

Java Jazz up

A B E T T E R W A Y T O L E A R N P R O G R A M M I N G

Mashup

- **Mashup Tools**
- **Mashups: The new breed of web app**



India's Cheapest web Service Provider



OUR SERVICES

- ERP Soluiotns
- Software Solutions
- Web Development
- Web Designing
- Web Redesigning
- Domain Registration
- Web Promotion
- SEO
- Article Writing
- Blog Writing
- News Writing
- SEO Copywriting
- Technical Documentation
- E-Commerce Solutions
- CRM
- Outsourcing

Extend your reach
with our solutions...

"Optimism with determination lets you hit the goal harder"

Published by

RoseIndia

JavaJazzUp Team

Editor-in-Chief

Deepak Kumar

Sr. Editor-Technical

Ravi Kant

Graphics Designer

Santosh Kumar

**Register with JavaJazzUp
and grab your monthly issue**

"Free"

Editorial

Dear Readers,

First of all, we would like to heartily thanks to all our readers for giving so much positive response to us. For enhancing and updating your knowledge, we are bringing a special web tools 'Mashup' tools in this July 2008 edition. This July edition is specially designed to enlighten you about the fundamentals of Mashup tools, which is used for creating video and photo, search and shopping, mapping and news websites.

In the special 'Mashup' issue, we have described all the Mashup tools used in the web like BEA Aqualogic pages, Aptar, Dapper, Extensio, RSSBus, Microsoft Popfly, Denodo, SnapLogic 2.0, JackBe's Presto, Proto, Yahoo Pipes, Kapow, WSO2, StrikeIron SOA Express and IBM's QEDWiki etc. Besides these, we have also covered the problems that programmers usually face during developing Mashup tool. In the new breed of web app, we have covered web protocol and screen scraping and component challenges. We hope that this new edition will provide you adequate knowledge of Mashup tools.

Mahup tools are being extremely popular and crucial web development tools in the present web world. We have also included a page of IBM's 'Mashup: The new breed of web apps' in our popular Java magazine and for this, we are highly grateful to IBM team members for granting us permission. Like earlier editions, we have categorized the whole topics in sections and each section is decorated with different colors and images that would certainly lure readers while reading technological stuffs. This will create interest in readers and users for learning it easily. We are also providing our online Java Jazzup magazine in PDF format that you can view and even download it as a whole or a part. This PDF version provides you a different experience. Please send us your valuable feedback about this issue and participate in the reader's forum with your problems and issues concerned with the topics you want us to include in our next issues.

Editor

Deepak Kumar
Java Jazz up

Content

- 05** [Mashup Tools](#) | In the mashup space, there are many tools that can be used to create mashups, like Yahoo pipes or Dapper. This section provides list of some Mashup tools as well as mashups that can be used as general-purpose tools.
- 20** [Mashup: The new bread of web app](#) | Mashups are an exciting genre of interactive Web applications that draw upon content retrieved from external data sources to create entirely new and innovative services. They are a hallmark of the second generation of Web applications informally known as Web 2.0. This introductory article explores what it means to be a mashup, the different classes of popular mashups constructed today, and the enabling technologies that mashup developers leverage to create their applications. Additionally, you'll see many of the emerging technical and social challenges that mashup developers face.
- 28** [Advertise with Us](#) | We are the top most providers of technology stuffs to the java community. Our technology portal network is providing standard tutorials, articles, news and reviews on the Java technologies to the industrial technocrats.
- 29** [Valued JavaJazzup Readers Community](#) | We invite you to post Java-technology oriented stuff. It would be our pleasure to give space to your posts in JavaJazzup.

Mashup Tools

Mashup Tools

In the mashup space, there are many tools that can be used to create mashups, like Yahoo pipes or Dapper. This section provides list of some Mashup tools as well as mashups that can be used as general-purpose tools.



The screenshot shows the Oracle website navigation menu with 'Solutions', 'Products', 'Services', 'Support', and 'C'. A sidebar on the left lists various BEA AquaLogic products, including 'BEA AquaLogic User Interaction', 'BEA AquaLogic BPM', 'BEA AquaLogic Integrator', 'BEA AquaLogic Service Bus', and 'BEA AquaLogic Data Services Platform'. The main content area features the heading 'BEA AquaLogic® Page' and the sub-heading 'Drag-and-drop simplicity for creating wil'. Below this, a snippet of text reads 'BEA AquaLogic Pages helps everyday participa'.

1. BEA's Aqualogic Pages

BEA AquaLogic Pages is one of the end-user mashup tools that helps the user creating web pages, blogs and powerful data driven web applications using drag and drop components. AqualLogic Pages offers many key capabilities like "In-place editing", "Data management", "DataSpaces", "LiveSpaces", "Page-building components", "Component connecting", "Page linking" etc. AqualLogic Pages provides benefits like "Everyday users become active creators", "IT distributes the burden of Web building to participants", "Avoid inefficiencies associated with using email and documents for everything", "Use the Web to solve urgent, situational challenges", "Boost productivity around ad hoc activities", "Apply the power of participation to every corner of your business" etc.

Read more at: <http://www.bea.com/framework.jsp?CNT=index.jsp&FP=/content/products/aqualogic/pages/>

Mashup Tools



The screenshot shows the Apatar website header with the logo "apatar beta connecting data" and a navigation menu with links for Home, Product, Services, Solutions, DataMaps, and Community. A sidebar menu on the left lists categories: For Mashup Data Integration, For Data Warehousing and Analytics, For Operational Integration, For Application Integration, and For Embedded Data. The main content area is titled "For Mashup Data Integration" and contains introductory text and a list of points (a, b, c) discussing middleware tools.

2. Apatar

This mashup product helps to join data sources with the the Web and Office 2.0 applications without coding. It helps in accessing data on a local network and on-demand applications and systems. Apatar can be the choice because of certain features like "Native connectivity to major Web 2.0 APIs", "Create and deploy enterprise-level data mashups", "Share, receive and re-use pre-built integration mappings", "Opens source Architecture". Apatar is open source mashup tool and runs on Windows and Linux.

View more at: http://www.apatar.com/for_structured_data_mashups.html



Get any content from the Web

Search for existing content feeds and web services

[how it works video](#) [or create a new Dapper](#)

3. Dapper:

Dapper is web based product for data mashup creation and use. It declares that user can use any web based content in any way. For example, creating RSS feed or Google Gadget for a site, receiving email to message Alexa ranking, putting a site's content on a map and many more. Any programmer can use Dapped content in its raw XML form. It can provide relevant data from anywhere into many formats of your choice(XML, RSS, Google Gadget, Netvibes Module, iCalendar, and more).

