

Grabbing Value from Big Data: The New Game Changer for Financial Services

How financial services companies can harness the innovative power of big data



People matter, results count.



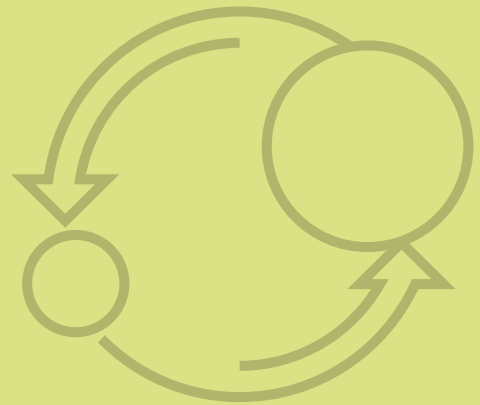
Table of Contents

1. The 3V's of Big Data in Financial Services	4
Volume: Big Data comes in one size: Large!	4
Variety: Lack of structure and mixed origin	5
Velocity: Coming from the real world in real-time	6
2. Calculating Value	7

From both a business and technology perspective, “big data” is generating a lot of discussion—so much so, that it can be hard to know what is hype and what is reality. Big data holds tremendous potential for financial services firms to develop new and innovative solutions that result in significant business value. To capture that value companies must leverage big data solutions to make better sense of the real data they have, get to it quickly and make valuable decisions. It requires congruence between business objectives and the big data storage and analytics approach.

A major factor in the creation of value is trust in the information used to make decisions. Lack of trust in the information sources and analytics can derail the success of an analytics project and, conversely, solid trust can be a huge benefit in appealing to executives and boards of directors for funding of big data analytics initiatives. Establishing trust in big data is paramount to value creation as the variety and number of information sources grows.

What do financial services (FS) companies need to know to establish trust in big data and drive the right business opportunities from it? This paper focuses on helping FS companies to understand the technology component of big data value generation. To better understand how companies can maximize the value they generate from big data from the business perspective, see *Grabbing Value from Big Data: Mining for Diamonds in Financial Services*.



The 3V's of Big Data in Financial Services

The three common characteristics of big data—volume, variety and velocity—are both relevant and challenging for FS companies. They must maintain a lot of data over time, the data lacks structure, and it is of mixed origin. In addition FS companies need the ability to process real data in near real-time.

Volume: Big Data comes in one size: Large!

For years enterprises planned terabytes of storage for enterprise data. Now, more and more frequently, storage needs are described in petabytes and exabytes as companies increasingly store and retrieve data of all forms: transactions and customer details, trading data, telemetrics, weblogs, audio files and more. By law in many countries financial services firms must be able

to recover data for 10 years. This requirement for historical information is unique for the financial services industry and adds complexity to FS firms' business processes. It could be an audio file or any form of documentation and a firm must be able to retrieve it for a decade.

Banks and insurers need technologies and methods to store, organize and retrieve a new volume and variety of data. Yet big data and analytics offer new business opportunities that leverage stored information far beyond record retention. Despite retention requirements most financial services firms are not working with the petabytes of data a company such as Google must handle. New big data technologies are able to deliver innovative solutions that deal with variety and velocity, even when very large volumes are not yet present.

Variety: Lack of structure and mixed origin

Data is no longer defined by traditional data types or found in traditional data warehouses or back offices. Banks and insurers are using multiple channels for their customer interactions. A customer can exchange email with the bank, call the branch with questions, gather information online and conduct transactions on a mobile phone. This results in a multitude of data types that don't fit traditional tabular (row, column) data structures. Wealth advisors must have access to large volumes of emails, know what's in these emails, and be able to search them to find the data they need. The structured and unstructured data when integrated offer a 360-degree view of customers—and enable access to that comprehensive information not just in a database in the back office but in all interactions the customer is having with the FS company's channels.

