



RAPID, COST-EFFECTIVE DEPLOYMENT OF DATA SERVICES IN A SERVICE-ORIENTED ARCHITECTURE

**INFORMATION YOU NEED, FROM DATA SOURCES YOU HAVE,
FOR APPLICATIONS THAT RUN YOUR ENTERPRISE**

This paper discusses the numerous challenges faced by both large companies and federal agencies seeking to embrace SOA and details a metadata-driven, model-based approach to overcoming these challenges.



3	INTRODUCTION: ENTERPRISE DATA AND SOA
4	THE PERENNIAL DATA ACCESS CHALLENGE
5	ADVANTAGES OF A SERVICE-ORIENTED ARCHITECTURE
5	LEVERAGING ENTERPRISE DATA ASSETS IN AN SOA
6	DATA VIRTUALIZATION
7	APPROACHING DATA SERVICES BOTTOM-UP AND TOP-DOWN
7	JBOSS ENTERPRISE DATA SERVICES: A METADATA-DRIVEN APPROACH
10	LEVERAGING DATA MODEL INTERCHANGE STANDARDS
12	AN INTEGRATED SERVICE-ORIENTED ARCHITECTURE FOR THE ENTERPRISE: PROCESS, RULES, AND DATA
14	THE OPEN SOURCE PLATFORM FOR ENTERPRISE SOA AND DATA SERVICES
14	CONCLUSION
15	ABOUT JBOSS DATA SERVICES PLATFORM



INTRODUCTION: ENTERPRISE DATA AND SOA

Service-oriented architecture (SOA), and particularly its most common implementation as web services, represents the latest stage of evolution in application architecture for the enterprise. Service-oriented architectures are enabling organizations to increase their agility in the face of change, improve operating efficiency, and reduce the cost of doing business—often significantly. In spite of these advances, many organizations with significant investments in data collection and storage technologies still struggle with how best to embrace and deploy the new architecture in a way that leverages data assets.

Most cannot afford to re-engineer or replicate the vast amounts of data in existing data stores spread throughout the enterprise. And as concerns about data security, confidentiality, and compliance grow, system architects must balance the need for information integration and sharing with the imperative to protect and secure critical information systems.

Organizations want and need to integrate business processes, applications, and data rapidly, cost-effectively, and incrementally while ensuring immediate and ongoing benefits. For service-oriented architectures to help organizations accomplish this goal, any solution must deploy not only application services that loosely couple certain types of processing with the overall business processes and applications that use them, but also data services that loosely couple existing data sources with applications and business processes while providing integrated access to an organization's data assets. Of course, such data access and integration must occur with an appropriate level of control, performance, and scalability.

The JBoss Enterprise Data Services Platform addresses these challenges with a powerful set of tools and runtime components. With JBoss Enterprise Data Services Platform, enterprises can rapidly deploy their business processes and applications into a service-oriented architecture, making it easy for applications and business processes to use data from many different data sources. Organizations can use JBoss Enterprise Data Services Platform to cost-effectively expose their varied, distributed, and heterogeneous data assets as federated, scalable, secure data services—services that present views of data consistent with the organization's standard data models and application requirements.

In this paper we discuss the numerous challenges faced by both large companies and federal agencies seeking to embrace SOA and detail a metadata-driven, model-based approach to addressing them.



