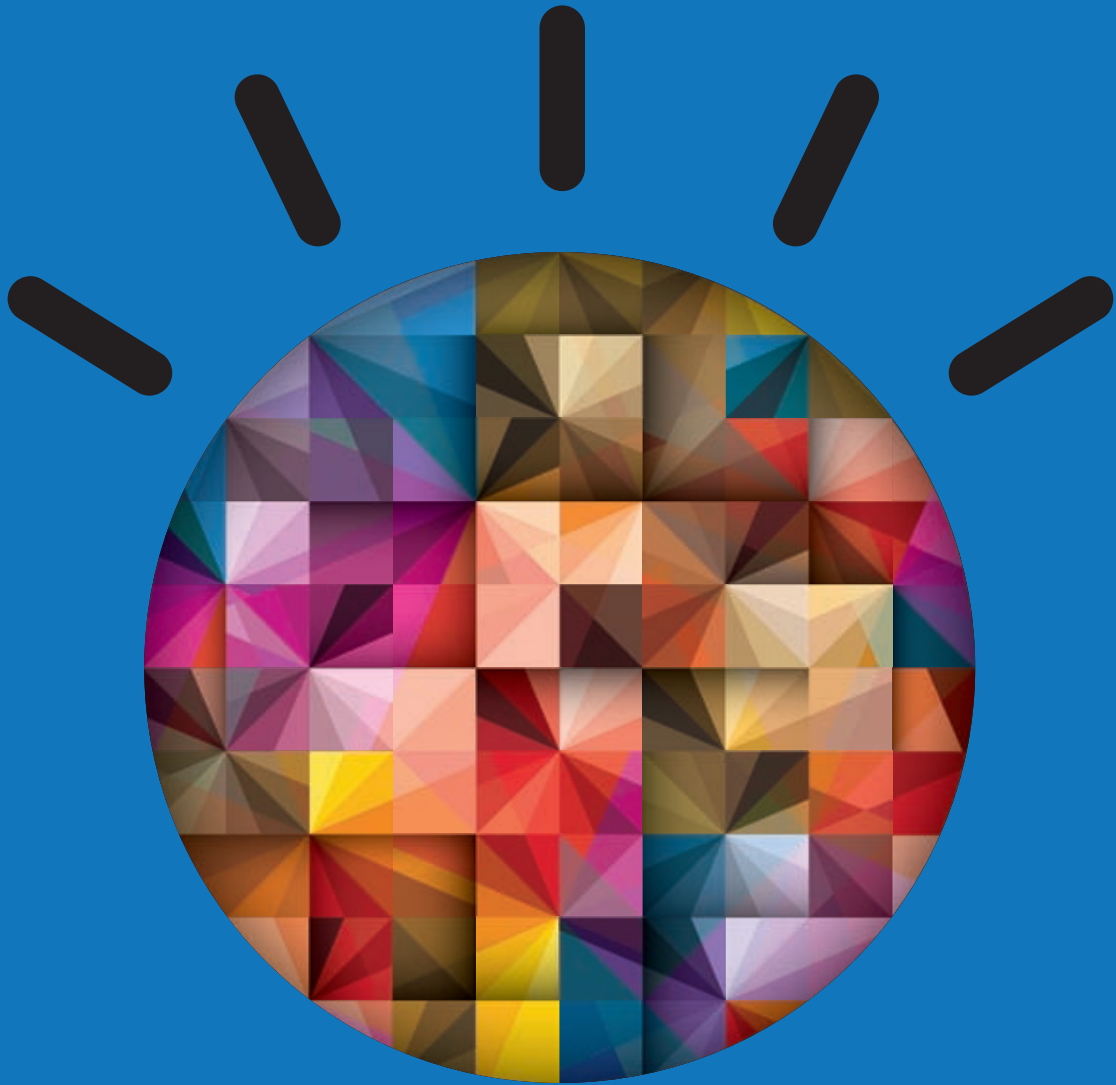
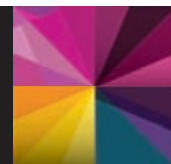


# IBM Big Data Success Stories





## A note from Rob Thomas



Big data. By now you have heard the term and it's easy to grasp what it means as the world continues to create 2.5 quintillion bytes daily. Or, maybe not; can you fathom one quintillion bytes? I can't. But I can relate to Vestas Wind Systems, a leader in the development of wind energy that uses their IBM big data solution and one of the world's largest supercomputers to analyze weather information and provide location site data in minutes instead of weeks, even while its wind library is increasing from 2.8 petabytes to as much as 24 petabytes of data - the equivalent of 1420 times the books in America's Library of Congress.

In your business, you have your own big data challenges. You have to turn mountains of data about your customers, products, incidents, etc., into actionable information. While the volume, variety and velocity of big data seem overwhelming, big data technology solutions hold great promise. The way I see it, we are on the mountain top with a vista of opportunity ahead. We have the capacity to understand; to see patterns unfolding in real time across multiple complex systems; to model possible outcomes; and to take actions that produce greater economic growth and societal progress. IBM is marshaling its resources to bring smarter computing to big data. With the IBM big data platform, we are enabling our clients to manage data in ways that were never thought possible before.

In this collection of Big Data Success Stories, we share a sample of our customers' successes including:

- [x+1], an end-to-end digital marketing platform provider for advertisers and agencies, is helping their clients realize a 20% growth in digital sales by analyzing massive volumes of advertising data in real-time using IBM Netezza
- KTH Royal Institute of Technology in Stockholm, which uses streaming data in their congestion management system, is already reducing traffic in the Swedish capital by 20 percent, lowering average travel times by almost 50 percent and decreasing the amount of emissions by 10 percent
- Researchers at the University of Ontario-Institute of Technology who are using streaming analytics to help neonatal care hospitals predict the onset of potentially fatal infections in premature babies

We are humbled at "miracles" our clients are achieving and are very proud of the role we are playing in making cities, commerce, healthcare and a full spectrum of additional industries smarter.

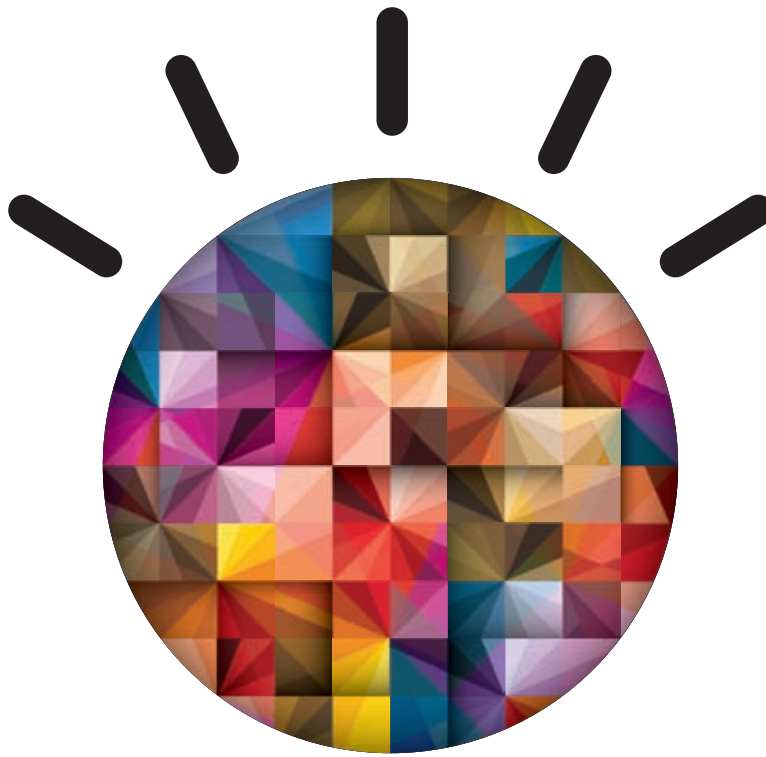
I hope you will enjoy reading these Big Data Success Stories and consider IBM when you take on big data challenges in your enterprise.

Sincerely,

Rob Thomas  
Vice President, Business Development  
IBM



<i>Bringing smarter computing to big data</i> . . . . .	1
<i>IBM Unveils Breakthrough Software and New Services to Exploit Big Data</i> . . . . .	2
<i>Customer Success Stories</i>	
<i>Beacon Institute</i> . . . . .	5
<i>Faces</i> . . . . .	8
<i>Hertz</i> . . . . .	12
<i>KTH – Royal Institute of Technology</i> . . . . .	16
<i>Marine Institute Ireland</i> . . . . .	18
<i>Technovated</i> . . . . .	22
<i>TerraEchos</i> . . . . .	24
<i>University of Ontario Institute of Technology</i> . . . . .	28
<i>Uppsala University</i> . . . . .	34
<i>Vestas</i> . . . . .	37
<i>Watson</i> . . . . .	44
<i>[x+1]</i> . . . . .	47
<i>IBM Business Partner Ecosystem</i> . . . . .	54
<i>Featured Business Partners</i>	
<i>Datameer</i> . . . . .	60
<i>Digital Reasoning</i> . . . . .	61
<i>Jaspersoft</i> . . . . .	63
<i>Karmasphere</i> . . . . .	67
<i>MEPS</i> . . . . .	69



Smarter computing builds a Smarter Planet

## Bringing smarter computing to big data.

To build a smarter planet, we need smarter computing—computing that is tuned to the task, managed through the cloud and, importantly, designed for big data.

How big? We're now creating 2.5 quintillion bytes daily—so much that 90% of the data in the world today has been created in the last two years alone.

This data is also big in another way—in its promise. We now have the capacity to understand, with greater precision than ever before, how our world actually works—to see patterns unfolding in real time across multiple complex systems; to model possible outcomes; and to take actions that produce greater economic growth and societal progress.

We can do more than manage information—we can manage vast information supply chains. They're made up of not only the ones and zeros of structured data that traditional computers love, but streams of unstructured text, images, sounds, sensor-generated impulses and more.

We can parse the real languages of commerce, processes and natural systems—as well as conversations from the growing universe of tweets, blogs and social media. We can also draw on advanced technologies such as stream computing, which filters gigabytes of data per second, analyzes these while still in motion and decides on the appropriate action for the data, such as a real-time alert or storing an insight in a data warehouse for later analysis.

But we can only do all of this if our computing systems are smart enough to keep up. According to the IBM Business Analytics and Optimization for the Intelligent Enterprise study, one in three business leaders frequently make decisions without the information they need. Half don't have access to the information they need to do their jobs. And that has significant competitive implications. The 2010 IBM Global

CFO Study by the IBM Institute for Business Value showed that companies that excel at finance efficiency and have more mature business analytics and optimization outperform their peers, with 49% higher revenue growth, 20 times more profit growth, and 30% higher return on invested capital.

With continuously analyzed data, organizations can be what they want to be, at all times. Consider the Memphis Police Department, which compiles volumes of crime records from a variety of sources and systems, and has reduced serious crime by more than 30%. Fresh food grower Sun World International is leveraging insights from their data to cut natural resource use by 20%. Research at the University of Ontario Institute of Technology is developing streaming analytics to help neonatal care hospitals. By analyzing 43 million streaming data points per patient, per day, they can improve patient outcomes by using all of the data available.

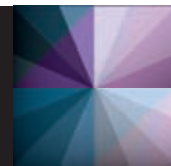
This list could go on. And at the leading edge of smarter computing, IBM's Watson—the computer that bested the two all-time champions on the television quiz show Jeopardy!—demonstrates the power of analytics to provide meaningful insights from an ever-increasing volume and variety of data, enabling correct answers and winning actions, in real time.

As our world gets smaller, our data keeps getting bigger—which is good news. Information that was once merely overload now lets us see our planet in entirely new ways and intervene to make it work better. Because computing systems designed for big data are systems designed for good decision making. Which is, after all, what being smarter is all about.

Let's build a smarter planet. Join us and see what others are doing at [ibm.com/smarterplanet](http://ibm.com/smarterplanet)

# IBM Unveils Breakthrough Software and New Services to Exploit Big Data

*Commits \$100 Million to Massive Scale Analytics Research*



**ARMONK, N.Y., - 20 May 2011:** As companies seek to gain real-time insight from diverse types of data, IBM (NYSE: IBM) today unveiled new software and services to help clients more effectively gain competitive insight, optimize infrastructure and better manage resources to address Internet-scale data. For the first time, organizations can integrate and analyze tens-of-petabytes of data in its native format and gain critical intelligence in sub-second response times.

IBM also announced a \$100 million investment for continued research on technologies and services that will enable clients to manage and exploit data as it continues to grow in diversity, speed and volume. The initiative will focus on research to drive the future of massive scale analytics, through advancing software, systems and services capabilities.

The news comes on the heels of the 2011 IBM Global CIO Study where 83 percent of 3,000 CIOs surveyed said applying analytics and business intelligence to their IT operations is the most important element of their strategic growth plans over the next three to five years.

Today's news further enables Smarter Computing innovations realized by designing systems that incorporate Big Data for better decision making, and optimized systems tuned to the task and managed in a cloud.

According to recent IT industry analyst reports, enterprise data growth over the next five years is estimated to increase by more than 650 percent. Eighty percent of this data is expected to be unstructured.

The new analytics capabilities pioneered by IBM Research will enable chief information officers (CIOs) to construct specific, fact-based financial and business models for their IT operations. Traditionally, CIOs have had to make decisions about their IT operations without the benefit of tools that can help interpret and model data.

With today's news, IBM is expanding its portfolio and furthering its investments in analytics with:

- New, patented software capabilities to analyze massive volumes of streaming data with sub-millisecond response times and Hadoop-based analytics software to offer scalable storage to handle tens-of-petabytes level data. These capabilities complement and leverage existing IT infrastructure to support a variety of both structured and unstructured data types.
- 20 new services offerings, featuring patented analytical tools for business and IT professionals to infuse predictive analytics throughout their IT operations. The services enable IT organizations to assess, design and configure their operations to address and take advantage of petabytes of data.

"The volume and velocity of information is generated at a record pace. This is magnified by new forms of data coming from social networking and the explosion of mobile devices," said Steve Mills, Senior Vice President and Group Executive, IBM Software & Systems. "Through our extensive capabilities in business and technology expertise, IBM is best positioned to help clients not only extract meaningful insight, but enable them respond at the same rate at which the data arrives."

## New Services Address Analytics for IT Infrastructure

Leveraging years of intellectual capital in managing data centers and IT departments, as well as over 30 patented technologies from IBM Research, the new IT services feature dozens of analytical tools to help IT professionals use server, storage and networking technologies more efficiently, improving security and insight into planning major IT investments. Examples of services that help clients with analytics include:

- **Cloud Workload Analysis** -- The new analysis tool maps your IT workload characteristics and current capabilities to prioritize cloud deployment and migrations plans. This allows IT managers to identify cloud opportunities 90 percent faster to reduce costs.
- **Server and Storage** -- New server optimization and analysis tools achieve up to 50 percent reduced transformation costs and up to 80 percent faster implementation time. New storage services help create self-service to provision explosive growth while reducing architects time by 50 percent.
- **Data Center Lifecycle Cost Analysis Tool** -- Identifies how to reduce total data center costs by up to 30 percent by assessing total cost plus including environmental impact over a 10 to 20 year life.
- **Security Analytic services** -- Analytic systems identify known events and automatically handle them; This results in handling of more than 99 percent of critical events without human intervention.

## IBM Big Data Software Taps into Hadoop

IBM is making available new InfoSphere BigInsights and Streams software that allows clients to gain fast insight into information flowing in and around their businesses. The software, which incorporates more than 50 patents, analyzes traditional structured data found in databases along with unstructured data -- such as text, video, audio, images, social media, click streams -- allowing decision makers to act on it at unprecedented speeds.

# IBM Unveils Breakthrough Software and New Services to Exploit Big Data



BigInsights software is the result of a four-year effort of more than 200 IBM Research scientists and is powered by the open source technology, Apache Hadoop. The software provides a framework for large scale parallel processing and scalable storage for terabyte to petabytes-level data. It incorporates Watson-like technologies, including unstructured text analytics and indexing that allows users to analyze rapidly changing data formats and types on the fly.

Additional new features include data governance and security, developer tools, and enterprise integration to make it easier for clients to build a new class of Big Data analytics applications. IBM also offers a free downloadable BigInsights Basic Edition for clients to help them explore Big Data integration capabilities.

Also born at IBM Research, InfoSphere Streams software analyzes data coming into an organization and monitors it for any changes that may signify a new pattern or trend in real time. This capability helps organizations to capture insights and make decisions with more precision, providing an opportunity to respond to events as they happen.

New advancements to Streams software makes it possible to analyze Big Data such as Tweets, blog posts, video frames, EKGs, GPS, and sensor and stock market data up to 350 percent faster than before. BigInsights complements Streams by applying analytics to the organization's historical data as well as data flowing through Streams. This is an ongoing analytics cycle that becomes increasingly powerful as more data and real-time analytic results are available to be modeled for improvement.

As a long time proponent of open source technology, IBM has chosen the Hadoop project as the cornerstone of its Big Data Strategy. With a continued focus on building advanced analytics solutions for the enterprise, IBM is building upon the power of these open source technologies while adding improved management and security functions, and reliability that businesses demand. Hadoop's ability to process a broad set of information across multiple computing platforms, combined with IBM's analytics capabilities, now makes it possible for clients to tackle today's growing Big Data challenges. IBM's portfolio of Hadoop-based offerings also include IBM Cognos Consumer Insight which integrates social media content with traditional business analytics, and IBM Coremetrics Explore which segments consumer buying patterns and drills down into mobile data. Additionally, Hadoop is the software framework the IBM Watson computing system uses for distributing the workload for processing information, which supports the systems breakthrough ability to understand natural language and provide specific answers to questions at rapid speeds.

## University of Ontario Institute of Technology Expands Neo-Natal Research to China

Dr. Carolyn McGregor, Research Chair in Health Informatics at the University of Ontario Institute of Technology has been exploring new approaches for the last 12 years to provide specialists in neonatal intensive care units better ways to spot potentially fatal infections in premature babies.

Changes in streams of real-time data such as respiration, heart rate and blood pressure are closely monitored in her work and now she is expanding her research to China. "Building upon our work in Canada and Australia, we will apply our research to premature babies at hospitals in China. With this new additional data, we can compare the differences and similarities of diverse populations of premature babies across continents," said Dr. McGregor. "In comparing populations, we can set the rules to optimize the system to alert us when symptoms occur in real time, which is why having the streaming capability that the IBM platform offers is critical. The types of complexities that we're looking for in patient populations would not be accessible with traditional relational database or analytical approaches."

