

EMML Changes Everything:

Profitability, Predictability & Performance
through Enterprise Mashups

Sponsored by:



mashupalliance.org

Table of Contents

Introduction	ii
How Do Mashups Work & Why Should I Care?	iv
Why Enterprise Mashups Now?	v
Enter EMML	vi
EMML Brings Profitability, Predictability & Performance to Business	ix
Reduced Costs	x
Increased Maturity	xi
Improved Quality	xiii
Our Digital Universe - Treasures of Data (if we can consume it)	xiv

Introduction

Immediacy is power. Today, businesses have access to exponentially increasing amounts of real-time, near-real time, and historic data. Technologies, markets, and product lines can prove vital or irrelevant at a moment's notice, and organizations operating within the 2.0 environment need to respond to opportunities and challenges alike with agility and speed. Such acceleration of data creation - coupled with the market's natural evolution - requires stakeholders to leverage state of the art business intelligence (BI) in order to remain competitive. Over the last five years the enterprise mashup has steadily become BI's go-to technology, integrating and repurposing often untapped resources to enable well-informed solutions, foster productivity, streamline processes, facilitate innovation, invent new markets, and ultimately transform the way organizations conduct their business analysis.

What is a mashup? Simply put, a mashup is a customizable application that takes seemingly disparate data sets - both static and real-time - and integrates them to create a new data set. Mashups allow the user to tap into siloed information within or outside the organization, turning what was formally a complex, time consuming, and expensive process into one of relative ease and speed, and at minimal cost.

For many organizations, the enterprise mashup is becoming the weapon of choice in their BI arsenal. It runs on a dedicated and secure mashup server, which features tools for collecting and analyzing mashup data within the security of a corporate firewall.

Within almost any industry requiring knowledge work, mashups open infinite possibilities for analysis and operations. The financial services sector was one of the first to leverage mashup technology for immediate results. In 2006 at JP Morgan Chase, rudimentary mashup technology was used to bring together real-time commodities performance data and integrate it with analytical tools that allowed securities traders to track as many as 500 portfolios at once. This mashup - called the Trading Algorithm Optimizer, or "TAO" - not only afforded traders the

ability to make informed decisions, it also gave them the agility to respond and collaborate in new and interesting ways.¹

To be sure, the true power of TAO didn't come from its technology or its data, it came from the ease with which powerful new tools could be created. The self-described "skunkworks" project allowed professionals within JP Morgan Chase to experiment with combinations of technologies, data, and algorithms, tuning the mashup just-in-time for each problem. This flexibility allowed TAO to move from concept to a fully vetted and realized competitive advantage within seven months.

Mashups replace composite applications creating real business value with low initial investment in minutes, money, or manpower. Today, mashups are becoming big business by providing companies with powerful analysis and launching stand-alone products. As of October 2009, Programmable Web, a website dedicated to tracking mashups for general use, listed nearly 1,500 public and real-time data sources and over 4,000 publicly available and viable mashups.²

The hybridization of information is quickly becoming a corporate staple. The mashup market itself is projected to grow tenfold over the next 5 years as companies seek to create software solutions more rapidly and less expensively while capturing business opportunities that weren't previously possible.³ Fueling this growth is the explosion in available data and the subsequent reduction in mashup creation cost.

¹ <http://www.mediakat.com/JPMorgan.pdf>

² <http://www.programmableweb.com/>

³ http://www.businesswire.com/portal/site/google/?ndmViewId=news_view&newsId=20090921005858&newsLang=en

How Do Mashups Work & Why Should I Care?



Mashups allow both technical and non-technical users to quickly and easily access data, mix it together, test out ideas, validate concepts, and view current status of data or business processes.

To illustrate: You have an idea for a new product. You consider who its customer might be, gather some quick and freely available data such as demographics, income, and topography, run the data through some simple tests, and vet the idea in a matter of minutes. Mashups can drive innovation quickly by letting new ideas be tried at low expense. An effective mashup can be refined and improved, and an idea that doesn't work can be discarded with minimal loss in both cost and effort.