Read more at: <http://www.dapper.net/>

Mashup Tools

Home | Products | Case Studies

Extensio

- Hot Downloads
- Test Drive
- White Papers
- View Demo
- Analyst Speak
- Extensio in the News
- Our Customers



Press Releases



Extensio® offers the indust

4. Extensio

Extensio is new mashup tool serving as end to end information delivery platform. With the help of this tool data from many data sources can be extracted, data can be assembled. The collected information can be delivered on multiple user front-ends. Extensio provides SOA based information delivery mechanism.

According to Richard Monson-Haefel (Sr. Analyst, Burton Group), "Extensio's Information Delivery framework provides end-to-end enterprise widget (a.k.a. Desktop Gadgets) functionality that is very compelling and includes ERP connectors, recombination (mashups) at the server level, and an excellent widget engine that allows user to capture exactly the information they need at the time they need it. Support for legacy Window platforms is also a huge bonus as many organizations continue to use Windows 2000 for their desktop."

Read more at: <http://extensio.com/>

Mashup Tools

The screenshot shows the RSSBus website. At the top left is the RSSBus logo. Below it is a search bar with the word "SEARCH" above it. To the right of the search bar is a blue play button icon. Below the search bar is a "PRODUCTS" section with the text "Generate, manage, orchestrate, and pipeline Feeds." and three links: "RSSBus Feed Server", "RSSBus Desktop", and "RSSBus Connectors". Below that is a "TECHNICAL OVERVIEW" section with the text "For Architects, Developers, and Script Builders:" and a link "RSSBus Architecture". To the right of these sections is a large orange graphic with the text "INTRODUCING THE RSSBUS FEED" and three bullet points: "Easily create dynamic RSS feeds for Spreadsheets, Directories", "Choose from hundreds of connectors for various systems and applications", and "Use RSS feeds as Simple APIs to communicate with your partners". Below the graphic is the heading "What does RSSBus do?" and the text "RSSBus provides a consistent interface for accessing documents, network protocols, and other data sources."

5. RSSBus

RSSBus works well as mashup tool. It offers many connectors to make easy to mashup the data from various sources like pre-existing RSS feeds, Flickr, Amazon. RSSBus provides accessing SQL databases, office documents, email systems, credit card companies, web resources like Google, Yahoo, Amazon and many information sources.

Read more at: <http://rssbus.com/>

Mashup Tools

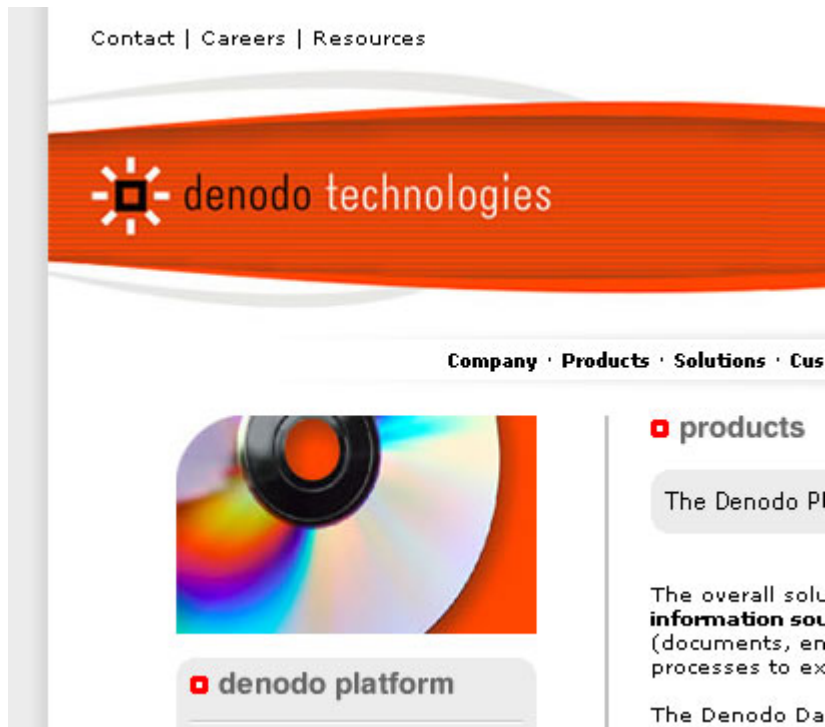


6. Microsoft's Popfly

Microsoft provides a new mashup development tool called Popfly. Popfly is an easy to use and good looking tool. It provides rich user interface based on Silverlight. You can take photos, RSS feeds and various other information with popfly's mashup creator and personalize the view of the web by combining the data. It declares that you can create a mashup without writing a line of code and combine different web sites together for new creations.

Read more at: <http://www.popfly.com/>

Mashup Tools



7. Denodo

Denodo is one of the data mashup product suite that provides highly innovative architecture which can be used to integrate existing data (structured, unstructured, semi structured or internal data) from various resources. It has the ability to collect data from wide range of source like local intranet, e-documents, database, XML repositories, e-mails, SAP, Seibel etc. It provides key benefits like "Non-intrusive alignment of business processes with data", "360° enterprise-wide single point of data integration", "Incremental, easy and rapid development of solutions based on just-in-time composite data".

Read more at: <http://www.denodo.com/english/products.html>

[Wiki](#)[Forums](#)[Blog](#)[Roadmap](#)[Download](#)[Mailing Lists](#) | [Login](#) | [Account](#)

Welcome to The SnapLogic Commu

SnapLogic is an Open Source Data Integration framework that interfaces to solve today's most pressing problems in data inte

SnapLogic 2.0

The release of SnapLogic 2.0 is [now available in the dow](#) 1.0 release, and incorporates a number of architectural enhan