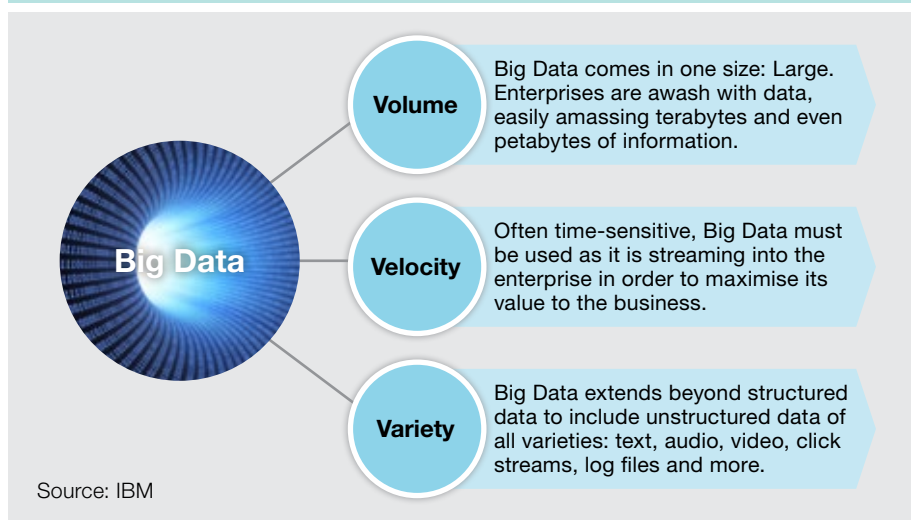
To support the above scenario the enterprise must store these communications but also understand the content of each one. This requires use of search engine technology (such as Natural Language Processing and Text analytics) that gives the company the ability to search unstructured data, aggregate this information and provide meaningful integration to present to the teller, wealth advisor or call center operator the best information updated with all assets inside and outside the enterprise.

An additional dimension of variety is that data comes from both inside and outside the organization. Historically a bank has drawn its reports and information needs from data that resides inside the organization. Yet, with ever increasing volumes of relevant information now residing outside the company, banks and insurers are challenged to manage inside and outside sources and marshal that information in a timely and relevant manner. It requires powerful content management tools to do this.

For example, in France 300,000 data sets are now available to the public on www.data.gouv.fr, including statistics about all aspects of the country, its people and its government. More and more companies are able to use this outside information to help market their business and service their customers. Consider sources such as YAGO¹, a knowledge base developed at the Max Planck Institute for Computer Science in Saarbrücken. As of 2012, YAGO2s has knowledge of more than 10 million entities such as corporations and businesses and contains more than 120 million facts about these entities. The information in YAGO is extracted automatically from several sources like Wikipedia, Wordnet and Geonames. The accuracy of YAGO was manually evaluated to be above 95 percent on a sample of facts. Copies of the whole database are available, as well as thematic and specialized subsets. It can also be queried through various online browsers. YAGO has been used in the [Watson artificial intelligence system of IBM](#).

In addition, social media is giving companies data from outside the company for their own business intelligence but also represents a source of information for the company to better service customers and improve product innovation. Financial services firms can leverage what the consumer is saying about their product outside the company with other customers. For example, a bank may direct a customer to an outside blog or online customer community for more information. More and more banks and insurers will have to manage their networks and communications with communities of customers using cloud solutions that connect to social media. By listening to these customers and gathering peer-to-peer product feedback they will be able to innovate products more rapidly and better meet customer needs.

Figure 1. The Three Dimensions of Big Data



¹ Watson artificial intelligence system of IBM.

Velocity: Coming from the real world in real-time

No longer is data something that is compiled and processed. It is real-time from the real world. By analyzing volumes of real-time tweets in multiple languages, the United States Geological Survey's Twitter Earthquake Dispatch (TED) is able to use time, geo-tagging and location data contained within the tweets to pinpoint earthquakes in anywhere from 30 seconds to two minutes. About half the time, TED provides earthquake alerts before seismometers can confirm them. Financial services firms can now conduct real-time analytics on a variety of sources such as mobile and social data to distinguish between fraudulent and normal credit card activity.