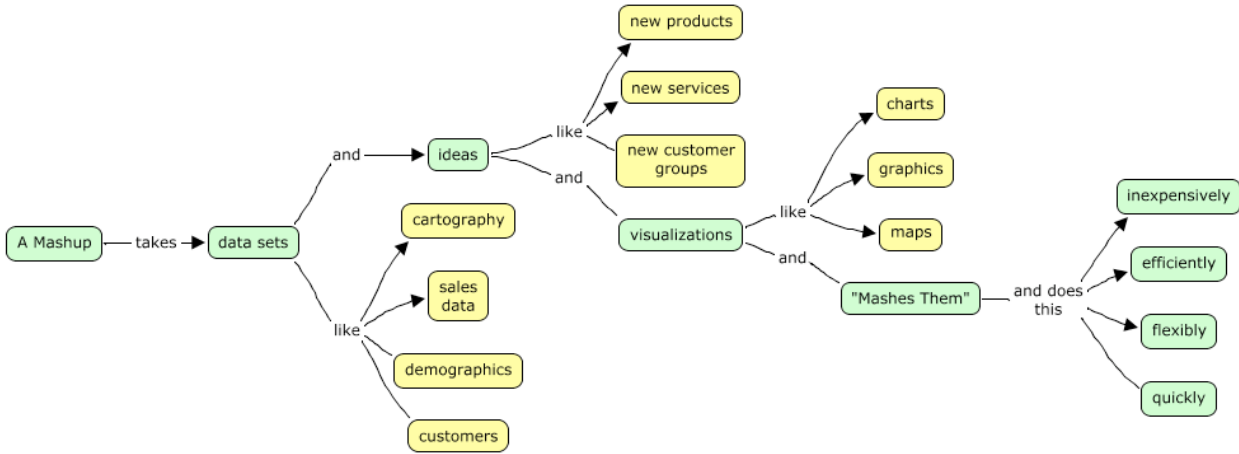
Another example: Your company's fleet vehicles are already feeding location information back to your distribution point for routing and instructions. You could easily repurpose that location data, mashing it up with a map and staffing information from your HR department, to track employee performance.

However, discussing such typical uses for mashups might be missing the point. The mashup's strength lies in discovering the **atypical**, in exploiting data in new ways. The fact is, any information your business needs can be analyzed with a mashup, often more quickly, with minimal effort, and at much lower expense than hiring consultants or using traditional and more time-consuming SOA approaches to do the same work. Rapid experimentation with data leads to invention.

Data can originate from data files (static data) or data feeds (real-time data), and feed into analysis tools as simple as Microsoft Excel or as complex as a BI suite

from Oracle, SAP, or Cisco. Enterprise mashups deal directly with the data, easily obtaining it and feeding it into any system for analysis. The mashups can then be analyzed to drive business goals such as reducing waste, increasing sales, understanding competition, developing new products, or making strategic decisions.

Businesses are generating new data at an unbridled pace. A swelling data inventory both from within and beyond the company often precludes the need for expensive proprietary data. It also means that data is being created - an invaluable commodity in today's knowledge economy - that depends on tools like mashups for rapid analysis in order to see ROI on generating the data in the first place.



Why Enterprise Mashups Now?

