zapthink
white paper

METADATA: THE LIFEBLOOD OF THE SOA

ZapThink, LLC • 11 Willow Street, Suite 200 • Waltham, MA 02453 • info@zapthink.com • www.zapthink.com

METADATA: THE LIFEBLOOD OF THE SOA

April 2004

Analyst: Jason Bloomberg

Abstract

Enterprises need information and application visibility across their IT environment, regardless of how far along they are in their plans to implement an enterprise-wide Service-Oriented Architecture (SOA). Without visibility into the workings of the systems, applications, and other elements of their IT infrastructure, businesses are unable to manage or improve their IT environment, and most importantly, meet their business requirements.

Today, however, only a few companies have implemented SOAs. As companies build such architectures, the need for visibility becomes even more acute, because an SOA represents an abstraction layer that masks the complexity of the underlying technology while at the same time providing greater power and agility to the business user. Thus, in an SOA, both IT and business users require visibility into the workings of the SOA at some level.

The key to this visibility in both types of situations is metadata: information about the elements of the SOA. To provide adequate IT visibility, companies must follow basic metadata best practices for discovering and organizing metadata, encapsulating business logic in metadata, managing with metadata, and modeling with metadata. While there are many metadata management solutions on the market, few are able to provide the visibility needed for IT personnel in traditional integration environments, as well as those transitioning to and running SOAs. The IQ Server from Metallect is just such a solution.

Table of Contents

I. SOAs in Action

Human beings have a tremendous capability to deal with complexity. When first presented with complexity, it fazes us. Yet, when we study the particulars of a situation, and if we have sufficient visibility into the elements that form the strands of complexity, we are somehow able to weave an abstraction—a useful mental construct that both masks the complexity and provides us with the tools we need to deal with the situation at hand. Of all the tools we might use to work with complex situations, foremost are those tools that provide us with the visibility we need to take action. For without visibility, we can have no control.

Without visibility, we can have no control.

In the complex world of distributed computing, there is a new approach to building such an abstraction as a means to deal with complexity. This approach is known as *Service-oriented computing*, which leverages a new approach to organization and access to IT assets. Fundamental to the Service-oriented approach is a separation between the business requirements and logic, defined in the form of business processes and rules—and the technology, consisting of the infrastructure that underlies the Services layer of abstraction. What is the secret to building, running, and managing a Service-oriented architecture (SOA)? The answer lies in how software can support the capabilities necessary to offer business users the power they need to achieve the goals of the business. The secret is metadata.

There is a sea change facing the world of IT today as approaches to distributed computing shift to Service Orientation (SO), based upon open standards-based computing. Service orientation is a way of understanding how software works—fundamentally, the perspective of IT functionality being available as discoverable Services on the network. Essentially, Service orientation provides business users with understandable business Services they can call upon and incorporate into business processes as needed. The Service orientation vision is therefore one of agility and flexibility for users of technology, coupled with an abstraction layer that hides the complexity of today's heterogeneous IT environments from those users.

Service orientation is a way of understanding how software works—fundamentally, the perspective of IT functionality being available as discoverable Services on the network.

SOAs have been around for many years, but the difference with the SOAs we talk about today is that they are based on Web Services. Web Services are standards-based interfaces to software functionality. Producers of these Services may publish information about them in a Service registry, where Service consumers can then look up the Services they need and retrieve the information about those Services they need to bind to them.

In a properly architected SOA, business Services represent the data, business logic, and processes available to the business and the core functionality of the underlying systems. Line-of-business users can then compose any set of Services into more abstract SO processes that they can configure to meet applicable

business rules, and then expose as Services to be composed into other processes. In a fully developed SOA, the SO processes, rather than the underlying applications, contain all of the business logic in the enterprise.. Business users will create, configure, and compose SO processes without traditional programming languages, but instead will use tools appropriate for those users. None of this grand vision for the operation of an SOA would be possible, however, without visibility across the architecture.

The Challenge of Visibility

IT users must have visibility into the functioning of systems, network devices, applications, and all the other elements of a distributed computing environment.