Highlights of the 2.0 Release

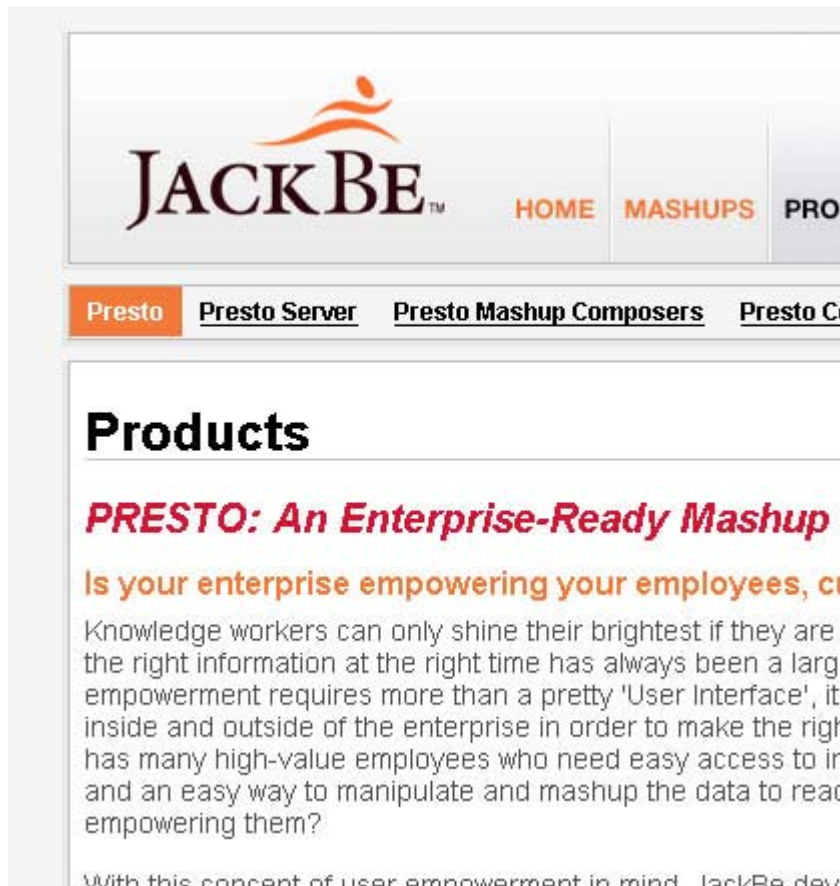
- Improved Designer and Explorer
- New SnapLogic API (SnAPI)
- Independent Server and Component Containers
- New Component API

8. SnapLogic

SnapLogic is one of the open source mashup solution and provides a community-based data mashup server. SnapLogic provides graphical IDE and support for JSON and RSS. SnapLogic provides support for SugarCRM, QuickBooks, and Salesforce connectors. SnapLogic supports Windows and Linux.

Read more at: <https://www.snaplogic.org>

Mashup Tools



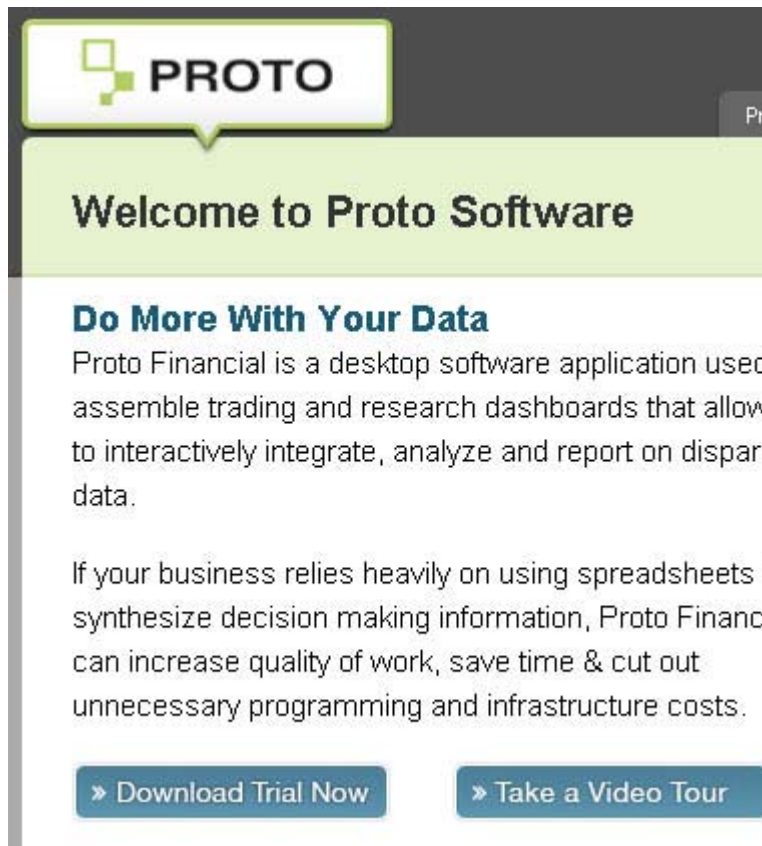
The screenshot shows the JackBe website. At the top left is the JackBe logo, which consists of a stylized orange figure above the text "JACK BE™". To the right of the logo are navigation links: "HOME", "MASHUPS", and "PROI". Below the navigation is a secondary menu with links: "Presto", "Presto Server", "Presto Mashup Composers", and "Presto Co". The main content area features a large heading "Products" followed by a sub-heading "PRESTO: An Enterprise-Ready Mashup". Below this is a question: "Is your enterprise empowering your employees, cu". The text continues: "Knowledge workers can only shine their brightest if they are the right information at the right time has always been a large empowerment requires more than a pretty 'User Interface', it inside and outside of the enterprise in order to make the right has many high-value employees who need easy access to in and an easy way to manipulate and mashup the data to reac empowering them?". At the bottom of the visible text, it says "With this concept of user empowerment in mind JackBe dev".

9. JackBe's Presto

"JackBe's Presto series of enterprise mashup solutions is one of the most complete and compelling solutions available today for organizations that want to move the excitement and results of the consumer mashup story into their day-to-day business.", Dion Hinchcliffe, Web 2.0 industry analyst and founder of Web 2.0 University. It provides many benefits like "Empowers the Business User", "Faster and Flexible User-driven Information Access", "faster, easier, and flexible access to business data resulting in better self-sufficiency and ROI", "Allows IT to Maintain Enterprise Governance", "Accelerates Use & Adoption of Re-usable IT assets", "Satisfies Business User Information Needs without Causing IT Anarchy".

Read more at: <http://www.jackbe.com/Products/index.php>

Mashup Tools



The screenshot shows the Proto Software website. At the top left is the Proto logo, which consists of a stylized 'P' made of three squares (two white, one green) followed by the word 'PROTO' in a bold, sans-serif font. Below the logo is a green banner with the text 'Welcome to Proto Software'. Underneath the banner is a section titled 'Do More With Your Data' in bold blue text. The text below this title describes Proto Financial as a desktop software application used for assembling trading and research dashboards. It mentions that the software allows users to interactively integrate, analyze, and report on disparate data. A second paragraph states that if a business relies heavily on spreadsheets to synthesize decision-making information, Proto Financial can increase the quality of work, save time, and cut out unnecessary programming and infrastructure costs. At the bottom of the section are two blue buttons with white text: '» Download Trial Now' and '» Take a Video Tour'.