IBM's Big Data software and services reinforces IBM's analytics initiatives to deliver Watson-like technologies that help clients address industry specific issues. On the heels of The IBM Jeopardy! Challenge, in which the IBM Watson system demonstrated a breakthrough capability to understand natural language, advanced analytical capabilities can now be applied on real client challenges ranging from identifying fraud in tax or healthcare systems, to predicting consumer buying behaviors for retail clients.

Over the past five years, IBM has invested more than \$14 billion in 24 analytics acquisitions. Today, more than 8,000 IBM business consultants are dedicated to analytics and over 200 mathematicians are developing breakthrough algorithms inside IBM Research. IBM holds more than 22,000 active U.S. patents related to data and information management.

To hear how IBM clients are using analytics to transform their business visit: <http://www.youtube.com/user/ibmbusinessanalytics>.

For more information on IBM Big Data initiatives, visit: [www.ibm.com/bigdata](http://www.ibm.com/bigdata).

For more information on IBM's full set of new analytics services, visit: [www.ibm.com/services/it-insight](http://www.ibm.com/services/it-insight).

# Customer Success Stories



***Beacon Institute***



***Faces***



***Hertz***



***KTH – Royal Institute of Technology***



***Marine Institute Ireland***



***Technovated***



***TerraEchos***



***University of Ontario Institute of Technology***



***Uppsala University***



***Vestas***



***Watson***



***[x+1]***





## Beacon Institute, Clarkson University and IBM



*Managing the environmental impact on rivers  
by streaming information*

---

### Overview

#### The need

Scientists need new technology to study complex environmental interactions to better understand how communities and ecosystems interact.

#### The solution

IBM InfoSphere Streams software and high-performance computing system collect and analyze data in real time as it streams in from environmental data sources to support predictive analysis and decision making.

#### The benefit

Streaming real-time data technology helps resource management programs respond more effectively to chemical, physical and biological alterations to local water resources.

---

Most of the world's population lives near a river or estuary. Yet, there is typically no way to gain a clear understanding of what is happening below the surface of the water to help predict and manage changes in the river that could impact local communities that rely on the waterway.

The River and Estuary Observatory Network (REON) project is a joint effort between Beacon Institute for Rivers and Estuaries, Clarkson University and IBM® Research. REON is the first technology-based, real-time monitoring network for rivers and estuaries of its kind, and allows for continuous monitoring of physical, chemical and biological data from points in New York's Hudson, Mohawk and St. Lawrence Rivers by means of an integrated network of sensors, robotics, mobile monitoring and computational technology deployed in the rivers.

“Imagine predicting environmental impacts the way we forecast and report the weather,” says John Cronin, Founding Director of Beacon Institute and Beacon Institute Fellow at Clarkson University. “With that technological capability we can better understand the effects of global warming, the movements of migrating fish or the transport of pollutants. The implications for decision-making and education are staggering.”





---

Solution components:

**Software**

- IBM® InfoSphere® Streams
- 

*“Imagine predicting environmental impacts the way we forecast and report the weather. . . . The implications for decision-making and education are staggering.”*

— John Cronin, Founding Director of Beacon Institute for Rivers and Estuaries and Beacon Institute Fellow at Clarkson University

---

## Applying real-time technology to help understand the environment

REON is a test bed for the IBM System S stream computing system. A team of IBM engineers and scientists work on the REON collaboration and have access to IBM’s extensive analytical and computational resources from the IBM Watson Research Lab. The IBM Global Engineering Solutions team executed the fundamental design elements of the data streaming pilot. This high-performance architecture rapidly analyzes data as it streams in from many sources.

A networked array of sensors in the river provides the data necessary to locally observe spatial variations in such variables as temperature, pressure, salinity, turbidity, dissolved oxygen and other basic water chemistry parameters. All of these sensors, transmitting information in real time, results in massive amounts of data.

Using real-time, multi-parameter modeling systems helps develop a better understanding of the dynamic interactions within local riverine and estuarine ecosystems. Making real-world data easily accessible to outside systems, researchers, policymakers and educators helps foster increased collaboration. The ultimate benefit is helping resource management programs respond more effectively to chemical, physical and biological alterations to local water resources.

## REON—New technology for the smarter water management

“The Hudson River is the pilot river system for REON, and the 12 million people who live within its watershed will be the first beneficiaries of our work,” says Cronin.

Helping to make sense of all that data is IBM InfoSphere® Streams software, part of IBM’s big data platform. InfoSphere Streams provides capabilities that collect and analyze data from thousands of information sources to help scientists better understand what is happening in the world—as it happens. Eventually, REON data could be applied to visualize the movement of chemical constituents, monitor water quality, and protect fish species as they migrate, as well as provide a better scientific understanding of river and estuary ecosystems.



---

*“The Hudson River is the pilot river system for this groundbreaking initiative, and the 12 million people who live within its watershed will be the first beneficiaries of our work.”*

— John Cronin

---

“As water resource management expert Doug Miell has said, you can’t manage what you can’t measure. . . Society and business are facing increasingly complex challenges when it comes to understanding and managing water resources on this planet,” says John E. Kelly III, Senior Vice President and Director, IBM Research. “Getting smart about water is important to all of us for one simple reason: water is too precious a resource to be wasted.”

### **Positively Impacting the Environment Worldwide**

Cronin concludes, “This new way of observing, understanding and predicting how large river and estuary ecosystems work ultimately will allow us to translate that knowledge into better policy, management and education for the Hudson River and for rivers and estuaries worldwide.”

### **For more information**

To learn more about IBM InfoSphere Streams, visit:

[ibm.com/software/data/infosphere/streams](http://ibm.com/software/data/infosphere/streams)

To learn more about IBM big data, visit:

[ibm.com/software/data/bigdata](http://ibm.com/software/data/bigdata)

To increase your big data knowledge and skills, visit:

[www.BigDataUniversity.com](http://www.BigDataUniversity.com)

To get involved in the conversation, visit:

[www.smartercomputingblog.com/category/big-data](http://www.smartercomputingblog.com/category/big-data)

For more information on Beacon Institute for Rivers and Estuaries, visit: [www.bire.org/home](http://www.bire.org/home)

## IBM

*Applies emerging technologies to deliver instantaneous people searches*

---

### Overview

#### The need

With over 600,000 names in BluePages, IBM's employee directory, and over 500,000 queries daily, the average search session takes two minutes. IBM needed a faster, more efficient application.

#### The solution

Using Apache open source technologies, the IBM CIO Lab Analytics team developed a new people-search application that allows flexible queries and returns as many results as possible, as fast as possible. Additional capabilities include quick browsing and photo images.

#### The benefit

The new *Faces* application offers instantaneous response time, saving on average over a minute for each search session—and thousands of hours daily for IBM employees.

---

With an enterprise population of over 600,000 people worldwide, how do IBM® employees find and connect with their colleagues? For over a decade, IBM BluePages has been the primary source. This high-demand, intranet application provides information on all IBM employees and contractors, including areas of expertise and responsibilities. And with IBM's focus on innovation and emerging technologies, positive changes are always on the horizon.

“BluePages is one of the most used applications at IBM,” says Sara Weber, manager of IBM's CIO Lab Analytics team. “At one time, BluePages was state-of-the-art; however, over the years it was not updated to keep up with new advances in Internet technology. With over 500,000 BluePages searches done every day, and with BluePages accessing huge volumes of data, an average search session can take up to two minutes. When multiple results are returned they do not show individual photo images, and incorrect spelling may yield no results. My team was tasked with addressing the question: ‘How can we build a better and faster people search?’”

The goals for this project, aptly named *Faces*, were to support flexible queries and return as many results as possible, as fast as possible. Results that more closely matched the query would appear first. Additional capabilities would permit quick browsing and photo images.



---

*“At IBM, when we find an open source technology that has potential, we experiment with it to understand how to use it to bring the most business value to IBM. For example, IBM InfoSphere BigInsights is a new class of analytics platform based on Hadoop and innovation from IBM. It can store raw data ‘as-is’ and help clients gain rapid insight through large scale analysis.”*

—Sara Weber, Manager, IBM's CIO Lab Analytics team

---

## Applying emerging technologies to deliver innovation

Weber's CIO Lab Analytics team identifies problems that IBM employees are experiencing and finds ways to apply emerging technologies to develop solutions. “We had to process tremendous amounts of data, and then store it in a way that it could be accessed quickly,” says Weber. “For this project, we selected Apache Hadoop and Apache Voldemort; both are open source technologies. My development team has extensive expertise in using Hadoop technology. The *Faces* application was developed by two members of our team over a five month period.”

Apache Hadoop allows developers to create distributed applications that run on clusters of computers. Organizations can leverage this infrastructure to handle large data sets, by dividing the data into “chunks” and coordinating the data processing in the distributed, clustered environment. Once the data has been distributed to the cluster, it can be processed in parallel. Apache Voldemort is a distributed key-value storage system that offers fast, reliable and persistent storage and retrieval. Specific keys return specific values. If no additional query power is needed, a key value store is faster than a database.

“At IBM, when we find an open source technology that has potential, we experiment with it to understand how to use it to bring the most business value to IBM,” says Weber. “For example, IBM InfoSphere® BigInsights is a new class of analytics platform based on Hadoop and innovation from IBM. It can store raw data ‘as-is’ and help clients gain rapid insight through large scale analysis.”

For *Faces*, Hadoop preprocesses data from the IBM Enterprise Directory and Social Networks and sends this information to the Voldemort Person Store (2.2 GB). Voldemort, in turn, sends data to Hadoop processing for the Person ID fetcher, Reports Loader, Query Expander, and Location Expander. These results are saved to Voldemort's Query Store (5.5 GB). Hadoop also receives images from BluePages that are saved in Voldemort's image store to remain available for Hadoop's montage generator.

---

## Solution components

### Servers

- IBM® BladeCenter® servers

### Software

- Apache Hadoop
  - Apache Voldemort Key Value Storage System
- 

*“We could not have developed Faces without the distributed processing capabilities Hadoop provides. The Faces application has really highlighted the power of Hadoop and has helped us address a major pain point for all IBMers.”*

—Sara Weber

---

“We placed all 600,000 names into memory for immediate access,” says Weber. “Preprocessing with Hadoop directly improves performance. Each time you type a letter in a name, results are immediate. We have precomputed the search process to retrieve every employee name that matches what is entered. Every time you type another letter, scoring retrieves people who are more relevant to the search criteria. The information is available and, from a performance perspective, everything is ready to go. Memory and storage are inexpensive and nightly processing takes only a few hours.”

Weber adds, “We run Hadoop on ten, five-year-old IBM BladeCenter® servers. These Blades are low powered, but Hadoop distributes the workload and takes advantage of the hardware to the fullest. If more computation is needed, we can add machines and improve performance without modifying the code.”

## Measuring business value

According to Weber, the new *Faces* application enables employees to receive instantaneous search results. “Conservatively speaking, we are saving on average over a minute for each search session,” says Weber. “Searches are faster and easier. The information is timely and accurate. With over 500,000 searches daily, IBMers are saving thousands of hours each day.”

For IBM employees, the improvement is noticeable. “To gain user acceptance or change user behavior, we know any new solution we create has to be significantly faster and better,” says Weber. “As far as I know, *Faces* is the fastest growing innovation ever introduced at IBM. In the first two weeks, *Faces* went from zero to 85,000 users with continued viral growth throughout the entire IBM organization. What used to take minutes now takes milliseconds. We provide a feedback button on all our applications so users can report errors or issues. With *Faces*, IBMers were using the feedback button to say, “Thank you for making my job so much easier.”

Weber concludes, “We could not have developed *Faces* without the distributed processing capabilities Hadoop provides. The *Faces* application has really highlighted the power of Hadoop and has helped us address a major pain point for all IBMers.”

## For more information

To learn more about IBM Information Management solutions, please contact your IBM sales representative or IBM Business Partner, or visit the following website: [ibm.com/software/data](http://ibm.com/software/data)

To learn more about IBM InfoSphere BigInsights, visit:  
[ibm.com/software/data/infosphere/biginsights](http://ibm.com/software/data/infosphere/biginsights)

Additionally, financing solutions from IBM Global Financing can enable effective cash management, protection from technology obsolescence, improved total cost of ownership and return on investment. Also, our Global Asset Recovery Services help address environmental concerns with new, more energy-efficient solutions. For more information on IBM Global Financing, visit: [ibm.com/financing](http://ibm.com/financing)



---

© Copyright IBM Corporation 2011

IBM Corporation  
Software Group  
Route 100  
Somers, NY 10589  
U.S.A.

Produced in the United States of America  
October 2011  
All Rights Reserved

IBM, the IBM logo, ibm.com, InfoSphere, and BladeCenter are trademarks of International Business Machines Corporation in the United States, other countries or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.



Please Recycle



## Hertz, Mindshare Technologies and IBM

*Analyzing huge volumes of customer comments in real time delivers a competitive edge*

---

### Overview

#### The need

Improving service means listening to customers and gathering thousands of comments via web, email and text messages. Each comment is viewed and categorized manually for customer service reporting. Inconsistencies were at an unacceptable level.

#### The solution

Using feedback management and content analytics software, customer comments are captured in real time to be transformed into actionable intelligence. Linguistic rules automatically analyze and tag unstructured content into meaningful service reporting categories.

#### The benefit

Automated tagging increased report consistency, freed Hertz field managers from tagging comments, and roughly doubled what the managers had achieved manually.

---

As the world's largest airport car rental brand with more than 8,300 locations in 146 countries, Hertz continually requests and receives feedback from its customers. To retain a competitive edge, the feedback is analyzed so that issues can be identified in real-time and problems can be addressed and resolved quickly.

“Hertz gathers an amazing amount of customer insight daily, including thousands of comments from web surveys, emails and text messages. We wanted to leverage this insight at both the strategic level and the local level to drive operational improvements,” says Joe Eckroth, Chief Information Officer, The Hertz Corporation.

### Leveraging unstructured data to improve customer satisfaction

Hertz and Mindshare Technologies, a leading provider of enterprise feedback solutions, are using IBM® Content Analytics software to examine customer survey data, including text messages. The goal is to identify car and equipment rental performance levels to enable pinpointing issues and making the necessary adjustments to improve customer satisfaction levels.

IBM Content Analytics allows for deep, rich text analysis of information, helping organizations gain valuable insight from enterprise content regardless of source or format. This technology can help reveal undetected problems, improve content-centric process inefficiencies, and take customer service and revenue opportunities to new levels, while helping to reduce operating costs and risks.







---

Solution components:

**Software**

- IBM® Content Analytics
- 

*“Hertz gathers an amazing amount of customer insight daily, including thousands of comments from web surveys, emails and text messages. We wanted to leverage this insight at both the strategic level and the local level to drive operational improvements.”*

— Joe Eckroth, Chief Information Officer,  
The Hertz Corporation

---

Using Content Analytics together with a sentiment-based tagging solution from Mindshare Technologies, Hertz introduced a “Voice of the Customer” analytics system that automatically captures large volumes of information reflecting customer experiences in real-time, and helps transform the information into actionable intelligence. Using a series of linguistic rules, the “Voice of the Customer” system categorizes comments received via email and online with descriptive terms, such as Vehicle Cleanliness, Staff Courtesy and Mechanical Issues. The system also flags customers who request a callback from a manager or those who mention #1 Club Gold, Hertz’s customer loyalty program.

“Working closely with the IBM-Mindshare team, we are able to better focus on improvements that our customers care about, while removing a time-consuming burden from our location managers. This has greatly improved the effectiveness of our ‘Voice of the Customer’ program and has helped build on our reputation for delivering superior customer service.”

**Improving speed and accuracy of processing customer feedback**

In the ultra-competitive world of vehicle and equipment rental, Hertz recognizes that understanding customer feedback and adapting the business accordingly is what drives market share and success. However, most of this valuable information is trapped inside free-form customer feedback surveys.

Prior to working with IBM and Mindshare Technologies, Hertz location managers read each customer comment submitted online via email or by phone, and then manually categorized it for basic reporting and analysis. This approach proved to be labor-intensive and inconsistent, as comments were categorized based on a manager’s personal interpretation. Automating the task of tagging customer comments has increased report consistency and roughly doubled what the managers had achieved manually.



---

*“Working closely with the IBM-Mindshare team, we are able to better focus on improvements that our customers care about, while removing a time-consuming burden from our location managers.”*

– Joe Eckroth

---

IBM Content Analytics software has improved the accuracy and speed of the tagging and analyzing process, setting the stage for more reliable analytics. Free from manually tagging comments, Hertz field managers can now focus attention on performing deep-dive analysis on the information, quickly identifying trends or issues and adjusting operational service levels accordingly.

For instance, wait times at car rental locations can be a contentious issue. The faster and more efficient the car rental/return process, the more likely the customer will do repeat business. Using analytics software, Hertz location managers are able to effectively monitor customer comments to deliver top customer satisfaction scores for this critical level of service. In Philadelphia, survey feedback led managers to discover that delays were occurring at the returns area during certain parts of the day. They quickly adjusted staffing levels and ensured a manager was always present in the area during these specific times.

### **Hertz remains focused on customers and providing superior service**

The Internet and new social media technologies have made consumers more connected, empowered and demanding. The average online user is three times more likely to trust peer opinions over retailer advertising, underlining the importance for retailers to tap new technologies that pay close attention to what customers are saying.

This effort with Hertz reflects IBM’s focus on helping organizations use analytics to get the most value from their information. IBM has a Business Analytics & Optimization services organization, with 7,000 consultants who can help clients get up and running with deep analytics capabilities.

## For more information

To learn more about IBM Content Analytics, visit:

[ibm.com/software/data/content-management/analytics](http://ibm.com/software/data/content-management/analytics)

To learn more about IBM Business Optimization and Analytics services, visit: [ibm.com/services/us/gbs/bao](http://ibm.com/services/us/gbs/bao)

To increase your big data knowledge and skills, visit:

[www.BigDataUniversity.com](http://www.BigDataUniversity.com)

To get involved in the conversation, visit:

[www.smartercomputingblog.com/category/big-data](http://www.smartercomputingblog.com/category/big-data)

For more information on Hertz, visit:

[www.hertz.com](http://www.hertz.com)



---

© Copyright IBM Corporation 2011

IBM Corporation  
Software Group  
Route 100  
Somers, NY 10589  
U.S.A.