Naturally, visibility into the workings of the various IT systems and applications is important to any user of those systems, regardless of whether they have an SOA. IT users must have visibility into the functioning of systems, network devices, applications, and all the other elements of a distributed computing environment, regardless of the role of the IT worker. Likewise, the business-oriented knowledge worker also needs visibility into the workings of the business that are relevant for their particular task at hand—and it is up to the technology to provide that visibility in most cases.

Some examples of the visibility that an IT worker might need include:

- What applications, databases, and other information services exist in the IT environment and how they relate to each other
- How IT personnel can reuse existing applications to fulfill new or changing business requests
- How developers can combine and expose existing application functionality as reusable business Services

Correspondingly, the business user might need the following examples of visibility:

- What business Services are available to automate processes
- Which Services people can reuse to address a business need
- The impact changes in business requirements make on existing processes

For most companies, the primary IT visibility challenges are for IT personnel. In particular, most enterprises who have not yet implemented an SOA struggle every day with their heterogeneous systems, applications, and data sources. In particular, architects have a broad need for visibility across the scope of their responsibility, both before they begin to implement SOAs, as well as during the SOA rollout process.

The need for business visibility into IT then also increases as companies begin to add Services that abstract underlying technical complexity—as they begin to build an SOA. After all, abstraction layers mask underlying details, requiring careful planning to ensure that business users have the visibility they need, without extraneous information that might create noise that interferes with their work. In other words, it is not sufficient to provide visibility to users; users must have the *right* visibility.

Users must have the right *visibility.*

Users require two different kinds of information for the right kind of visibility. First, there is the business content the user is after, whether in the form of structured data or unstructured information like documents, email messages, and other forms of file-based information. The second kind of information is *metadata.*

Metadata are information about the business processes and rules that the business users work with, as well as information about the systems, applications, data stores, and Services that users work with.

Metadata are information about data, or more broadly, information about the business processes and rules that the business users work with, as well as information about the systems, applications, data stores, and Services that users work with. Metadata include information about those processes as well. In fact, most systems produce metadata that contain valuable information. The challenge with those metadata is often how to turn them into insight into the IT environment or the business.

Instead of hard-coding business logic into the programming code, the business logic itself appears in the form of metadata, which the programs must be able to deal with. The underlying software must therefore be able to process the metadata for those processes. Programmers must understand that the role of the software infrastructure is to deal with data—moving them, processing them, storing them, making them available, and not coding to specific business requirements, since, after all, metadata are data. The specifics of what those data and metadata include in an SOA, however, is in the realm of the business user.

II. The Many Roles of SOA

SOA is a particular approach to enterprise architecture. While SOA does require a shift in perspective regarding the role of IT in an organization, it is nevertheless important to remember that SOA is an evolutionary change to the practice of enterprise architecture; many of the established approaches and principles of the existing practice of enterprise architecture apply to SOA.

It is important, therefore, to bring to bear many of the existing best practices of enterprise architecture when practicing SOA, with an eye toward those areas of SOA whose best practices are not yet well established. The best way to delineate the visibility requirements that metadata can satisfy is by identifying the various different roles of the people that work within the architecture, and views of the architecture that they take. Much like the fable of the four blind men and the elephant, who each touched the elephant in a different place and thus had very different perspectives on the nature of the animal, people in different roles can have very different perspectives on the nature of an SOA.

People in different roles can have very different perspectives on the nature of an SOA.

The Roles in an SOA

Because SOA is enterprise architecture, it cuts across the work of all employees within a firm, as well as customers, suppliers, and partners of that firm. This section will consider seven roles (or role categories) that this wide range of people may have, and discuss their visibility requirements in turn. Taking each of the seven roles one at a time, we will then explain how people in each of these roles gets their work done, what sort of visibility the roles that take that view must have, and what metadata are needed to provide that visibility.