The demand for better BI has existed since ancient merchants relied on an abacus to calculate transaction costs. Recently, major BI tool vendors (SAP's Business Objects, Corda Technologies, InetSoft, LogiXML, Microsoft, Microstrategy, SAS and Strategy Companion) funded a study exploring what drove or impeded the use of BI packages. Their findings revealed an 18% growth in the use of BI tools

between 2005 and 2008, but most of this use was relegated to business analysts because the software was simply too complicated for untrained staff.⁴

Historically, data storage was expensive, contained in specialized, encoded formats by the programs that generated them. However, with recent advances in technology, storage cost has dropped considerably. According to hard drive manufacturer Seagate, the price of a megabyte of storage has fallen from \$2000 in 1980 to 2 cents today. This significant reduction per megabyte means we can reconsider the way we store and re-deploy information.⁵

Considering how the price of data storage currently costs much less than a piece of bubblegum, it is no longer a major budgetary factor, allowing businesses to shift their focus to data usability. What was preventing this line of thinking before? High storage costs translated into compact, proprietary files no one could read. The goal was to save money through efficient storage.

Now, with low data storage costs the goal is to store data in formats that facilitate reuse. This has led to open XML-based data storage, but there are many schema and many ways to use XML. Now we have the data - and it's readable - but it isn't predictable. This lack of predictability raises complexity and costs, inhibiting the speed of mashup creation. What's needed is a new data format that is open and predictable.

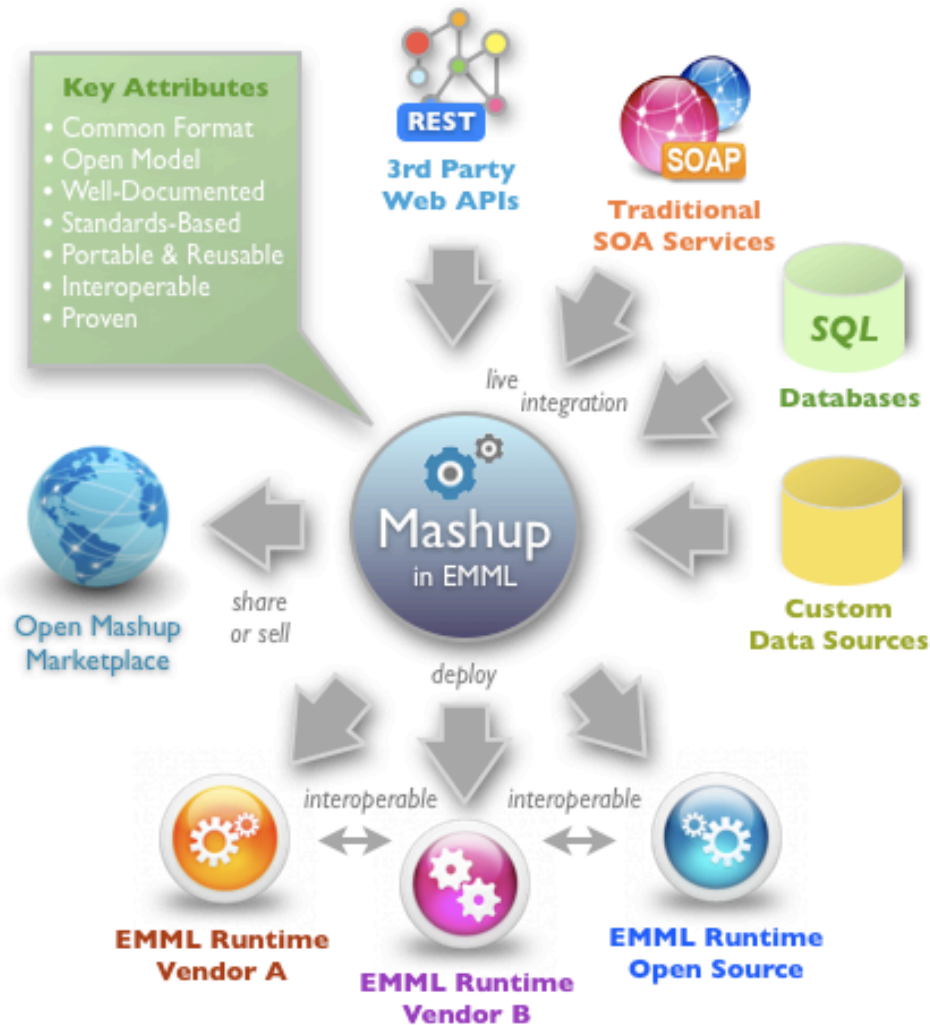
Enter EMMML

Enterprise Mashup Markup Language, or "EMML," is an open "markup language" designed to make data universally readable. Open formats like XML greatly increase the readability and usability of information. The suffix "ML" indicates a markup language, which includes XML (data formats), HTML (the basic language behind web formatting pages), and EMMML.

