A Dozen Ways Insurers Can Leverage Big Data to Extract Business Value

The voluminous amount of structured and unstructured, internal and external data coursing into every organization is increasing exponentially. Data today comes via varied and disparate sources that include customer interactions in channels such as call centers, Telematics devices, social media, agent conversations, smart phones, emails, faxes, police reports, day-to-day business activities, and others. Gartner predicts an 800 per cent growth in data over the next 5 years (1).

Organizations actually process only about 10 per cent to 15 per cent of available data, almost all of it in structured form. While managing this overwhelming data flow can be challenging, insurers that can capture, store, search, aggregate, and eventually analyze the data can find themselves reaping very real benefits like increased productivity, improved competitive advantage and enhanced customer experience. This value does not necessarily come from simply managing big data, but rather, from harnessing the actionable insights from it. Insurers that can glean objective-driven business value by applying science to their data and effectively mine data for insights will maintain competitive advantage and stay ahead of the curve in this information age.

About the Authors

Ajay Bhargava, Director, Analytics & Big Data, TCS

Ajay Bhargava has more than 23 years of Industry, Research, and Teaching experience in the areas relating to Databases, Enterprise Data Management, Data Warehousing, Business Intelligence, and Advanced Analytics. Over the years, he has provided business and technology oriented strategic thinking, mentoring, and customer-centric solutions to his clients.

Ajay heads up the Analytics & Big Data Center of Excellence for TCS for Insurance & Healthcare customers. He publishes articles, and is a speaker & panelist at various conferences on Enterprise Data Management, Business Intelligence, and Advanced Analytics. As an Adjunct Faculty, he has also taught at The University of Texas, Austin and College of Engineering Pune, India.

Ajay holds an M.S. in Computer Science and M.S. in Aerospace Engineering from The University of Texas at Arlington. He obtained his B.Tech in Aeronautical Engineering from Indian Institute of Technology (I.I.T.), Mumbai, 1984.

Table of Contents

Big Data, Defined

Big data provides opportunities in existing environments, but also creates new opportunities for business stakeholders that were not possible by dealing with structured content in traditional ways. Although there is certainly a lot of industry hype surrounding big data, it can be described using a combination of the following five characteristics:

■ **Volume:** Big data refers to the enormous – and exponentially growing – amount of data flooding into and out of every insurance enterprise. Examples of these can be found in a variety of sources including:

- The structured granular call detail records (CDR) in a call center;

- Detailed sensor data from telematics devices;

- External information including weather, geo & peril, traffic, demographic and psychographic behavior data;

- Unstructured data from social media, adjustor notes, police reports, and medical records. There are currently more than 100 vendors providing external data services to insurers such as Insurance Services Organization (ISO), Choicepoint, and Lexus/Nexus.

■ **Variety:** Proliferating channels have led to burgeoning types of data. This data explosion has presented insurance organizations with the challenge of extracting data from disparate sources beyond the usual structured environment of data warehouses and source systems that include mobile, online, agent-generated, adjustor notes, medical reports, social media, text, audio, video, log files, and more.

- **Velocity:** Insurance firms must be able to process, access, and analyze huge volumes of information as quickly as possible in order to make timely decisions. Insurers also need to:

  - Reduce latency to optimize cross-selling and up-selling in a call center environment;

  - Provide quick enterprise intranet documents search to study the impact of different events;

  - Reduce business delivery time for reports in a data warehousing environment.

- **Veracity:** The reliability and consistency of the data – its dependability and quality – is a critical issue for insurers looking to derive meaningful insights. This runs true for both big data and "small data" as well. In some cases, such as in voice to text conversions, we find that even not-so-perfect data quality can result in meaningful insights, especially if insurers are trying to analyze macro-level phenomenon such as sentiment analysis.

- **Value:** Insurers that adopt business objective-driven valuable insights will come out ahead of the pack in this challenging environment. However, in order for insurers to derive true value from big data, they must enable new business models and be able to perform analytics on the data faster and more cost-efficiently. Value is the most important of these five characteristics.

Harnessing and Harvesting Big Data

There are two fundamental aspects to big data. The first is harnessing, which involves collection, administration and management of big data. The second is harvesting, the artful skills and techniques required to apply science to the data, in order to derive actionable and meaningful insight from it.

Harnessing Big Data

At the most basic level, harnessing is the amassing of big data, how insurers manage big data and how they create an ecosystem that can not only create big data but sustain it as well. Years ago, harnessing data was much easier than it is today – and benefits of using this data were more limited as well. It's a combination of additional sources of data like social media as well as technology that exists today to give insurers access to that data, as well as the ability to analyze it.

Gartner estimates that between 80 per cent and 90 per cent of all data produced today is unstructured. Today, insurers can tap into a treasure trove of unstructured data of all varieties: text, audio, video, adjustor notes, click streams, and log files, for instance, and combine it with other structured types such as weather, traffic, peril, and geographic data.

Insurers can no longer manage big data with traditional technologies. Instead, insurance organizations must leverage a whole new class of platforms, like the open source Hadoop ecosystem and its commercial variants offered by IBM, Cloudera, Hortonworks, MapR, etc., which is a distributed file system with the Map Reduce paradigm and technologies built on top of them. The Hadoop based frameworks represent a paradigm shift in not only being able to handle different kinds of data, but also provide speedy processing capabilities on huge volumes of data. Additionally, programming languages such as "R" and Python are finding adoption in IT departments that typically are responsible for administering and managing the big data environment. While they originated in the context of handling big data at Yahoo, Google, Amazon, Facebook, etc., these programming languages can be used to handle many traditional data processing tasks as well.