Produced in the United States of America  
October 2011  
All Rights Reserved

IBM, the IBM logo and [ibm.com](http://ibm.com) are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.



Please Recycle



Stockholm, Sweden  
[www.kth.se/?l=en\\_UK](http://www.kth.se/?l=en_UK)

*“Analyzing large volumes of streaming data in real time is leading to smarter, more efficient and environmentally friendly traffic in urban areas.”*

— Haris N. Koutsopoulos,  
Head of Transportation and Logistics,  
Royal Institute of Technology,  
Stockholm, Sweden

## KTH – Royal Institute of Technology

*Analyzes real-time data streams to identify traffic patterns*

The Royal Institute of Technology (abbreviated KTH) is a university in Stockholm, Sweden. KTH was founded in 1827 as Sweden's first polytechnic and is with Aalto University School of Science and Technology in Espoo, depending on definition, Scandinavia's largest institution of higher education in technology and one of the leading technical universities in Europe.

### The Opportunity

Researchers at KTH, Sweden's leading technical university, gather real-time traffic data from a variety of sources such as GPS from large numbers of vehicles, radar sensors on motorways, congestion charging, weather, etc. The integration and analysis of the data in order to better manage traffic is a difficult task.

### What Makes It Smarter

Collected data is now flowing into IBM InfoSphere Streams software—a unique software tool that analyzes large volumes of streaming, real-time data, both structured and unstructured. The data is then used to help intelligently identify current conditions, and estimate how long it would take to travel from point to point in the city, offer advice on various travel alternatives, such as routes, and eventually help improve traffic in a metropolitan area.

### Real Business Results

- Uses diverse data, including GPS locations, weather conditions, speeds and flows from sensors on motorways, incidents and roadworks
- Enters data into the InfoSphere Streams software, which can handle all types of data, both structured and unstructured
- Handles, in real time, the large traffic and traffic-related data streams to enable researchers to quickly analyze current traffic conditions and develop historical databases for monitoring and more efficient management of the system



---

## Solution Components

- IBM® InfoSphere™ Streams
  - IBM BladeCenter® HS22
  - IBM BladeCenter H Chassis
  - IBM System Storage® DS3400
  - Red Hat Linux®
- 

## For more information

Please contact your IBM sales representative or IBM Business Partner.  
Visit us at: [ibm.com/education](http://ibm.com/education)

To learn more about KTH – Royal Institute of Technology visit:  
[www.kth.se/?l=en\\_UK](http://www.kth.se/?l=en_UK)



---

© Copyright IBM Corporation 2011

IBM Corporation  
1 New Orchard Road  
Armonk, NY 10504  
U.S.A.

Produced in the United States  
March 2011  
All Rights Reserved

IBM, the IBM logo, ibm.com, BladeCenter and InfoSphere are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product or service names may be trademarks or service marks of others.

The information contained in this documentation is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this documentation, it is provided “as is” without warranty of any kind, express or implied. In addition, this information is based on IBM’s current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this documentation or any other documentation. Nothing contained in this documentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM (or its suppliers or licensors), or altering the terms and conditions of the applicable license agreement governing the use of IBM software.



Please Recycle



## Marine Institute Ireland

*Putting real-time data to work and providing a platform for technology development*

---

### Overview

#### The need

The Marine Institute sought to establish SmartBay as a research, test and demonstration platform for new environmental technologies—paving the way to commercialization and the development of new markets for Irish-based companies.

#### The solution

The Institute, working with IBM, developed a pilot information system to feed environmental data into a data warehouse, where it is processed, analyzed and displayed in new ways.

#### What makes it smarter

The project yields greater insight into the bay environment, as well as providing practical value—from understanding how water quality impacts fisheries to predicting hazard locations and more.

---

When sensors become pervasive, entirely new and unexpected uses for the flood of information they produce often arise, yielding benefits far beyond those originally envisioned. Seeing the world in a new way—via technology—generates an inventive spark, prompting people to devise new uses for information that they may never have considered before.

That's exactly what is happening in Ireland's Galway Bay, as part of the SmartBay project initiated by the Marine Institute Ireland. In support of its advanced technology platform, which seeks to make Ireland a major player in the development of smart ocean technologies, the project's initial purpose was to develop a platform for testing environmental monitoring technologies, and the idea was simple: Deploy a series of radio-equipped "smart buoys" in the bay containing sensors that could collect data such as sea state (wave height and action) and other weather conditions, water data such as salinity, and similar environmental information.

### A basis for economic transformation

When the Marine Institute learned of the IBM Big Green Innovations initiative to find ways to use technology to promote and enable environmental science, the idea of a collaboration on the SmartBay project was born. The IBM Advanced Water Management Centre Dublin built upon the domain expertise of the Marine Institute, complimenting it with its deep computing intelligence.

While the synergy with the IBM Smarter Planet™ strategy's drive towards Smart Green technology was clear, the real impetus behind the decision to expand SmartBay is largely economic. Beginning in the 1990s, the Irish economy became a global growth powerhouse. Wise policy decisions and forward-thinking investment had transformed Ireland into a manufacturing phenomenon.

More recently, with the global economy encountering difficulty, Ireland's prosperity began to wane. The government saw the need to change course, moving the country towards a knowledge-based economy. Investment in projects that showcase Ireland as a technological leader would not only create new commercial opportunities,





---

## Business benefits

- Enables the creation of a vast array of diverse applications that goes far beyond the original purpose of the project, from technical research to tourism promotion
  - Real-time access via the web delivers valuable insight quickly to remote users
  - Open architecture enables new applications to be brought on line easily, combining data from both SmartBay sensors and other sources, such as geographical information systems
  - Add-on effect of the project promotes education and stimulates economic development in the Irish economy
- 

attract talent and additional capital investment, but also prompt a new generation of Irish citizens to pursue careers in knowledge-based industries.

## Taking SmartBay to a new level

The Marine Institute, working in conjunction with government agencies, research institutions and the private sector, is working together to leverage the significant R+D capacity that exists in Ireland to help drive economic development. There is clear potential to expand SmartBay into an international platform demonstrating new approaches to environmental challenges and delivering new technological solutions for a range of global markets.

IBM is working with the Marine Institute to speed the process of innovation, starting with an assessment of existing capabilities. The team saw that if the data could be centralized, processed and accessed in the right way, it could become far more useful—the information already available could be turned into intelligence and put to work to create real practical value that impacts the lives of citizens directly.

IBM designed and deployed an enterprise-scale data warehouse using IBM InfoSphere™ Warehouse, that is connected to the SmartBay sensors, as well as external sources such as mapping databases and sensors beyond the bay. An open-standards application layer processes and analyzes the data in a variety of ways, making it available via a Web interface enabled by IBM WebSphere® Portal and WebSphere Application Server. Additional WebSphere products, including WebSphere MQ and WebSphere Sensor Events, provide a key middleware layer that integrates the sensors with the data warehouse. To ensure reliability and scalability, the system is housed on IBM System x® 3950 servers.

---

## Smarter water:

## Creating new value from environmental data



### Instrumented

Sensors deployed on buoys in Galway Bay transmit key data on ocean conditions and water quality.



### Interconnected

Sensor data is fed into a central data warehouse for aggregation and processing, and can be accessed by diverse groups using customized web applications to generate targeted value.



### Intelligent

Combining real-time data with a flexible technology platform creates near-limitless new uses for information—from environmental research to predictive monitoring, technology validation and much



---

## Solution components

### Software

- IBM DB2® Alphablox® v9.5
- IBM DB2 Enterprise Server Edition v9.5
- IBM InfoSphere™ Streams
- IBM WebSphere® Application Server v6.1
- IBM WebSphere MQ v5
- IBM WebSphere Sensor Events
- IBM WebSphere Portal Server v6.1

### Servers

- IBM System x® 3950

### Services

- IBM Global Business Services®
- 

*“The immediate benefits of SmartBay, whether it’s helping and supporting industrial development or promoting marine safety, are tangible, direct and worthwhile.”*

—John Gaughan, project coordinator,  
SmartBay

---

The system design makes it easy to combine data from the sensors with other online databases—such as geographical information—as needed to create new functionality. Rapid development, enabled by IBM DB2® Alphablox® is an important feature, giving project managers the ability to deploy new applications quickly and easily.

The project yields greater insight into the bay environment and can provide real-time information feeds to a range of stakeholders, while at the same time enabling commercial technology developers to test new environmental product and service offerings. The project is now moving into a new phase, with higher bandwidth and powered cabled sensors being deployed that will enable more information to be gathered. IBM is also working with Irish-based companies on an advanced initiative to add stream (i.e., real-time) computing capabilities to the project, with the goal of increasing its capacity utilizing the real-time analytical processing capacity of InfoSphere Streams.

## Applications limited only by imagination

As the IBM and Marine Institute team began to map out the possibilities for delivering information and services via the SmartBay portal, more and more potential new uses began to spring up. Stakeholders—the harbormaster, fishermen, researchers, tourism officials and others—were all part of the brainstorming process. The SmartBay vision was quickly expanding far beyond its initial goals.

The variety of applications either deployed or under consideration for SmartBay is strong testament to the power of creative thinking enabled by the right technological tools. The critical element is the ability to analyze, process and present the data in a useful form, tailored to the needs of specific users. For example:

- Technology developers can conduct a variety of sophisticated studies remotely and in near real time, instead of retroactively. Climate researchers, using sensors on land paired with sensors in the bay, can learn about the exchange of CO<sub>2</sub> across the land-sea interface, and marine biologists can use acoustic sensors deployed throughout the bay to assess marine mammal populations.
- Alternative energy developers can access real-time wave data and use it to determine the effectiveness of prototype wave-energy generators, and developers of new sensor technologies can deploy prototypes on the buoys to find out how well the hardware holds up in a harsh marine environment, with continuous monitoring.
- The project can also promote commercial interests. Fishermen can use environmental data to tell them when to put to sea. Fishery managers can monitor and track water quality issues, gaining a comprehensive view of actual conditions throughout the bay.



- Applications developed as part of the SmartBay project can also help increase public safety. Mariners who spot floating objects that pose a hazard to navigation can report the location, and the system will combine this information with geographic data, real-time weather, current, and tide data to predict the path and position of the hazard hours in advance. Collaboration with the Galway harbormaster has also enabled the creation of an expert system based on human expertise that can issue flood warnings more promptly and accurately than he can himself, based on real-time weather, sea state and tidal information.

Gaughan says the project provides a positive benefit in many areas. “The immediate benefits of SmartBay, whether it’s helping and supporting industrial development or promoting marine safety, are tangible, direct and worthwhile.”

### For more information

To learn more about how IBM can help you transform your business, please contact your IBM sales representative or IBM Business Partner.

Visit us at:

- [ibm.com/government](http://ibm.com/government)
- [ibm.com/smarterplanet/water](http://ibm.com/smarterplanet/water)



---

© Copyright IBM Corporation 2010

IBM Corporation  
1 New Orchard Road  
Armonk, NY 10504  
U.S.A.

Produced in the United States of America  
November 2010  
All Rights Reserved

IBM, the IBM logo, [ibm.com](http://ibm.com), Let’s Build A Smarter Planet, the planet icons, AlphaBlox, DB2, Global Business Services, InfoSphere, System x and WebSphere are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

This case study illustrates how one IBM customer uses IBM products. There is no guarantee of comparable results.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.



Please Recycle



a jStart™  
case study



using Big Data to identify Big Opportunities in retail

# helping companies deliver the web experience their customers want.

## At a Glance

There is a “Big Data” challenge in the e-commerce industry with the explosive growth of social networking sites. With 700 million users on Facebook—expected to reach 1 billion in 2011, and Twitter up to 140 million tweets per day, retailers are trying to reach their customers and understand their shopping habits better using these channels. Without social analytics,



*“We are able to vastly improve the online shopping experience by responding almost instantly to customers and delivering the products they want to purchase at a very attractive price point.”*

-Gareth Knight  
CEO, Technovated

online retailers risk becoming a victim to this deluge of data – unable to make sense out of the massive volume of product data and customer feedback, or even able to respond to it in a timely way.

Working with IBM's jStart™ team, Technovated created a system that uses IBM BigSheets to reduce manual processes while simultaneously tackling the “Big Data” challenges that many online retailers experience.

## Providing a Big Data Edge

Technovated is able to respond to shoppers instantly based on customers' latest product searches, blog posts and tweets about recent purchases. Using this valuable consumer insight, Technovated can automatically set up new online stores in a matter of days to deliver shoppers with the products they are searching for at a competitive price point. It used to take six

See how IBM using analytics to create [Smarter Retail](#)

[ibm.com/jstart](http://ibm.com/jstart)



## About Techovated

jStart works with a wide variety of clients and customers, but frequently, we find some of the best partnerships to be with startups. Technovated is very much a partner in that vein. With offices in London and Johannesburg, Technovated describes itself this way: “we are able to vastly improve the online shopping experience by responding almost instantly to customers and delivering the products they want to purchase at a very attractive price point.” The Technovated team is focused on leveraging the latest technologies to give them—and their customers—a competitive edge. In this case, utilizing IBM Big Data technologies, like BigSheets, to provide capabilities and business opportunities that simply didn't exist for SMB's until today.



weeks to put products up for sale online. Now, using IBM technology combined with Technovated's know-how, it takes a few days.

## Enter Big Data Analytics

By using IBM BigSheets, Technovated plans to jump-start its business growth. Starting off its Web stores with a few thousand product stock-keeping units (SKUs), Technovated will quickly be able to cull through terabytes of data to set up niche e-commerce sites ranging from office chairs to running shoes.

**IBM BigSheets** is a system developed by IBM's Emerging Internet Technologies group to allow for the easy and quick exploration of big data. If you're wondering what your data may be trying to tell you, BigSheets is a great place to start—since any line-of-business professional can manipulate the tool to identify and take action on opportunities which may reside in the data, itself. Since BigSheets can merge data from numerous sources, your company can obtain a high level overview of what's possible with the data available—and the opportunity to act on those insights.



## get started with jStart:

### David Sink

Program Director, jStart Team  
IBM Emerging Technologies  
dsink@us.ibm.com  
Tel: 919.254.4648

### Ed Elze

Manager, Bus. Dev., Strategy & Client Engagement  
jStart Team, IBM Emerging Technologies  
eelze@us.ibm.com  
Tel: 360.866.0160

### Jim Smith

Manager, Client Engagements, Chief Architect  
jStart Team, IBM Emerging Technologies  
jamesmi@us.ibm.com  
Tel: 919.387.6653

### John Feller

Manager, Development  
jStart Team, IBM Emerging Technologies  
fellerj@us.ibm.com  
Tel: 919.543.7971

### Learn More:

[ibm.com/jstart/bigsheets](http://ibm.com/jstart/bigsheets)  
[ibm.com/jstart/bigdata](http://ibm.com/jstart/bigdata)  
[ibm.com/jstart/textanalytics](http://ibm.com/jstart/textanalytics)  
[ibm.com/jstart/portfolio/technovated.html](http://ibm.com/jstart/portfolio/technovated.html)  
[jstart@us.ibm.com](mailto:jstart@us.ibm.com)

The jStart team also has extensive experience with IBM **data analytics** technologies and solutions as well.

By leveraging these technologies, your business could extract information from publicly available sources, internal data sources, and partner resources, and use them to identify patterns, markets, and opportunities to make the sale. In the end, big data can help identify big opportunities for retail. Ready to get started? jStart is. Contact us today.

## Who is jStart?

jStart is a highly skilled team focused on providing fast, smart, and valuable business solutions leveraging the latest technologies. The team typically focuses on emerging technologies which have commercial potential within 12-18 months. This allows the team to keep ahead of the adoption curve, while being prepared for client engagements and partnerships. The team's focus in 2011 includes: big data, text analytics, and the commercialization of IBM's Watson technologies.





## TerraEchos and IBM

*Streaming data technology supports covert intelligence and surveillance sensor systems*

---

### Overview

#### The need

U.S. Department of Energy (DOE) Research lab needed a solution to protect and monitor critical infrastructure and secure its perimeters and border areas.

#### The solution

IBM Business Partner, TerraEchos, implemented an advanced security and covert surveillance system based on the TerraEchos Adelos S4 System with IBM InfoSphere Streams software and IBM BladeCenter hardware.

#### The benefit

Captures and analyzes huge volumes of real-time, streaming, acoustical data from sensors around research lab perimeters and borders, providing unprecedented insight to detect, classify, locate, track, and deter potential threats.

---

A leading provider of covert intelligence and surveillance sensor systems, TerraEchos, Inc., helps organizations protect and monitor critical infrastructure and secure borders. One TerraEchos client is a science-based, applied engineering national laboratory dedicated to supporting the U.S. Department of Energy in nuclear and energy research, science and national defense.

One of the lab's initiatives is to be the first to develop safe, clean and reliable nuclear power. Another is to investigate and test emerging capabilities for the production, manufacturing, conveyance, transmission and consumption of renewable energy, such as solar and wind power. Securing the scientific intelligence, technology and resources related to these initiatives is vital. Protecting and sustaining the resiliency and operational reliability of the country's power infrastructures—from natural disasters, cyber attacks and terrorism—are matters of national and homeland security.

Protecting its work and securing America's energy future are responsibilities the lab takes seriously. To this end, it needed a technology solution that would detect, classify, locate and track potential threats—both mechanical and biological; above and below ground—to secure the lab's perimeters and border areas. This solution would provide scientists with more situational awareness and enable a faster and more intelligent response to any threat.

### Distinguishing the sound of a whisper from the wind even from miles away

The requirements of the ideal solution were considerable. The solution would have to continuously consume and analyze massive amounts of information-in-motion, including the movements of humans, animals and the atmosphere, such as wind. In addition, because scientists lacked time to record the data and listen to it later, the solution had to gather and analyze information simultaneously.





---

## Solution components:

### Software

- IBM® InfoSphere® Streams

### Server

- IBM BladeCenter® servers
- 

Once analyzed, scientists could extract meaningful intelligence, as well as verify and validate the data, such as distinguishing between the sounds of a trespasser versus a grazing animal. To put the sophistication of the needed technology into perspective, the data consumption and analytical requirements would be akin to listening to 1,000 MP3 songs simultaneously and successfully discerning the word “zero” from every song—within a fraction of a second.