⁴ <http://www.tdwi.org/research/display.aspx?ID=9152>

⁵ <http://storageeffect.media.seagate.com/2009/03/storage-effect/a-brief-history-of-the-disk-drive/>

The Enterprise Mashup Markup Language (EMML) Creates A Unified Framework for Mashups



From <http://blogs.zdnet.com/Hinchcliffe>

EMML is an open format that naturally encourages data interoperability. Today, data comes from many places in the enterprise including traditional Web services (SOAP), 3rd party Web APIs (REST), and SQL databases. This means that data stored in one application can be easily leveraged in another. Mashups described in EMML exist in a common and predictable format, allowing them to be shared, swapped, or even sold. For mashups, this is the holy grail, as data stored in IT systems as well as on the web can be accessed and integrated via a mashup.

While the browser - the place where mashups were invented - remains a good place to display the data that results from mashups, the model itself is limited when it comes to all the places mashups need to run in a typical enterprise. EMMML-based mashups can run well beyond the browser, including on the server-side.

EMMML's server-side architecture has some inherent benefits:

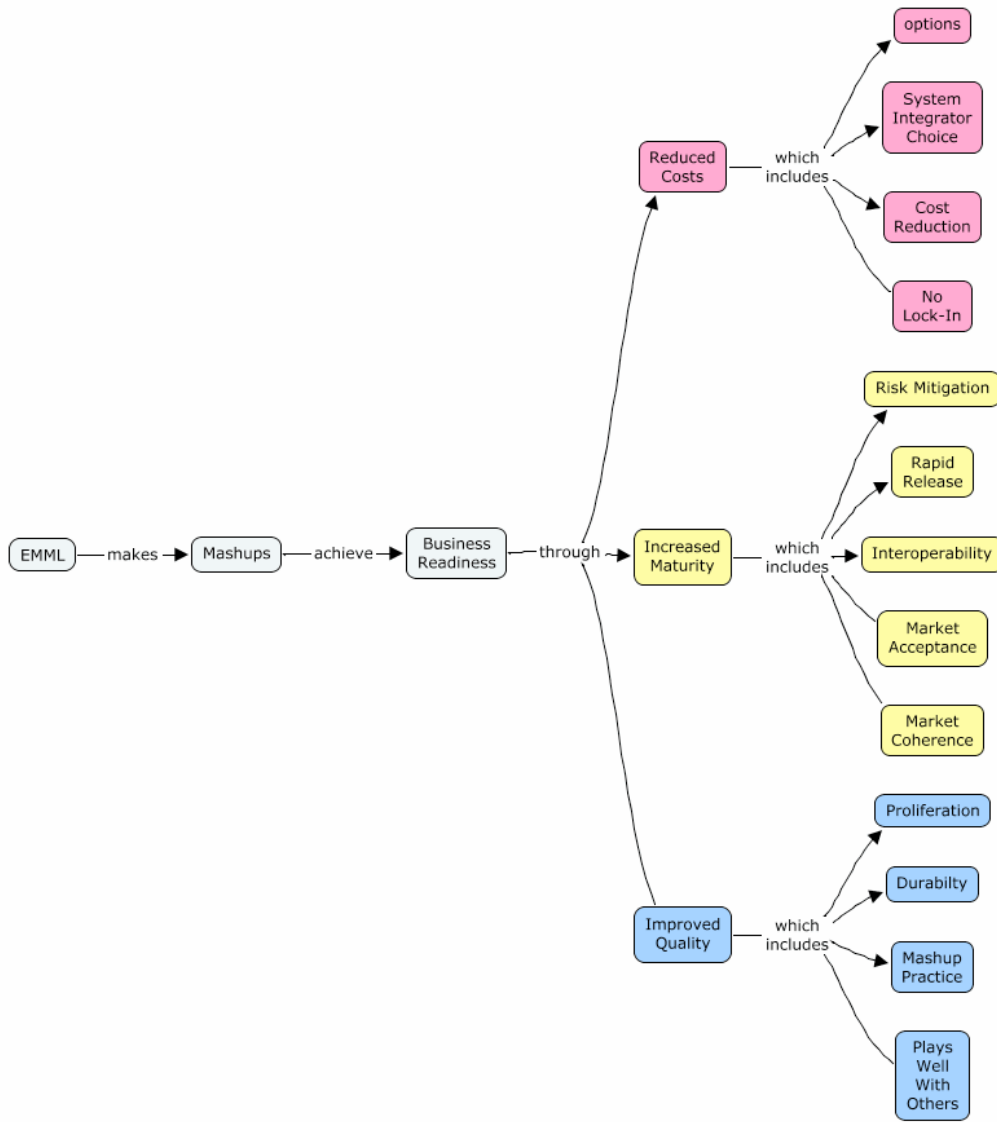
Security - Mashups that run on the server can take advantage of the controlled and regulated server environment. This enhances security through tighter controls and tools available on the server.