The seven roles we will consider are systems experts, component architects and developers, data experts, information specialists, business analysts, line-of-business experts, business stakeholders, and enterprise architects, as shown in Figure 1 below:

Figure 1: The Roles in an SOA and their Views

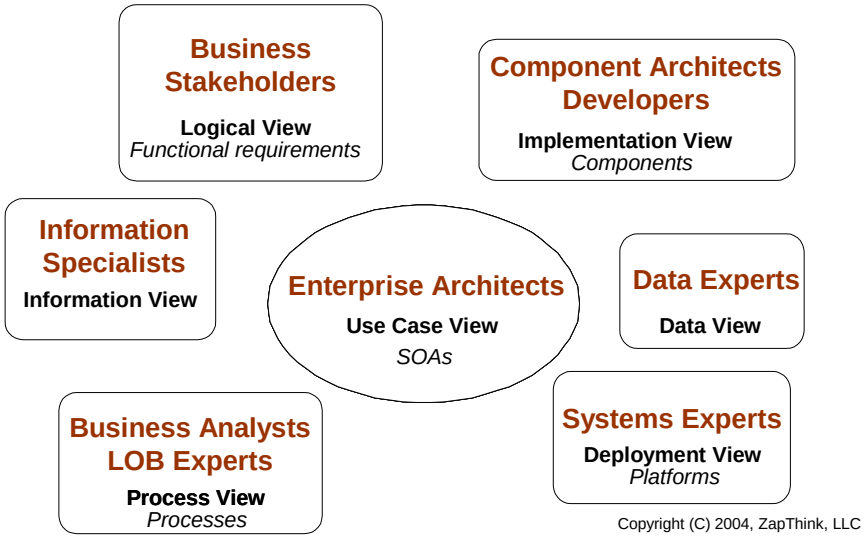


Figure 1 illustrates the different perspectives that participants in an architecture project must take in order to build the architecture. When building, running, or managing an SOA, users in each of the specified roles find their activities falling into the seven views represented above, as shown in the table below:

Table 1: Metadata Needs of the Roles in an SOA

Role	Description	Visibility Needs	Metadata Needs	Sources of Metadata
Systems experts	Focus on the underlying deployment aspects of the distributed infrastructure.	Information about the systems and applications under their purview.	System-level metadata and application metadata.	System management interfaces, application APIs.
Component architects and developers	View of the architecture from the applications perspective.	The component architect must design Service models that represent the design and current state of the SOA. Developers require visibility into code.	Information about the contents of source code repositories, including versioning information and specifics about the interfaces of objects and Services.	Source code repositories, asset management systems, code documentation, application APIs.

Role	Description	Visibility Needs	Metadata Needs	Sources of Metadata
Information specialists	Focuses on the meaning of the information that moves through the company, who is responsible for it, and what people do with it.	Who uses information, how they use it, and what people do with it as they create, transmit, and use that information.	Metadata about individuals and their use of information.	Directories, asset management systems, knowledge management systems.
Data experts	Focus on the taxonomies that the company will use. Normalize the various vocabularies across the enterprise.	The format, source, and purpose of data, as well as semantic information regarding the context of those data.	Basic information about tables, columns, and fields within databases, as well as schemas, tags, and namespaces within XML document stores and message formats.	Databases, data modeling tools.
Business stakeholders	Work to identify the functional requirements for the project or task at hand.	Internal stakeholders require visibility into the overall operation of the business, and key performance indicators. External stakeholders monitor their relationships.	High-level information about how the technology infrastructure is meeting the needs of the business or the business relationship.	Business activity dashboard, asset management system, business reporting tool

Role	Description	Visibility Needs	Metadata Needs	Sources of Metadata
Business analysts and line-of-business experts	Work with the business requirements and the business-oriented Services and craft (or optimize) business processes .	Visibility into the processes and the Services that make up those processes.	Information about the creation and status of processes and Services.	Business activity dashboard, asset management system, Web Services management system.
Enterprise architects	Coordinating the collection of business requirements, the definition of Services, and the planning of the IT infrastructure.	Visibility into all aspects of the SOA, including both the business and technical perspectives.	All of the above.	All of the above.

Because SOAs are a layer of abstraction that hides the complexity of the underlying implementation, many of the systems issues considered from the deployment view perspective are relatively unaffected by the fact that a project is an SOA. The primary exceptions are considerations of the necessary security and management infrastructure. Other differences, like the implementation of a Service registry, are relatively straightforward from the deployment view.