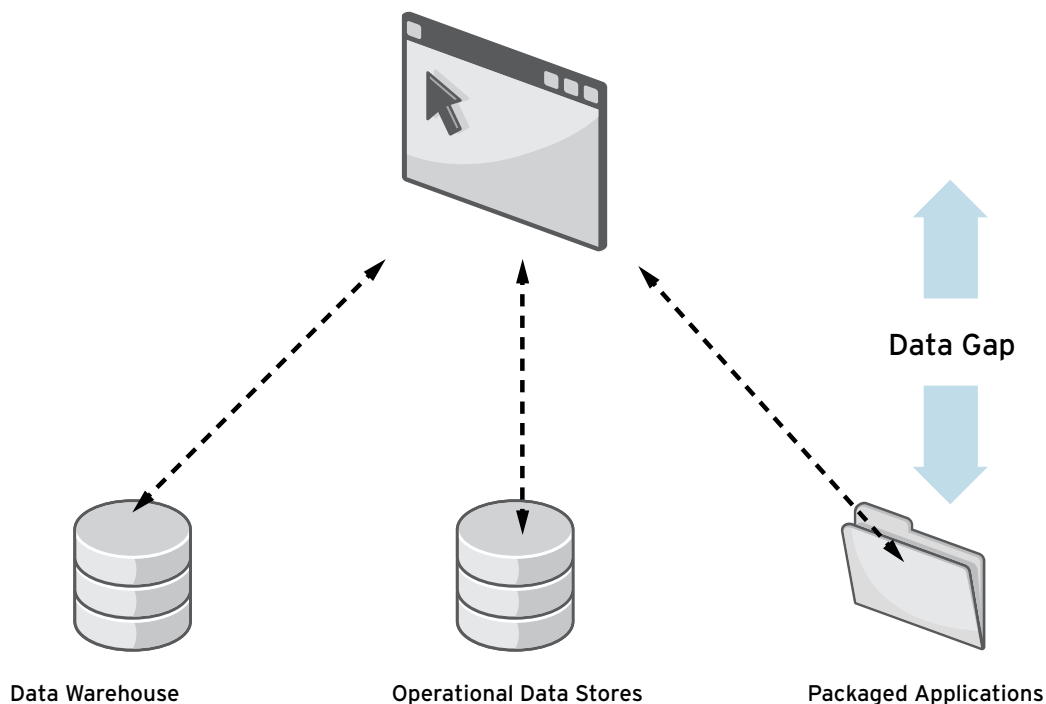
THE PERENNIAL DATA ACCESS CHALLENGE

Today's organizations are faced with a compelling need to share their data assets with stakeholders far beyond those for whom those assets were originally designed, whether the goal is integrating merged companies, putting comprehensive information in front of customer-facing personnel, or enabling decision-makers to understand current operations in real time.

As organizations strive to address these requirements, financial constraints dictate that they maximize the use of existing systems, rather than make broad investments in new information assets.

Enterprises maintain information in many different systems, and typically of many different types. Often, these systems have grown organically, rather than in a structured manner, in order to respond to rapidly changing business needs. This organic process inevitably results in a set of information silos, disconnected and seemingly unrelated to each other. And yet much of this information is in fact related, and, if integrated and leveraged, can deliver tremendous value.

FIGURE 1: THE CHALLENGE OF DATA SILOS



The challenge for the enterprise is to achieve data, business, and process integration rapidly, cost-effectively, and incrementally while ensuring immediate and ongoing benefits. Today's service-oriented architectures are making this possible.



ADVANTAGES OF A SERVICE-ORIENTED ARCHITECTURE

Service-oriented architectures (SOAs) encompass a set of policies, practices, principles, and frameworks that allow for the encapsulation of data and processes as a set of software services with standard interfaces and protocols that can be accessed by a growing and ever-changing community of information consumers. By leveraging the industry-standard XML, SOAP, WSDL, REST, and UDDI protocols, services can be published, discovered, and used in a technology-neutral, standard way. This architecture can begin to bridge the gaps between the information silos of the enterprise—and between the information enterprises have today and the forms of information needed by new, more integrated applications and business processes.

A service-oriented architecture has some key advantages, including:

- **Opportunities for reuse.** Reusing existing services, rather than copying code or implementation, enables faster time to solution.
- **Published interfaces.** The service is made available and its interface is explicitly and precisely described.
- **Formal definitions.** The provider and consumer understand and accept the rules of interaction.
- **Abstraction.** All aspects of the service implementation are hidden from its user.
- **Functional relevance.** Functionality can be presented at a granularity that is appropriate and useful for the consumer.

All of the above benefits give companies and government agencies increased agility to cope with changing conditions and requirements, increased visibility into current operations by identifying and encapsulating key business functionality, and reduced complexity, risk, and cost of integrating systems.

LEVERAGING ENTERPRISE DATA ASSETS IN AN SOA

Even with all the components of an enterprise-ready SOA architecture in place, creating the services to populate that architecture can be an enormous undertaking. The information assets in an enterprise, as mentioned above, tend to exist in a series of silos, each potentially managed, maintained, and persisted in a distinct locale. Examples of these assets include modern and legacy relational databases, proprietary applications and services, text files, spreadsheets, ERP and CRM systems, and data warehouses and data marts. Significant roadblocks to accessing and integrating the data in these systems include:

- Access technologies that vary from system to system
- Source-specific security protocols and policies
- Wide variation in data semantics, syntax, and structure
- Absence of any complete view of information within the enterprise, which severely limits opportunities for reuse, sharing, and standardization
- Varying levels of data quality and integrity

These roadblocks to data and information sharing will always exist. A means must be found to integrate disparate data assets while at the same time recognizing the need for and accommodating their autonomy.



