding data quality in this company, and we are very committed to this project.” This limited view should raise an alarm. What issues? Which issues are important? What is the business impact? And, most importantly, which high priority business initiatives will benefit? Contrast that with this approach: “We are implementing a data profiling practice in June as part of our data quality initiative. We are going to start by profiling customer data as part of the customer journey initiative. The customer experience program leadership understands the value and is counting on us. There are a number of customer attributes important to segmentation, and we want to know the data quality gaps ahead of time so we can take action and manage risk. We’ve assigned a customer data steward from marketing who will help interpret profiling results from a business perspective. This profiling practice will be the first step in a broader approach to data quality management and data stewardship. Let me explain other aspects of data quality management we are implementing and talk about other projects that will benefit...”

The EDM program will help make those efforts more successful while also providing significant cost sharing opportunities.

Without having any dependency between the capability (data quality management in this example) and a real project or initiative (customer journey in this case), it really doesn’t matter how long the data quality initiative is delayed or if it disappears altogether.

Link EDM to other IT strategies to increase political will and ensure an “enterprise fit”

To succeed with EDM, it should become a normal, natural part of the way the organization does business. Therefore, it is useful to consider functions outside of the EDM program and think of ways to partner for mutual benefit while embedding EDM concepts directly into other processes. For example, as stated earlier, data is not the only victim of the siloed approach to solution development and deployment. The enterprise architecture function arose primarily because there had been a splintering of business processes, applications, infrastructure, and data. Governing bodies such as Architecture Review Boards have been created within organizations to promote the use of shared resources and standards. EDM concepts and constructs should be included in this process. Solution architects who represent enterprise architecture within individual projects should be aware of the role of the EDM structures and standards and help projects leverage them as part of solutions. There are many other functions that can be leveraged for planning and execution of EDM, such as embedding data management functions into the Solution Development Life Cycle (SDLC) methodology, or including a review of common data needs as projects are proposed for funding, preferably as a part of a broader enterprise architecture review of project proposals. If there are program governance structures in place, they should be used to govern inter-dependent projects originating from the EDM program, with appropriate links and dependencies to other programs. And to create the right business alignment from the outset, a chief data officer or similar role should be in place to link the data strategy with the company strategy so they can be created, evolved, and implemented together in a mutually supportive relationship.

Steps

While the specific priorities for an organization will depend to a large degree on the current state of the enterprise, the following steps can be translated into tangible actions for any organization that wants to move from a stovepipe approach to data deployment to a pragmatic, enterprise-wide strategy. Note that although these steps are listed sequentially, the timing can and should overlap to some degree.

STEP 1. Identify top business initiatives to support and recruit executive sponsorship

It is often stated that executive support is the most important ingredient for a successful EDM program (or any major program for that matter). But support is about more than just getting the right person to say “yes.” It’s also more than securing funding.

In this step, it is critical to connect the EDM strategy to what matters to the business. This will help gain executive support and, more importantly, sustain support even in the absence of any one individual down the road. To gain sponsorship:

- Review top company initiatives and show how EDM can help.** Remember, in this step we are developing a message to gain executive support, not building a detailed roadmap (yet). Start by reviewing the known projects in the portfolio. Also review functional business area roadmaps and have conversations with IT/business liaisons. This is like playing the card game “Concentration” where you are looking for things that are alike, at least from a data perspective, even if proposed projects have very different objectives. Even with limited information, you can predict which projects will require data from the same sources. You may find projects with objectives like labor scheduling, automated inventory replenishment, and pricing that all require a different view of the same sales and product data. You don’t need the detailed requirements to make this connection. Also, connect the program to core company philosophies and strategies. For example, you can show how a data and analytics environment can serve as a repository of metrics required for the Balanced Scorecard initiative, or how data quality management is another instance of Six Sigma, applied to data.
- Find examples of previous challenges.** These challenges can include excessive costs due to deploying the same data multiple times and an inadequate focus on good data management practices. For example, it should not be difficult to point to a series of databases that are so complex, they can only be navigated by highly skilled developers and end users who have built up this capability over many years. Look for previous projects to point out what synergies could have been realized if the analytical data had been deployed for them as a shared service, with shared data management capabilities (e.g., data quality management, metadata management, master data management). Also, find customer experience and other issues related to multiple points of entry and disconnected data in the transactional environment. Be careful here as you are not trying to place blame. You’re only trying to show real-world examples of where a more pragmatic EDM strategy would have helped and, if nothing changes, how these challenges will only get worse.
- Partner with at least one executive who is passionate about using data effectively.** With the information gathered thus far, meet with an executive who has an enterprise mindset (finding one with a lot of stock in the company doesn’t hurt). At a minimum, this person should care about the success of projects across multiple departments, and perhaps understand the need to analyze data across domains. If you clearly show how EDM will directly support multiple top business initiatives, this meeting will go very well. If you only talk about the wonders and benefits of data management in general, or you talk about value unrelated to top business initiatives, you will probably have a different experience. After this meeting, tweak the message based on feedback, and get the executive to agree to attend a future meeting to help gain broader support.

This is like playing the card game “Concentration” where you are looking for things that are alike, at least from a data perspective, even if proposed projects have very different objectives.

- **Identify a small group of representative business leaders to gain wider support.** Using the same message and a proposal for a set of recommended next steps (see steps 2 through 4 for the next steps), meet with a wider group of business leaders to explain the strategy. These should be leaders who are closely associated with the business objectives of the projects and programs that will get the benefit. It should become evident that the data being deployed by department A in its isolated project would come in awfully handy for the later initiative that department B would like to do. This way, you are setting it up so that this group will help with “enforcement.” If the EDM strategy begins to fizzle out during implementation by splintering or becoming parochial, there should be some upset business leaders because they won’t get what they need. They will help provide the political will to keep it on track.

STEP 2. Identify data and analytics to support initiatives.

In this step, you will drive deeper into the connections you made in step 1. To understand what goes into the roadmap, there will be two major inputs. The first, as discussed in step 1, is to understand the common data needs from the projects already planned, including any functional roadmaps being developed.


The second input should be a set of interviews to further understand the data needs of projects, but also to understand general analytical needs. Interview end user analysts, decision makers, power users, and others at all levels of the organization to understand how people access information, what information they would like to have, what data issues they encounter, and the actions they would take and business value that would result if they had easy access to the information they needed—but be sure to understand which needs are associated with top business initiatives and which are “wish list” items.

Especially look for common data needs and analysis requiring information across domains (from multiple sources). This roadmap should be documented at a summarized level for executives and a detailed level to serve as a program plan. The links should be clear between business initiatives, their information needs, applications to deliver the analytics, systems needed to deploy technologies, and enablement capabilities.

Depending on what is being implemented, there may be more than one roadmap. There may be other roadmaps for building out other structures, such as a mechanism for operational data integration and the associated “canonical” message structures. Note that these roadmaps do not create a harsh dividing line between operational systems and analytical systems, because analytic data is often used in the operational environment. Also, the analytic ecosystem can be a recipient of interfaces created in the transactional environment.

STEP 3. Identify data management capabilities and readiness

This step can overlap significantly with step 2. Here you will plan to build or enhance the organization’s “enabling” capabilities to manage and distribute data effectively. Capabilities include functions such as data governance and stewardship, data quality management, metadata management, security and privacy, master data management, data architecture, and others. The first activity is to assess the current state capabilities. All organizations perform these functions at some level, even if they use different names or no names at all.



There are many dimensions to consider when initiating or improving an enterprise data management (EDM) program and the specific initiatives within the program

For example, when collecting data from a source, there is some set of quality checks, even if it is implemented as a band-aid to fix a set of critical issues. Also, there is usually at least some business involvement in data management activities, even if it is not called “data stewardship.”