When building an SOA, however, it's important to in mind that the underlying technology components will be exposed as Web Services. These Services then form an abstraction layer, which architects represent in the Service model. The Service model acts as the central coordinating model that represents the business-oriented Services to the business, so that these Services can be orchestrated into business process flows. Fundamentally, such models are metadata that provide visibility into the design of the SOA as well as an essential view of the current state of the SOA.

Service models are metadata that provide visibility into the design of the SOA as well as an essential view of the current state of the SOA.

The bulk of the work for the Service-oriented enterprise architect, then, involves Service definition: identifying which business-oriented Services the SOA should expose, how those Services should be discovered and accessed, and what interfaces they should have. Because the Service abstraction interface essentially separates the business users from the technologists in an SOA, the enterprise architect must work to keep both groups aligned.

III. Metadata Best Practices

Even in a siloed IT environment, in which heterogeneous systems and applications integrate poorly, the use of metadata is important for the effective operation of those disparate applications. However, in an SOA, where the architecture provides application and data integration via an abstraction layer, the proper use and management of metadata are exceptionally important to the effective operation of the SOA. Therefore, adopting enterprisewide metadata

Adopting enterprisewide metadata best practices is a critical step in leveraging the power of Service orientation for any company.

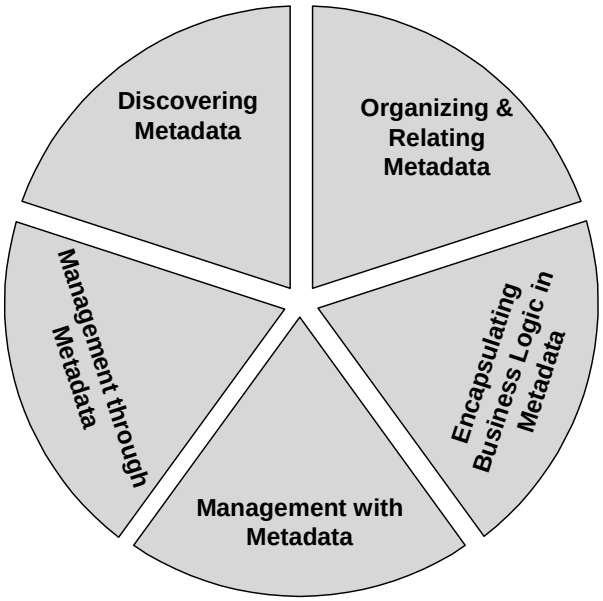
best practices is a critical step in leveraging the power of Service orientation for any company.

In fact, there are three questions that every company asks that metadata best practices can answer:

- What data and capabilities are available in the IT environment?
- Which Services or processes are the appropriate choices for any particular business task?
- What is the impact of change within the SOA, and how should users adapt to those changes?

Depending upon the particular task at hand, the best practices that can help answer these questions fall into five categories: discovering metadata, organizing metadata, encapsulating business logic in metadata, managing with metadata, and modeling with metadata, as shown in Figure 2 below:

Figure 2: Metadata Best Practices



The first step in any SOA project is for users to understand what exists in the IT environment today.

Discovering metadata – The first step in any SOA project, whether integration or process automation, is for users to understand what exists in the IT environment today. IT users must understand the technology available to them, and business users must understand the Services and processes they can access and manage. However, it is not sufficient simply to be able to find large quantities of disparate metadata; users must be able to discover just the metadata they need. Therefore, it is essential that users have access to sophisticated metadata search capabilities.

In other cases, the necessary metadata simply do not exist. In those instances, the IT user may be called upon to create metadata manually. For example, when data are stored in flat files, the information about what data appear where in the files may not exist. The user must then look at the files in order to create the necessary metadata from the context of the data.

To gain visibility into their IT infrastructure, companies must create a catalog containing descriptions of the functionality inside of applications, databases, and middleware products across the enterprise.