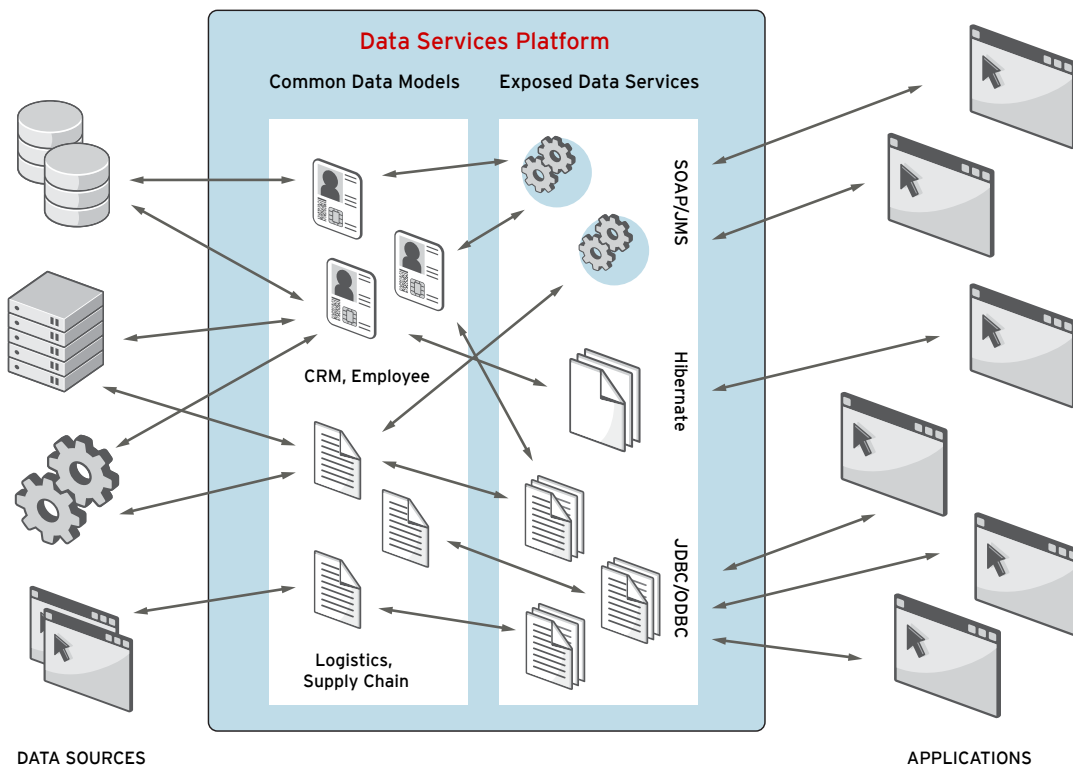
DATA VIRTUALIZATION

To accommodate autonomy of existing data sources while providing the integrated and standard views needed by enterprise applications, some organizations are embracing a virtualized approach to data. Not unlike the system virtualization that is streamlining infrastructure management and enabling cloud computing and on-demand resource allocation, data virtualization enables applications and information analysts to work with data at a level of abstraction appropriate for an application, business process, or enterprise view, while at the same time supporting seamless, controlled access to existing data sources.

Virtual views of data may access an existing source, combine data from multiple sources, and resolve semantic differences. As an authoritative source of information, these new business artifacts ease the reuse of data, giving the enterprise tremendous agility for adapting to changing strategies and requirements. Large cost reductions and increased operating efficiencies can be achieved because existing data sources and services can be repurposed and reused, avoiding the need to build ever more data store silos to serve ever-changing communities of interest with changing requirements. The process of crafting these views of data also facilitates the discovery and potential elimination of redundant or duplicate data stores. Finally, integration can be accomplished in less time, and with fewer resources, particularly if the solution makes use not only of the data resources in the enterprise, but of the metadata describing them.

FIGURE 2: DATA VIRTUALIZATION BETWEEN SOURCES OF DATA AND SERVICES

HOW IT WORKS





APPROACHING DATA SERVICES BOTTOM-UP AND TOP-DOWN

Data services are artifacts that bring the benefits of service-oriented architectures to the problem of data virtualization. JBoss Enterprise Data Services Platform is a powerful set of tools and runtime components that make it easy for applications and business processes to integrate and use data from many data sources and deploy disparate and legacy data sources into a service-oriented architecture. Data services go beyond the traditional concept of a federated data object and include rich semantic and syntactic mediation capabilities, as well as real-time mapping and transformation to structured XML documents for use in a web service.

Today's organizations need a practical approach to solving both bottom-up (data-driven) and top-down (business-process-driven) information problems within an SOA. For the bottom-up problem, the solution needs to provide access to information in the existing physical data sources, in a secure manner, as quickly as possible, and with the least amount of effort. For the top-down challenge, the solution must also be able to provide higher levels of data abstraction to support the longer-term goals of SOA and decouple the data from the applications.

Many organizations face short-term imperatives to allow existing information to be shared by other parts of the organization, regardless of the physical limitations imposed by the data source's structure and technology. Data services can solve this problem and can provide a much-needed abstraction layer that enables eventual migration from existing data environments to new ones.

To truly align IT with business requirements and make SOA initiatives more useful and agile, data services must become part of an organization's long term strategy—a focus of data management. On the processing side, web services would ideally be based on coarse-grained business processes, rather than fine-grained components. To achieve this agility, organizations need a distributed development process in which business analysts drive process articulation, architects conceive of aggregate services, and developers may easily discover, publish, and implement business services. For this approach to be effective, enterprise-wide business data must be associated with each service, without being hard-coded into the service. With the ability to implement data services, organizations can manage a collection of both data and application services that are consumable by the business.

