zapthink zapnote

Doc. ID: ZTZN-1227
Released June 19, 2008

ZAPTHINK ZAPNOTE™

KAPOW TECHNOLOGIES: TURNING THE WEB INTO THE WORLD'S LARGEST DATABASE

Analyst: David S. Linthicum

Abstract

The use of unstructured data within enterprises today has been more art than science, and in most cases an opportunity that many people have yet to understand. Truth-be-told, there are terabytes and terabytes of useful information throughout the enterprise and the World Wide Web that if placed into the proper context could have an ROI much higher than existing data mining approaches.

Access to this unstructured information has been out of reach of many enterprises today due to the cost and complexity of the technology. However, with the Kapow Technologies offering, enterprises now have an on-demand mechanism for searching, grabbing, refining, and understanding unstructured data, and can place them in the context of enterprise data for even more value.

In addition, in the emerging world of Service-Oriented Architecture (SOA), Kapow is also providing customers with the ability to address unstructured data as Services, thus providing an on-ramp for any data source for use within mashups in the SOA environment.

The Value of Unstructured Data

You can think of unstructured data (or unstructured information) as masses of computerized information which do not have a data structure or have one that is machines cannot easily read. In fact, most of the data people deal with is unstructured. We can also refer to data with some form of structure as unstructured data if the structure is not helpful for the desired processing task. For example, an HTML Web page is highly structured, but this structure is often oriented toward formatting, not the use of the data. To that point, most Weblogs are created with format in mind, and not how the information will be useful to others who wish to process it.

On the Web, the wide distribution of information is the most complex problem to solve, since you must consider information feeds from hundreds, sometimes thousands of sources. The number of information sources drives the need for a common access method, creating a consistent data presentation that spans the system. The complex and diverse ways to access that information goes to the reality that the ways we store and access information varies greatly from source to source. While some information sources use very well designed and efficient information access mechanisms, most don't.

Moreover, you need to consider the wide array of information structures, including text, object, sequencing, etc., as well as enabling technology. Indeed, the structures, formats, and enabling technologies (such as databases) are as wide ranging as the types of information they manage. Accessing information contained within a wide array of structures requires a programmatic approach to data access. This programmatic approach allows you to place the volatility of the information into a single domain which is controlled programmatically to account for the differences between semi-structured and unstructured data.

Kapow Technologies is approaching this problem by both understanding the differences between unstructured and semi-structured data, and then understanding the proper approaches to manage the underlying information through programmatic abstraction. Also, Kapow provides an mechanism to programmatically access structured and unstructured data, with the ability to leverage that information as Services in the context of SOA and also for mashups. Moreover, Kapow delivers both an on-premise or an on-demand version of their product, thus allowing the customer to select the best approach for their architecture.

The Kapow Technology

Kapow provides an approach to unstructured data that provides the user with the ability to cull only the information they want and, in essence, place a logical structure around it to separate it from data that are largely unstructured. This task is accomplished through the use of pattern matching and other programmatic techniques that can locate, identify, validate, and abstract information from an unstructured format, as if it were highly structured.

Beyond unstructured data there is semi-structured data. Indeed, with the exception of text documents, the formats of these documents generally conform to a standard that offers the

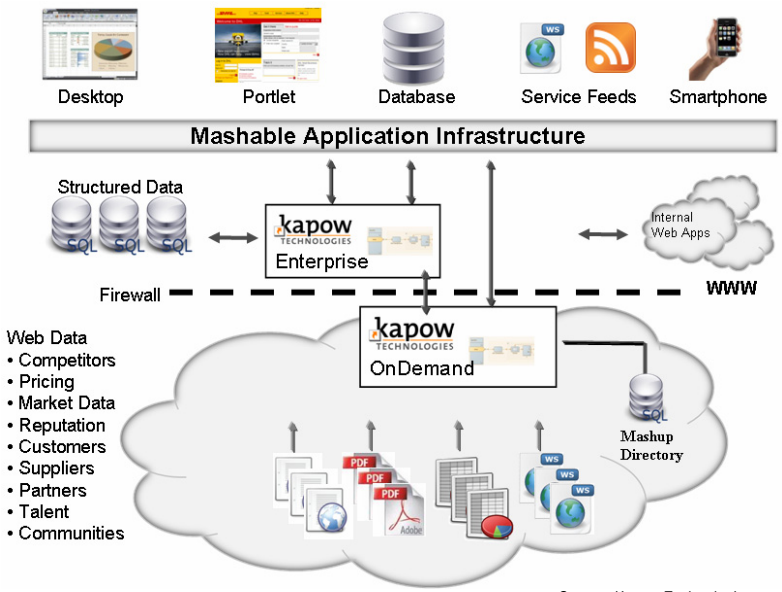
option of metadata. Thus, the core information could include information such as author and time of creation, and it can be easily stored in a structured data management system.