Organizing and relating metadata – To gain visibility into their IT infrastructure, companies must create a catalog containing descriptions of the functionality inside of applications, databases, and middleware products across the enterprise. This integration registry should be populated with metadata without changing the systems the metadata describe. In some cases, metadata already exist but are distributed across the IT environment in range of different formats and locations. In other cases, systems may not provide sufficient metadata at all. Companies must organize such disparate metadata into centralized repositories that provide visibility across the IT environment.

Such organization also includes the aggregation of metadata as well as the normalization of those metadata. The aggregation step involves pulling disparate metadata together in one place, but those metadata may be in different formats and follow different organizing conventions. As a result, the IT user must often normalize those metadata to make them consistent across multiple sources. In many cases, this normalization involves the reconciliation of the semantics across disparate metadata sources, for example, quantities may be in boxes in one data source, but pallets in another. The user must make the necessary semantic conversion as part of the normalization process.

Furthermore, it's one thing to collect and organize metadata from various disparate sources, but to achieve true visibility users must link Services semantically. Users must be able to identify the common threads across metadata, even when those metadata have differences in format and context. Relating metadata is one of the key best practices that companies must get right, because only by relating metadata semantically can people get a true understanding of the overall IT environment.

Only by relating metadata semantically can people get a true understanding of the overall IT environment.

Encapsulating business logic in metadata –For most large organizations (and many midsize ones, as well), business logic resides in many places across the company—in applications, management tools, presentation programs, and more. Where it doesn't reside is in the hands of the business people, which presents a contradiction: how can business logic be business logic if it's locked away in the technology, rather than in the hands of the business?

The unfortunate truth is that advances in software architectures like client/server and n-tier have been little more than a business logic shell game, moving the hard-coded logic from this piece of compiled code to this JavaBean or that distributed app. Companies need Service-oriented technology that separates the business logic from the underlying software—encapsulating that logic in metadata. The software, then, should focus on dealing with metadata instead of containing business logic directly.

In order to make this vision a reality, it is vital that IT and business users have solution roadmaps that provide them with the visibility they need to assemble application functionality into Services and business processes. The reality of today's IT environment is that business logic is scattered throughout the IT infrastructure, and that fact will change slowly, at best. Therefore, most companies need to temper the vision of business logic encapsulated in metadata with solution roadmaps that provide visibility into existing business logic across the enterprise.

Managing with metadata – The comprehensive nature of SOAs present exacerbated management challenges to IT shops as compared to the management headaches they had before adopting an SOA. Now, to guarantee that a particular Service meets its required Service level, management tools must provide visibility into any number of underlying systems and infrastructural elements. The only way such management tools can work is through the

sophisticated collection and analysis of metadata that describe the operation of the SOA, from the highest process level, through the Service level, all the way down to the hardware.

Modeling with metadata – Gone are the days where models were little more than complex diagrams put together by high-priced consultants, only to sit on a shelf and become obsolete. To properly manage an SOA, enterprise architects must maintain active representations of the Services available to the enterprise—and those representations are models. Such models, then, essentially consist of metadata, because they represent information about the underlying data. In this way, such metadata truly are the lifeblood of the SOA.

To properly manage an SOA, enterprise architects must maintain active representations of the Services available to the enterprise.

IV. Metallect: Metadata Visibility across the SOA

One startup vendor focused on metadata-driven visibility is Richardson, Texas-based Metallect. Metallect provides a software solution for application visibility. The Metallect *IQ Server* enables IT personnel to rapidly discover, relate, and search existing application infrastructure, increasing the productivity of integration and accelerating SOA adoption.