JBoss ENTERPRISE DATA SERVICES: A METADATA-DRIVEN APPROACH

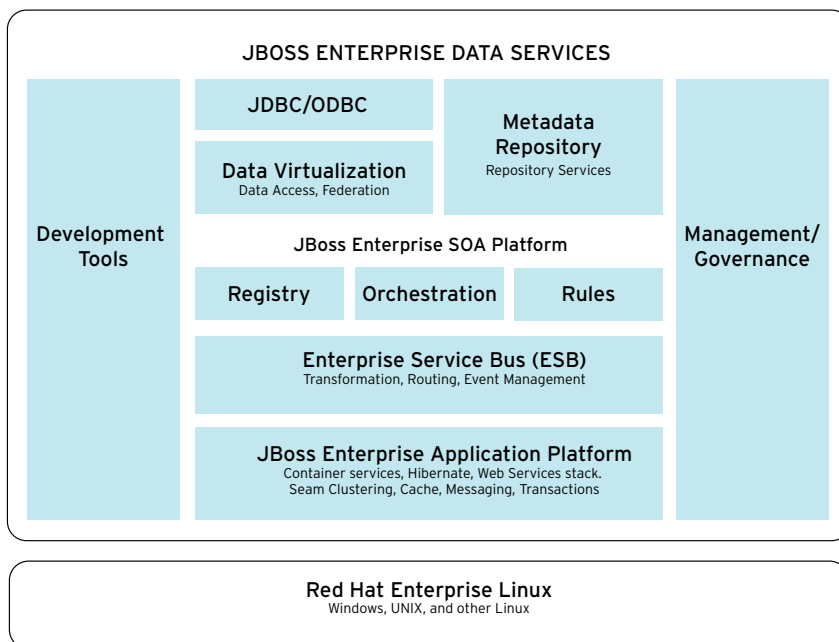
The JBoss Enterprise Data Services Platform from Red Hat uses a unique, model-based, metadata-driven approach to creating data services—an approach that addresses both bottom-up and top-down requirements. By taking advantage of what organizations already know about the underlying structure and semantics of their data, in both their data sources and the applications that consume them, organizations can define their data services rapidly and cost-effectively, deploying them easily into a web services architecture. Offering a significant advance over embedding data access and integration logic in application code—an approach that can be both costly and brittle—JBoss Enterprise Data Services Platform leverages the power of a metadata-driven architecture to provide extensible and scalable infrastructure software for defining, managing, and accessing enterprise data and for seamlessly integrating data into your SOA. By simultaneously accessing multiple sources of data (databases, files, applications, services) and presenting the data in unified views, JBoss Enterprise Data Services Platform insulates applications from the underlying data source details. Data may be accessed through JDBC or ODBC or may be published as services through the UDDI registry that makes the data service known through the ESB.



JBoss Enterprise Data Services Platform consists of:

- Tools for creating data views that are accessible through standard protocols (including an Eclipse-based GUI modeling environment)
- A repository for storing metadata
- A robust runtime environment that provides enterprise-class performance, data integrity, and security
- A full-use version of JBoss Enterprise SOA Platform, including ESB, registry, workflow, and business rules

FIGURE 3: THE COMPONENTS OF JBOSS ENTERPRISE DATA SERVICES PLATFORM



JBoss Enterprise Data Services Platform allows organizations to capture, model, store, and manage metadata for enterprise information systems and then dynamically deploy the models to drive optimized, high-performance, real-time data integration. One logical starting point is to create representative models of the physical data sources—a step that can be performed largely automatically. By creating new models that represent combined, transformed views of those sources, teams build abstractions of the underlying data sources—virtual models that are application- and service-oriented, rather than data-centered. Mappings between the models are then defined, including unions, joins, selection criteria, procedural functions and transformations. This is where the details of the data integration are defined, including such commonly needed data reconciliation functions as name, attribute, and data type conversions. During this process, teams can specify transformations, mappings, and reconciliations to various standard data models. Once the models and their associated transformations are defined, they are deployed as data services in the runtime environment, where they can be directly executed to provide rapid, real-time access to data across the enterprise. These data services can be automatically deployed as web services via registering the WSDL describing them into any industry-standard UDDI registry. Well-documented and visible, the services also become reusable components for the creation of new services and functionality.



A particular strength of this metadata-driven, model-based approach is that organizations can easily define the mappings from relational and other sources into the structured XML documents required for web services. For example, in Figure 4, an XML schema representing a portion of the Global Force Management data model has been imported into JBoss Enterprise Data Services Platform for use as a template to create compliant XML document responses to a web service request. The data fields comprising a fragment of the hierarchical document (in this case, the Equipment object) are automatically mapped to a specialized virtual model called a Mapping Class. A transformation node (the yellow "T") is associated with each Mapping Class and is used to define the syntactic and semantic mediation necessary to populate the data fields from the back end data sources. The models representing physical sources are shown as blue. Note the Input Set parameter that contributes to the transformation. This allows for identifiers and other fields higher up in the hierarchical document to be propagated to the XML fragment, allowing for the selection of the parts of the XML document to return. In this way, mappings can be performed to highly complex XML schemas (including simple or complex datatypes, sequences, choices, association groups, and recursions) without any custom coding. As the sources or schemas change, the mappings can be easily and rapidly modified, leading to remarkable agility on the part of the enterprise. The ANSI-standard SQL that comprises the transformation (which is generated automatically by the JBoss Enterprise Data Services Platform) is shown in the Transformation Editor, where it can be examined or edited.