10. Proto

Proto is one of the new mashup tools. Proto is a commercial, non-web-based product and is available free for personal use. According to Jeffrey Pham, Management Consultant, "Proto is an impressive tool. It's a visual approach to data mining and analysis that can increase your efficiency. It can actually make mashup tools valuable, not just fun."

Read more at: <http://www.protosw.com/>

Mashup Tools



[Home](#) [My Pipes](#) [Browse](#) [Discuss](#) [

Featured Pipe: eBay Price Watch



NEW! [Pipes badges in the wild and cool blog posts](#) | [New Yahoo](#)

About Pipes

Pipes is a powerful composition tool to aggregate, manipulate, and mashup content from around the web.

Like Unix pipes, simple commands can be combined together to create output that meets your needs:



11. Yahoo's Pipes

Yahoo Pipes is data mashup tool that can be used to aggregate, manipulate and mashup content from wide variety of sources around the web. Pipe word is taken from UNIX, where pipe is for transferring data between applications, but in this case it is for transferring data from one or more sources. It can be used to combine many feeds into one, sort, filter and translate. It can grab pipe's output as RSS, JSON, KML and other formats. It can be used to geocode your favorite feeds and remix your favorite data sources.

Read more at: <http://pipes.yahoo.com/pipes/>



PRODUCTS

SOLUTIONS

SUPPORT

Be smart

Kapow Technologies harvests market
web intelligence you need to make
that will move

12. Kapow

Kapow is one of the most capable tools in the mashup world. It has an open community (OpenKapow) and a powerful desktop-based graphic mashup builder. Kapow can collect data from various sources including web services and XML, CMS, PDFs, etc. Kapow provides robust error handling and exception processing capabilities.

Read more at: <http://kapowtech.com/>

Mashup Tools

The screenshot shows the WSO2 Oxygen Tank website. The header includes the logo "WSO2 OXYGEN TANK" and the tagline "THE DEVELOPER PORTAL FOR SOA". A navigation bar contains links for Home, Projects, Downloads, Library, and Forum. The main content area is divided into two columns. The left column, titled "Project Resource List", contains four items: "Documentation" (Latest Product Documentation), "Source Code" (SVN), "Library" (Read Articles, Presentations, Tutorials etc), and "Mailing Lists" (Participate & be Notified of Project Activities). The right column, titled "WSO2 Mashup", features a "WSO2 Mashup Se" logo, a description "The WSO2 Mashup S services Mashups in", and a "Features" section with a list of bullet points: "Support for languages", "Trivial deplo", "Automatic a schema, pol", "A set of gate POX/REST V", and "Human con".

13. WSO2

WSO2 is one of the open source mashup solutions. It offers data mashup service and a platform for creating, deploying, and consuming Web services Mashups. It provides some features like "Support for consuming and deploying services", "Automatic and UI-based generation of Web services artifacts", "Getting results various user interfaces as web pages, portals, e-mail, Instant Messenger service, SMS".

Read more at: <http://wso2.org/projects/mashup>

Mashup Tools



MARKETPLACE

SOLUTIONS

PRODUCTS

DEVELOPERS

PRODUCTS | IRONCLOUD PLATFORM | **DESKTOP TOOLS** | APPLICATIONS |

StrikeIron SOA Express™ for Excel

"Bringing SOA to the Desktop"

New! Now includes support for Microsoft® Office Excel® 2007 and Windows Vista®!

Bring your SOA to life in minutes! Build rich composite applications. Traditionally this has only been possible by hard-wiring Web services. First drag-and-drop composite application builder available for Microsoft Office Excel 2007 available within any SOA, enabling rich "live" powerful applications.

Now, I.T. can put the power of live data directly to the desktop with "live" Excel workbooks with real-time data available all over the org: Web.

"It's tough to imagine an easier way to hook things up short of the Gunderloy, ADT Magazine"

14. StrikeIron SOA Express

SOA Express is the mashup tool which is fully based on Microsoft Excel. This tool is good to create rich composite applications within Excel. It eliminates the needs of purchasing complex software and creates lightweight enterprise mashups in Excel. It uses simple Excel functionality for web services data management. It claims that with SOA Express there is no need of programming, has a minimal learning curve who already knows Excel, provides easy access for service-enabled legacy applications and increases ROI on SOA investments.

Read more at: http://strikeiron.com/tools/tools_soaexpress.aspx

Mashup Tools

IBM

Home | Products | Services & industry solutions | Support & download

alphaWorks > alphaWorks Services

QEDWiki
A powerful browser-based mashup tool

OVERVIEW | **FAQ** | FORUM | RATE IT

→ **Try It Now** (Registration required.) - [View Demo](#)

QEDWiki 1.3 is now available

Update Oct. 9, 2007: Try the new downloadable version of [IBM@ Mashup Hub](#), IBM DAMIA, and QEDWiki.

Important Note
The following application may not be fully accessibility-compliant or bug-prototype or proof of concept currently under development in IBM research.

15. IBM's QEDWiki

QEDWiki is one of the most capable mashup tool. QEDWiki is a browser-based tool to create mashups. QEDWiki is a lightweight mashup maker written in PHP 5 and hosted on a LAMP, WAMP, or MAMP stack. It provides several benefits to the different type of users like "Mashup assemblers", "Content providers", "Mashup enablers", "Mashup maker administrators".

Benefits to Mashup assemblers are: "simplicity of virtual workspace construction", "ease of situational application development", "rich user experience", "freeform data input with revision tracking" etc. Benefits to Content providers are: "simple content sharing", "detailed access control", "quick, DIY content aggregation" etc. Benefits to Mashup enablers are: "encapsulation and invocation of external services", "viral application development and deployment", "extensibility", "ability to import or stream external data sources", "lower programmer skill required", "increased productivity" etc. Benefits to Mashup maker administrators are: "an enabler for SaaS repository consumption", "ease of installation", "minimal start-up cost for infrastructure" etc.