Next, prioritize the activities based on the gaps found, comparison to best practices, and the needs of the roadmap. For example, if it is known that customer data will be deployed to the data and analytics environment for a customer experience initiative, then security and privacy will take on added importance. If that customer data is known to be in multiple systems that must be reconciled, then master data management will also be a priority. In this way, you are not implementing capabilities for their own sake; you are implementing them because they are needed by important projects. At this point, you are ready to create a roadmap that links the capabilities and data to be deployed to the projects dependent upon them.

STEP 4. Implement the roadmap.

The roadmap should now be translated into a program plan, to be managed like any program, including actively managing the dependencies with other programs. If there is a program governance process in place (usually part of a project management office), it should be leveraged. With each data deployment project, you will also incrementally grow and mature the organization’s capabilities.

For example, if the first data and analytics project is to integrate sales and customer data, then you should ensure that there are data stewards and associated networks assigned to these data domains. As they are assigned, you will have an established set of responsibilities that will also increase over time. Will the sales and customer data stewards establish data quality rules for ongoing monitoring? Are the mechanisms in place to do this, or will they be in place in time for the project? This is all part of implementing the roadmap.

Architecture and design principles should be followed so that new and unknown future requirements might necessitate additional work, but won’t need significant rework (modeling data at the lowest level of detail, creating right-time and adjustable data loads, implementing scalable infrastructure, etc.)

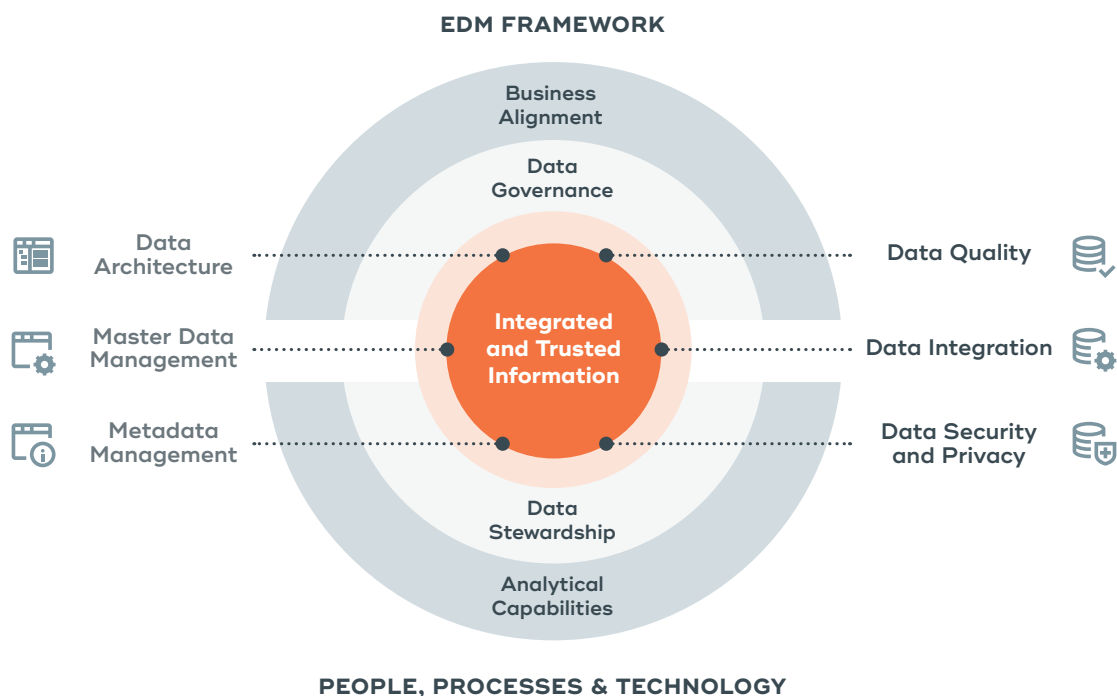
As each new capability is established or matured, the SDLC and other functions should be modified to make them a standard way of delivering projects. With each new data domain, the analytic environment will become more robust, and present an opportunity to develop and mature analytical functions, such as an Analytic Competency Center or other similar function. The status of data deployment, dependent projects, capability development, and escalated issues should be communicated regularly with the incrementally maturing Data Governance Council (hint: consider the group of business leaders from step 1 when forming this council). Finally, celebrate and communicate each success widely, expanding on the message developed in Step 1.