FIGURE 4: DESIGNER TOOLING

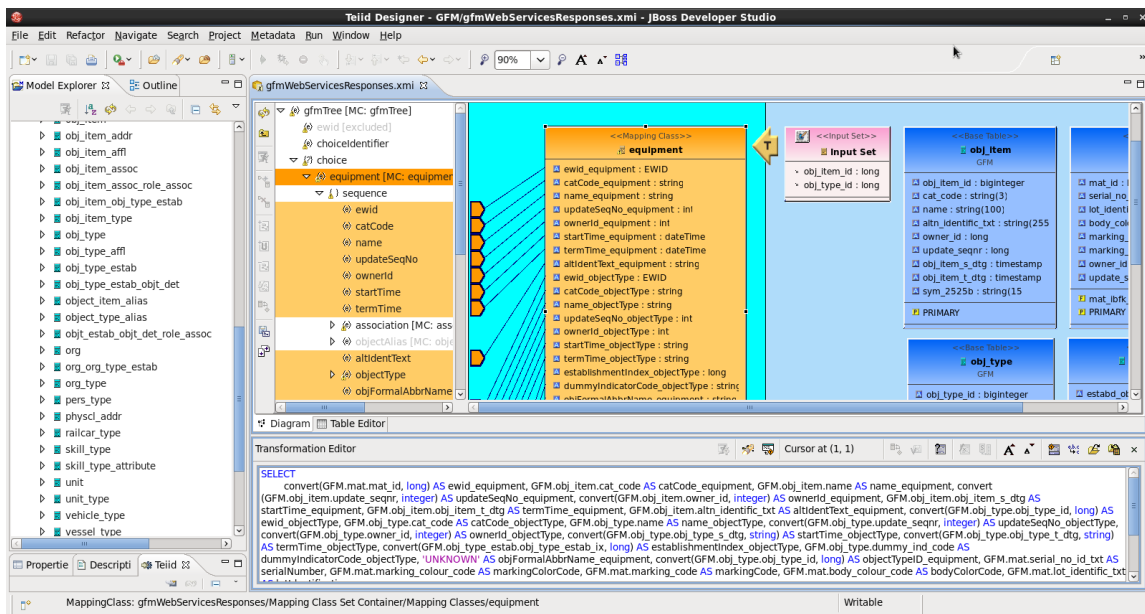


Figure 4: JBoss Enterprise Data Services Platform, showing the mapping of a portion of a structured XML document to the underlying data sources via a user-specified transformation. The Transformation Editor shows the SQL comprising the transformation, which can perform semantic as well as syntactic mediation.



LEVERAGING DATA MODEL INTERCHANGE STANDARDS

Using models to drive integration has another substantial benefit—enabling organizations to leverage data model standards. Organizations are realizing there is great benefit to be had in standardizing the terminology and semantics they use to describe important entities within their business processes and workflow. Examples of such models in the federal government include C2IEDM, GJXDM, JC3IEDM, GFM, DDMS, and NIEM. By creating domain data models with agreed-on standards for data element names, data types, and semantics, organizations can deploy a common semantic layer that can be used to rationalize and reconcile disparate contributing data sources, as well as help define interfaces to various consumers.

Organizations typically face numerous challenges both in defining such a standard and applying it to the systems comprising the enterprise. One of the biggest challenges is determining how to mandate adherence to the new (and typically evolving) standard from the contributing sources and systems. Rather than force every data source in the enterprise to change to adopt these standards, JBoss Enterprise Data Services Platform allows the necessary semantic mediation and transformation to occur in the integration layer. Because of the model-driven approach, JBoss Enterprise Data Services Platform is uniquely positioned to allow for the rapid use of data models (e.g., XSDs, UML models, etc.) to serve as templates for data integration, reconciliation, and data service creation. The data exchange model is represented by a virtual model layer in JBoss Enterprise Data Services Platform, where mappings and transformations between sources and services can be easily specified. You can think of the models and transformations thus created as a “future-proof” abstraction layer—new mappings can be incorporated quickly and models can be modified as current standards evolve and new standards appear.