The solution would also serve as the lab’s central nervous system and would have to meet strict technical requirements, including:

- Interoperability, allowing sensors to work with other sensor types—such as video data—and enabling scientists to collect an array of data and create a holistic view of a situation.
- Scalability to support new requirements as the lab’s fiber-optic arrays, surveillance areas, and security perimeters change.
- Extensibility, serving as a framework to fit into the lab’s existing IT architecture and integrating with signal processors and mobile and mapping applications.

To meet these requirements, the lab sought to implement and deploy an advanced security and surveillance system.

## Advanced fiber-optics combine with real-time streaming data

The lab turned to IBM® Business Partner, TerraEchos, to implement an advanced security and covert surveillance system based on its TerraEchos Adelos S4 System, IBM InfoSphere® Streams software and IBM BladeCenter® servers. InfoSphere Streams is part of the IBM big data platform.

TerraEchos selected InfoSphere Streams as the engine that processes approximately 1,600 megabytes of data in motion continually generated from fiber optic sensor arrays. The processing capacity of InfoSphere Streams enables Adelos to analyze all of the data streaming from the sensors. In addition, the technology enables Adelos to match the sound patterns against an extensive library of algorithms, giving TerraEchos the most robust classification system in the industry.

The Adelos S4 solution is based on advanced fiber-optic acoustic sensor technology licensed from the United States Navy. Using InfoSphere Streams as the underlying analytics platform, the Adelos S4 solution analyzes highly unstructured audio data in real time before the audio signals are stored in the database. InfoSphere Streams allows multiple sensor types and associated streams of structured and unstructured data to be integrated into a fused intelligence system for threat detection, classification, correlation, prediction and communication by means of a service-oriented architecture (SOA).



---

*“Given our data processing and analytical challenges associated with the Adelos Sensor Array, InfoSphere Streams is the right solution for us and our customers. We look forward to growing our strategic relationship with IBM across various sectors and markets to help revolutionize the concept of Sensor as a Service.”*

– Dr. Alex Philp, President and CEO,  
TerraEchos, Inc.

---

Adelos S4 technology comprises a fiber-optic sensor array buried in the ground to gather real-time acoustic information. These data are analyzed, and the sound patterns are matched against complex algorithms to determine what made the noise. Incorporating InfoSphere Streams technology, the Adelos S4 system can instantly identify, distinguish and classify a variety of objects detected by the fiber-optic sensor array, such as a human whisper, the pressure of a footstep and the chirping of a bird.

### **Distinguishing between true and false threats**

The solution captures and transmits volumes of real-time, streaming acoustical data from around the lab premises, providing unprecedented insight into any event. Specifically, the system enables scientists and security personnel to “hear” what is going on—even when the disturbance is miles away. In fact, the solution is so sensitive and the analytics so sophisticated that scientists can recognize and distinguish between the sound of a human voice and the wind. In this way, the lab can confidently determine whether a potential security threat is approaching—and prepare for action—or whether it is simply a storm.

Using miles of fiber-optic cables and thousands of listening devices buried underground, the lab collects and analyzes gigabytes of data within seconds and then classifies that data. These capabilities enable the lab to extend its perimeter security and gain a strategic advantage. It not only enables security to make the best decisions about apprehending the trespassers—such as how many officers to deploy and which tactics to use—but also thwarts any plans the intruders may have had to breach the property.

### **Meeting data processing and analytical challenges**

The solution is part of a more comprehensive security system. With the ability to integrate and collect data from video and airborne surveillance systems, scientists gain a holistic view of potential threats and issues—or nonissues. For instance, by cross-analyzing the acoustic data collected by the solution with the video data of another, the lab can eliminate or minimize unnecessary security actions, such as dispatching crews to investigate sounds made by a herd of deer or a fallen tree.

Finally, in addition to meeting the lab’s requirements for extensibility, interoperability and scalability, the solution saves the lab costs associated with data storage because data does not have to be stored before being analyzed.

“Given our data processing and analytical challenges associated with the Adelos Sensor Array, InfoSphere Streams is the right solution for us and our customers,” says Dr. Alex Philp, President and CEO of TerraEchos, Inc. “We look forward to growing our strategic relationship with IBM across various sectors and markets to help revolutionize the concept of Sensor as Service.”

## For more information

To learn more about IBM InfoSphere Streams, visit:

[ibm.com/software/data/infosphere/streams](http://ibm.com/software/data/infosphere/streams)

To learn more about IBM big data, visit:

[ibm.com/software/data/bigdata](http://ibm.com/software/data/bigdata)

To increase your big data knowledge and skills, visit:

[www.BigDataUniversity.com](http://www.BigDataUniversity.com)

To get involved in the conversation, visit:

[www.smartercomputingblog.com/category/big-data](http://www.smartercomputingblog.com/category/big-data)

For information on TerraEchos visit:

[www.terraechos.com](http://www.terraechos.com)



---

© Copyright IBM Corporation 2011

IBM Corporation  
Software Group  
Route 100  
Somers, NY 10589  
U.S.A.

Produced in the United States of America  
October 2011  
All Rights Reserved

IBM, the IBM logo, [ibm.com](http://ibm.com), InfoSphere and BladeCenter are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.



Please Recycle



# University of Ontario Institute of Technology

*Leveraging key data to provide proactive patient care*

---

## Overview

### The need

To better detect subtle warning signs of complications, clinicians need to gain greater insight into the moment-by-moment condition of patients.

### The solution

A first-of-its-kind, stream-computing platform was developed to capture and analyze real-time data from medical monitors, alerting hospital staff to potential health problems before patients manifest clinical signs of infection or other issues.

### What makes it smarter

Early warning gives caregivers the ability to proactively deal with potential complications—such as detecting infections in premature infants up to 24 hours before they exhibit symptoms.

---

The rapid advance of medical monitoring technology has done wonders to improve patient outcomes. Today, patients are routinely connected to equipment that continuously monitors vital signs such as blood pressure, heart rate and temperature. The equipment issues an alert when any vital sign goes out of the normal range, prompting hospital staff to take action immediately, but many life-threatening conditions do not reach critical level right away. Often, signs that something is wrong begin to appear long before the situation becomes serious, and even a skilled and experienced nurse or physician might not be able to spot and interpret these trends in time to avoid serious complications.

Unfortunately, the warning indicators are sometimes so hard to detect that it is nearly impossible to identify and understand their implications until it is too late. One example of such a hard-to-detect problem is nosocomial infection, which is contracted at the hospital and is life threatening to fragile patients such as premature infants.

According to physicians at the University of Virginia,<sup>1</sup> an examination of retrospective data reveals that, starting 12 to 24 hours before any overt sign of trouble, almost undetectable changes begin to appear in the vital signs of infants who have contracted this infection. The indication is a pulse that is within acceptable limits, but not varying as it should—heart rates normally rise and fall throughout the day. In a baby where infection has set in, this doesn't happen as much and the heart rate becomes too regular over time. So, while the information needed to detect the infection is present, the indication is very subtle; rather than being a single warning sign, it is a trend over time that can be difficult to spot, especially in the fast-paced environment of an intensive care unit.







---

### Business benefits

- Holds the potential to give clinicians an unprecedented ability to interpret vast amounts of heterogeneous data in real time, enabling them to spot subtle trends
  - Combines physician and nurse knowledge and experience with technology capabilities to yield more robust results than can be provided by monitoring devices alone
  - Provides a flexible platform that can adapt to a wide variety of medical monitoring needs
- 

The monitors continuously generate information that can give early warning signs of an infection, but the data is too large for the human mind to process in a timely manner. Consequently, the information that could prevent an infection from escalating to life-threatening status is often lost.

“The challenge we face is that there’s too much data,” says Dr. Andrew James, staff neonatologist at The Hospital for Sick Children (SickKids) in Toronto. “In the hectic environment of the neonatal intensive care unit, the ability to absorb and reflect upon everything presented is beyond human capacity, so the significance of trends is often lost.”

### Making better use of the data resource

The significance of the data overload challenge was not lost on Dr. Carolyn McGregor, Canada Research Chair in Health Informatics at the University of Ontario Institute of Technology (UOIT). “As someone who has been doing a lot of work with data analysis and data warehousing, I was immediately struck by the plethora of devices providing information at high speeds—information that went unused,” she says. “Information that’s being provided at up to 1,000 readings per second is summarized into one reading every 30 to 60 minutes, and it typically goes no further. It’s stored for up to 72 hours and is then discarded. I could see that there were enormous opportunities to capture, store and utilize this data in real time to improve the quality of care for neonatal babies.”

With a shared interest in providing better patient care, Dr. McGregor and Dr. James partnered to find a way to make better use of the information produced by monitoring devices. Dr. McGregor visited researchers at the IBM T.J. Watson Research Center’s Industry Solutions Lab (ISL), who were extending a new stream-computing

---

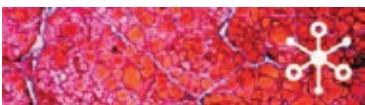
## Smarter healthcare: Using streaming data to help clinicians spot infections

---



### Instrumented

Patient’s vital-sign data is captured by bedside monitoring devices up to 1,000 times per second.



### Interconnected

Monitoring-device data and integrated clinician knowledge are brought together in real time for an automated analysis using a sophisticated, streamlined computing platform.



### Intelligent

Detecting medically significant events even before patients exhibit symptoms will enable proactive treatment before the condition worsens, eventually increasing the success rate and potentially saving lives.



---

## Solution components

### Software

- IBM InfoSphere™ Streams
- IBM DB2®

### Research

- IBM T.J. Watson Research Center
- 

*“I could see that there were enormous opportunities to capture, store and utilize this data in real time to improve the quality of care for neonatal babies.”*

—Dr. Carolyn McGregor, Canada Research Chair in Health Informatics, University of Ontario Institute of Technology

---

platform to support healthcare analytics. A three-way collaboration was established, with each group bringing a unique perspective—the hospital focus on patient care, the university’s ideas for using the data stream, and IBM providing the advanced analysis software and information technology expertise needed to turn this vision into reality.

The result was Project Artemis, part of IBM’s First-of-a-Kind program, which pairs IBM’s scientists with clients to explore how emerging technologies can solve real-world business problems. Project Artemis is a highly flexible platform that aims to help physicians make better, faster decisions regarding patient care for a wide range of conditions. The earliest iteration of the project is focused on early detection of nosocomial infection by watching for reduced heart rate variability along with other indications. For safety reasons, in this development phase the information is being collected in parallel with established clinical practice and is not being made available to clinicians. The early indications of its efficacy are very promising.

Project Artemis is based on IBM InfoSphere™ Streams, a new information processing architecture that enables near-real-time decision support through the continuous analysis of streaming data using sophisticated, targeted algorithms. The IBM DB2® relational database provides the data management required to support future retrospective analyses of the collected data.

## A different kind of research initiative

Because SickKids is a research institution, moving the project forward was not difficult. “The hospital sees itself as involved in the generation of new knowledge. There’s an expectation that we’ll do research. We have a research institute and a rigorous research ethics board, so the infrastructure was already there,” Dr. James notes.

Project Artemis was a consequence of the unique and collaborative relationship between SickKids, UOIT and IBM. “To gain its support, we needed to do our homework very carefully and show that all the bases were covered. The hospital was cautious, but from the beginning we had its full support to proceed.”

Even with the support of the hospital, there were challenges to be overcome. Because Project Artemis is more about information technology than about traditional clinical research, new issues had to be considered. For example, the hospital CIO became involved because the



system had to be integrated into the existing network without any impact. Regulatory and ethical concerns are part of any research at SickKids, and there were unique considerations here in terms of the protection and security of the data. The research team's goal was to exceed provincial and federal requirements for the privacy and security of personal health information—the data had to be safeguarded and restricted more carefully than usual because it was being transmitted to both the University of Ontario Institute of Technology and to the IBM T.J. Watson Research Center.

After the overarching concerns were dealt with, the initial tests could begin. Two infant beds were instrumented and connected to the system for data collection. To ensure safety and effectiveness, the project is being deployed slowly and carefully, notes Dr. James. “We have to be careful not to introduce new technologies just because they're available, but because they really do add value,” says Dr. James. “It is a stepwise process that is still ongoing. It started with our best attempt at creating an algorithm. Now we're looking at its performance, and using that information to fine tune it. When we can quantify what various activities do to the data stream, we'll be able to filter them out and get a better reading.” The ultimate goal is to create a robust, valid system fit to serve as the basis for a randomized clinical trial.

### **Merging human knowledge and technology**

The initial test of the Project Artemis system captured the data stream from bedside monitors and processed it using algorithms designed to spot the telltale signs of nosocomial infection. The algorithm concept is the essential difference between the Artemis system and the existing alarms built into bedside monitors. Although the first test is focused on nosocomial infection, the system has the flexibility to handle any rule on any combination of behaviors across any number of data streams. “What we've built is a set of rules that reflects our best understanding of the condition. We can change and update them as we learn more, or to account for variations in individual patients. Artemis represents a whole new level of capability,” Dr. James notes.

The truly significant aspect of the Project Artemis approach is how it brings human knowledge and expertise together with device-generated data to produce a better result. The system's outputs are based on algorithms developed as a collaboration between the clinicians themselves and programmers. This inclusion of the human element is critical,



because good patient care cannot be reduced to mere data points. Validation of these results by an experienced physician is vital since the interpretation of these results has to do with medical knowledge, judgment, skill and experience. As part of the project, the rules being used by Project Artemis are undergoing separate clinical research to support evidence-based practice.

Artemis also holds the potential to become much more sophisticated. For example, eventually it might integrate a variety of data inputs in addition to the streaming data from monitoring devices—from lab results to observational notes about the patient's condition to the physician's own methods for interpreting information. In this way, the knowledge, understanding and even intuition of physicians and nurses will become the basis of the system that enables them to do much more than they could on their own.

“In the early days, there was a lot of concern that computers would eventually ‘replace’ all health care providers,” Dr. James says. “But now we understand that human beings cannot do everything, and it's quite helpful to develop tools that enhance and extend the physicians' and nurses' capabilities. I look to a future where I'm going to receive an alert that provides me with a comprehensive, real-time view of the patient, allowing me to make better decisions on the spot.”

### **Broadening the impact of Artemis**

The flexibility of the platform means that in the future, any condition that can be detected through subtle changes in the underlying data streams can be the target of the system's early-warning capabilities. Also, since it depends only on the availability of a data stream, it holds the potential for use outside the ICU and even outside the hospital. For example, the use of remote sensors and wireless connectivity would allow the system to monitor patients wherever they are, while still providing life-saving alerts in near-real time.

“I think the framework would also be applicable for any person who requires close monitoring—children with leukemia, for example,” says Dr. James. “These kids are at home, going to school, participating in sports—they're mobile. It leads into the whole idea of sensors attached to or even implanted in the body and wireless connectivity. Theoretically, we could ultimately monitor these conditions from anywhere on the planet.”

## For more information

To learn more about how IBM can help you transform your business, contact your IBM sales representative or IBM Business Partner.

Visit us at: [ibm.com/smarterplanet/healthcare](http://ibm.com/smarterplanet/healthcare)



---

© Copyright IBM Corporation 2010

IBM Corporation  
1 New Orchard Road  
Armonk, NY 10504  
U.S.A.

Produced in the United States of America  
December 2010  
All Rights Reserved.

IBM, the IBM logo, [ibm.com](http://ibm.com), Let's Build A Smarter Planet, Smarter Planet, the planet icons, DB2 and InfoSphere are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

This case study illustrates how one IBM customer uses IBM products. There is no guarantee of comparable results.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.

<sup>1</sup> P. Griffin and R. Moorman, "Toward the early diagnosis of neonatal sepsis and sepsis-like illness using novel heart rate analysis," *Pediatrics*, vol. 107, no. 1, 2001.



Please Recycle



## Uppsala University, Swedish Institute of Space Physics and IBM

*Streaming real-time data supports large scale study  
of space weather*

---

### Overview

#### The need

Plasma eruptions from the sun adversely affect energy transmission over power lines, communications via radio and TV signals, airline and space travel, and satellites. Collecting huge amounts of data has surpassed the ability to store or analyze it.

#### The solution

IBM InfoSphere Streams software collects huge volumes of data to be analyzed in real time. Data filtering capabilities separate meaningful data from “noise” to reduce data storage requirements.

#### The benefit

Predictive analysis warns when a magnetic storm on the sun will reach the earth; preventive changes to sensitive satellites and power grids can minimize damage caused by energy bursts from the sun.

---

Uppsala University, the Swedish Institute of Space Physics and IBM® are collaborating on major new Stream Computing project to analyze massive volumes of information in real time to better understand “space weather.” By using IBM InfoSphere® Streams to analyze data from sensors that track high frequency radio waves, endless amounts of data can be captured and analyzed on the fly. This project offers the capability to perform analytics on at least 6 gigabytes of data per second or 21,600 gigabytes per hour—the equivalent of all the web pages on the Internet. InfoSphere Streams software is part of IBM’s big data platform.

### Analyzing large volumes of space weather data in real time

Scientists sample high frequency radio emissions from space to study and forecast “space weather” or the effect of plasma eruptions on the sun that reach the earth and adversely affect energy transmission over power lines, communications via radio and TV signals, airline and space travel, and satellites. However, the recent advent of new sensor technology and antennae arrays means that the amount of information collected by scientists has surpassed the ability to intelligently analyze it. IBM InfoSphere Streams, software derived from the IBM Research project System S, enables large volumes of data to be analyzed in real time making an entirely new level of analytics possible.

“IBM InfoSphere Streams is opening up a whole new way of doing science, not only in this area, but any area of e-Science where you have lots of data coming in from external sources and sensors, streaming at such high data rates you can’t handle it with conventional technology,” says Dr. Bo Thide, Professor and Head of Research, Swedish Institute of Space Physics and Director of the LOIS Space Center in Sweden. “It has helped create a paradigm shift in the area of online observation of the earth, space, sun and atmosphere.”





---

Solution components:

**Software**

- IBM® InfoSphere® Streams
- 

*“IBM InfoSphere Streams is opening up a whole new way of doing science, not only in this area, but any area of e-Science. It has helped create a paradigm shift in the area of online observation of the earth, space, sun and atmosphere.”*

– Dr. Bo Thide, Professor and Head of Research, Swedish Institute of Space Physics, and Director of the LOIS Space Center in Sweden

---

Sunspot activity, electromagnetic storms, and other types of solar activity can impact communications signals. As critical infrastructure such as power grids and telecommunications networks become more digitally aware, instrumented and interconnected, it is increasingly important to understand how these can be affected by influences such as electromagnetic interference or other changes in the atmosphere.

