MASHUPS

Strategies for the Modern Enterprise

J. JEFFREY HANSON

Mashups

- **Personnel recruitment—**Open job requisitions can be matched to potential candidates using semantically rich data extracted from online resumes, blogs, and articles.

Potential Uses of Mashups for Your Enterprise

The dynamic and semantically rich nature of a mashup infrastructure can be applied in a number of different uses for an enterprise:

- **Identify potential customers—**Information pertaining to a given customer or company extracted from publicly available sources such as news feeds, stock reports, and search results can be combined with contact information and past sales efforts to more accurately identify potential customer needs.
- **Provide effective customer service—**Information about orders (for example, status, tracking, and delays) can be placed in the hands of your customer service personnel to enable them to provide accurate information and help to customers. Most shipping companies supply resource URIs and/or service APIs that provide up-to-the-minute information about a package in transit. This information can be combined with geomapping components and internal order-tracking systems in an aggregated web page to allow customer service personnel a precise view of packages as they are being shipped.
- **Enable your human resources departments—**Active job requisition requirements can be matched semantically with potential candidates' skill-sets, location, and experience level using data pulled from online profiles, blogs, job sites, or personal web sites. This leads to a deeper talent pool from which to draw as well as a more current pool.
- **Enable your IT departments—**Product defects and feature requests can be semantically aggregated with issue tracking software, online user-community data, customer support reports, and system monitoring data to discover bugs and feature requests earlier. Bugs can be more easily solved with a larger, more semantically accurate collection of data to use for analysis and bug re-creation.
- **Empower your R&D departments—**Research and development departments can exploit research data from a vast number of resources when using semantically enhanced resource mashup data. Data can be derived from

RSS feeds, screen-scraped sites, user forums, search results, government agency services, and semantically enhanced internal documents. Using data from different sources external and internal creates a virtual mindset that empowers your R&D departments with a far greater asset than just adding more head count.

- **Compete more effectively—**Provide your people with powerful tools with which to compete. Using data extracted from RSS feeds; search results; and competitor data such as financials, product prices, product news, acquisition announcements, and strategic partner agreements can be combined and delivered to your sales departments, marketing departments, and engineering departments, enabling them to keep a more watchful eye on the competition.

Uses of Mashups for Specific Enterprises

Mashups benefit enterprises in most every industry. Some benefits are evident. However, a number of industry applications are just now realizing the benefits of an agile mashup infrastructure.

The following items illustrate some of the unique industry benefits realized by a mashup environment:

- **Trucking and shipping—**Trucking and shipping companies can gather and analyze data from manufacturing companies, weather reporting sites, geo-mapping services, and road condition agencies to coordinate capacity, loading, and delivery for products.
- **Economic analysis—**Financial institutions and government agencies can combine and analyze data pulled from sources such as unemployment reports, interest rates, blue-chip company financials, foreign currency values, and mortgage-lending reports to make accurate assessments of the economy on macro and micro levels.
- **Intelligence gathering and analysis—**Government and defense agencies can gather and analyze data from external sources and internal sources to predict potential trouble situations and prepare more effectively.

Determining Relevant Application Patterns for Your Mashups

Applications built on a mashup infrastructure can take many forms. You should determine how your particular business needs will be best suited before beginning the design work for your mashup infrastructure.

A good way to start is by identifying the common business processes that are in place now and how they will be best served by UI artifacts, services, resources, and data formats.

Also, perform analysis of the target user audience for the mashups that will be created. Try to envision how each business process will be addressed from a user's perspective or from the perspective of another internal process.

Some of the common items to address when analyzing the potential uses of your mashup infrastructure are

- Serve specific business purposes with your mashup's perspective to meet market demands and solve business problems.
- Address user needs with your mashups to solve specific problems and to address necessary feature enhancements.
- Apply similar techniques and concepts used by other mashups in your industry that have solved the same problems.
- Realize ROI potential offered by a successful mashup infrastructure deployment.
- Solve integration issues using the semantic richness of an effective mashup infrastructure.
- Analyze existing internal systems and where they fall short of solving your business process challenges.
- Determine logical abstractions on your existing business logic that can be refactored as service APIs.
- Envision how future mashups will be used to form aggregate services and other applications or mashups by combining mashups.
- Examine your data sources to determine how they can best serve content in a semantically rich fashion.

A thorough analysis of the items described in the preceding list helps you to form rough designs of the primary components and artifacts to include in your

mashup infrastructure, as well as helps you to understand the use cases that can be solved by the final product. Once this analysis is complete, you can start to identify sources of information from internal and external resources.