FIGURE 5: STANDARD DATA MODELS INTO JBOSS ENTERPRISE DATA SERVICES PLATFORM

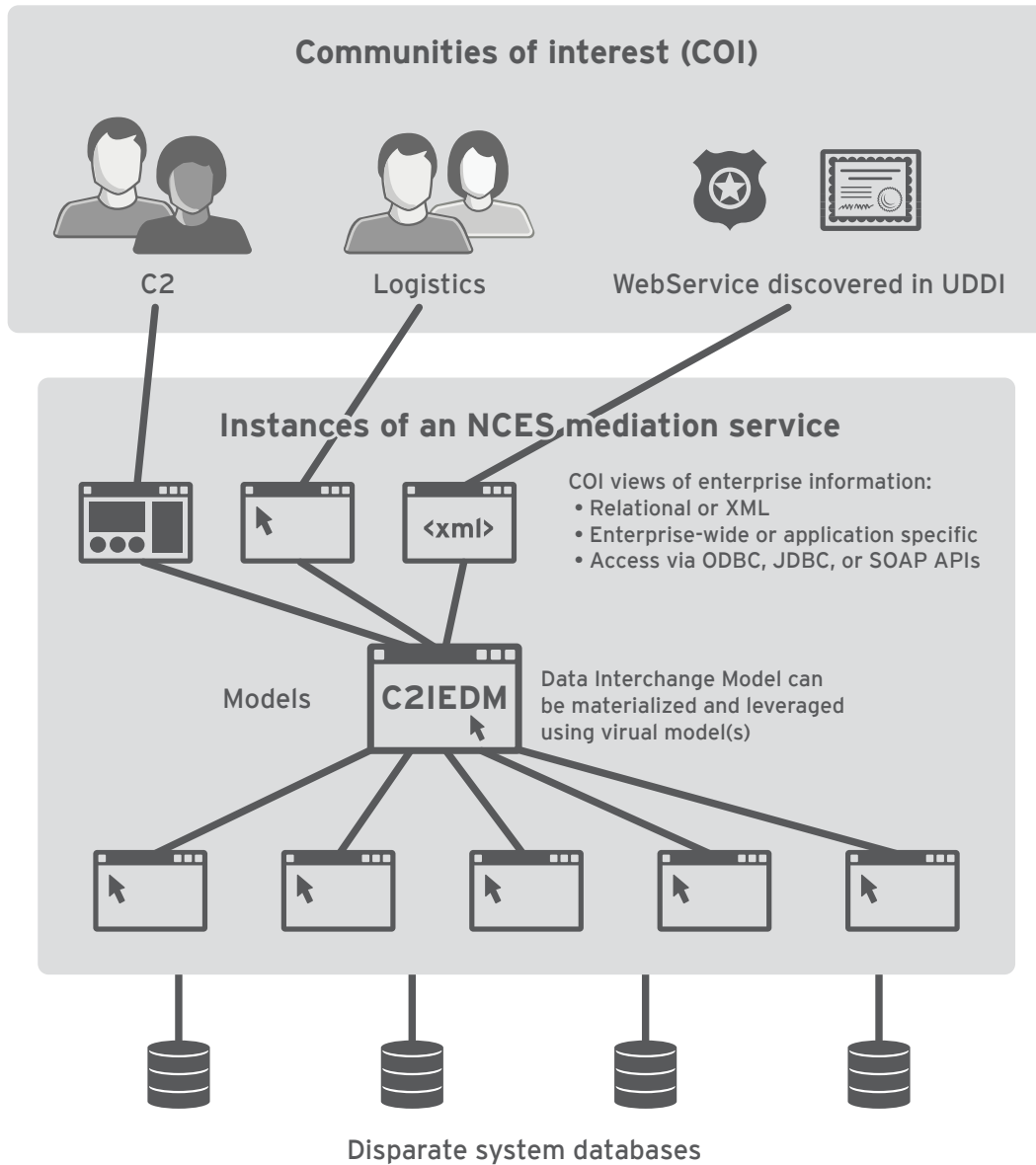


Figure 5. Standard data models can be imported directly into JBoss Enterprise Data Services Platform to be used for semantic rationalization and integration, without requiring changes in the source data.

Federal government agencies in addition to enterprises have lots of data sources that can be more productively used as virtualized data services. For example, in Figure 5 the C2IEDM data model has been brought into JBoss Enterprise Data Services Platform and is being used both to rationalize the differing semantics used by the underlying data sources and to consolidate the disparate data in those sources to COIs. (This system can form the basis of a core mediation service as described by the NCES architecture.)



AN INTEGRATED SERVICE-ORIENTED ARCHITECTURE FOR THE ENTERPRISE: PROCESSES, RULES, AND DATA

All would agree that the ability to provide access to enterprise data is an important capability for today's service-oriented architectures. The question is how data will be integrated into the SOA. Will access and integration be coded into application services? Will data simply be replicated for new applications? Modern SOA platforms such as JBoss Enterprise SOA Platform incorporate powerful capabilities for defining business processes and workflows, capturing business rules in a structure independent of specific applications of those rules, and exposing services for reuse. Organizations also need to be able to create higher-level abstractions that accomplish the same goals for enterprise data—making data use as agile as the use of business logic services and business processes in an SOA. Realizing that they can neither keep creating more copies of data nor creating custom code that is difficult to discover and maintain, enterprises are embracing data services and deploying them in their service-oriented architectures.

