Consistency amid the chaos: an enterprise approach to data integration.
Introduction

In 2001, a regional logistics company set out to achieve its goal of becoming a world leader in its industry. Step one was a series of acquisitions in various countries. Step two was the transformation of these separate companies into one unified, integrated, consistent whole.

Innovative information technology solutions had a big role to play in this transformation. Each of the seven previously independent companies had its own systems, applications, processes, infrastructure—and data silos. Chaos reigned. How could all of this data be unified, verified and turned into business-aligned information that could be accessed on demand?

Enterprise data integration was the answer. By implementing an enterprisewide information integration platform that leveraged shared metadata to provide data transparency, standardized processes and methods to provide efficiency, and deployed a service-oriented architecture (SOA) to provide repeatability, the CIO created consistency from the chaos. The result is an IT organization that can deliver the real-time operational and business intelligence that provides the company with a long-term, sustainable competitive advantage—helping increase profits by approximately €1.4 billion per year.

This now-global logistics company is not the only organization looking to leverage data to dramatically change its business processes and improve cost-effectiveness, flexibility, competitiveness and value to its customers. Unfortunately, other organizations have not been as successful in their efforts—primarily because their integration efforts have focused on a single data system or silo, rather than the enterprise as a whole.
Only when data integration efforts are designed to serve the entire enterprise can the full benefit of these efforts be realized—for the company, its customers and its bottom line.

### Why enterprise data integration matters

A recent Gartner presentation described the need for enterprise data integration this way:

> Changing conditions in the business environment are creating a new set of drivers for organizations to focus on data integration from a strategic perspective. Business drivers, such as the imperative for speed to market as well as the agility to change business processes and models quickly, are forcing organizations to manage their data assets differently. Simplification of processes and simplification of the IT infrastructure are necessary to achieve transparency, and transparency requires a consistent and complete view of the data which represents the performance and operation of the business.¹

Gartner also predicts that “through 2008, data integration will be an increasing percentage of overall integration work, consuming more than 50% of integration investment in large enterprises (0.7 probability).”¹ A recent IBM survey of CIOs from around the world also revealed that data integration and delivery of business intelligence were top priorities for CIOs in the short and long term.²

The emphasis on data integration is a direct result of pressing IT and business needs. IT is faced daily with the need to manage an ever-increasing amount of data more effectively—and align the integration of that data with business
Consistency amid the chaos: an enterprise approach to data integration.

Highlights

needs—under static or decreasing budgets. Line-of-business and other business executives need consistent access to accurate, high-quality information that can aid in their decision-making efforts and shorten time to market and reaction times.

Under these circumstances, a consistent, enterprisewide solution to the current information chaos is essential.

The chaos in previous methods

Up to now, many data integration efforts have resulted in only incremental improvements, instead of breakthrough innovations. Considering the amount of money many companies have spent on their data integration efforts, incremental improvement is disappointing, to say the least.

The limited scope of these largely unsuccessful integration efforts is most likely the cause of their lackluster results. Typically, organizations have designed and deployed data integration solutions for a single business system or data silo. With each new system or silo, the data integration efforts begin anew. There is also often a lack of standards or best practices in place to ensure that the information resulting from the data integration effort meets end-user expectations—and there are no viable processes to track the quality of the information from source systems over time.

The limited scope of data integrations that focus on a single business system or data silo typically produces disappointing results.
Consistency amid the chaos: an enterprise approach to data integration.

Figure 1
Current system-specific data integration solutions can be redundant, inconsistent and costly.

Fully customized integration can reduce productivity and may not improve the bottom line.

If this approach to data integration continues, the chaos can only get worse as the amount of data collected and processed continues to grow dramatically. Completely customized integration of all the data for each of an organization’s myriad systems and silos—as is required in the efforts illustrated above—demands the creation of new tools for each integration project and carries a potentially crippling productivity penalty. What’s more, completely customized integration can rarely provide the positive bottom-line contribution needed to allow the organization to remain agile and competitive.
To enable data integration efforts to live up to their business-transforming potential, each integration development team should have access to common, shared data integration services that are callable from any system and any application on demand. This is what enterprise data integration is all about.

**Putting the “enterprise” in data integration**

What business and IT need is a consolidated view of the enterprise that is capable of supporting management strategy and optimizing business processes by delivering enterprisewide, real-time information to the right people at the right time. Enterprise data integration achieves this by building repeating processes that use best practices, shared metadata and reusable integration services to support the entire data lifecycle.