As the world moves toward digital transformation banks and insurers need to use multiple distinct channels for customer interaction to grow their business. Each channel needs real-time access to customer information to support the interaction and each touch point adds new information to the customer's profile.

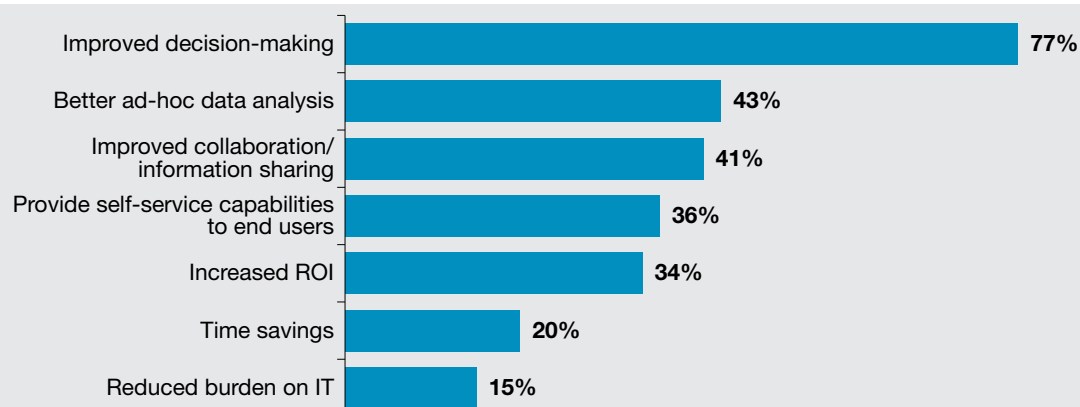
Every year more chips are being put in products letting companies collect data to combat fraud as well as a multitude of other possibilities. Big data has changed insurance pricing since Insurance companies began putting a chip in cars to collect information on what the car is doing. Insurers can now offer incentives based on the behavior of the car and customer driving patterns. The next opportunity is to overlay other data sources such as traffic, weather and geographical data to gain greater insight into a specific car and driver combination.

Banks and insurance companies need new capabilities to make business decisions with real data in near real time. For example, banks used to take three to four days to respond to credit applications. Now some are completing risk management processes and getting responses to customers in 24 hours. Insurers are providing quotes in five minutes along with a price comparison against other firms. More and more you can initiate an insurance claim on a mobile device, including taking a photo of the damage from a car accident, getting roadside assistance, transferring the GPS coordinates for the claim and more. In the future, it's likely that by harnessing real-time analytics capabilities banks will not have to worry about Basel liquidity risk because it could be recalculated in real-time with each credit decision based on data that is in memory and accessible in nanoseconds.

What kind of data will you manage tomorrow? What kind of services will you be able to deliver?

Traditional data modeling is typically based on a data sample because the modeling tools used cannot handle all the data at once. As a consequence, potential errors due to sampling biases, is a common concern. However, big data technologies can handle entire data sets in models and churn through hundreds of scenario combinations and thus help avoid a sampling bias. Furthermore, technology tools for big data are not only about processing big volumes. They also process data that is in memory, which means it can be accessed in nanoseconds and process large volumes in a very short time. Rules engine-based technologies enable companies to design into processes business rules that will allow the process to decide itself (without human intervention) the decisions for 85 to 90 percent of cases.

Figure 2. Value Levers for Big Data Initiatives



Source: IDG Research Services, August 2012

Calculating Value

Volume, variety and velocity define what the technology toolset must address to enable FS businesses to establish organizational trust and harness the innovation possibilities of big data. To maximize value from big data, FS companies should begin their big data initiatives by agreeing on the business objectives they are trying to accomplish. Business objectives should govern the area where Big Data should apply. This helps to scope the data collection effort and prevents gathering and managing data that is not needed or usable.