As a superset of JBoss Enterprise SOA Platform, JBoss Enterprise Data Services enables the power of SOA to expand to incorporate enterprise data. It provides comprehensive capabilities for defining at a meaningful level and implementing in robust detail both the processing and information components of today's business processes. Taken together, these capabilities facilitate a level of abstraction and integration that is fulfilling long-standing IT visions for business views of data and processing—fully integrated with implementation.

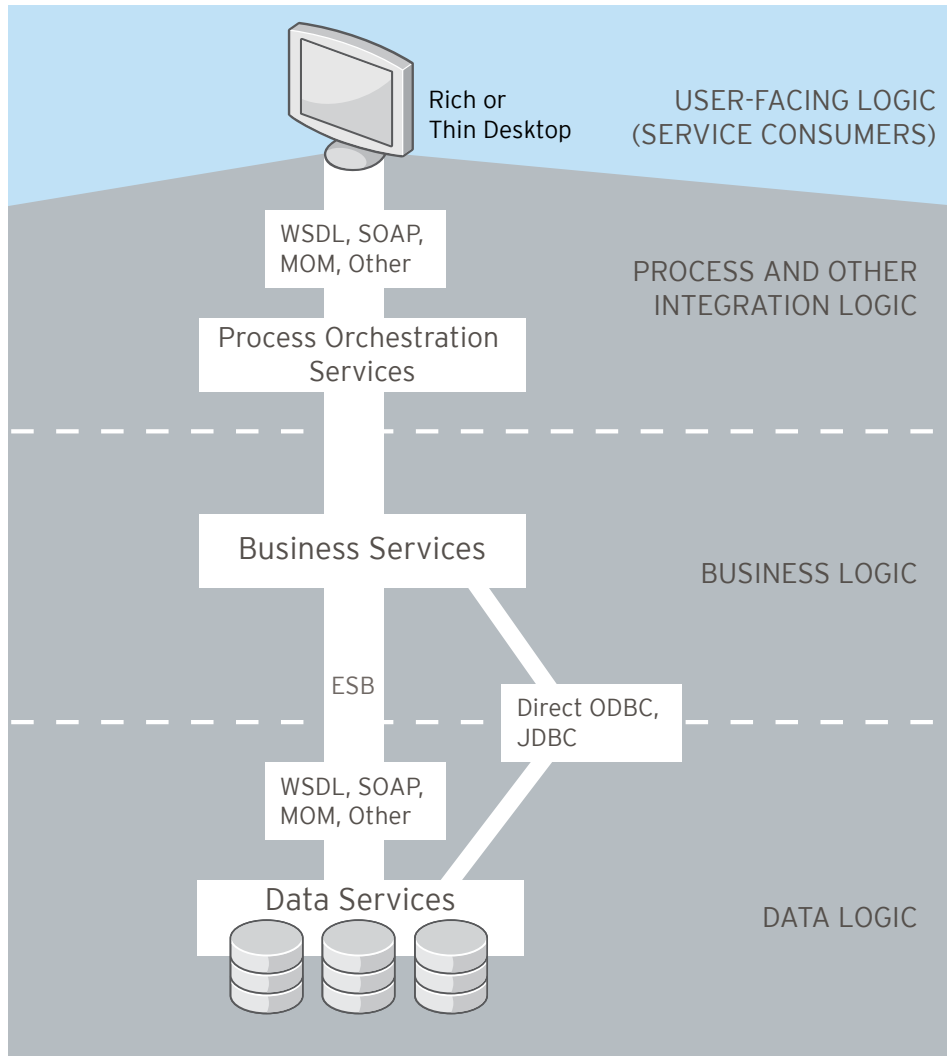
Today's business processes are increasingly characterized by a higher degree of automation driven by real-time events, processing that reaches beyond the walls of the enterprise, and the need to support more stringent real-time SLAs. As a result, applications rely more heavily on integration and on business rules. The ability to integrate data services with business rules—providing the facts the rules engine needs in order to execute a rule—can streamline development of advanced applications, while enhancing transparency and opportunities for reuse.

Creating virtual sources for the authoritative version of enterprise data increases consistency of information and processing within and across applications, as well as creating a central point of access that eases security, compliance, and auditing.

The ability to work with both data services and application services—all exposed through a registry—enables a level of understanding, abstraction, and reuse that accelerates project cycles and creates a more agile and responsive organization. Building on well-defined existing artifacts, teams can complete projects in hours or days rather than weeks and months.



FIGURE 6: RELATIONSHIP BETWEEN DS, RULES, WORKFLOW, ETC.





THE OPEN SOURCE PLATFORM FOR ENTERPRISE SOA AND DATA SERVICES

JBoss Enterprise Data Services Platform is a comprehensive, robust platform for implementation of service-oriented architectures that encompass powerful capabilities for defining and executing the processing and information components of business processes.