Read more at: <http://services.alphaworks.ibm.com/qedwiki/>

Mashups: The new breed of web app

An introduction to mashups

Level: Introductory

Duane Mernll (duane@duanemernll.com), Writer, Freelance

08 Aug 2006

Updated 16 Oct 2006

Mashups are an exciting genre of interactive Web applications that draw upon content retrieved from external data sources to create entirely new and innovative services. They are a hallmark of the second generation of Web applications informally known as Web 2.0. This introductory article explores what it means to be a mashup, the different classes of popular mashups constructed today, and the enabling technologies that mashup developers leverage to create their applications. Additionally, you'll see many of the emerging technical and social challenges that mashup developers face.

Introduction

A new breed of Web-based data integration applications is sprouting up all across the Internet. Colloquially termed mashups, their popularity stems from the emphasis on interactive user participation and the monster-of-Frankenstein-like manner in which they aggregate and stitch together third-party data. The sprouting metaphor is a reasonable one; a mashup Web site is characterized by the way in which it spreads roots across the Web, drawing upon content and functionality retrieved from data sources that lay outside of its organizational boundaries.

This vague data-integration definition of a mashup certainly isn't a rigorous one. A good insight as to what makes a mashup is to look at the etymology of the term: it was borrowed from the pop music scene, where a mashup is a new song that is mixed from the vocal and instrumental tracks from two different source songs (usually belonging to different genres). Like these "bastard pop" songs, a mashup is an unusual or innovative composition of content (often from unrelated data sources), made for human (rather than computerized) consumption.

So, what might a mashup look like?

The ChicagoCrime.org Web site is a great intuitive example of what's called a *mapping mashup*. One of the first mashups to gain widespread popularity in the press, the Web site mashes crime data from the Chicago Police Department's online database with cartography from Google Maps. Users can interact with the mashup site, such as instructing it to graphically display a map containing pushpins that reveal the details of all recent burglary crimes in South Chicago. The concept and the presentation are simple, and the composition of crime and map data is visually powerful.

In Mashup genres, you'll survey the popular genres of mashups, including mapping mashups. Related technologies overviews the technology landscape that relates to the construction and operation of mashups. Technical challenges and Social challenges present the eminent technical and social challenges, respectively, affecting mashups.

Mashup genres

In this section, I give a brief survey of the prominent mashup genres.

Mashups: The new breed of web app

Mapping mashups

In this age of information technology, humans are collecting a prodigious amount of data about things and activities, both of which are wont to be annotated with locations. All of these diverse data sets that contain location data are just screaming to be presented graphically using maps. One of the big catalysts for the advent of mashups was Google's introduction of its Google Maps API. This opened the floodgates, allowing Web developers (plus hobbyists, tinkerers, and others) to mash all sorts of data (everything from nuclear disasters to Boston's CowParade cows) onto maps. Not to be left out, APIs from Microsoft (Virtual Earth), Yahoo (Yahoo Maps), and AOL (MapQuest) shortly followed.

Video and photo mashups

The emergence of photo hosting and social networking sites like Flickr with APIs that expose photo sharing has led to a variety of interesting mashups. Because these content providers have metadata associated with the images they host (such as who took the picture, what it is a picture of, where and when it was taken, and more), mashup designers can mash photos with other information that can be associated with the metadata. For example, a mashup might analyze song or poetry lyrics and create a mosaic or collage of relevant photos, or display social networking graphs based upon common photo metadata (subject, timestamp, and other metadata.). Yet another example might take as input a Web site (such as a news site like CNN) and render the text in photos by matching tagged photos to words from the news.

Search and Shopping mashups

Search and shopping mashups have existed long before the term mashup was coined. Before the days of Web APIs, comparative shopping tools such as BizRate, PriceGrabber, MySimon, and Google's Froogle used combinations of business-to-business (b2b) technologies or screen-scraping to aggregate comparative price data. To facilitate mashups and other interesting Web applications, consumer marketplaces such as eBay and Amazon have released APIs for programmatically accessing their content.

News mashups

News sources (such as the New York Times, the BBC, or Reuters) have used syndication technologies like RSS and Atom (described in the next section) since 2002 to disseminate news feeds related to various topics. Syndication feed mashups can aggregate a user's feeds and present them over the Web, creating a personalized newspaper that caters to the reader's particular interests. An example is Diggdot.us, which combines feeds from the techie-oriented news sources Digg.com, Slashdot.org, and Del.icio.us.

Related technologies

This section gives an overview of the technologies that are facilitating the development of mashups. For further information about any of these technologies, consult Resources at the end of this article.

The architecture

A mashup application is architecturally comprised of three different participants that are logically and physically disjoint (they are likely separated by both network and organizational boundaries): API/content providers, the mashup site, and the client's Web browser.

Mashups: The new breed of web app

- The API/content providers. These are the (sometimes unwitting) providers of the content being mashed. In the ChicagoCrime.org mashup example, the providers are Google and the Chicago Police Department. To facilitate data retrieval, providers often expose their content through Web-protocols such as REST, Web Services, and RSS/Atom (described below). However, many interesting potential data-sources do not (yet) conveniently expose APIs. Mashups that extract content from sites like Wikipedia, TV Guide, and virtually all government and public domain Web sites do so by a technique known as *screen scraping*. In this context, screen scraping connotes the process by which a tool attempts to extract information from the content provider by attempting to parse the provider's Web pages, which were originally intended for human consumption.
- The mashup site. This is where the mashup is hosted. Interestingly enough, just because this is where the mashup logic resides, it is not necessarily where it is executed. On one hand, mashups can be implemented similarly to traditional Web applications using server-side dynamic content generation technologies like Java servlets, CGI, PHP or ASP.

Alternatively, mashed content can be generated directly within the client's browser through client-side scripting (that is, JavaScript) or applets. This client-side logic is often the combination of code directly embedded in the mashup's Web pages as well as scripting API libraries or applets (furnished by the content providers) referenced by these Web pages. Mashups using this approach can be termed *rich internet applications* (RIAs), meaning that they are very oriented towards the interactive user-experience. (Rich internet applications are one hallmark of what's now being termed "Web 2.0", the next generation of services available on the World Wide Web.) The benefits of client-side mashing include less overhead on behalf of the mashup server (data can be retrieved directly from the content provider) and a more seamless user-experience (pages can request updates for portions of their content without having to refresh the entire page). The Google Maps API is intended for access through browser-side JavaScript, and is an example of client-side technology.

Often mashups use a combination of both server and client-side logic to achieve their data aggregation. Many mashup applications use data that is supplied directly to them by their user base, making (at least) one of the data sets local. Additionally, performing complex queries on multiple-sourced data (such as "Show me the average purchase price for real estate bought by actors who have co-starred in movies with Kevin Bacon") requires computation that would be infeasible to perform within the client's Web browser.