Identifying Sources of Information for Your Enterprise Mashups

Once you have identified the applicable uses for your mashups, you need to identify the sources of information from which your mashups will draw. Conversely, new uses will inevitably appear after sources are identified. The sources can include

- **Existing relational databases**—Data sources that are already in use by your organization.
- **Organizational documentation**—Documents, spreadsheets, sales reports, presentations, and forums that are available across servers and data stores throughout your organization.
- **Existing applications**—Information can be drawn from existing applications as event data, monitoring statistics, user trends, and so on.
- **Search results**—Search results can be filtered and structured to establish a semantic context that can be organized and stored as structured data.
- **External commentary**—Commentary from external sources about your organization. This can include analyst reports, news feeds, news sites, financial reports, competitor articles, online trade magazines, and user forums.

Identifying Services for Your Enterprise Mashups

Service APIs exposing functionality pertaining to the parts of your enterprise for which you want a public interface should be identified for your mashup infrastructure. These can be identified from a number of different mechanisms:

- **Legacy APIs**—Existing APIs that have no public interface can be refactored to expose a publicly available service API.
- **Composite services**—New functionality can be derived from the combination of multiple services or software modules. This functionality can then be exposed as service APIs.

- **Data access operations—**Data access operations often contain business semantics that may serve better as a service API. This can help to reduce couplings between data consumers and data sources, as well.
- **Search semantics—**Semantics of any search operations targeting content internal to your organization often indicate a need for business logic that is not yet available.
- **Events—**Events and notifications published by monitoring tools and software components can be combined to form complex events and service APIs.

Enterprise Mashup Design Tips

Designing your mashup infrastructure can entail some common principles as well as some unique techniques. Understanding the scope of these concepts can help you make the most effective use of resources and personnel. Some of these are as follows:

- **Optimize reuse—**Reuse is the primary benefit of a mashup environment. Promoting reuse promotes creativity. Seek to promote reuse by reducing couplings in code and by establishing effective internal documentation and policies in which software modules and services are clearly defined and promoted.
- **Always consider bandwidth—**Mashups tend to promote a high degree of network traffic due to the use of resources scattered across many diverse regions and sites. Be smart about making remote invocations and seek to keep payloads at a minimum. Use caching where applicable and try to transmit just changes in data rather than transmitting entire datasets over and over.
- **Promote the use of standards—**Doing this enables you to avoid proprietary and ad hoc data formats and protocols. This helps to present your mashups to more potential consumers, development tools, and client devices.
- **Establish a consistent look-and-feel—**Even though a mashup environment is dynamic and free-form, a consistent look-and-feel can be established for UI artifacts from which mashup builders can choose. Try to promote the look-and-feel even with external UI artifacts using CSS techniques, dynamic scripting techniques, and even refactoring data before it is presented.

- **Separate data models from presentation logic—**Keep data models and presentation logic separated. Markup languages provide a quick and easy way to present information in exciting views and forms. However, this can lead to the tendency to combine data and presentation logic in a tightly coupled page. Many presentation frameworks go to great lengths to keep logic and data separated, but they are only effective if this separation of data and logic is exploited in a strict manner.

Separation of data and presentation logic is essential for building agile teams of developers with different skill-sets. If data and presentation logic are separated, teams can work concurrently to produce deliverables much faster than a linear development approach.

One of the most effective technologies used by web page developers to realize separation of data and presentation logic is CSS. CSS allows look-and-feel for web pages to be completely separated from the actual data being presented. To modify the look-and-feel for single items, multiple items, or an entire web site, CSS classes and tags can be adjusted without any interaction with the presentation data.

Building the Foundation for an Enterprise Mashup Infrastructure

This section applies the concepts of this chapter to the construction of a foundation for an enterprise mashup infrastructure. The basis for the architecture used in this chapter is a multilayered platform as illustrated in Figure 3.1.

In the diagram shown in Figure 3.1, the layers for the mashup infrastructure will be embodied as interconnected service engines or kernels that will act as the primary segregation points of scalability and performance.

The service engines in this scenario will be implemented using OSGi objects, where each kernel can run independently from one another and manage service registrations, service invocations, and service lifecycles. Each kernel can be managed using JMX-based instrumentation, and each kernel can publish its registered services on which other kernels and software modules can make service invocations.

Implementing Infrastructure Layers Using OSGi

The foundation technology for each kernel will be the OSGi Service Platform (http://www.osgi.org). OSGi technology is ideally suited for any project that is