Unlike other SOA platforms, many of which are proprietary, JBoss Enterprise Data Services Platform is delivered as open source software that blends the collaborative efforts of an open source community with the unparalleled expertise of Red Hat software professionals. JBoss Enterprise Data Services Platform is thoroughly tested, hardened, and productized through a process that anticipates enterprise challenges—and it is backed by Red Hat's expert technical support. Because Red Hat's open source approach brings benefits in implementation, troubleshooting, and collaborative problem solving at all levels—while also offering a budget-friendly business model—more companies are standardizing on Red Hat enterprise open source software as they upgrade their systems and application infrastructures.



CONCLUSION

IT organizations are eager to reap the full benefits of service-oriented architectures using practical, incremental, and enterprise-ready technology. With the JBoss Enterprise Data Services Platform, a metadata-driven, model-based technology for rapidly and cost-effectively turning existing data assets into SOA data services, enterprises now have the capability to address today's challenging integration needs.

The JBoss Enterprise Data Services Platform enables companies to build successful data services—rapidly and cost-effectively—from the valuable data already maintained by the enterprise. Information services can be provided at the granularity appropriate for the information consumer, hiding the changing landscape of data implementations. Domain data models used by the enterprise to assist in conformance efforts can be seamlessly incorporated into the architecture via JBoss Enterprise Data Services Platform's unique, model-driven approach to integration and mediation. Breaking through the traditional barriers of data location, structure, semantics, and context, JBoss Enterprise Data Services Platform dramatically enhances organizations' ability to reap the many promised benefits of service-oriented architectures.

ABOUT JBOSS ENTERPRISE DATA SERVICES PLATFORM

JBoss Enterprise Data Services Platform is a powerful set of tools and runtime components that makes it easy for your applications and business processes to integrate and use data from many different sources—to turn the data you have into the information you need. JBoss Enterprise Data Services Platform includes tools for creating data views that are accessible through standard protocols, a repository for storing metadata, and a robust runtime environment that provides enterprise-class performance, data integrity, and security. By simultaneously accessing multiple sources of data (databases, files, applications, services) and presenting the data in unified views, JBoss Enterprise Data Services Platform insulates applications from the underlying data source details. Applications can access data in real time, without copying or moving it to a new location. Data is accessed via JDBC or ODBC, or you can publish a service in the UDDI registry, making the data service known through the ESB. JBoss Enterprise Data Services Platform helps companies and government agencies leverage their data assets rapidly and effectively, so they can maximize returns on investments in data and applications.

JBoss Enterprise Data Services Platform is an open source solution offered by Red Hat on a subscription basis and is part of the JBoss Enterprise Middleware family. JBoss Enterprise Data Services Platform contains a full-use version of JBoss Enterprise SOA Platform, with its enterprise service bus (ESB), registry, workflow, and business rules capabilities. As a result, organizations can more quickly and effectively integrate data services into applications and business processes to realize the promised benefits of service-oriented architectures.

With the JBoss Enterprise Data Services Platform, organizations can:

- Maximize return on assets by reusing existing data in new applications and business processes
- Speed time to implementation by quickly repurposing data assets into data services that meet new or changed requirements—without creating more data stores to manage, without expensive and brittle hand coding, and without disrupting existing data stores
- Enhance the value of business intelligence, reporting, and dashboard initiatives by gathering and integrating operational and other data from any location—on-premises and cloud-based data, for example



- Streamline compliance with new formats and standards (for industry models, regulatory schemas, enterprise data models, etc.) and improve transparency to ease data governance
- Service-enable data for use by higher-level business processes defined in your SOA, and integrate application-independent data services with application-independent business rules

JBoss Enterprise Data Services Platform is built using open standards, interoperates with a range of application infrastructures including enterprise application integration (EAI) and enterprise service bus (ESB) technologies, and offers the advanced flexibility and scalability needed to address any organization's enterprise data. Current customers include numerous US and foreign government agencies, as well as Fortune 1000 companies, such as Merrill Lynch, SAP, Credit Suisse First Boston, and Motorola. See (<http://www.redhat.com/products/jbossenterprisemiddleware/data-services/>) for product details, white papers, analyst reports, and information on how to arrange a free trial.

ABOUT RED HAT

Red Hat, the world's leading open source solutions provider, is headquartered in Raleigh, NC with more than 50 satellite offices spanning the globe. CIOs have ranked Red Hat first for value in enterprise software for four consecutive years in the CIO Insight Magazine Vendor Value study. Red Hat provides high-quality, affordable technology with its operating system platform, Red Hat Enterprise Linux, together with applications, management and service-oriented architecture (SOA) solutions, including JBoss Enterprise Middleware. Red Hat also offers support, training, and consulting services to its customers worldwide.

Learn more at redhat.com.

RED HAT SALES AND INQUIRIES

NORTH AMERICA
1-888-REDHAT1
www.redhat.com
sales@redhat.com

**EUROPE, MIDDLE EAST
AND AFRICA**
00800 7334 2835
www.europe.redhat.com
europe@redhat.com

ASIA PACIFIC
+65 6490 4200
www.apac.redhat.com
apac@redhat.com

LATIN AMERICA
+54 11 4329 7300
www.latam.redhat.com
info-latam@redhat.com