- The client's Web browser. This is where the application is rendered graphically and where user interaction takes place. As described above, mashups often use client-side logic to assemble and compose the mashed content.

Ajax

There is some dispute over whether the term Ajax is an acronym or not (some would have it represent "Asynchronous JavaScript + XML"). Regardless, Ajax is a Web application model rather than a specific technology. It comprises several technologies focused around the asynchronous loading and presentation of content:

- XHTML and CSS for style presentation
- The Document Object Model (DOM) API exposed by the browser for dynamic display and interaction
- Asynchronous data exchange, typically of XML data
- Browser-side scripting, primarily JavaScript

Mashups: The new breed of web app

When used together, the goal of these technologies is to create a smooth, cohesive Web experience for the user by exchanging small amounts of data with the content servers rather than reload and re-render the entire page after some user action. You can construct Ajax engines for mashups from various Ajax toolkits and libraries (such as Sajax or Zimbra), usually implemented in JavaScript. The Google Maps API includes a proprietary Ajax engine, and the effect it has on the user experience is powerful: it behaves like a truly local application in that there are no scrollbars to manipulate or translation arrows that force page reloads.

Web protocols: SOAP and REST

Both SOAP and REST are platform neutral protocols for communicating with remote services. As part of the service-oriented architecture paradigm, clients can use SOAP and REST to interact with remote services without knowledge of their underlying platform implementation: the functionality of a service is completely conveyed by the description of the messages that it requests and responds with.

SOAP is a fundamental technology of the Web Services paradigm. Originally an acronym for Simple Object Access Protocol, SOAP has been re-termed *Services-Oriented Access Protocol* (or just SOAP) because its focus has shifted from object-based systems towards the interoperability of message exchange. There are two key components of the SOAP specification. The first is the use of an XML message format for platform-agnostic encoding, and the second is the message structure, which consists of a header and a body. The header is used to exchange contextual information that is not specific to the application payload (the body), such as authentication information. The SOAP message body encapsulates the application-specific payload. SOAP APIs for Web services are described by WSDL documents, which themselves describe what operations a service exposes, the format for the messages that it accepts (using XML Schema), and how to address it. SOAP messages are typically conveyed over HTTP transport, although other transports (such as JMS or e-mail) are equally viable.

REST is an acronym for *Representational State Transfer*, a technique of Web-based communication using just HTTP and XML. Its simplicity and lack of rigorous profiles set it apart from SOAP and lend to its attractiveness. Unlike the typical verb-based interfaces that you find in modern programming languages (which are composed of diverse methods such as `getEmployee()`, `addEmployee()`, `listEmployees()`, and more), REST fundamentally supports only a few operations (that is POST, GET, PUT, DELETE) that are applicable to all pieces of information. The emphasis in REST is on the pieces of information themselves, called resources. For example, a resource record for an employee is identified by a URI, retrieved through a GET operation, updated by a PUT operation, and so on. In this way, REST is similar to the document-literal style of SOAP services.

Screen scraping

As mentioned earlier, lack of APIs from content providers often force mashup developers to resort to screen scraping in order to retrieve the information they seek to mash. *Scraping* is the process of using software tools to parse and analyze content that was originally written for human consumption in order to extract semantic data structures representative of that information that can be used and manipulated programmatically. A handful of mashups use screen scraping technology for data acquisition, especially when pulling data from the public sectors. For example, real-estate mapping mashups can mash for-sale or rental listings with maps from a cartography provider with scraped "comp" data obtained from the county records office. Another mashup project that scrapes data is XMLTV, a collection of tools that aggregates TV listings from all over the world.

Screen scraping is often considered an inelegant solution, and for good reasons. It has two primary inherent drawbacks. The first is that, unlike APIs with interfaces, scraping has no specific programmatic

Mashups: The new breed of web app

contract between content-provider and content-consumer. Scrapers must design their tools around a model of the source content and hope that the provider consistently adheres to this model of presentation. Web sites have a tendency to overhaul their look-and-feel periodically to remain fresh and stylish, which imparts severe maintenance headaches on behalf of the scrapers because their tools are likely to fail.

The second issue is the lack of sophisticated, re-usable screen-scraping toolkit software, colloquially known as *scrAPIs*. The dearth of such APIs and toolkits is largely due to the extremely application-specific needs of each individual scraping tool. This leads to large development overheads as designers are forced to reverse-engineer content, develop data models, parse, and aggregate raw data from the provider's site.

Semantic Web and RDF

The inelegant aspects of screen scraping are directly traceable to the fact that content created for human consumption does not make good content for automated machine consumption. Enter the Semantic Web, which is the vision that the existing Web can be augmented to supplement the content designed for humans with equivalent machine-readable information. In the context of the Semantic Web, the term information is different from data; data becomes information when it conveys meaning (that is, it is understandable). The Semantic Web has the goal of creating Web infrastructure that augments data with metadata to give it meaning, thus making it suitable for automation, integration, reasoning, and re-use.

The W3C family of specifications collectively known as the Resource Description Framework (RDF) serves this purpose of providing methodologies to establish syntactic structures that describe data. XML in itself is not sufficient; it is too arbitrary in that you can code it in many ways to describe the same piece of data. RDF-Schema adds to RDF's ability to encode concepts in a machine-readable way. Once data objects can be described in a data model, RDF provides for the construction of relationships between data objects through subject-predicate-object triples ("subject S has relationship R with object O"). The combination of data model and graph of relationships allows for the creation of ontologies, which are hierarchical structures of knowledge that can be searched and formally reasoned about. For example, you might define a model in which a "carnivore-type" as a subclass of "animal-type" with the constraint that it "eats" other "animal-type", and create two instances of it: one populated with data concerning cheetahs and polar bears and their habitats, another concerning gazelles and penguins and their respective habitats. Inference engines might then "mash" these separate model instances and reason that cheetahs might prey on gazelles but not penguins.

RDF data is quickly finding adoption in a variety of domains, including social networking applications (such as FOAF — Friend of a Friend) and syndication (such as RSS, which I describe next). In addition, RDF software technology and components are beginning to reach a level of maturity, especially in the areas of RDF query languages (such as RDQL and SPARQL) and programmatic frameworks and inference engines (such as Jena and Redland).