**Metadata is the cornerstone**

A unified, metadata-driven infrastructure facilitates a shared understanding across business and technical domains. It can reduce the time between specification and build. Furthermore, it can provide a transparent and persistent record of understanding that can drastically reduce future project delivery times and improve overall insight and confidence in information.

**What is metadata?**

Briefly, metadata is data about data. Metadata can describe an individual data unit or a collection of data. Depending on the type of data being described, metadata may include a description of the content of the data, authoring person or system, creation date, or physical location of the data.
A good, metadata-driven enterprise data integration solution uses the three primary types of metadata as it is generated by the business:

- **Operational metadata** provides authoring information, creation date, physical location and other operations-based information.
- **Business metadata** identifies the business processes and analytical applications to which raw data should be mapped. Business metadata is critical for end users—or consumers—of information so they can be confident that the data they rely on for making business decisions is exactly what they expected.
- **Technical metadata** provides information on the applications and systems that are being used to store, verify, aggregate and cleanse raw data. Technical metadata helps companies understand what information they have today and the reliability of that information, and streamlines development efforts by providing technical users with information about the data elements and how they are implemented currently across various systems. Additionally, this metadata can help establish a metadata map of source systems that can be leveraged to build federation queries, improve auditability and provide visibility through impact analysis and direct lineage reporting.

The four critical integration capabilities

A true enterprise platform for data integration should deliver all the functions required to integrate, enrich and deliver information that you can trust for your key business initiatives. To provide a strong foundation for enterprise information architecture, the platform would include rich functionality, broad connectivity to heterogeneous sources and a unified, metadata-driven approach.
A unified platform helps integrate information for access and use in new ways to drive innovation, increase operational efficiencies and lower risk.

Understand your data

A unified platform for data integration will help you understand and analyze the meaning, relationships and lineage of information. By automating data profiling and data quality auditing within systems, you can:

- Create an understanding of data sources and relationships
- Eliminate the risk of using or proliferating inaccurate or inconsistent data
- Improve productivity
- Leverage existing IT investments while increasing capabilities.
Consistency amid the chaos: an enterprise approach to data integration.

Cleanse your information
Data cleansing capabilities help ensure auditable data quality and consistency by standardizing, validating, matching and merging information to create comprehensive and authoritative information for multiple uses. The data quality logic should be deployed universally, ensuring data consistency and accuracy across the enterprise. By improving the quality of information, your organization can:

- Make more efficient and effective business decisions with trustworthy data
- Reduce IT and supply chain costs with a single, accurate view of customers, products and suppliers
- Improve customer service and revenue-generating opportunities
- Assemble the auditable, trusted information needed to comply with regulations such as the Sarbanes-Oxley Act and Basel II.

Transform your data into information
An enterprise data integration platform must also transform and enrich information for new uses in new contexts. Hundreds of prebuilt, metadata-driven transformation functions combine, restructure and aggregate information from its current application-centric form into entirely new contexts, allowing information to be used in new ways to suit new business needs. Using transformation, you can:

- Remove the complexity of integrating data from heterogeneous data sources
- Derive important and relevant information out of complex, heterogeneous data
- Ensure information is in a form appropriate for its intended use
- Provide an enterprisewide view of the business at any time, to anyone.

Data transformation encompasses combining, restructuring and aggregating information from application-centric formats into entirely new contexts that allow new uses.
Deliver your information
With the ability to virtualize, synchronize or move information to deliver it in-line to the people, processes or applications that need it, a unified platform for integration provides information when you need it and how you need it. Whether the information is delivered on demand, through federation or on a timed or event-driven basis, it can be moved in bulk from location to location or accessed in place reusing the same core logic. Information delivery helps enable you to:

• Ensure information is always available, when and where it’s needed
• Improve data accessibility and consistency for improved self-service operations for customers
• Reduce latency for real-time visibility into operational information.

SOA as the enabler
A service-oriented architecture takes all of the metadata, methods and processes inherent in a good enterprise data integration solution and makes them work at their most effective and efficient level. An SOA is capable of working with custom, third-party and legacy systems to extract and deliver data. It does this by turning the integration processes, methods and tools into “services” that can be used and reused for virtually every system, every application, every type of data a business creates.