Manageability - An enterprise mashup also takes advantage of a single location for data and tools making it easier to administer.

Social - Enterprise mashups facilitate the sharing of data, tools, and results, meaning that enterprise mashups can take advantage of network effects. This enhances the ability to peer review, reuse previous work, and obtain cross-silo expertise.

Performance - Data and processing occur on a powerful, optimized central server. The results can be viewed through a browser or other viewer on the desktop. Server-side processing provides a faster, more stable environment for business mashups.

Reliability - Server side storage of data and tools means that regularly scheduled system maintenance and backups occur more reliably. If an analyst drops their laptop or loses a hard drive, no business critical information will be lost.



EMML Brings Profitability, Predictability & Performance to Business

EMML's standardized data formats reduce costs and increase the quality and maturity of mashups. Business then benefits from more robust analysis, achieved in less time and for less money.

Reduced Costs

EMML strikes at the heart of what makes technology-based projects expensive. By creating a common and predictable description of mashup data, analysis tools can become less complex, mashup designers can be less specialized (and therefore less expensive), and end users are immediately empowered to create systems on their own. All of this contributes to mashups being less expensive and easier to create, experiment with, and maintain.

Cost reductions can come from:

Enterprise Adoption - As mashups become more commonplace throughout the organization, standards like EMML will result in streamlined and mutually supportive applications that will lower costs by removing complexity, increasing efficiency and - as they spread through the organization - through economies of scale.

No Vendor Lock-In - Without standardization, selected data needs to closely match the structure of other data being introduced to the mashup. This often results in the selection of a single vendor for data. Single-source data requirements create a cascade effect, where companies become locked into one data provider, one tool provider, and specific developers who are experienced with those platforms. In such case options are then limited, and costs are fixed by the data and software providers.

Options - EMML gives businesses more choices. Developers no longer have to build software to address multiple data formats. Users can select analytical tools based on their accuracy and analysis. From the run-time to the mashups themselves, mashup components are interchangeable. By bringing a broader marketplace to bear, EMML provides more options.

Rapid Release - EMML takes existing, easy-to-use mashup technologies and removes the uncertainties from the data formats, a major impediment to rapid release. Previously, people creating a mashup spent a great deal of time "taming" various data sources. With EMML however, all data comes with a predictable

format. Now mashups can be created and released rapidly. Focus shifts to the quality of analysis and not the state of the data.