The Kapow product offering falls into two categories, providing a different set of value propositions for customers:

- Enterprise Products, including the *Kapow Mashup Server* family of products, available in three different product editions. Customers install, operate, and manage these products on-site. *Kapow Mashup Server* provides visual programming language that automates Web integration tasks. Using this technology, customers can take advantage of the unstructured information creating an information access mechanism that spans hundreds, perhaps thousands of Web pages. Moreover, Kapow provides an API to unlock unstructured data sources, including use of a fine grained data element access mechanism. In addition, Kapow provides data extraction, migration, and transformation capabilities that are deployable as batch or real-time operations.
- *OnDemand* Service, which is an inclusive hosted service for the collection of real time Web intelligence. Kapow On-Demand is a Web-based, hosted service that enables automated, high-volume collection of web intelligence and market data to help customers make informed business decisions. Web content integrates into existing applications and IT infrastructure, extending its current capabilities and creating new business value. The method of consuming this technology is particularly useful when architecture requirements lead to the Web as a platform or for companies that find it much more productive and cost effective to place the features that Kapow provides outside of the firewall.

All Kapow products support standards, including the ability to leverage information abstraction (typically converted unstructured information), as a service using SOAP or REST. Moreover, Kapow provides a complete solution, including the ability to leverage these Services using traditional office automation applications such as Microsoft Excel. Also, the ability to generate data feeds for mashups or integration that leverages key industry standards such as .NET or Java. Customers can publish data feeds as a portlet into a portal server, as well as create RSS/ATOM mashups feeds, or using Web Services. An illustration of Kapow's architecture is shown in the figure below.

Kapow Technologies' OnDemand Architecture



Source: Kapow Technologies

Kapow Technologies Features

Kapow Technologies

Overview:

Kapow Technologies offers both on-premise and on-demand technology for harvesting mashable public and private Web intelligence Based on their visual scripting method, Kapow OnDemand powers solutions in reputation management, competitive intelligence, asymmetric intelligence, content aggregation, content syndication, content migration, SOA enablement and business automation.

Features:

- **Visual programming language** – automates the task of leveraging Web Integration allowing end users to take advantage of the unstructured information creating an information access mechanism that spans hundreds of Web pages.
- **Application Programming Interface** – unlocks unstructured data sources, including use of a fine grained data element access mechanism. The API provides programmatic and automated access to information. In addition, Kapow provides data extraction, migration, and transformation capabilities that are deployable as batch or real-time operations.
- **Universal Data Access** – Includes the ability to create APIs for unstructured information, and enable programmatic data access features that are reusable from mashup to mashup, application to application. Furthermore, Kapow supports automated data flow, including direct data migration between applications, ERP, CRM, and CMS.

Value Proposition:

- Gives companies the ability to create enterprise mashup applications that provide structured Web intelligence in a matter of minutes.
- Turns the Web into the world's largest database.

The ZapThink Take

ZapThink believes that the use of unstructured data is a valuable component of Service compositions, including mashups. The ability to gather information from blogs, e-tailers, online articles, or other Web-based information allows end users to create powerful and useful applications. For example, applications could include analyzing stock information with blog posts about that stock, or the ability to access hundreds of Web sites with price and inventory information to instantly find a vendor who will provide the most value in the supply chain. The number and types of applications that would find value in leveraging unstructured data is quite large.