In short, SOA is the enabler that can transform yesterday’s incremental gains into tomorrow’s data integration breakthrough.
Making enterprise data integration work

How do the metadata, processes, methods and SOA fit together? The following structural solution has already proved its worth in organizations large and small all over the world—including the global logistics company referenced at the beginning of this white paper.

This optimal enterprise data integration solution has three parts:

1. **People**—a data integration Center of Excellence (CoE) that provides mentoring, training and expert consultative problem solving
2. **Method**—consistent, proven methodology and repeatable processes
3. **Technology**—metadata-driven infrastructure and architecture

The enterprise data integration CoE

The CoE is a core team among the IT staff that has been specially trained in effective enterprise data integration techniques. These domain experts in enterprise data integration promulgate a standard methodology and best practices to bring consistency and reusability to development projects—and enable project teams to benefit from the experience and knowledge gained from every other data integration project within the organization.

The CoE may provide the resources to staff each of these components:

- **Architecture**
- **Methodology**
- **Best practices and standards**
- **Education programs**
- **Support services**
By aggregating data integration expertise and making it available to integration efforts throughout the enterprise as needed, the CoE can greatly reduce the risks associated with each additional project and help minimize each project’s time to value.

Figure 3
The enterprise data integration CoE and its dedicated IT staff provide a central point of expertise—and the consistent, repeatable methodology, practices and processes that speed implementation, reduce redundancies and lower integration costs.

Methodology
To function most effectively, the CoE staff needs to build its efforts on a standard methodology that will bring consistency and repeatability to development projects. This methodology should be consistent, proven and comprehensive—providing guidance and structure from initial project planning and strategy all the way through to tactical implementation. To ensure that the methodology can be adapted to any data integration project throughout the organization, it should not be based on any one set of integration tools.
A consistent, proven methodology that can be repeated as often as necessary for enterprise data integration efforts large and small can reduce risks and costs and enable an organization to design, develop and deploy enterprise information infrastructure rapidly—including support for enterprise-scale data warehouses, complex legacy migrations or application consolidations, and the creation of master data stores.

**Unified technology platform**

Without a unified platform for data integration, a significant amount of time and money can be spent on integrating the various integration tools, rather than the data itself. The ideal unified data or information integration platform comprises a common set of metadata services, a scalable parallel processing engine, rich connectivity and capabilities to support the entire data lifecycle.

---

**Figure 4**

An effective enterprise data integration solution gives each integration project access to the same methodology, best practices, tools—and opportunities for cost and time efficiencies.
Consistency amid the chaos: an enterprise approach to data integration.

Highlights

The benefits of consistency
When a consistent, repeatable solution for enterprise data integration is put to work throughout an organization, the entire company can realize significant benefits. Data analysts receive data that specifically answers business information needs. Business analysts gain immediate access to precisely the verified, accurate information they need to support business decision making.

Of most immediate importance to the CIO, the IT department can gain an enhanced ability to:

- Deliver trusted information across the organization
- Provide data on demand to any application or person
- Respond quickly to changes in business requirements
- Work more effectively with business units and partners
- Maximize the amount of resources available for new projects.

Taking the first step
Although enterprise data integration is, by its very nature, enterprisewide in its scope, introducing an enterprise data integration solution does not have to be daunting or disruptive. Implementing this type of solution can be done in manageable stages.

Establishing a data integration CoE can be made significantly easier when you work with an external vendor who can provide your staff with valuable training and expertise.
Once the CoE, methodology and unified technology platform are in place, enterprise data integration efforts can begin in earnest with an initial focus on newly introduced or the most critical applications and systems. Subsequent projects can be prioritized based on business needs and strategies. In this way, enterprise data integration can be achieved without undue disruption to the enterprise.

Conclusion
Enterprise data integration brings consistency to enterprise data chaos. Without an enterprisewide data integration solution, an organization may never be able to go beyond the incremental improvements of system-specific integration efforts to achieve breakthrough benefits. With an enterprise data integration solution, an organization can effectively support key business initiatives by integrating and transforming all types of corporate data and content to deliver authoritative, consistent, timely and complete information, while governing its quality throughout its lifecycle. The solution can also help the organization respond more quickly to changing marketplace conditions and maintain or even extend its competitive advantage.

For more information
Help is available. Whether your organization is large or small, IBM can help you explore the alternatives and implement a scalable, comprehensive enterprise data integration solution that works for your business. To learn more, contact your IBM sales representative or IBM Business Partner, or visit:

ibm.com/cio/leverage