Researchers at Uppsala University and the Swedish Institute of Space Physics worked with the LOIS Space Center facility in Sweden to develop a new type of tri-axial antenna that streams three-dimensional radio data from space, extracting a magnitude more physical information than any other type of antennae array before. Since researchers need to measure signals from space over large time spans, the raw data generated by even one antenna quickly becomes too large to handle or store.

“We’ve embarked upon an entirely new way of observing radio signals using digital sensors that produce enormous amounts of data,” Thide adds. “With this type of research, you have to be able to analyze as much data as possible on the fly. There is no way to even consider storing it. InfoSphere Streams is playing a pivotal role in this project. Without it, we could not possibly receive this volume of signals and handle them at such a high data rate because until now, there was not a structured, stable way of analyzing it.”



---

*“InfoSphere Streams is playing a pivotal role in this project. Without it, we could not possibly receive this volume of signals and handle them at such a high data rate because until now, there was not a structured, stable way of analyzing it.”*

– Dr. Bo Thide

---

### **Predicting events in space and on the sun**

The technology addresses this problem by analyzing and filtering the data the moment it streams in, helping researchers identify the critical fraction of a percent that is meaningful, while the rest is filtered out as noise. Using a visualization package, scientists can perform queries on the data stream to look closely at interesting events, allowing them not only to forecast, but to nowcast events just a few hours away. These capabilities will help predict, for example, if a magnetic storm on the sun will reach the earth in 18 to 24 hours.

The ultimate goal of the project at Uppsala University with IBM InfoSphere Streams is to model and predict the behavior of the uppermost part of the atmosphere and its reaction to events in surrounding space and on the sun. This work could have lasting impact for future science experiments in space and on earth. With a unique ability to predict how plasma clouds travel in space, new efforts can be made to minimize damage caused by energy bursts or make changes to sensitive satellites, power grids or communications systems.

### **For more information**

To learn more about IBM InfoSphere Streams, visit:

[ibm.com/software/data/infosphere/streams](http://ibm.com/software/data/infosphere/streams)

To learn more about IBM big data, visit:

[ibm.com/software/data/bigdata](http://ibm.com/software/data/bigdata)

To increase your big data knowledge and skills, visit:

[www.BigDataUniversity.com](http://www.BigDataUniversity.com)

To get involved in the conversation:

[www.smartercomputingblog.com/category/big-data](http://www.smartercomputingblog.com/category/big-data)

For more information on Uppsala University, visit

[www.uu.se](http://www.uu.se)

For more information on the Swedish Institute of Space Physics, visit:

[www.irfu.se](http://www.irfu.se)

For more information on the LOIS Space Center, visit:

[www.lois-space.net](http://www.lois-space.net)





---

## Smart is...

### ***Pinpointing the optimal location for wind turbines to maximize power generation and reduce energy costs***

Precise placement of a wind turbine can affect its performance and its useful life. For Vestas, the world's largest wind energy company, gaining new business depends on responding quickly and delivering business value. To succeed, Vestas uses one of the largest supercomputers worldwide along with a new big data modeling solution to slice weeks from data processing times and support 10 times the amount of data for more accurate turbine placement decisions. Improved precision provides Vestas customers with greater business case certainty, quicker results and increased predictability and reliability in wind power generation.

---

## Vestas

### *Turning climate into capital with big data*

For centuries, sailors have seen how fickle the wind can be. It ebbs and flows like the tide and can allow ships to travel great distances or remain becalmed at sea.

But despite the wind's capricious nature, new advances in science and technology enable energy producers to transform the wind into a reliable and steadfast energy source—one that many believe will help alleviate the problems of the world's soaring energy consumption.

“Wind energy is one of today's most important renewable energy sources,” says Lars Christian Christensen, vice president, Vestas Wind Systems A/S. “Fossil fuels will eventually run out. Wind is renewable, predictable, clean and commercially viable. By 2020 as much as 10 percent of the world's electricity consumption will be satisfied by wind energy and we believe that wind power is an industry that will be on par with oil and gas.”

### **Producing electricity from wind**

Making wind a reliable source of energy depends greatly on the placement of the wind turbines used to produce electricity. The windiest location may not generate the best output and revenue for energy companies. Turbulence is a significant factor as it strains turbine components, making them more likely to fail. Avoiding pockets of turbulence can extend the service life of turbines and lower operating costs, which reduces the cost per kilowatt hour of energy produced.

---

*“We can now show our customers how the wind behaves and provide a solid business case that is on par with any other investment that they may have.”*

– Lars Christian Christensen, Vice President, Vestas Wind Systems A/S

---



---

## Business benefits

- Reduces response time for wind forecasting information by approximately 97 percent—from weeks to hours—to help cut development time
  - Improves accuracy of turbine placement with capabilities for analyzing a greater breadth and depth of data
  - Lowers the cost to customers per kilowatt hour produced and increases customers' return on investment
  - Reduces IT footprint and costs, and decreases energy consumption by 40 percent—all while increasing computational power
- 

Selecting wind turbine sites is a science that Vestas understands well. Since 1979, this Danish company has been engaged in the development, manufacture, sale and maintenance of wind power systems to generate electricity. The company has installed more than 43,000 land-based and offshore wind turbines in 66 countries on six continents. Today, Vestas installs an average of one wind turbine every three hours, 24 hours a day, and its turbines generate more than 90 million megawatt-hours of energy per year—enough electricity to supply millions of households.

“Customers want to know what their return on investment will be and they want business case certainty,” says Christensen who heads the company’s division responsible for determining the placement of wind turbines. “For us to achieve business case certainty, we need to know exactly how the wind is distributed across potential sites, and we need to compare this data with the turbine design specifications to make sure the turbine can operate at optimal efficiency at that location.”

What happens if engineers pick a sub-optimal location? According to Christensen, the cost of a mistake can be tremendous. “First of all, if the turbines do not perform as intended, we risk losing customers. Secondly, placing the turbines in the wrong location affects our warranty costs. Turbines are designed to operate under specific conditions and can break if they are operating outside of these parameters.”

For Vestas, the process of establishing a location starts with its wind library, which incorporates data from global weather systems with data collected from existing turbines. Combined, this information helps the company not only select the best site for turbine placement, but also helps forecast wind and power production for its customers.

---

## Smarter Energy:

## Increases wind power generation through optimal turbine placement

---



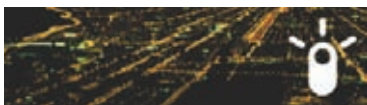
### Instrumented

Determines the optimal turbine placement using weather forecasts and data from operational wind power plants to create hourly and daily predictions regarding energy production.



### Interconnected

Combines turbine data with data on temperature, barometric pressure, humidity, precipitation, wind direction and velocity from the ground level up to 300 feet.



### Intelligent

Precisely models wind flow to help staff understand wind patterns and turbulence near each wind turbine and select the best location to reduce the cost per kilowatt hour of energy produced.



---

## Solution components:

### Software

- IBM® InfoSphere® BigInsights Enterprise Edition

### Hardware

- IBM System x® iDataPlex® dx360 M3
  - IBM System Storage® DS5300
- 

*“In our development strategy, we see growing our library in the range of 18 to 24 petabytes of data. And while it’s fairly easy to build that library, we needed to make sure that we could gain knowledge from that data.”*

– Lars Christian Christensen

---

“We gather data from 35,000 meteorological stations scattered around the world and from our own turbines,” says Christensen. “That gives us a picture of the global flow scenario. Those models are then cobbled to smaller models for regional level called mesoscale models. The mesoscale models are used to establish our huge wind library so we can pinpoint a specific location at a specific time of day and tell what the weather was like.”

The company’s previous wind library provided detailed information in a grid pattern with each grid measuring 27x27 kilometers (about 17x17 miles). Using computational fluid dynamics models, Vestas engineers can then bring the resolution down even further—to about 10x10 meters (32x32 feet)—to establish the exact wind flow pattern at a particular location.

However, in any modeling scenario, the more data and the smaller the grid area, the greater the accuracy of the models. As a result, Christensen’s team wanted to expand its wind library more than 10 fold to include a larger range of weather data over a longer period of time. Additionally, the company needed a more powerful computing platform to run global forecasts much faster. Often company executives had to wait up to three weeks for feedback regarding potential sites—an unacceptable amount of time for Vestas and its customers in this competitive industry.

“In our development strategy, we see growing our library in the range of 18 to 24 petabytes of data,” says Christensen. “And while it’s fairly easy to build that library, we needed to make sure that we could gain knowledge from that data.”

## Turning climate into capital

Working with IBM, Vestas today is implementing a big data solution that is slicing weeks from data processing time and helping staff more quickly and accurately predict weather patterns at potential sites to increase turbine energy production. Data currently stored in its wind library comprises nearly 2.8 petabytes and includes more than 178 parameters, such as temperature, barometric pressure, humidity, precipitation, wind direction and wind velocity from the ground level up to 300 feet, along with the company’s own recorded historical data. Future additions for use in predictions include global deforestation metrics, satellite images, historical metrics, geospatial data and data on phases of the moon and tides.



---

## Journey to Smarter Computing

### **Designed for Data**

Implementing a big data solution enables Vestas to create a wind library to hold 18 to 24 petabytes of weather and turbine data at various levels of granularity and reduce the geographic grid area used for modeling by 90 percent for increased accuracy.

### **Tuned to the Task**

Working with IBM, Vestas can increase computational power while shrinking its IT footprint and reducing server energy consumption by 40 percent. Today, twice the number of servers can be run in each of its supercomputer's 12 racks.

### **Managed for Rapid Service Delivery**

Processing huge volumes of climate data and the ability to gain insight from that data enables Vestas to forecast optimal turbine placement in 15 minutes instead of three weeks. This in turn shortens the time to develop a wind turbine site by nearly a month.

---

“We could pose the questions before, but our previous systems were not able to deliver the answers, or deliver the answers in the required timeframe,” says Christensen. “Now, if you give me the coordinates for your back yard, we can dive into our modeled wind libraries and provide you with precise data on the weather over the past 11 years, thereby predicting future weather and delivering power production prognosis. We have the ability to scan larger areas and determine more quickly our current turbine coverage geographically and see if there are spots we need to cover with a type of turbine. We can also assess information on how each turbine is operating and our potential risk at a site.”

IBM® InfoSphere® BigInsights software running on an IBM System x® iDataPlex® system serves as the core infrastructure to help Vestas manage and analyze weather and location data in ways that were not previously possible. For example, the company can reduce the base resolution of its wind data grids from a 27x27 kilometer area down to a 3x3 kilometer area (about 1.8x1.8 miles)—a nearly 90 percent reduction that gives executives more immediate insight into potential locations. Christensen estimates this capability can eliminate a month of development time for a site and enable customers to achieve a return on investment much earlier than anticipated.

“IBM InfoSphere BigInsights helps us gain access to knowledge in a very efficient and extremely fast way and enables us to use this knowledge to turn climate into capital,” says Christensen. “Before, it could take us three weeks to get a response to some of our questions simply because we had to process a lot of data. We expect that we can get answers for the same questions now in 15 minutes.”

For customers, the detailed models mean greater business case certainty, quicker results and increased predictability and reliability on their investment.

“Our customers need predictability and reliability, and that can only happen using systems like InfoSphere BigInsights,” says Christensen. “We can give customers much better financial warranties than we have been able to in the past and can provide a solid business case that is on par with any other investment that they may have.”



---

*“IBM InfoSphere BigInsights helps us gain access to knowledge in a very efficient and extremely fast way and enables us to use this knowledge to turn climate into capital.”*

– Lars Christian Christensen

---

## **Smarter Computing by design**

### **Tackling big data challenges**

Vestas and IBM worked together to implement IBM InfoSphere BigInsights software, designed to enable organizations to gain insight from information flows that are characterized by variety, velocity and volume. The solution combines open source Apache Hadoop software with unique technologies and capabilities from IBM to enable organizations to process very large data sets—breaking up the data into chunks and coordinating the processing across a distributed environment for rapid, efficient analysis and results.

“IBM gave us an opportunity to turn our plans into something that was very tangible right from the beginning,” says Christensen. “IBM had experts within data mining, big data and Apache Hadoop, and it was clear to us from the beginning if we wanted to improve our business, not only today, but also prepare for the challenges we will face in three to five years, we had to go with IBM.”

### **Maintaining energy efficiency in its data center**

For a company committed to addressing the world's energy requirements, it's no surprise that as Vestas implemented its big data solution, it also sought a high-performance, energy efficient computing environment that would reduce its carbon footprint. Today, the platform that drives its forecasting and analysis comprises a hardware stack based on the IBM System x iDataPlex supercomputer. This supercomputing solution—one of the world's largest to date—enables the company to use 40 percent less energy while increasing computational power. Twice the number of servers can be run in each of the system's 12 racks—reducing the amount of floor space required in its data center.

“The supercomputer provides the foundation for a completely new way of doing business at Vestas and combined with IBM software delivers a smarter approach to computing that optimizes the way we work,” says Christensen.



---

*“Before, it could take us three weeks to get a response to some of our questions simply because we had to process a lot of data. We expect that we can get answers for the same questions now in 15 minutes.”*

– Lars Christian Christensen

---

### ► The inside story: getting there

According to Christensen, the idea for this project began with the collaboration among his team, the company’s global research department and its sales business units.

“We needed to know where the goldmines of wind are hidden, and we needed to have more information to aid our decisions,” says Christensen. “We quickly formed a project group that took the idea forward and set out some key performance indicators that had to be met in order to proceed to the stage where we are today.”

For Vestas, the opportunity that a big data solution could provide made the decision easy. “Once we had the business potential of having these capabilities, it was fairly easy to gain acceptance,” says Christensen. “We were able to show the cost of a system alongside the near-term and long-term benefits, so it was really a no brainer.”



Let's build a smarter planet

### **For more information**

To learn more about how IBM can help you transform your business, please contact your IBM sales representative or IBM Business Partner.

To learn more about big data solutions from IBM, visit:

[ibm.com/software/data/bigdata](http://ibm.com/software/data/bigdata)

To learn more about IBM InfoSphere BigInsights, visit:

[ibm.com/software/data/infosphere/biginsights](http://ibm.com/software/data/infosphere/biginsights)

To increase your big data knowledge and skills, visit:

[www.BigDataUniversity.com](http://www.BigDataUniversity.com)

To get involved in the conversation:

[www.smartercomputingblog.com/category/big-data](http://www.smartercomputingblog.com/category/big-data)

For more information about Vestas Wind Systems A/S, visit:

[www.vestas.com](http://www.vestas.com)



---

## Highlights

- IBM's Watson—the computing system that competed with human contestants on Jeopardy!—illustrates how managing “Big Data” and applying analytics can help businesses gain meaningful insights
  - Watson shows how we can confidently make decisions through ranking answers, and handle structured and unstructured data by running hundreds of different kinds of analytical queries across all different kinds of information
  - Applying those innovations from Watson to an organization can help transform business models
- 

# Transform insights into action

## *IBM's Watson and the future of data*

Watson, named after IBM founder Thomas J. Watson, was built by a team of IBM scientists who set out to accomplish a grand challenge—build a computing system that rivals a human's ability to answer questions posed in natural language with speed, accuracy and confidence. The Jeopardy! format provides the ultimate challenge because the game's clues involve analyzing subtle meaning, irony, riddles, and other complexities in which humans excel and computers traditionally do not.

But Watson's breakthrough is not in natural language processing alone. Its ability to ingest massive amounts of data, apply hundreds of analytical queries to come up with an answer, and then put confidence behind that answer, represents an advance for the kinds of problems that are emerging in business.

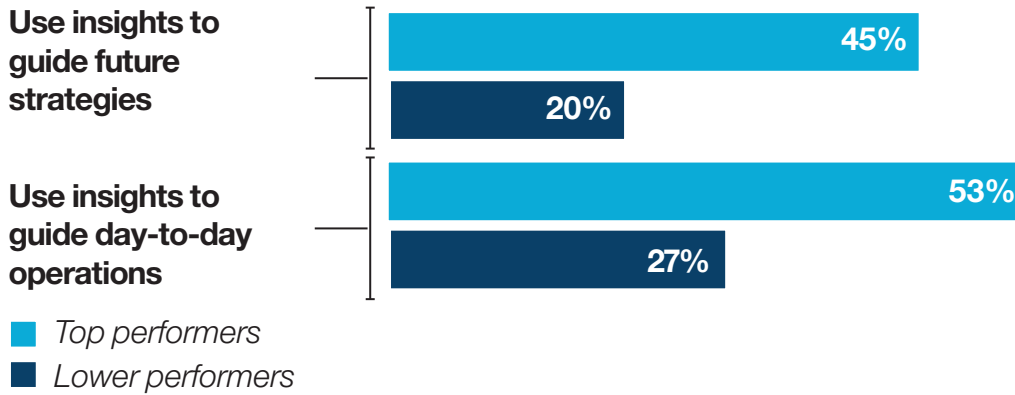
Today, computing is increasingly instrumenting business, underlying every process that runs operations—from supply chain management, to human resources and payroll, to financial management, security and risk. And now, as more of the world becomes instrumented—everything from roadways, power grids, consumer goods and food—businesses need the ability to analyze the data coming from these sources in real-time.

Traditional computing systems are built to analyze only structured data, or to run analytics in batch reporting jobs. But today's businesses require the same kind of information consumption, advanced analytics and real-time response that is needed to answer questions on Jeopardy!





*Insights to drive business decisions*



Note: Respondents were asked to rate how well their business unit or department performed the noted tasks. Chart represents answers from those who selected “very well” using a five-point scale from “not well at all” to “very well.”

Source: Analytics: The New Path to Value, a joint MIT Sloan Management Review and IBM Institute for Business Value study. Copyright © Massachusetts Institute of Technology 2010.

Figure 1: More than twice as many top performers as lower performers used analytics to guide day-to-day operations and future strategies.

*While Watson represents a technological milestone, the real pioneers will be the people and organizations that embrace this innovation and turn its potential into results.*

*How can Watson-like analytics capabilities transform your business? How does your organization’s use of “Big Data” management and business analytics compare to that of top-performing companies?*

The performance of these computing systems—the hardware and software that manages the information and runs both analytics and the business processes—is increasingly associated with the performance of the business. Watson is one example of the new kind of workloads that businesses will apply to achieve their business goals.