The IQ Server (illustrated in Figure 3 below) helps IT managers, system architects, and business analysts answer critical integration planning and SOA questions such as:

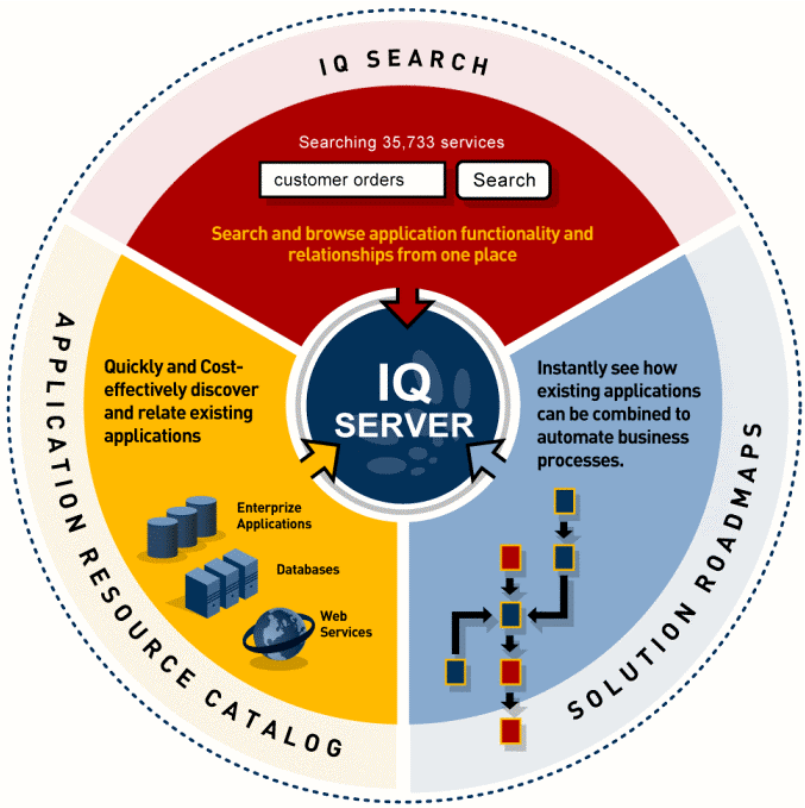
- What resources are available in my IT environment?
- How can I combine existing application functionality to fulfill a business request?
- What are the relationships among various applications, databases, and other information services?
- Where are the gaps in my current systems?
- What are my options for integrating two systems?
- How can I adopt Web Services and SOAs incrementally and justify the cost?

The Metallect IQ Server offers the following capabilities:

- It provides users visibility into the operation of their IT environments (regardless of whether they have an SOA in place) by offering insights and answering difficult IT questions.
- It increases developer productivity by automating manual tasks and reducing the elapsed time between receipt of a business request and delivery of an appropriate solution.
- It increases the reuse of existing applications.
- It helps users determine the impact of changes to the IT environment, especially as companies tackle the task of transitioning to an SOA.

The Metallect IQ Server offers users visibility into the operation of their IT environments (regardless of whether they have an SOA in place) by offering insights and answering difficult IT questions.

Figure 3: The Metallect IQ Server



The core of the IQ Server is the *Application Resource Catalog* (ARC), which is a metadata registry that spans the IT environment, federates application metadata, and provides tools that help users reconcile semantic differences among application metadata. The ARC enables users to discover, translate, and aggregate enterprise metadata into a unified view of application capabilities. The ARC spans enterprise applications, databases, Web Services, middleware, and application servers.

The ARC first creates a catalog of the application assets distributed across the enterprise. The IQ Server then discovers, translates, and aggregates metadata into the ARC. During discovery, the IQ Server uses inference algorithms to create semantic links between application parameters.

The IQ Server also contains a suite of visibility applications that provide insights to users and automate difficult manual tasks. Metallect provides two visibility applications:

- *IQ Search* – Provides the sophisticated metadata search capabilities that are necessary to adopt the metadata discovery best practice. IQ Search enables IT personnel to browse and search all application assets from one place.
- *Solution Roadmaps* – Allows users to automatically determine the best way to use existing application assets to solve a particular business request. The roadmaps automate planning for application integration and SOA projects. System architects, business analysts, and developers

can view solution alternatives graphically and perform rapid scenario planning in an interactive graphical environment.

Metallect also makes the important distinction between *application* metadata, which describes available functionality and Services, and metadata that describes *data* and the corresponding data structures. Most competing metadata solutions focus on the data layer, helping users understand implementation issues such as the flow of data, how to transfer data from one format to another, and details about the format of the data. Metallect, however, primarily focuses on the application layer, helping users understand design and planning questions related to what application and Service functionality exists, and how to best use that functionality, both for IT and business purposes. This focus on application functionality uniquely positions Metallect to help companies follow metadata best practices now, and as they move toward SOAs.