Increased Maturity

While data mashups have been around for several years, their widespread adoption has been hampered by unnecessary complexities brought about by trying to get many different data formats to play nice. EMMML changes the ground rules.

Interoperability - EMMML's benefits may well entirely rest on interoperability. Beyond just being able to use information created in one application in another, EMMML adopters can use different runtimes that meet their needs, or create mashups that work interchangeably. EMMML creates a new component model and common language for mashups.

EMMML also promotes data interoperability, without which there is only proprietary data. By their very nature, mashups require open data in useable formats. Interoperable data greatly simplifies mashup creation, speeding development time and increasing cost effectiveness.

Market Coherence - Greater faith in the data and mashup results will create a coherent market for data and analysis. As more people trust the data they are receiving, more mashups can result. The greater the number of mashups made available, the greater the amount of data being presented for other mashups.

The mashup world tends to focus on sharing. As mashups become more prevalent, the market will spawn a new generation of mashup tools, making it even easier to quickly analyze data.

Risk Mitigation - Closed data standards cost businesses time and money. Obtaining data from proprietary systems requires the help of experts. Open formats like EMMML can be quickly and efficiently employed and extended by a much larger open and free community.

Mashup products like JackBe's Presto and IBM Mashup Server already allow users to simply upload data into the system and begin analyzing the data. The resulting mashups are often easier to use than common programs such as Microsoft Excel. With a common format like EMMML, in-house applications will be more common and even more useful.

This mitigates potential risks in mashup adoption or experimentation by lowering costs of deployment. It also mitigates risks in the adoption of enterprise software by ensuring that information in those large and expensive rollouts won't become trapped in the application and not generate further value for the business.

Risk is further mitigated through plug-and-play data sets. People conducting the analysis can try various data sets with impunity. Swapping out various similar data sets from various sources becomes trivial. For example, if you have neighborhood transportation data from each of the last five years, you can examine each year to determine which is the best for your purposes (e.g. one that isn't skewed by a lengthy construction project). Spending more time with the actual content of the data - rather than the format - increases an understanding of the data better and results in building more defensible products.

Market Acceptance - The ability to quickly gather and analyze mission-critical information is obviously attractive to business. The current problems rest around enterprise IT standards and job descriptions. Mashup tools have largely been a novelty. Only recently has there been sufficient data and tools robust enough to truly and rapidly create useful business-class mashups.

Open formats will greatly increase adoption by facilitating both the amount of data and the number of peers using mashups.

Innovation in Assembly - EMMML's predictable data format will spawn predictable mashup components making assembly simple. So simple in fact, as to be elementary. What previously consumed most of the time and esoteric expertise in mashup creation has been removed, suggesting true innovation is driven where it belongs: the analysis and the visualization. The work is now carried out by the people needing the results of the analysis - not the IT department - at the precise moment they need it. They are then able to directly experiment with the data and

the analysis. Experimentation here supercharges the mashup process. Being able to grab several different data sets and piece them together in new and interesting ways gives you an unprecedented avenue to innovation.

Improved Quality

As cost decreases and facility of creation increases, mashup quality will correspondingly improve. Data quality can now become a priority. Ease of creation will advance the state of the art, increasing the number of mashups being provided, and the environments in which they can operate.

Mashup Practice - Due to the variability in data quality and construction, the mashup creation workflow currently includes time to examine the data and massage it for use. EMMML's standards remove much of the uncertainty from mashup creation, significantly reducing the complexity of the process. In the absence of complexity, focus can remain on the quality of the analysis, which immediately translates into the durability and acceptability of the results.

Increased Mashup Adoption - Mashups are tools which rapidly analyze data. Their adoption nevertheless, relies on acceptance of both the Mashup concept and the product that is created by the mashup. By increasing both the quality of output and ease of creation, EMMML can help to accelerate enterprise adoption of mashups.