RSS and ATOM

RSS is a family of XML-based syndication formats. In this context, syndication implies that a Web site that wants to distribute content creates an RSS document and registers the document with an RSS publisher. An RSS-enabled client can then check the publisher's feed for new content and react to it in an appropriate manner. RSS has been adopted to syndicate a wide variety of content, ranging from news articles and headlines, changelogs for CVS checkins or wiki pages, project updates, and even audiovisual data such as radio programs. Version 1.0 is RDF-based, but the

Mashups: The new breed of web app

most recent, version 2.0, is not.

Atom is a newer, but similar, syndication protocol. It is a proposed standard at the Internet Engineering Task Force (IETF) and seeks to maintain better metadata than RSS, provide better and more rigorous documentation, and incorporates the notion of constructs for common data representation.

These syndication technologies are great for mashups that aggregate event-based or update-driven content, such as news and weblog aggregators.

Technical Challenges

Like any other data integration domain, mashup development is replete with technical challenges that need to be addressed, especially as mashup applications become more feature- and functionality-rich. This section touches on a handful of these challenges, some of which you can address and mitigate, while others are open issues.

Data Integration Challenges: Semantic Meaning and Data Quality

Qualitative surveys suggest that the number one enterprise IT concern today is data integration within the enterprise virtual organization. (In this context, I use the term *virtual organization* to mean a composition of federated business units, each contained within its own administrative domain.) Like many enterprise IT managers who find themselves up to the task of integrating legacy data sources (for example, to create corporate dashboards that reflect current business conditions), mashup developers are faced with the analogous challenges of deriving shared semantic meaning between heterogeneous data sets. Therefore, to get an idea for what mashup developers have in store, you need look no further than the storied integration challenges faced by enterprise IT.

For example, translation systems between data models must be designed. When converting data into common forms, reasonable assumptions often have to be made when the mapping is not a complete one (for example, one data source might have a model in which an address-type contains a country-field, whereas another does not). Already challenging, this is exacerbated by the fact that the mashup developers might not be domain experts on the source data models because the models are third-party to them, and these reasonable assumptions might not be intuitive or clear.

In addition to missing data or incomplete mappings, the mashup designer might discover that the data they wish to integrate is not suitable for machine automation; that it needs cleansing. For example, law enforcement arrest records might be entered inconsistently, using common abbreviations for names (such as "mkt sqr" in one record and "Market Square" in another), making automated reasoning about equality difficult, even with good heuristics. Semantic modeling technologies, such as RDF, can help ease the problem of automatic reasoning between different data sets, provided that it is built-in to the data-store. Legacy data sources are likely to require much human effort in terms of analysis and data cleansing before they can be availed to semantic modeling technologies.

Mashup developers might also have to contend with several issues that IT integration managers might not, one of which is data pollution. As part of their application design, many mashups solicit public user input. As evidenced in the wiki application domain, this is a double-edged blade: it can be quite powerful because it enables open contribution and best-of-breed data evolution, yet it can be subject to inconsistent, incorrect, or intentionally misleading data entry. The latter can cast doubts on data trustworthiness, which can ultimately compromise the value provided by the mashup. Another host of integration issues facing mashup developers arise when screen scraping techniques

Mashups: The new breed of web app

must be used for data acquisition. As discussed in the previous section, deriving parsing and acquisition tools and data models requires significant reverse-engineering effort. Even in the best case where these tools and models can be created, all it takes is a re-factoring of how the source site presents its content (or mothballing and abandonment) to break the integration process, and cause mashup application failure.

Component Challenges

The Ajax model of Web development can provide a much richer and more seamless user experience than the traditional full-page-refresh, but it poses some difficulties as well. At its fundamentals, Ajax entails using the browser's client-side scripting capabilities in conjunction with its DOM to achieve a method of content delivery that was not entirely envisioned by the browser's designers. (Perhaps this hack-like nature of Ajax lends to its appeal.) However, this subjects Ajax-based applications to the same browser compatibility issues that have plagued Web designers ever since Microsoft created Internet Explorer. For example, Ajax engines make use of an XMLHttpRequest object to exchange data asynchronously with remote servers. In Internet Explorer 6, this object is implemented with ActiveX rather than native JavaScript, which requires that ActiveX be enabled.

A more fundamental requirement is that Ajax requires that JavaScript be enabled within the user's browser. This might be a reasonable assumption for the majority of the population, but there are certainly users who use browsers or automated tools that either do not support JavaScript or do not have it enabled. One such set of tools are the robots, spiders, and Web crawlers that aggregate information for Internet and intranet search engines. Without graceful degradation, Ajax-based mashup applications might find themselves missing out on both a minority user base as well as search engine visibility.

The use of JavaScript to asynchronously update content within the page can also create user interface issues. Because content is no longer necessarily linked to the URL in the browser's address bar, users might not experience the functionality that they normally expect when they use the browser's BACK button, or the BOOKMARK feature. And, although Ajax can reduce latency by requesting incremental content updates, poor designs can actually hinder the user experience, such as when the granularity of update is small enough that the quantity and overhead of updates saturate the available resources. Also, take care to support the user (for example, with visual feedback such as progress bars) while the interface loads or content is updated.

As with any distributed, cross-domain application, mashup developers and content providers alike will also need to address security concerns. The notion of identity can prove to be a sticky subject, as the traditional Web is primarily built for anonymous access. Single-signon is a desirable feature, but there are a multitude of competing technologies (ranging from Microsoft Passport to the Liberty Alliance), thus creating disjointed identity namespaces that you must integrate as well. Content providers are likely to employ authentication and authorization schemes (which require the notion of secure identity or securely identifiable attributes) in their APIs to enforce business models that involve paid subscriptions or sensitive data. Sensitive data is also likely to require confidentiality (that is, encryption), and you must take care when you mash it with other sources to not put it at risk. Identity will also be crucial for auditing and regulatory compliance. Additionally, with data integration happening both on the server and client-side, identity and credential delegation from the user to the mashup service might become a requirement.

Social Challenges

In addition to the technical challenges described in the previous section, social issues have (or will) surface as mashups become more popular.

Mashups: The new breed of web app

One of the biggest social issues facing mashup developers is the tradeoff between the protection of intellectual property and consumer privacy versus fair-use and the free flow of information. Unwitting content providers (targets of screen scraping), and even content providers who expose APIs to facilitate data retrieval might determine that their content is being used in a manner that they do not approve of. For a good review of Web aggregation and regulations, see [Recourses](#).

The mashup Web application genre is still in its infancy, with hobbyist developers who produce many mashups in their spare time. These developers might not be cognizant of (or concerned with) issues such as security. Additionally, content providers are only beginning to see the value in providing APIs for machine-based content access, and many do not consider them a core business focus. This combination can yield poor software quality, as priorities such as testing and quality assurance take the backseat to proof-of-concept and innovation. The community as a whole will have to work together to assemble open standards and reusable toolkits in order to facilitate mature software development processes.