The focus on application functionality uniquely positions Metallect to help companies follow metadata best practices now, and as they move toward SOAs.

The ZapThink Take

Metallect understands that while many organizations are currently executing a roadmap for SOA adoption, few have actually built SOAs. Therefore, the Metallect IQ Server meets the needs of companies struggling with visibility issues in traditional integration environments as well as SOAs. In fact, IQ Server can be useful in helping companies make the transition to SOAs.

One key differentiator that distinguishes the IQ Server from related products, including asset management and traditional integration tools, is that it works entirely with metadata. This reliance on metadata gives IQ Server the power to offer broad visibility across many elements of the IT environment, without the programming and integration overhead of more invasive solutions. Therefore, the Metallect approach is well positioned to provide substantial value to companies who are running complex SOAs, but can also prove quite useful within traditional IT environments that also require increased IT visibility.

In summation, metadata are becoming the new lifeblood of distributed computing infrastructure, because of their importance to corporate visibility, both for line-of-business and IT personnel. In fact, the more dynamic the business environment, the more important metadata become, because the alternative to building visibility solutions based on metadata is to create tightly coupled links to static information. Such brittle solutions simply do not pass muster in today's turbulent business environment.

Metadata serve an important unifying role within an SOA.

Naturally, metadata are also the lifeblood of the SOA, as SOAs are called for in environments of frequent change. Metadata also serve an important unifying role within an SOA, as SOAs touch upon many different parts of the company, including heterogeneous systems environments, multiple lines of business, and processes that involve all aspects of the business. Many different people in many different roles must work in different ways within the SOA, and metadata helps to provide the common basis for visibility that such diverse teams require.

Copyright, Trademark Notice, and Statement of Opinion

About ZapThink, LLC

ZapThink is an IT market intelligence firm that provides trusted advice and critical insight into XML, Web Services, and Service Orientation. We provide our target audience of IT vendors, service providers and end-users a clear roadmap for standards-based, loosely coupled distributed computing—a vision of IT meeting the needs of the agile business.

ZapThink's role is to help companies understand these IT products and services in the context of SOAs and the vision of Service Orientation. ZapThink provides market intelligence to IT vendors who offer XML and Web Services-based products to help them understand their competitive landscape and how to communicate their value proposition to their customers within the context of Service Orientation, and lay out their product roadmaps for the coming wave of Service Orientation. ZapThink also provides implementation intelligence to IT users who are seeking guidance and clarity into how to assemble the available products and services into a coherent roadmap to Service Orientation. Finally, ZapThink provides demand intelligence to IT vendors and service providers who must understand the needs of IT users as they follow the roadmap to Service Orientation.

ZapThink's senior analysts are widely regarded as the "go to analysts" for XML, Web Services, and SOAs by vendors, end-users, and the press. They are in great demand as speakers, and have presented at conferences and industry events around the world. They are among the most quoted industry analysts in the IT industry.

ZapThink was founded in October 2000 and is headquartered in Waltham, Massachusetts. Its customers include Global 1000 firms, public sector organizations around the world, and many emerging businesses. ZapThink Analysts have years of experience in IT as well as research and analysis. Its analysts have previously been with such firms as IDC and ChannelWave, and have sat on the working group committees for standards bodies such as RosettaNet, UDDI, CPExchange, ebXML, EIDX, and CompTIA.

Call, email, or visit the ZapThink Web site to learn more about how ZapThink can help you to better understand how XML and Web Services impact your business or organization.

ZAPTHINK CONTACT:
ZapThink, LLC
11 Willow Street, Suite 200
Waltham, MA 02453
Phone: +1 (781) 207 0203
Fax: +1 (786) 524 3186
info@zapthink.com

ZapThink, LLC • 11 Willow Street, Suite 200 • Waltham, MA 02453 • info@zapthink.com • www.zapthink.com