The fact of the matter is that most information that has value is also unstructured, and the ability to access that information provides a core strategic advantage. Kapow's approach to leveraging unstructured information, using either an on-premise or on-demand version of their technology creates a valuable layer between the unstructured data and the applications that need to process this

information. This layer includes management of this information, including abstraction and complete programmatic access.

.

Profile: Kapow Technologies June 2008

Funding:
Kennet Venture Partners, NorthCap Partners, Steamboat Ventures, Morgan Staley, SAL Holding, Impress Group

CEO:
John Cimral

Products:
Kapow OnDemand
- Data Collection Edition
- Web 2.0 Edition
- Portal Content Edition

Address:
260 Sheridan Avenue
Suite 420
Palo Alto, CA 94306

URL:
http://www.kapowtech.com

Phone:
800-805-0828

Contact:
sales@kapowtech.com

Related Research

- *Composing Services into Enterprise Mashups* White Paper (WP-0144)
- *JackBe* ZapNote (ZTZN-1226)
- *IDV Solutions* ZapNote (ZTZN-1224)
- *Corizon* ZapNote (ZTZN-1216)
- *NexaWeb* ZapNote (ZTZN-1139)

About ZapThink, LLC

ZapThink is an Enterprise Architecture (EA) strategy advisory firm. As a recognized authority and master of Service-Oriented Architecture (SOA) and EA, ZapThink provides its audience of IT practitioners, consultants, and technology vendors with practical advice, guidance, education, and mentorship solutions that assist companies in leveraging SOA to meet their business needs and presenting viable SOA solutions to the market. We provide this audience a clear roadmap for standards-based, loosely coupled distributed computing – a vision of IT meeting the needs of the agile business.

ZapThink provides IT practitioners strategic insight and practical guidance for addressing critical agility and change management issues leveraging the latest EA and SOA best practices. ZapThink helps these customers put EA and SOA into practice in a rational, well-paced, and best practices-driven manner and helps to validate or recover architecture initiatives that may be heading down an unknown or incorrect path. ZapThink assists with solution vendor, technology, and consultant selection based on in-depth, objective evaluation of the capabilities, strengths, and applicability of the solutions to meet customer needs as they relate to EA initiatives and as they map against emerging best practices. ZapThink enhances its customer's skills by providing education, credentialing, and training to EAs to develop their skills as architects.

ZapThink helps to augment consulting firms' EA offerings and intellectual property by providing guidance on emerging best practices and access to information that supports those practices. ZapThink provides frameworks for product-based consulting based on ZapThink insight and research, such as SOA Implementation Roadmap guidance, Governance Framework development, and SOA Assessments, and provides a means to endorse and validate consulting firm offerings. ZapThink also accelerates consulting firms' efforts to attract, retain, and enhance the skills of EA and SOA talent by providing education and skills development

For solutions vendors, ZapThink provides retained advisory for guidance on product strategy, as well as marketing, visibility, and third-party endorsement benefits through its marketing activities, lead generation activities, and subscription services. ZapThink enables vendors to leverage ZapThink knowledge to transform their offerings in a cost-effective manner.

ZapThink's Managing Partners are widely regarded as the "go to advisors" and leading experts on SOA, EA, and Enterprise 2.0 by vendors, end-users, and the press. Respected for their candid, insightful opinions, they are in great demand as speakers, and have presented at conferences and industry events around the world. They are among the most quoted experts in the IT industry.

ZapThink was founded in October 2000 and is headquartered in Baltimore, Maryland. Its customers include Global 1000 firms and government organizations, as well as many emerging businesses. Its Managing Partners have worked at such firms as IDC, marchFIRST, and ChannelWave, and have sat on the working group committees for standards bodies such as RosettaNet, UDDI, and ebXML.

Call, email, or visit the ZapThink Web site to learn more about how ZapThink can help you to better understand how SOA and Enterprise 2.0 will impact your business or organization.

ZAPTHINK CONTACT:
ZapThink, LLC
108 Woodlawn Road
Baltimore, MD 21210
Phone: +1 (781) 207 0203
Fax: +1 (815) 301 3171
info@zapthink.com