**Putting the power of Watson to work**

For many companies, business analytics has emerged as a strategic priority throughout the C-suite. In fact, top-performing organizations use analytics five times more than lower performers, according to a 2010 report by the IBM Institute for Business Value and MIT Sloan Management Review.

Organizations already benefiting from advanced analytics include:

- **The New York State Department of Taxation and Finance**—The organization, which processes 24 million business and personal tax returns annually, is using IBM analytics software and services to transform its approach from “pay and chase” to “next best case”.

The system identifies the next refund requests most likely to be questionable and focuses precious audit resources on these. In its five years of operation, the system has preserved more than \$1.2 billion against fraudulent requests.

- **Cincinnati Zoo**— Located in Cincinnati, Ohio, the zoo features more than 500 animal and 3,000 plant species, making it one of the largest collections in the country. To keep the facility running in a sustainable fashion and maximize resources, the Cincinnati Zoo implemented IBM analytics software. As a result, the zoo’s growing amount of information was turned into knowledge for their staff to improve operations.

The zoo was able to increase in-park spending by as much as 25 percent by utilizing 360 degree customer views. They turned that information into customized offers and perks for visitors to keep them happy and coming back, and the zoo is now able to arm their managers with real-time data that allows them to react to a dynamic and fluid business driven by seasonal weather patterns.

Business analytics has also allowed the zoo to integrate the operations and run a more sustainable business. This has helped free up their staff’s time so they can focus on the day-to-day operations in a more meaningful way, while also focusing on the larger picture of ensuring the zoo’s animals continue to receive the best care. Further, the zoo’s revenue has increased \$350,000 per year, which enables them to dedicate more resources to the well-being of the animals.

---

*“Almost immediately after going live with IBM analytics software, we were able to increase our in-park spending by as much as 25 percent by utilizing 360 degree customer views. We now have the ability to see and analyze data in all corners of our business—presented in the way we want to see it whenever we need it—and be more responsive to our customers.”*

—John Lucas, Director of Operations, Cincinnati Zoo & Botanical Garden

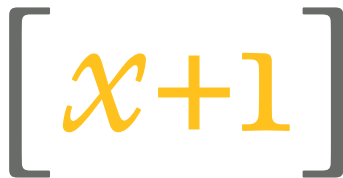
---

### For more information

IBM can provide the same kind of system, information management and analytics capabilities that power Watson for your organization. The experts who built Watson are on hand to help you chart a path to get more value out of your IT systems.

To learn more about Watson and how advanced analytics can be applied to optimize business outcomes, visit one of our IBM Analytic Solution Centers or ask about coordinating an IBM briefing at a location of your choice. Contact your IBM sales representative or IBM Business Partner for more information, or visit: [ibm.com/bao/](http://ibm.com/bao/)

Additionally, financing solutions from IBM Global Financing can enable effective cash management, protection from technology obsolescence, improved total cost of ownership and return on investment. Also, our Global Asset Recovery Services help address environmental concerns with new, more energy-efficient solutions. For more information on IBM Global Financing, visit: [ibm.com/financing](http://ibm.com/financing)



Make every interaction count

---

## Overview

### The need

Need for stronger computing power to accommodate real-time analysis on massive data volumes of online and offline data

### The solution

IBM Netezza 1000 data warehouse appliance

### The benefit

- 20% growth in digital sales – the clients see more revenue from more customers
  - Ability to gauge online and offline marketing impact
  - More robust view of the consumer
  - Break down of data silos
- 

## [x+1]

*Helping clients reach their marketing goals with analytics powered by IBM Netezza*

Digital marketers are good at collecting data, but often find it challenging to derive actionable insights from the massive volumes of information they gather online. When buying ads, for example, many marketers base their decisions on the last click from a previous campaign. This leaves them unable to identify potent indicators revealed earlier in the purchase funnel, such as in-market readiness.

This strategy is far from perfect. Some consumers are barraged with ad messages, others are under-exposed, and as a result they do not fully understand the product or offer message. The bottom line is that advertising dollars aren't being spent optimally and the business opportunity is not maximized.

How does a company manage its messaging and media channels to effectively propel consumers through the purchase funnel? The answer lies in the application of complex but essential advertising analysis on massive volumes of data in real-time. This is a capability offered by [x+1] and enabled by IBM® Netezza®.

## [x+1] and IBM Netezza

Founded in 1999, [x+1] helps marketers and agencies to maximize prospect and customer interactions across multiple digital channels through [x+1] ORIGIN, its digital marketing hub and a suite of advanced analytics. The process begins with finding consumers and by “flagging key data elements that tell you if they're in your target audience,” says Leon Zemel, [x+1]'s chief analytics officer. Then, by delivering messages based on the segment and the consumer's place in the purchase-decision funnel – along with the right exposure range (called Optimal Frequency Range, or OFR) – all calculated in real time, success is achieved.



---

*“Historically, we talked about lift in the response rate or the conversion rate. Now we’re talking about lift in total digital sales. And we’re seeing a big year-over-year impact – 20 percent growth. Net-net, the client is seeing more revenue from more customers.”*

– Leon Zemel  
Chief Analytics Officer, [x+1] Inc.

---

[x+1] ORIGIN enables the management of audience interactions through the following products and services:

- **Media+1** – An audience targeting and bidding Demand Side Platform (DSP) for pre-purchased and exchange-based digital media.
- **Site+1** – A website personalization management tool that assembles data about prospects and customers, which chooses the statistically optimal mix of offers or content to show each site visitor.
- **Landing Page+1** – A service for delivering tailored landing pages based on visitor profiles and traffic sources. When paired with Media+1, it becomes a highly effective media-aware landing page.
- Analytics tools and services, including the 2011 release of **Reach/Frequency Manager**, which provides packaged and custom reporting and insights to track and improve digital marketing across the customer purchase decision funnel.
- **Open Data Bridge DMP (Digital Management Platform)** to collect, store and manage all first and third party data for in-bound and out-bound marketing.

POE™, [x+1]’s proprietary Predictive Optimization Engine which is at the heart of [x+1] ORIGIN, is engineered to leverage sophisticated mathematical models to test, optimize and scale marketing return on investment.

The strategic and tactical marketing, and media outputs made possible by [x+1]’s technology and tools, are driven by data that spans the massive Internet population. Though it’s not about volume alone; effective use depends on the analysis of the right elements.

As Zemel sees it, too many firms rely on small-data approaches – such as attribution analysis based on the last click – which fail to track the impact of offline media. [x+1] tracks attributions across both digital and offline channels and delivers effective, predictive analysis.

It takes granular data to complete this task and the data points have to be “organized so they can be analyzed and leveraged for marketing value,” according to Zemel. As many firms have learned the hard way, massive data capture cannot be effectively leveraged with traditional database marketing technology.

### **Big computing power**

Enter IBM Netezza. [x+1] had decided to replace its legacy MySQL database with a data warehouse appliance that would provide the needed horsepower, scalability and ease of use.

---

## Solution Components

### Hardware

- IBM® Netezza® 1000

Previously [x+1] used Oracle, SAS, and in-house developed ETL processes, which put flat files directly into solutions like SAS. Data volumes were growing and the analytics team had to perform increasingly complex ad-hoc analysis to serve clients and help them grow their businesses. That meant moving from a traditional relational database management systems (RDBMS) to proprietary analytical tools.

“We used to look at every impression individually as opposed to taking a comprehensive view of that user,” Zemel says. “We had to take a more longitudinal look. But we couldn’t support that level of complexity.”

What [x+1] needed was processing power, the kind that facilitates data-intensive analysis in a real-time environment. Having heard from partners and other firms in the space, [x+1] turned to IBM Netezza. While other solutions were also considered, “We compared IBM Netezza to our Oracle environment more than anything,” Zemel says.

Based on this review, [x+1] chose the IBM Netezza data warehouse appliance and deployed it with minimal effort. One deciding factor was speed – IBM Netezza facilitates real-time analytics. Additionally [x+1] was impressed with IBM Netezza’s scalability and price/performance ratio.

The IBM Netezza data warehouse appliance architecturally integrates database, server and storage into a single, easy to manage system which requires minimal set-up and ongoing administration. It delivers high performance, out-of-the-box, with no indexing or tuning required, and it simplifies business analytics dramatically by consolidating all analytic activity in the appliance, right where the data resides.

Data is now run through TIBCO® Spotfire and placed in visualization outputs for the convenience of the end users – namely media planners and analytics professionals at digital marketing firms and their agencies. IBM Netezza helps marketers cut through the digital exhaust and respond more quickly to consumer needs. In short, it helps them synchronize large data volumes into meaningful marketing.

By installing the IBM Netezza data warehouse appliance, [x+1] was able to provide its analytics team with a simple SQL interface that could handle massive volumes of data. The analysts can focus on gleaning insights, and the engineering team can focus on the company’s core products.

---

*“For this single client, we collect five billion cross-channel marketing impressions per month from all its marketing activities. This is where we really use the power of IBM Netezza.”*

— Leon Zemel

---

At the same time, clients can now move quickly up the maturity curve – they can leverage increasingly sophisticated types of data analysis to create business value. Firms that climb the maturity curve the fastest are the ones most likely to win.

### **A client's story**

With the IBM Netezza engine empowering [x+1]’s solutions, [x+1] is helping marketers solve seemingly insoluble problems. For example, one client had a “mass of uncultivated user interactions – log files, web site analytic data, customer data,” says Zemel. “But it had trouble fully monetizing this sprawling virtual metropolis of digital customers.”

They had the typical problem of bombarding some consumers with the same ad over and over just because they visited a web site. Meanwhile, other consumers who needed multiple touches, simply didn’t get them. “Last-view attribution analysis leads us to believe that this might actually be working,” Zemel says. “But consumers are not going to switch brands just because they saw one display ad.”

The result for this client: “The audience composition was way below where it needed to be,” Zemel says. Even worse, the firm didn’t know the full impact of its marketing. “There was a disconnect between the digital investment and digital P&L.”

### **Multi-dimensional data**

To solve this problem, [x+1] applied two core customer-centric, data-driven marketing precepts:

- Define the consumer and their needs.
- Determine the messages and investment that will move the consumer along the purchase funnel.

This required a multi-dimensional data approach: The company had to update the consumer’s record with every interaction – in real-time. They also needed to access demographic and lifestyle data from third-party sources. This was needed to determine who the consumer is and their personal profile segment, as well as behavioral data based on all the touches that are being supplied to that consumer. These included banner-clicks, search activity, site visits, product signups and comparison shopping.

### How do these different data elements work together?

Prospect segmentation does not tell the business owner enough information regarding the person who is preparing to make a purchase. An audience prospect segment for a car dealer (e.g. urban dweller, head of household, student) won't reveal that he or she is in the market to buy a car, but it will when combined with his or her behavior. "If he or she has searched or visited a car shopping site, we have a strong indication of how likely he or she is to buy a car," says Zemel.

He warned, though, that it takes at least a half-dozen data sources to create a robust consumer profile, and that the marketer must judge the accuracy of each source to decide which ones to use for modeling and targeting.

At this point, having applied predictive segmentation to the data, the client was able to decide the message and the Optimal Frequency Range (OFR). "The OFR is a critical lever for creating marketing success," says Zemel. "The family guy with two cars may require more message exposure to get him to consider to switch brands than a person buying their first car."

OFR analysis looks at the entire marketing picture by segment and user. It is based not on the last impression, but on all interactions from the start of the relationship – thus, it is a broader and far more effective gauge of consumer intent than last-view attribution. "We bid higher for people that were below the OFR and got impressions in front of them," Zemel says. "And we reduced our bids for people who were beyond that range or not in the target audience. We shifted the entire media plan into that sweet spot."

That done, [x+1] built "look-a-like segments to expand the coverage and the size of our target audience," Zemel says. Then, during the calibration period, [x+1] analyzed all media sources and their audience impact, applying mathematical models to determine the spend and frequency cap on each one. The client could move dollars where they needed to go – within the OFR.

The client was now able to track – and more effectively use – traditional or negotiated media and, "at the same time, complementary to that, we were able to fill in the gaps in the real-time inventory exchanges," says Zemel.

You might wonder: Is it difficult to connect online and offline activity when the sale is offline? The answer, no. Take the case of the auto purchase. "If someone requests a quote or a dealer visit online, there are ways through lead management to optimize that," Zemel says. "Sometimes there isn't a direct connection, so it's a little bit more correlative at first."

### The benefit: digital sales growth

Armed with the power of IBM Netezza, [x+1] produced several benefits for its client. First, there was an attitudinal change. “We shifted the client’s whole view of how they were managing media in market,” Zemel says. “They went from a last-view, CPA performance-based optimization plan to a more meaningful and comprehensive approach.”

Based on this, the client determined how consumers were moving through the funnel – and the financial impact. “We had to prove that there was a causal effect – that we put dollars in and got total digital sales out,” Zemel says.

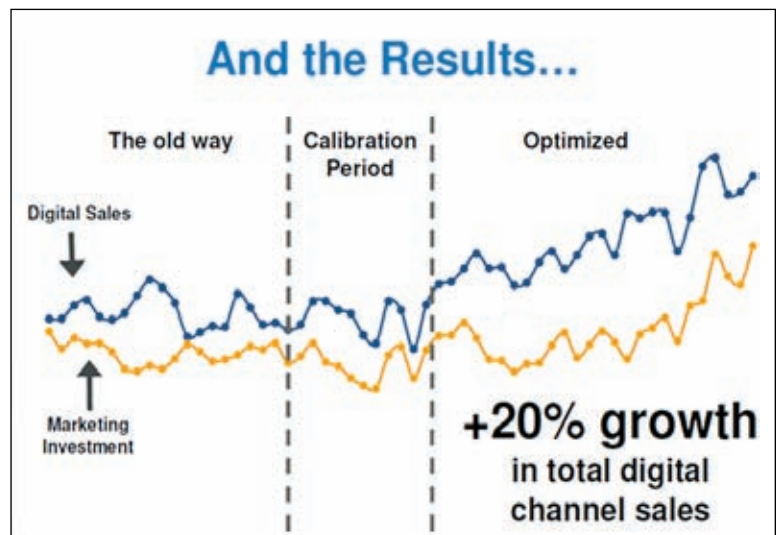
The firm also knocked down barriers separating brand and performance marketing. “Breaking down the silos didn’t take a hammer or a re-org,” Zemel says. All it took was “a marketing framework focusing on the audience.” People at the firm and its agency could see where they fit in, and work toward the same business goal.

Another benefit was control: The client is in full command of frequency and audience engagement. At the same time, the client has moved away from relying on near-term performance for analysis and can now see the total effect on its business. This has led to a better audience composition.

The result is that the company is now able to work with massive data volumes. “For this single client, we collect five billion cross-channel marketing impressions per month from all its marketing activities,” Zemel says. “This is where we really use the power of IBM Netezza.”

### And what about the most important barometer: revenue?

“Historically, we talked about lift in the response rate or the conversion rate,” Zemel says. “Now we’re talking about lift in total digital sales. And we’re seeing a big year-over-year impact – 20 percent growth. Net-net, the client is seeing more revenue from more customers.”





### About [x+1]

[x+1], the online targeting platform leader, maximizes the return on marketing investment (ROI) of websites and digital media using its patented targeting technology. Providing the first end-to-end digital marketing platform for advertisers and agencies, it optimizes engagement rates and lift conversion in both media and on websites. Its predictive marketing solutions enable automated, real-time decision making and personalization so the right advertisement and content is delivered to the right person at the right time. Top companies in financial services, telecommunications, online services and travel have significantly increased the performance of their digital marketing using the services of [x+1]. The company is headquartered in New York City. For more information, please visit [www.xplusone.com](http://www.xplusone.com); follow us on twitter @xplusone.

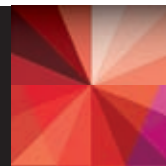
### About IBM Netezza

IBM Netezza pioneered the data warehouse appliance space by integrating database, server and storage into a single, easy to manage appliance that requires minimal set-up and ongoing administration while producing faster and more consistent analytic performance. The IBM Netezza family of data warehouse appliances simplifies business analytics dramatically by consolidating all analytic activity in the appliance, right where the data resides, for blisteringly fast performance. Visit [netezza.com](http://netezza.com) to see how our family of data warehouse appliances eliminate complexity at every step and lets you drive true business value for your organization. For the latest data warehouse and advanced analytics blogs, videos and more, please visit: [thinking.netezza.com](http://thinking.netezza.com).

### IBM Data Warehousing and Analytics Solutions

IBM provides the broadest and most comprehensive portfolio of data warehousing, information management and business analytic software, hardware and solutions to help customers maximize the value of their information assets and discover new insights to make better and faster decisions and optimize their business outcomes.

# InfoSphere BigInsights – Business Partner Ecosystem



Put the power of IBM Business Partners behind your business. Whether you are looking for solutions, tools or system integrators, you'll find the resources you require in IBM's BigInsights eco-system offerings outlined below. Explore the business partner websites as well to find more detail or call your local IBM representative for more information.



Buckley Data Group is a leading independent IT infrastructure authority offering comprehensive infrastructure services from assessment through implementation. With technical consultants specializing in storage, servers, security, virtualization and network management, Buckley provides expertise to your clients across industries globally. Using a channel-based sales model, Buckley builds your brand with your clients.



CCG Partners Inc. offers highly specialized Data Management resources providing value-add data services for the installation, integration and deployment of the IBM Big Data Platform using data quality processes and supporting best practices.

CCG Partners provides enterprise-class data quality management services and data governance frameworks enabling trusted enterprise analytics, risk mitigation, increased rate of adoption and improved ROI enhancing IBM's InfoSphere BigInsights deployment activities for big data initiatives.



ClickFox maps the complex maze of customer experience journeys formed by interactions at every touch point with a company. Unlike business intelligence tools, ClickFox links disjointed, cross-channel data to fully understand and analyze customer behavior in a holistic view. Without ClickFox, businesses see only siloed views and scattered pieces that make up the complete picture of the customer experience.



Concord is a specialty solution provider with extensive experience in process, data, and system integration. Concord is an established IBM Premiere Business Partner with a proven track record delivering industry solutions based on IBM's Information Management, and WebSphere product lines as well as Hadoop. In addition, Concord has created COMPLETE suite that complements and enhances the BigInsights platform by providing end-to-end business process visibility in mainframe & distributed environments as well as environments where establishing precise transaction relationships seems impossible. We offer true end-to-end correlation. The suite includes transaction monitoring, transaction trending, transaction analytics, event management and payload forensics.