Enterprise Data Management— Core Capabilities

There are many dimensions to consider when initiating or improving an enterprise data management (EDM) program and the specific initiatives within the program. The following functions have emerged as core capabilities that need to be carefully evaluated to succeed with EDM. As stated earlier, these functions should not be implemented for their own sake, but should instead be tied to the business initiatives that they will support.

- **Business alignment** is the degree to which data management capabilities are positioned to support top business initiatives.
- **Analytical capability** is the competency of the enterprise in leveraging its data assets to meet business goals, from basic static reporting through artificial intelligence and machine learning techniques.
- **Data governance** provides the political will to implement data management improvements to meet business objectives within and across business areas. Structures and processes are put in place to establish priorities, enable decision making, and ensure that departments cooperate to establish shared data resources.

- **Data stewardship** is the day-to-day business involvement within each of the data management capabilities. For example, data stewards help interpret data profiling results, establish data quality monitoring rules, and facilitate consistent business definitions and rules. Data stewards are assigned specific data domains and work with their network of contacts to ensure that the data deployed meets the needs of the enterprise, not just one business area.
- **Data quality management** is about managing the quality of data so it is “fit for purpose.” That is, data quality is “good” to the extent that it meets business objectives. With cooperation from data governance, data stewardship, and IT, data quality management profiles data to determine quality issues that may affect specific objectives, determines root causes of data issues, implements fixes to data and root causes, and monitors the quality of data on a regular basis.
- **Security and privacy** establishes and implements policies so that only authorized users have access to protected data. These policies are enforced through system mechanisms, business process controls, and communication/training. Data governance, data stewardship, IT, legal, compliance, human resources, information security, and other areas all work together to establish broad-based policies and implement policies for specific data domains.
- **Data integration** involves linking data across systems to minimize redundancies and inconsistencies. In the transactional environment, strategies are established to minimize redundant entry and maintenance of data. In the data and analytic environment, architectures are put in place to allow for cross domain analysis and provide read-intensive data to many applications. Links among these environments are also established.
- **Metadata management** ensures that IT and business users understand the data resource. Metadata includes “business” metadata, such as definitions, data lineage, and quality metrics, and also includes “technical” metadata such as table and column names, data types, and dependencies among structures for impact analysis.
- **Master data management (MDM)** processes “master” data such as customer, product, organization, and “reference” or look-up data. These data domains differ from transactional data in that they represent entities that change over time and are often managed in multiple systems. MDM uses a combination of reconciliation, cross referencing, and workflow to provide a consistent view of these critical data domains.



- **Data architecture** is a set of techniques, artifacts, and processes to gain a better understanding of the data resources and direct deployment activities. As a part of this capability, data modeling creates artifacts depicting data structures in logical (business) form and physical (technical) form. These artifacts are part of broader functions of information requirements and data design that gather additional information such as access patterns, service level expectations, and other factors that influence design decisions to meet business needs. Data architecture looks broadly across systems and structures to rationalize data across traditional boundaries.

Conclusion: Intelligent Strategy Development

Many organizations have challenges controlling the proliferation of data, leading to a lack of trust in data resources along with costly, labor-intensive efforts to leverage data for business needs. Rather than diving directly into specific data management practices, it is crucial to develop an overall strategy to align EDM with important company initiatives and the broader operating model. With this approach, EDM and its component practices can be positioned to support important business priorities while maturing foundational data management capabilities at the same time.

About Teradata

With all the investments made in analytics, it's time to stop buying into partial solutions that overpromise and underdeliver. It's time to invest in answers. Only Teradata leverages all of the data, all of the time, so you can analyze anything, deploy anywhere, and deliver analytics that matter most to your business. And we do it on-premises, in the cloud, or anywhere in between. We call this pervasive data intelligence. It's the answer to the complexity, cost and inadequacy of today's analytics. And how we transform how businesses work and people live through the power of data. Get the answer at [Teradata.com](https://www.teradata.com).