As an organization begins to incorporate mashups into their routine, the ability to base new presentations and ideas on data-driven mashup exercises becomes compelling.

Plays Well With Others - The intent of an open and common language like EMMML is to ensure the most universal use of the data. New programming languages can easily utilize EMMML data, existing programming languages are taken into account when creating the language.

Our Digital Universe - Treasures of Data (if we can consume it)

EMML revolutionizes enterprise mashups - and by extension the world of business intelligence - by creating a single, predictable format for information sources. This information can now be collected, analyzed, and acted upon at speeds previously unimaginable. This means that BI work can be produced at all levels of business with minimal cost and skill requirement.

Over the past 4 decades, National Archives and Records Administration has seen rampant growth in the generation of presidential administration data, from just a few thousand electronic files in the 1970s to over 326 million files in the GW Bush Administration.⁶ Such unprecedented data production is mirrored both in and beyond the enterprise. Publicly generated data such as blogs, ebooks, emails, text messages, images - in addition to mission-critical, business data - is multiplying at a rate that threatens to outpace Moore's law. Just as it is incomprehensible for the human brain to process the sheer magnitude of outer space, so too is it unfathomable for us to envision the ever-expanding nature of our "digital universe."

According to the technology research firm International Data Corp (IDC), "At nearly 500 exabytes - or 500 billion gigabytes - the Digital Universe, if converted to pages of text and assembled into books, would stretch to Pluto and back 10 times. You could wrap the entire earth 8 times with all that paper."⁷ In their annual "Digital Universe" report, IDC researchers found that last year alone individuals and businesses generated 161 billion gigabytes of data. IDC posits that number will continue to multiply, doubling every 18 months and that at this rate, "the universe of content created across the globe will grow fivefold from 486 exabytes to more than 2,500 exabytes by the end of 2010."

Never before has business had more data at its disposal, nor has it had the ability to mine, manage, and mobilize that data thru the combined advantage of lower costs and facility of use. Real-time information feeds, analytical tools, and collaborative systems all inform and empower the enterprise. Mashups, now made

⁶ <http://www.archives.gov/era/pdf/cacas042709.pdf>

⁷ <http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.htm>

economical through EMMML, can efficiently put this vital resource to work. Companies using this information to innovate can expand into new markets, make internal systems more efficient, and invent new products with alacrity.

The increase in the amount of information, the decrease in the cost of storage and technology, and the simplification of tools via the mashup to analyze data sends one clear message: The organization that understands that BI is now a company-wide responsibility and provides the infrastructure to do so - wins in the end.

About The Authors

Dion Hinchcliffe is a well-known business strategist and enterprise architect. He speaks and writes about the intersection of technology and business while working hands-on with clients in the Fortune 500, federal government, and the Internet startup community. He has extensive hands on experience building and deploying large-scale, next generation Web applications. Mr. Hinchcliffe also helps lead the industry by evolving the thinking around Web 2.0 in the enterprise on ZDNet and the Social Computing Journal. He has been quoted by BusinessWeek, CNET, Wired Magazine, CIO Magazine, and many other well-known industry periodicals.

Mr. Hinchcliffe is also a regular keynote speaker on the topics of Web 2.0, SOA, and Enterprise 2.0 and has presented or keynoted at Web 2.0 Expo, Business Integration Forum, Interop, JavaOne, SOA Web Services Edge, Collaboration Technologies Conference, AjaxWorld, Office 2.0, and other major business and software conferences. He is also the founder of Web 2.0 University, the world's leading education solution around Web 2.0 as well as The Enterprise 2.0 TV Show.

Jim Benson incorporates his background in cognitive psychology, government, and management to build community through policy and technology. His company [Modus Cooperandi](#), helps organizations change through the application of Lean principles, Agile methodologies, and social media. He is also the developer of the productivity tool [Personal Kanban](#), an adaptation of industrial kanban which helps individuals and small teams actualize.