The suite couples the power of Hadoop with in-memory MOLAP cubes embedded in our RETE rules engine to deliver the fastest real-time analytics & simulation platform on the market.



The Datameer Analytics Solution provides four key elements:

- Wizard-based data integration platform designed for IT users and BI analysts to integrate large datasets of structured and unstructured data
- Integrated analytics with familiar spreadsheet-like interface with more than 180 built-in analytic functions
- Drag and drop reporting and dash boarding visualization for business-users
- Big data scalability and cost-effectiveness of Hadoop together with IT management tools that overcome Hadoop's heavy technical burden



Datameer utilizes and runs on IBM's platform for big data which provides a dependable, enterprise-ready implementation of Apache Hadoop. Datameer provides a packaged business intelligence platform on IBM's platform for big data that helps overcome Hadoop's complexity and lack of end-user tools by providing business and IT users with business intelligence (BI) functionality across data integration, analytics and data visualization in the world's first BI platform for Hadoop.



Uncovering hidden connections by reading and processing data in advance, Synthesys empowers the data analyst to make smart decisions faster. Synthesys automates the understanding of cloud-scale data and uncovers the hidden connections of entities that lie within.

Synthesys® integrates with InfoSphere BigInsights by seamlessly operating in the scalable Hadoop environment. Synthesys brings unique value to the InfoSphere BigInsights solution by automatically transforming massive amounts of text into the underlying facts and connections.

By performing this knowledge extraction process without any prior definition of the meaning of words (e.g., no use of ontology, taxonomy, etc.) Synthesys uniquely identifies associations and non-obvious connections by digitally examining and comparing contexts around extracted facts.

This also allows Synthesys to continue to be useful in "dirty data" (all caps, machine translations, etc.) as well as coded language. Through our API, integration of the analysis results of Synthesys can be seamlessly integrated into IBM BigSheets and other emerging visualization and workflow solutions.



Fully integrated with the IBM InfoSphere BigInsights platform, Jaspersoft BI Suite provides BigInsights users with plug-and-play access to their organization's Big Data and the ability to combine this with information from a wide range of other sources, e.g. the web and subscription services. Jaspersoft's easy-to-use reporting, dashboard and analytic tools enable BI builders and business users to build, for example, a 360o view of a customer's history, website behavior and credit record for retail analytic and targeting applications. BigInsights and Jaspersoft are ideally suited for departmental applications within the enterprise or complete BI solutions for larger SMB customers.



All Karmasphere products are built on the Karmasphere Application Framework to unlock the power of Hadoop with unparalleled ease:

- Deliver dramatic productivity improvements to the big data job developer
- Make it easy for technical data analysts to discover value in their big data set
- Provide the framework for business intelligence analysts to drive valuable insights from big data

By working together to integrate IBM's implementation of Apache Hadoop with Karmasphere products, there is a seamless out-of-the-box experience for data professionals ensuring application development and analysis on the IBM platform for big data is completed quickly and productively, increasing the ROI of enterprise big data projects.



Kitenga provides the industry's first "big data" search & analytics platform with integrated information modeling & visualization capabilities - an entirely **new kind of insight engine for today's big data world.**



- Kitenga ZettaVox combines proven next-generation technologies like Hadoop for scalability and performance, Lucene/SOLR search, Mahout machine learning, 3D information modeling, and advanced Natural Language Processing in a fully integrated, configurable, cloud-enabled software platform that can be deployed quickly and cost effectively.
- ZettaVox is designed for non-programming professionals, empowering them to efficiently create customized, domain-specific analytics ecosystems supporting massive scale ingestion and processing of information resources with the ease of drag-and-drop widgets.
- Kitenga's solution is a radical improvement over traditional BI dashboards that support basic charting from static, transactional, structured data sources while ignoring the wealth of knowledge buried in mounds of unstructured information. Traditional analytics solutions based on databases inherently suffer from scalability limitations, are inflexible, offer an impoverished suite of analytical and visualization tools, and are outrageously expensive. Kitenga empowers organizations to extract unprecedented levels of actionable insights from their information universe.

Kitenga ZettaVox ships with out-of-the-box integration with IBM InfoSphere BigInsights Enterprise Edition. This not only minimizes customer risk, time and effort wasted in cobbling together one-off solutions, but ZettaVox customers can now benefit from significant add-value functionality of the IBM platform. Enterprise customers can now enjoy the legendary customer support from IBM combined with the power and flexibility of open source Hadoop.



Someone can live or die depending on the correct and authentic medications being dispensed. Hospitals and medical professionals clearly agree leveraging RFID technology for better tracking of a drug's expiration date, information about the drug administered, tracking and updating inventory levels, all performed with real-time visibility, would increase efficiency, reduce costs, and improve patient safety.

The Intelliguard Medication Management System consists of three components:

- Pharmacy Reader: By reading multiple tags within a tote or container, the Intelliguard Pharmacy Reader makes receiving distributor shipments at the hospital pharmacy efficient and accurate by eliminating the need for item-level scanning or manual counting.
- Real-time inventory control is maintained as medication is distributed within the hospital to an Intelliguard Automated Dispensing Cabinet. The Automated Dispensing Cabinet increases nursing efficiency by eliminating manual counting and item-level barcode scanning and through access to ambient and refrigerated medications in one location.
- The Intelliguard Patient Bedside Reader assists with the compliance and verification necessary to eliminate medication errors through The Five Rights of Medication Safety: Right Patient, Right Drug, Right Dose, Right Route and Right Time.



mLogica, a technology and product consulting company, was founded by senior managers from leading technology organizations. mLogica is headquartered in Orange County, California, with development centers and sales offices in California, Florida, Massachusetts, New Jersey, Toronto, UAE, India, Scotland and Malaysia, including an ISO 9000 certified development center.

We have designed, implemented and managed mission-critical business applications, databases and systems for large commercial enterprises and public sector organizations, as well as mid-market businesses. Our clients include major organizations in the financial services, entertainment, technology, education, health care, telecommunications, manufacturing, and transportation and logistics industries.



Persistent is a global company specializing in software product and technology innovation. For more than two decades, we have partnered closely with pioneering start-ups, innovative enterprises and the world's largest technology brands. We have utilized our fine-tuned product engineering processes to develop best-in-class solutions for customers in technology, telecommunication, life science, healthcare, banking, and consumer products sectors across North America, Europe, and Asia.

Thanks to our extensive technology product expertise, customers also turn to us for technology strategy and consulting services. Persistent's customers benefit from our deep knowledge of next-generation Cloud, BI and Analytics, Collaboration as well as Mobility-based computing platforms. By leveraging our strategic technology partnerships, IP-based accelerators, and agile development processes, companies can successfully navigate increasing time-to-market pressures and deliver the highest quality solutions, faster and more cost effectively.



Revolution Analytics delivers advanced analytics software at half the cost of existing solutions. By building on open source R—the world's most powerful statistics software—with innovations in big data analysis, integration and user experience, Revolution Analytics meets the demands and requirements of modern data-driven businesses. It now runs on top of the IBM InfoSphere BigInsights platform, get the power of this joint solution today!



Systech is a leading provider of services and solutions in the area of Business Intelligence, Data Warehousing and Corporate Performance Management solutions for companies large and small in most industries around the world for over 15 years. Utilizing an approved technology and a proven methodology, Systech reveals business opportunities across the enterprise. Systech's unique approach enables clients to make continuous, fact-based decisions to improve their revenue and create value.



Think Big Analytics is the leading professional services firm for big data and advanced analytics. We work with innovators to create solutions that tap into the power of Hadoop and NoSQL to process unstructured data, unlocking new insights and products that were never before possible.



Large scale, open-source information platforms

- Agile approach
- Advanced analytics and data science
- Integration patterns for Hadoop and NoSQL
- Harness unstructured data

Develop your big data capabilities

- Big data integration
- Analytic solutions
- Software development
- Cluster configuration

Your big data solution starts with a Brainstorm

- Solution roadmap
- Big data architecture
- Recommended infrastructure
- Proof of concept
- Delivery project plan



Built on commodity hardware, Zettaset is an out-of-the-box offering that integrates more than 30 services and dependencies into a single autonomous solution. Built-in self-management includes automated server provisioning, a fail-safe process for monitoring all pertinent processes and self-healing. Ease of deployment and support for small files all add to the Zettaset competitive advantage. Further, a simple licensing model leads to a significantly lower total cost of ownership

- Zettaset's architecture supports BigInsights' Application Programming Interfaces (APIs) and tools, using ZooKeeper and Thrift to perform reporting and management. Thrift supports most major programming and scripting languages and all of Zettaset's Thrift API's are open.
- Zettaset provides value in monitoring, provisioning, and management of the system as well as significantly lowering the cost of integration; allowing users to easily make Zettaset a part of their platforms, frameworks and User Interfaces (UI).
- Strong authentication using Kerberos, in conjunction with group and user level access control and data encryption, extends BigInsights' LDAP authorization so that users can fully customize their security model to further protect the safety and availability of their data.
- Zettaset's administration console fully integrates with BigInsights' Web console to allow easy administration and management of services, nodes, and jobs.
- Failover of the NameNode as well as all other critical components in the system, such as Oozie, Hive and ZooKeeper, mitigates the risk of data loss, data access, and failure to schedule and coordinate jobs and query datasets.

## Featured Business Partners



***Datameer***



***Digital Reasoning***



***Jaspersoft***



***Karamsphere***



***MEPS***

## Datameer, Inc.

### *Datameer Analytics Solution (DAS)*



#### Solution Description

The Datameer Analytics Solution (DAS) leverages the scalability, flexibility and cost-effectiveness of Apache Hadoop to deliver a business user focused BI platform for big data analytics. DAS overcomes Hadoop's complexity and lack of tools by providing business and IT users with business intelligence (BI) functionality across data integration, analytics and data visualization of structured and unstructured data.

#### Features and Benefits

- Wizard-based data integration designed for IT users and BI analysts to integrate large datasets of structured and unstructured data
- Integrated analytics with familiar spreadsheet-like interface and over 180 built-in analytic functions
- Drag and drop reporting and dashboarding visualization for business-users
- Big data scalability and cost-effectiveness of Hadoop together with IT management tools that overcome Hadoop's heavy technical burden

#### Value Proposition

The Datameer Analytics Solution (DAS) provides a complete business user focused BI solution for Hadoop including data integration, analytics and visualization without the need for extensive IT and programming resources. DAS utilizes wizard-based data access, 180+ pre-built analytic functions and drag and drop visualization via charts, graphs, maps and dashboards. The end result is a big data analytics solution with dramatic ease-of-use and unparalleled cost effectiveness and scalability.

#### Company Description

Based in Silicon Valley, Datameer offers the first data analytics solution built on Hadoop. Founded by Hadoop veterans in 2009, the company's breakthrough product, Datameer Analytics Solution (DAS), provides unparalleled access to data with minimal IT resources. DAS scales to 4,000 servers and petabytes of data and is available for all major Hadoop distributions including Apache, Cloudera, EMC GreenPlum, Yahoo!, IBM, and Amazon.

For more information, contact:

**(650) 286-9100**  
**[www.datameer.com](http://www.datameer.com)**



# Synthesys®

Entity Oriented Analytics  
for Cloud-Scale Data Understanding

## Digital Reasoning introduces a new era in data analytics with Synthesys.

Built to address the most complex data analytics challenges, Synthesys® excels at extracting, resolving, and linking entities and concepts from unstructured and structured data. Uncovering hidden connections by reading and processing data in advance, Synthesys empowers the analyst to make smart decisions faster. Synthesys automates the understanding of cloud-scale data and uncovers the hidden connections of entities that lie within.

### Entity Oriented Analytics

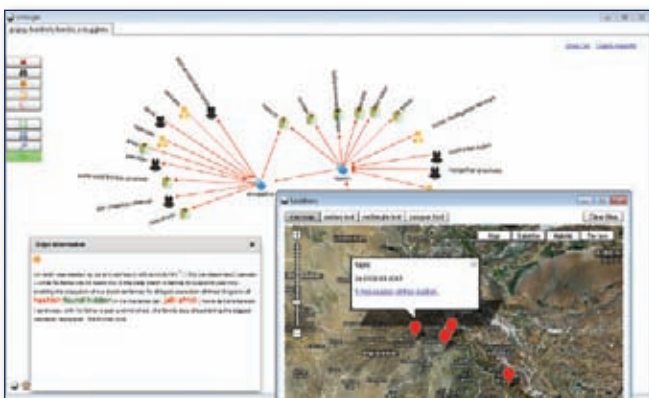
Synthesys takes a new approach to large scale data understanding by focusing analytics on the entity. By transforming documents and files into their underlying people, places, locations, and other entities, Synthesys reduces the reading burden for analysts and empowers new discovery and analytics. Entities and concepts are resolved into their unique characteristics while underlying connections are identified based on usage. Synthesys does not start with a preconception of the data model or the meanings of words. Instead, Synthesys learns the meaning of words

the way humans do — by analyzing the context around the entity and comparing that context signature across the entire corpus. In this way, Synthesys uniquely uncovers non-obvious connections and hidden meanings buried in spelling problems, dirty data or code words.

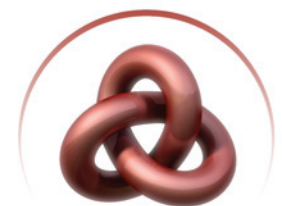
### Cloud-Scale Data Challenges

Enterprises and government agencies are dealing with data challenges that reach into the hundreds of millions of documents and more. Synthesys was built for these “big data” challenges. In

order to understand data in real time, Synthesys compares new data to the corpus already ingested and analyzed without re-indexing. Synthesys maintains all attributes about entities and context, continually comparing new data to the existing analysis. This allows Synthesys to constantly update the associations, similarities and the resulting link analysis. This allows Synthesys to maintain the associations, similarities and the resulting link analysis.



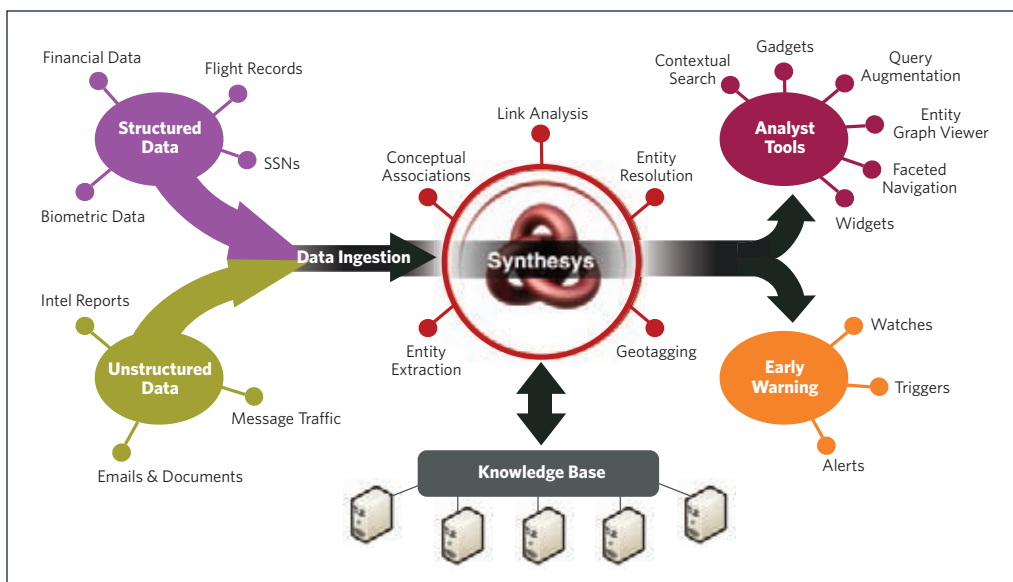
**Synthesys Analysis Tools**  
Entity Graph Viewer, Associative Net, GeoLocator



## Synthesys® 3.0

*“Synthesys is the culmination of 10 years of efforts working on the most critical data analytics challenges in the intelligence community.”*

Tim Estes  
Founder and CEO  
Digital Reasoning Systems



Synthesisys Architectural Diagram

## Knowledge Base

Synthesisys maintains data attributes in the Knowledge Base. The knowledge base is built on a horizontally scalable architecture including tight integration with Hadoop and Cassandra. By combining these best-of-breed Internet technologies, Synthesisys delivers advanced analytical capabilities with high performance and horizontal scalability.

## Associative Net

Associative Net is one of the most powerful and unique aspects of Synthesisys. It identifies synonyms or closely related entities as well as strength of relationship scores for entities in the corpus. For example, Associative Net would show "stinger missile" and "blow pipe" as synonymous because of their use in the corpus. Similarly, one person's connection to another person or place can be identified and the relationship strength scored. Associative Net provides confidence to the analyst that all connections, relationships and synonyms are being considered — including intentionally coded language

## Entity Graph Viewer (EGV)

The Entity Graph Viewer is a visualization tool that allows the analyst to view the connections and social "maps" identified by Synthesisys. Working in combination with GeoLocator and Associative Net, EGV provides the analyst with unique insight into the underlying facts in the data. EGV shows the connection of entities both in terms of "how" as well as the direction of the connection (i.e. who knows who). With this visualization, the analyst can clearly see how one entity is connected to another and can quickly drill into the abstract or context supporting the identification of this linkage. If the abstract is not sufficient, it is possible to drill further down to the original document where the evidence of the linkage originated. With this ability to show high-level linkage and drill down to the supporting data, Synthesisys simplifies the analyst's job by first identifying underlying facts and, only if needed, allowing the analyst to read the complete document. By pushing the time-intensive reading tasks later into their process, Synthesisys enables the analyst to spend more time interpreting and taking action."

