Electronic Business Registries

An Executive Overview

By JP Morgenthal

There are many different registries in the world — such as the phone book, bridal registries, and voter registration. The common trait is an entry that typically points to an associated object that's the real intended target, associated metadata that helps categorize the data, and one or more formal indices used to retrieve registry information.

Since the advent of the computer, one of its first uses was to provide naming registry services to organizations that ran them — typical "white pages" functionality that would list the company's employees, phone numbers, and room assignments. When the first electronic mail system was created, the registry was then updated to include the person's e-mail address.

The World Wide Web was formed as an information delivery medium made up of distinct, separate information domains called "sites." Each site is accessible through a Universal Resource Locator (URL). URLs let us access a particular site and all its associated links, but do little to let us find all the sites that address a particular issue. Thus, some of the first public registries were created to assist in organizing and cataloging Web content.

Of the types of registries discussed so far, each targeted human users and provided a supportive user interface. Use is an important issue with regard to registries because it defines the requirements for implementation and access. With registries designed for human consumption, the user interface is more important than the underlying implementation. How the person will access and retrieve the information has a higher design priority than how the data is stored.

However, more recent developments in the area of Enterprise Application Integration (EAI) and Business-to-Business (B2B) electronic commerce have forced a need for electronic business registries computers use to locate other computers and services. Here, since the computer is the consumer, the underlying implementation and associated access interface are a high design priority. Indeed, there's been a growth in the number of these registries — each with differing information models and application programming interfaces — lowering the opportunity for these registries to be used simultaneously and to be interoperable.

Current Registry Standards

Here's a brief overview of leading registry standards that satisfy the needs of machine-based consumption.

X.500 is an International Standards Organization (ISO) Open System Interconnection (OSI) directory standard; a directory service is a form of electronic registry. The benefit of the X.500 standard is that it can be part of a larger global directory without requiring an organization to maintain more than their local information. Additional features include:

- Strong query facilities with complex query support
- Single global namespace
- An information model framework that's extensible.

X.500 has a known programming interface that's extremely complex. Usually, X.500 directories are accessed through the Lightweight Directory Access Protocol (LDAP).

The **Domain Name Service (DNS)** is a standard issued by the Internet Engineering Task Force (IETF) and is perhaps the most widely used registry standard. This is because almost every Transmission Control Protocol/Internet Protocol (TCP/IP) transaction first queries the DNS to convert a domain name into a TCP/IP address. Since the DNS is an Internet registry, the data formats used to communicate with DNS have a higher priority than the information model it supports. DNS communicates with clients through a User Datagram Protocol (UDP) — the connectionless communications protocol that TCP/IP sits on top of. DNS queries are simple, but allow for some modest amount of filtering.

Electronic Business XML (ebXML) is a framework for conducting business electronically. It consists of many parts, with a central one being the ebXML registry/repository. As the name connotes, the ebXML registry/repository contains both monikers that point to final targets and the final targets themselves. For example, the registry environment facilitates queries for businesses involved in automotive manufacturing. The list of responses to that query are pointers into the repository environment, where those companies' Collaborative Protocol Profiles (CPPs) are stored.

The ebXML registry has its own information model and service provider interface. The information model defines the objects that can be found in the registry/repository and how to form relationships between those objects. This isn't much different from an entity-relationship diagram that might be used to define a relational database. The service provider interface defines the interface for performing maintenance, managing the registry/repository information set, and executing queries against it.

Universal Description, Discovery, and Integration (UDDI) is the specification put forth by a consortium of software vendors and non-software commercial sector companies. UDDI was designed as the registry for Web services. Web Services may ultimately provide a publicly exposed interface to saleable services. This drives the need for a large-scale registry for sharing Web services. While UDDI registries can be configured in a local manner, there's a global UDDI registry that's replicated according to the UDDI replication specification. UDDI has defined both an information model and services specification.

Microsoft .NET Passport is a Web-based registry service that provides a single, secure representation of a user over the Internet. Users register with Microsoft to get a .NET passport, which can then be used at Passport-enabled sites to provide security access and accelerate transaction processing. The .NET Passport site acts as the clearing authority to validate the user's Passport. This service leverages a proprietary Passport information format and programming interface.

Verisign hosts one of the largest certificate authorities on the Internet. A certificate is an electronically issued entity that represents either an organization or a person. The certificate can be used to digitally sign documents or create an encrypted communications pathway, as is the case with Secure-Sockets Layer (SSL). Much like Microsoft's .NET Passport, users register with Verisign for a digital certificate. Recipients can then validate it using the Verisign certificate authority registry. This registry leverages Public Key Infrastructure (PKI) standards and emerging W3C standards.

No one group is looking at the bigger picture — that the information stored in each of these registries is highly applicable to other applications

Stovepipe Registries: Problem or Not?

Each of the registries we've covered was created for a specific purpose. The purpose of X.500 was to support X.400 messaging standards. UDDI was designed to support the needs of a growing Web Services community. ebXML was designed to be the next-generation Electronic Data Interchange (EDI).

Unfortunately, no one group is looking at the bigger picture — that the information stored in each of these registries is highly applicable to other applications beyond those originally intended. This leaves the need to support multiple interfaces against the same data set or to define interoperability across all standards. We have an integration issue again because these standards were all defined with a narrow focus.

Because of the large body of installed registries, each with its own information models and service interfaces, it has become increasingly complex to think about consolidation. Vendors have put significant effort and dollars into their implementations and applications, which they wrote and deployed against a particular standard. Is this a problem or not?

Other than the tremendous waste of manpower and dollars invested in creating each of these different standards, the creation of stovepiped registries isn't as big a problem as other legacy issues such as Y2K. Issues such as these form the basis of the larger problems of information and knowledge management that plague the industry. Especially in the U.S., which has become a leading information-based economy, the inability to manage complementary sets of information in a consistent manner will lead to higher costs of information management and use downstream.

Alternatives

The World Wide Web Consortium (W3C) seeks to address information and metadata management with forthcoming standards such as:

- DAML+OIL
- Web Ontologies (WebOnt)
- Resource Description Framework (RDF).

Some pundits suggest combining these to create a declarative network of distributed information sets along with facilities to catalog and query them. Because these standards leverage the power and strength of the Web as a distributed medium for information dissemination and sharing, it follows that much of the existing registry work could be represented by these standards in a familiar, reusable manner. A secondary interface, in addition to the existing registry standard, might offer a consistent Web-oriented view. However, this remains speculation and only time and customer requirements will determine the future of these registries. eAI

About the Author

JP Morgenthal is chief services architect for Software AG. He is an internationally known expert on the design and implementation of enterprise applications and distributed computing technologies, and the author of two books and numerous articles on XML, EAI, B2B, ERP, XRM, and SCM. e-Mail: jp.morgenthal@softwareagusa.com; Website: www.softwareagusa.com.