Before mashups can make the transition from cool toys to sophisticated applications, much work will have to go into distilling robust standards, protocols, models, and toolkits. For this to happen, major software development industry leaders, content providers, and entrepreneurs will have to find value in mashups, which means viable business models. API providers will need to determine whether or not to charge for their content, and if so, how (for example, by subscription or by per-use). Perhaps they will provide varying levels of quality-of-service. Some marketplace providers, such as eBay or Amazon, might find that the free use of their APIs increases product movement. Mashup developers might look for an ad-based revenue model, or perhaps build interesting mashup applications with the goal of being acquired.

Summary

Mashups are certainly an exciting new genre of Web applications. The combination of data modeling technologies stemming from the Semantic Web domain and the maturation of loosely-coupled, service-oriented, platform-agnostic communication protocols is finally providing the infrastructure needed to start developing applications that can leverage and integrate the massive amount of information that is available on the Web. As mashup applications gain higher visibility, it will be interesting to see how the genre impacts social issues such as fair-use and intellectual property as well as other application domains that integrate data across organizational boundaries, such as grid computing and business-to-business workflow management.

For a deeper-dive into mashup development, stay tuned for the launching of a new series of tutorials on developerWorks that will teach you how to construct your own mashups. In fact, the series will even teach you how to use Semantic Web technology and ontologies to enable others to create their own mashups.

Advertise with JavaJazzUp

We are the top most providers of technology stuffs to the java community. Our technology portal network is providing standard tutorials, articles, news and reviews on the Java technologies to the industrial technocrats. Our network is getting around 3 million hits per month and its increasing with a great pace.

For a long time we have endeavored to provide quality information to our readers. Furthermore, we have succeeded in the dissemination of the information on technical and scientific facets of IT community providing an added value and returns to the readers.

We have serious folks that depend on our site for real solutions to development problems.

JavaJazzUp Network comprises of :

<http://www.roseindia.net>
<http://www.newstrackindia.com>
<http://www.javajazzup.com>
<http://www.allcooljobs.com>

Advertisement Options:

Banner	Size	Page Views	Monthly
Top Banner	470*80	5,00,000	USD 2,000
Box Banner	125 * 125	5,00,000	USD 800
Banner	460x60	5,00,000	USD 1,200
Pay Links		Un Limited	USD 1,000
Pop Up Banners		Un Limited	USD 4,000

The <http://www.roseindia.net> network is the "real deal" for technical Java professionals. Contact me today to discuss your customized sponsorship program. You may also ask about advertising on other Technology Network.

Deepak Kumar
deepak@roseindia.net

India's Cheapest web Service Provider

Web Packages

Package (in INR)

Package Name	Price
Starter Package	Rs 5,100 + Taxes Extra
Business Package	Rs 9,111 + Taxes Extra
Corporate Package	Rs 22,500 + Taxes Extra
Smart Package	Rs 45,500 + Taxes Extra

* Domain Registration free with every package

Packages Specifications

Starter Package

- 5 Web Pages
- 10 Stock Images
- 5 POP 3 Accounts
- 50 MB Hosting Space
- Unlimited Edits

Business Package

- 10 Web Page Design
- Flash Site Animation
- 25 Stock Images
- 10 POP3 Accounts
- 50 MB Hosting Space
- Unlimited Edits

Corporate Package

- 25 Web Page Design
- Flash Site Animation
- Flash Intro Page
- 50 Stock Images
- 25 POP3 Accounts
- 100 MB Hosting Space
- Unlimited Edit

Smart Package

- 200 Web Page Design
- Catalog with 200 items
- Flash Site Animation
- Flash Intro Page
- Unlimited Stock Images
- 50 POP3 Accounts
- 250 MB Hosting Space
- Unlimited Edits

 **RoseIndia**
Technologies Pvt. Ltd.

Logon to: <http://www.roseindia.net/services/>

Valued JavaJazzup Readers Community

We invite you to post Java-technology oriented stuff. It would be our pleasure to give space to your posts in JavaJazzup.

Contribute to Readers Forum

If theres something youre curious about, were confident that your curiosity, combined with the knowledge of other participants, will be enough to generate a useful and exciting Readers Forum. If theres a topic you feel needs to be discussed at JavaJazzup, its up to you to get it discussed.

Convene a discussion on a specific subject

If you have a topic youd like to talk about . Whether its something you think lots of people will be interested in, or a narrow topic only a few people may care about, your article will attract people interested in talking about it at the Readers Forum. If you like, you can prepare a really a good article to explain what youre interested to tell java technocrats about.

Sharing Expertise on Java Technologies

If youre a great expert on a subject in java, the years you spent developing that expertise and want to share it with others. If theres something youre an expert on that you think other technocrats might like to know about, wed love to set you up in the Readers Forum and let people ask you questions.

Show your innovation

We invite people to demonstrate innovative ideas and projects. These can be online or technology-related innovations that would bring you a great appreciations and recognition among the java technocrats around the globe.

Hands-on technology demonstrations

Some people are Internet experts. Some are barely familiar with the web. If you'd like to show others aroud some familiar sites and tools, that would be great. It would be our pleasure to give you a chance to provide your demonstrations on such issues : How to set

up a blog, how to get your images onto Flickr, How to get your videos onto YouTube, demonstrations of P2P software, a tour of MySpace, a tour of Second Life (or let us know if there are other tools or technologies you think people should know about...).

Present a question, problem, or puzzle

Were inviting people from lots of different worlds. We do not expect everybody at Readers Forum to be an expert in some areas. Your expertise is a real resource you may contribute to the Java Jazzup. We want your curiosity to be a resource, too. You can also present a question, problem, or puzzle that revolves around java technologies along with their solution that you think would get really appreciated by the java readers around the globe.

Post resourceful URLs

If you think you know such URL links which can really help the readers to explore their java skills. Even you can post general URLs that you think would be really appreciated by the readers community.

Anything else

If you have another idea for something youd like to do, talk to us. If you want to do something that we havent thought of, have a crazy idea, wed really love to hear about it. Were open to all sorts of suggestions, especially if they promote readers participation.

All Cool Jobs
<http://www.allcooljobs.com>

***SPOTLIGHT YOUR TALENT
IN THE BEST JOB SITE***





People see world through their eyes.
We see world through people's eyes.
“We see beyond boundary.”