**Synthesisys® — make better decisions, faster.**

## Product Features

- Entity Extraction
- Entity Resolution
- Link analysis
- Unstructured Data Analytics
- Analytics tools and visualizations
- Geolocation extraction
- Machine generated abstracts of documents
- Built on Cloudera Distribution of Hadoop (CDH3)
- Built on Cassandra v0.7

## Product Requirements

### Minimum requirements:

- 7 nodes of commodity servers
- Node details:
  - Memory - 8GB
  - CPU - 2 Cores
  - Storage - 850GB
  - Platform - 64 bit

### Typical requirements:

- 20 nodes of commodity servers
- Node details:
  - Memory - 16GB
  - CPU - 4 Cores
  - Storage - 1.5TB
  - Platform - 64 bit

## Operating Systems

- Red Hat® Enterprise Linux (or compatible)
- Runtime Platform - Java® 6



730 Cool Springs Blvd., Suite 110,  
Franklin, Tennessee 37067  
**+1 615 370 1860**

For more information  
visit our website at  
[www.digitalreasoning.com](http://www.digitalreasoning.com)

## The most widely used Business Intelligence Suite in the World:

- 14 Million Downloads
- 235,000 Community Members
- 165,000 Production Deployments
- 14,000 Commercial Customers

## Industry Recognition:



## Jaspersoft End-to-End BI Suite



Reporting



Dashboards



Analytics



Data Integration

- Complete Big Data analytic solution combining the strength of IBM with the world's most widely used BI suite
- Fully integrated, plug-and-play access to Big Data from internal, public and subscription services
- Easy-to-use reporting, dashboard and analytic tools
- Combine Big Data to build for example 360° customer view for retail analytic and targeting applications.
- Ideally suited to departmental BI or larger SMB customers needing ease of use and rapid ROI
- Powerful technical solution including full support for Hadoop Hive SQL interface, HDFS, Avro file format and Hbase

### Customer Overview Report

Name: Mary Francis Bengar  
Address: 9021 Sun View Ter  
Spokane WA 99248

Customer Card Number: 31761638100

197-665-2593  
306-665-5775  
M  
E90K - 100K  
M  
3  
1

Date Account Opened: 1980-05-27  
Member Card: Bronze  
Birth Date: 1947-03-18  
Education: Partial High School  
Occupation: Manual  
Homeowner: 78  
Num Cars Owned: 2

### Shopping Days

Feb: 4.02, Thu: 2.02, Fri: 2.02, Sat: 2.02, Sun: 2.02

### Brand Preferences

### Average Sales

All Customers: 23.24, Mary Francis: 44.33

YTD / YTD-1	Actual	Plan	% Var
Total Revenue	\$ 272,200	\$ 282,000	-3.5%
Cost of Sales	\$ 158,710	\$ 161,700	-1.8%
Gross Profit	\$ 113,490	\$ 120,300	-6.4%
Gross Profit Margin	41.7%	42.6%	-1.9%
General & Administrative	\$ 11,120	\$ 10,800	2.9%
Information Systems	\$ 22,297	\$ 24,200	-7.4%
Marketing	\$ 20,500	\$ 19,200	6.8%
Lease	\$ 10,351	\$ 10,500	-1.5%
Total Operating Expenses	\$ 64,268	\$ 64,800	-0.8%
Operating Income (at Loss)	\$ 49,222	\$ 55,500	-10.4%
Other Expenses	\$ 4,000	\$ 4,000	0.0%
Member Before Taxes	\$ 45,222	\$ 51,500	-10.4%

### Average Conversion Rate by Territory (Days)

MD: 60, NJ: 40, NY: 50, TX: 50

### Lead Source Across Territory - Avg Days

### Average Conversion Rate by Lead Source (Days)

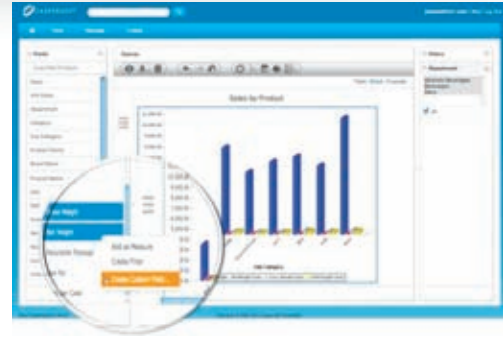
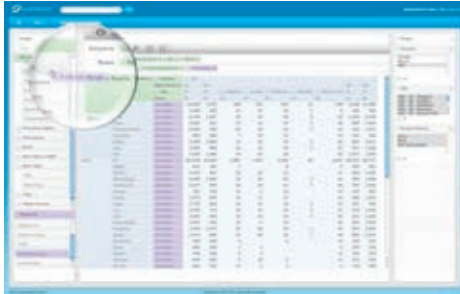
Campaign - Fall Special = 60, Campaign - Enterprise Clients = 64, Web Site = 53, Conference = 57

Date	Product Name	Recyclable Package	Low Fat	Promotion Name	Sales (\$)
13-Jan	1st Wheat Parfait Cookies	✓	✗	No Promotion	\$ 4.4
13-Jan	1stst Noodle Soup	✓	✗	No Promotion	\$ 3.1
13-Jan	Blue Products Conditioning Shampoo	✓	✗	No Promotion	\$ 3.1
13-Jan	1stst Apple Jam	✓	✗	No Promotion	\$ 3.1
13-Jan	1ststst Cheese Spread	✓	✗	No Promotion	\$ 3.1
13-Jan	1stst Cheese Filled Cookies	✓	✗	No Promotion	\$ 1.0
13-Jan	1ststst Bagels	✓	✗	No Promotion	\$ 1.1
13-Jan	1stst 1stst 1stst Smooth Hot Creamer	✓	✗	No Promotion	\$ 3.3
13-Jan	1ststst 1stst 1ststst 1ststst	✓	✗	No Promotion	\$ 3.3
13-Jan	1ststst 1stst 1ststst	✓	✗	Special Events	\$ 11.4

Counts and Totals	90	29	11	3	\$ 63.83
18-May-11	Unique Product	Recyclable Package	Low Fat	Promotion Name	Sales

## Business User

- Web-Based Ad Hoc report designer
- Metadata simplifies data access
- Chart, Table, Filters, Sorting, & more



## Data Analyst

- Web-based Ad Hoc analysis UI
- Speed-of-thought response time
- Advanced analytic queries via MDX

## IT and Power User

- Secure, auditable, scalable
- Highly formatted reports & dashboards
- Interactive reports for casual users



©2011 Jaspersoft Corporation. Proprietary and Confidential

# Karmasphere Analyst

Get graphical SQL access to IBM InfoSphere BigInsights from the desktop.

**Karmasphere Analyst** provides quick, efficient SQL access to big data on IBM InfoSphere BigInsights from a familiar graphical desktop environment running on Windows, MacOS or Linux.

Karmasphere Analyst expands the capabilities of Apache Hive, so that technical analysts, SQL programmers, data developers and DBAs can easily create and manage tables, access data on Hadoop with SQL, visualize and integrate results with other desktop applications and data stores – all from a familiar graphical desktop environment.

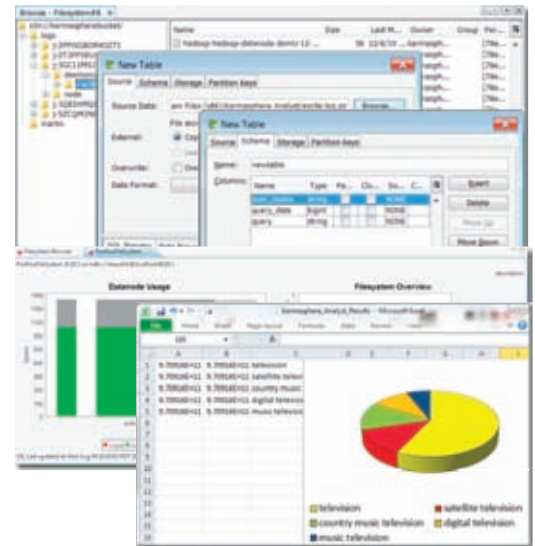
“ Karmasphere has significantly reduced our development time for MapReduce jobs ”

Jeff Ellin  
Vice President, Technology, TidalTV

Karmasphere Analyst works with structured and unstructured data, automatically discovers schema, and can access any Hadoop cluster in private data centers or in the cloud.

**Analyze all your Big Data**  
Supports IBM InfoSphere BigInsights

**Works on any Desktop**  
Windows, MacOS, Linux



Karmasphere Analyst gives you easy SQL access to your data in Hadoop.

Discover Data, Create and Manage Tables	Access any Hadoop cluster, its data, and create schemas for use with Hadoop and Hive	<ul style="list-style-type: none"> <li>Automatic discovery of Hadoop data structures including structured and unstructured data</li> <li>Unified view of multiple Hadoop data stores from the desktop</li> <li>Easy creation and manipulation of new tables and existing Hive tables</li> <li>Drag and drop access to Hadoop (HDFS) file system from the desktop</li> <li>Support for local metadata stores and remote, shared metadata stores via JDBC</li> </ul>
Write & Prototype SQL	Visually develop, optimize and debug SQL queries for any Hadoop environment from the desktop	<ul style="list-style-type: none"> <li>Query syntax checking</li> <li>Visual query plans</li> <li>Query explanations</li> <li>Embedded Hive and Hadoop for desktop prototyping</li> <li>More than 100 User Defined Functions (UDFs) and common SerDe's</li> <li>Customization with User Defined Functions (UDFs) and SerDe's</li> </ul>
Profile and Diagnose	Visually monitor, profile, manage and diagnose Hive-based SQL jobs	<ul style="list-style-type: none"> <li>Graphical query plan progress display</li> <li>Job profiling with calendars, I/O charts, Histograms, etc.</li> <li>Job diagnostics leveraging Apache Vaidya project</li> <li>Visual log file access of job task and mapper progress on a Hadoop cluster</li> </ul>
Generate, Visualize and Explore	View, store and integrate query results in multiple ways	<ul style="list-style-type: none"> <li>Out-of-the-box tabular and page-able display of results</li> <li>Out-of-the-box support to store results on Hadoop cluster</li> <li>Support for storage in other data stores via UDFs</li> <li>One button visualization within familiar desktop applications including Microsoft Excel and Tableau</li> </ul>
Keep Data Secure	Safely communicate with clusters behind firewalls	<ul style="list-style-type: none"> <li>SSH access to clusters behind firewalls</li> </ul>
Get Priority Support	Get priority technical support	<ul style="list-style-type: none"> <li>From the leader in Hadoop developer and analyst tools</li> </ul>



**Big Analytics for Big Data on Hadoop**

info@karmasphere.com • www.karmasphere.com • 1-650-292-6100

# Karmasphere Studio

Graphically develop Hadoop jobs for IBM InfoSphere BigInsights. Fast.

**Karmasphere Studio** is a graphical environment to develop, debug, deploy and monitor applications for Hadoop. It accelerates the development process for experienced Hadoop developers and reduces the learning curve for those new to Hadoop. By making it easy to learn and implement MapReduce jobs, Karmasphere Studio increases productivity by shielding users from the intricacies of Hadoop, enabling them to do more in fewer steps. Jobs can be deployed from any operating system, through any proxy and firewall, and to any version of Hadoop in private or public clouds.

" Karmasphere is beneficial because it gives the developer tools that they are familiar using in other environments, plus it brings in tools critical to working in a Hadoop environment, which allows users to quickly package and launch jobs without having to get their hands dirty inside Hadoop "

Will Duckworth, Vice President,  
Software Engineering, comScore, Inc.

Karmasphere Studio provides value to developers just starting with Hadoop and to experienced developers of Java, Cascading and Streaming jobs for Hadoop.

**Develop for IBM's Big Data Platform**  
Supports IBM InfoSphere BigInsights

**Develop and test from the Desktop**  
Windows, MacOS, Linux

**Use with your favorite IDE**  
Eclipse, NetBeans



Karmasphere Studio allows you to quickly and easily graphically develop and debug Hadoop applications.

## Community and Professional Versions

Get going quickly with the free Karmasphere Studio Community Edition. When you're ready to profile, optimize, package and debug production jobs, reach for the Professional Edition.

		Community Edition	Professional Edition
<b>Learn and Prototype</b>	<ul style="list-style-type: none"> <li>Simplify and reduce the learning curve with guided MapReduce development</li> </ul>	■	■
<b>Develop &amp; Debug</b>	<ul style="list-style-type: none"> <li>Visually build Hadoop applications quickly</li> <li>Debug locally without lengthy deployment and fixing cycles</li> <li>Understand every MapReduce application in detail</li> </ul>	■	■
<b>Monitor &amp; Access the Hadoop Cluster</b>	<ul style="list-style-type: none"> <li>Monitor the cluster, HDFS and jobs on the cluster</li> <li>Access local and HDFS files including log files with familiar drag and drop system</li> </ul>	■	■
<b>Profile and Optimize Jobs for Production</b>	<ul style="list-style-type: none"> <li>Graphically monitor and profile application performance and behavior in-depth</li> <li>Investigate and diagnose the behavior of any job</li> <li>Identify and fix problems</li> </ul>		■
<b>Package and Export for Production</b>	<ul style="list-style-type: none"> <li>Package and export jobs from the development environment</li> <li>Automatically package the MapReduce job into a JAR file to hand over to production cluster job schedulers</li> <li>Control parameter generation to limit configuration problems</li> </ul>		■
<b>Deploy and Manage on Production Clusters Securely</b>	<ul style="list-style-type: none"> <li>Profile, optimize, diagnose, and fix through firewalls</li> <li>Access Hadoop clusters through SSH</li> </ul>		■
<b>Get Priority Support</b>	<ul style="list-style-type: none"> <li>Get priority technical support from the leader in Hadoop developer and analyst tools</li> </ul>		■




**Big Analytics for Big Data on Hadoop**

info@karmasphere.com • www.karmasphere.com • 1-650-292-6100



**Location**       MEPS Real-Time, Inc. is headquartered in Carlsbad, CA.

**Company History**

 In 2001, MEPS was founded and, in 2006, was spun-off and incorporated as a wholly owned subsidiary of Howard Energy. Like so many great American corporate stories, the core intellectual property of MEPS Real-Time was developed in 2001...in an airport...on a napkin. Seriously.



Two key managers of Safety Syringes, Inc. asked themselves, “How can we better utilize technology to track medications in SSI syringes throughout the hospital?” Ultimately, the two concluded that this would be a valuable tool for all medications distributed to patient’s bedside....a **Medication Error Prevention System** with increased visibility of inventory ...MEPS Real-Time was conceived that day.

To say it was a commitment from our investors to get from 2001 to today would be an understatement.

The RFID industry was just evolving. There were no standards. In 2004, there was a brief thought that Wal-Mart would move the industry forward. But, their suppliers rejected the technology advancement. And so, the RFID industry languished. But, MEPS Real-Time didn’t stand still and our investors didn’t withdraw support. We learned and they stayed committed.

From 2001-2003, our early systems were based on passive 13.56 MHz high-frequency (HF) RFID tags. These tags operated well when affixed to packages of liquid medicines, however, only 30 to 40 HF tags could be reliably read when attached to drug products and stored in close proximity to each other inside the cabinet and this did not meet our requirements.

From 2004-2006, we then tested passive tags operating at 2.45 GHz, which functioned well during a hospital pilot test at MD Anderson Cancer Center. However, the 2.45 GHz tags utilized proprietary, soon to be obsolete, technology and we decided we wanted to offer only standardized hardware.

The technology was spun-off from SSI in 2006 and MEPS Real-Time, Inc. was incorporated as a wholly owned subsidiary of Howard Energy.

We redeveloped our system, in 2008, to utilize EPC Gen 2/ISO 18000-6c UHF tags and readers because the hardware is standardized and the tags can be read reliably and in required quantities—approximately 100 tags per drawer...an Intelliguard™ Automated Dispensing Cabinet (ADC) can have as many as eight drawers.

In 2009, we began the critical task of bringing together the right team to lead MEPS Real-Time into the future. We introduced our Intelliguard™ product at the American Society of Health System Pharmacists Mid-year Meeting in Las Vegas and received much interest from industry and from end-users.

A pilot project with Sharp Memorial Hospital was initiated in 2010 to manage the expiration dates of high-cost, slow-moving inventory in the pharmacy department. Previously, this was a labor intensive, time-consuming, critical task... a perfect opportunity to demonstrate the capabilities of RFID and Intelliguard™.

Today, the Intelliguard™ product is positioned as “RFID Solutions for Critical Inventory.” We hope you’ll be a part of our future.

# Corporate Fact Sheet

MEPS Real-Time, Inc.



---

## Company Background

- Initial interest in ability to utilize RFID to simply track pharmaceutical products with the Safety Syringes, Inc. Needle Guards™. Quickly recognized counterfeit prevention, patient safety and inventory management benefits of RFID as well as time management and nursing efficiency.

---

## Management Team

- Shariq Hussain, President and CEO
- Jim Caputo, Vice President, Corporate Strategy
- Jay Williams, Vice President, Marketing and Business Development
- Tom Hall, Vice President, Operations
- Paul Elizondo, Director, Engineering and R&D

---

## Technology Partners

- Impinj: The world's leading developer of UHF RFID.
- ThingMagic: A leading provider of UHF reader engines, development platforms and design services for a wide range of applications.
- Ethertronics: The leading developer and manufacturer of high performance embedded antennas for wireless devices.

---

## Products

- Intelliguard™ RFID Solutions for Critical Inventory offering: Expiration Date Control, Lot Number Control, NDC Control, ePedigree Capability, Counterfeit/ Diversion Prevention, and Medication Error Prevention.

---

## Industry Facts

- According to several national studies, there are 400,000 preventable medication injuries every year in America's hospitals.

Of 4 billion US prescriptions in 2007, up to 40 million may have been filled with counterfeits, up to 10% in California.

Counterfeit prescriptions projected to cost \$75 billion worldwide by 2010.

In 2009, California passed ePedigree legislation that will require all medications to have item level serialization by 2015-16. RFID is the most pragmatic solution for ePedigree when integrated into existing workflow and business practices.

While barcodes have been used to manage medication distribution for some time, by providing real-time visibility of inventory with RFID, hospitals and the pharmaceutical supply chain can implement inventory management efficiency and capabilities beyond all barcode systems.

---

## Contact

- 2841 Loker Ave. East, Carlsbad, CA 92010
- O: 760-448-9500 F: 760-448-9599 E: [info@mepsrealtime.com](mailto:info@mepsrealtime.com) [www.mepsrealtime.com](http://www.mepsrealtime.com)





© Copyright IBM Corporation 2011

Produced in the United States of America  
October 2011  
All Rights Reserved

IBM and the IBM logo are trademarks  
or registered trademarks of International  
Business Machines Corporation in the  
United States, other countries, or both.

Other company, product and service names  
may be trademarks or service marks of  
other companies.