Existing and upcoming regulations are also a key driver for much of the big data activity within FS companies. These regulations are putting greater emphasis on firms to increase governance, transparency and risk reporting, driving the need to go beyond traditional data analysis.

Beyond what is required for regulatory compliance, FS companies should define their own customer-focused data policy to serve as a strict guideline for data management. Having such a policy in place will help to achieve big data's intended value.

From a technology perspective, big data has the potential to substantially lower the total cost of ownership of technology solutions. Most big data technologies rely on inexpensive, commoditized hardware and therefore scale rapidly and very economically. They also make use of open source software avoiding licensing fees. This aspect in turn lowers the barriers to adoption and incorporation of analytics throughout an organization.

More broadly, finding the actual financial value of the data (ROI) can be a challenge. Value definition can be done using mathematical assumptions or real calculations where the component values are available. If structured properly, the ROI computation can be used to determine the big data value in

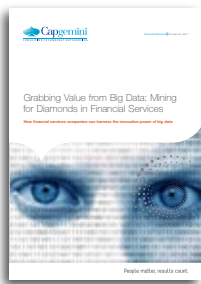
respect to the data owner's relative configuration and monetary data asset value, thus removing the arbitrary aspect of "It Depends." It is also beneficial to determine the breakeven price point on the purchase of new hardware and software, resource allocations, and project priorities.

To reduce capital expenditures and risk, many FS firms are looking to the Cloud. End-to-end Cloud-based analytics solutions such as Capgemini's [Elastic Analytics](#) enable FS firms to take full advantage of a consumption-based model while maintaining the same look and feel to applications as if they were running in their own data centers. For organizations that are struggling with the deluge of data and the ability to rapidly respond to new demand for insight from their business users, Cloud offers FS firms a way to access an end-to-end Business Intelligence and Big Data Analytics solution with a much shorter 'time-to-value'.

Banks, capital market firms and insurance companies all realize today that enterprise information is one of their most strategic corporate assets. Some financial institutions have already embarked on a path to transform their disparate operations and data repositories into an enterprise-level program that elevates information to its deserved status of strategic asset. They are reshaping the business into a truly information-centric enterprise where both data quality and consumption are aggressively and consistently managed by the leadership team. But most organizations also agree that they lack the focus, skills, competency and leadership to manage this strategic asset as effectively and efficiently as they would like. For many, the biggest big data problem they face is "How do I use it?"

While this paper examined big data value from a technology and tool set perspective, there is also a critical business component to the discussion. At its core, big data is about enabling business users to make decisions that create value. That means that FS companies must target actionable areas that impact business outcomes. To learn more about what

FS companies need to know to drive the right business opportunities from big data, please see *Grabbing Value from Big Data: Mining for Diamonds in Financial Services*.



Also, view these quick videos for more direct commentary from our Big Data experts:

- **Big Data for FS – Fact or Fiction:** What you need to know about the characteristics of big data to make the best decisions about your technology toolset.
- **Big Data “Diamond” Opportunities for FS:** Identifying the real, operational opportunities that are likely to drive specific, positive business outcomes in most banks and insurance companies today.
- **Calculating Value from Big Data:** Integrating the technology and business perspectives on value to harness the true innovation power from big data initiatives.

For more information, contact us at: financialservices@capgemini.com
or visit: www.capgemini.com/financialservices



About Capgemini

With 130,000 people in 44 countries, Capgemini is one of the world's foremost providers of consulting, technology and outsourcing services. The Group reported 2012 global revenues of EUR 10.3 billion.

Together with its clients, Capgemini creates and delivers business and technology solutions that fit their needs and drive the results they want.

A deeply multicultural organization, Capgemini has developed its own way of working, the Collaborative Business Experience™, and draws on Rightshore®, its worldwide delivery model.

For further information visit
www.capgemini.com

The information contained in this document is proprietary. ©2013 Capgemini. All rights reserved.
Rightshore® is a trademark belonging to Capgemini.

All products or company names mentioned in this document are trademarks or registered trademarks of their respective owners.

People matter, results count.