One example is the use of the Hadoop platform as an ETL enabler for business agility. Today, insurers are adopting this use case, often as their first experience leveraging big data to meet business service level agreements (SLAs). Even though the traditional data warehousing environment still involves structured data, insurers that invest in this platform know they will be able to support unstructured data as well, now and in the future.

Harvesting Big Data

Unstructured data cannot be consumed in its raw form. It must be processed into a consumable form before it can be both interpreted and acted upon.

Harvesting utilizes technology and algorithms that enable insurers to analyze, deliver actionable insights and derive real value from big data. Skill sets such as statistics, data mining, econometrics, business analytics, visualization techniques, and more are in high demand as they provide a solid foundation for deriving useful insights from the data. Universities have started trying to fill the supply demand gap by offering various graduate programs in business analytics to prepare for the next generational skills needed to mine actionable insights.

While the ability to successfully harness and harvest data is critical to a big data strategy, the harvesting aspect is where insurers derive the true value from their data, with the help of analytics.
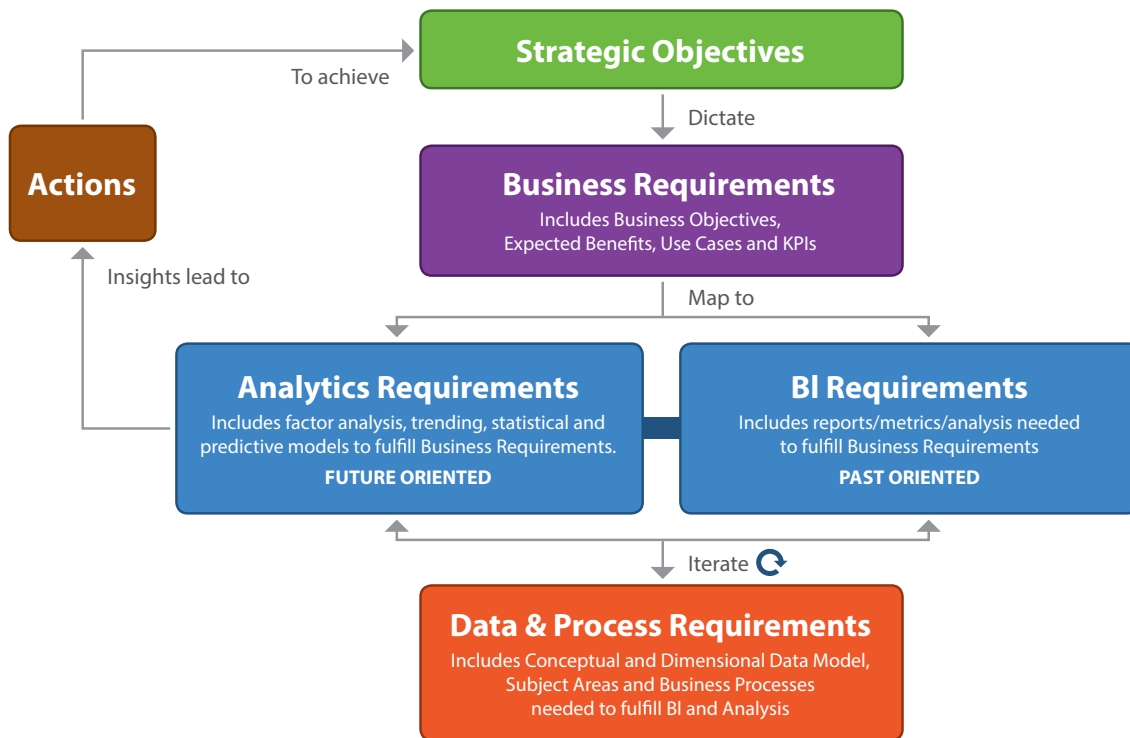
Defining use cases and hypotheses becomes crucial when following a focused "top-down" approach to creating actionable insights (see figure above).

Although this is a focused approach, many times insurers need to do some initial work in order to perform data exploratory analysis to even come up with the use cases utilizing big data in the first place. This initial "bottom up" approach becomes a prerequisite for determining and prioritizing use cases for which big data Proof of Concepts (PoCs) should be pursued.

Real value is derived when actionable insights can make a positive difference in achieving the strategic objectives. Some of the other comparisons between harnessing and harvesting are shown in the figure below.

Harvesting and harnessing activities are complementary to one another, two sides of the big data coin. Big data platforms do not replace existing traditional data management and analytics platforms. Instead, they merely complement, add, mature, and improve upon existing environments and capabilities.

Value Creation from Big Data — Use Cases

According to McKinsey & Co. (1), big data creates value in five ways:

- First, managing big data can increase transparency, making data more easily accessible to relevant stakeholders.

- Second, as they create and store more transactional data in digital form, organizations can collect accurate, detailed performance data in real-time or near real-time, enabling experimentation to identify needs and improve performance.

- Third, big data gives organizations the means to improve customer segmentation and then better develop and tailor products, services, and promotions to each specific segment.

- Fourth, a big data strategy can include sophisticated analytics to provide actionable insights that minimize risks and improve decision-making.

- And finally, big data can be indispensible for organizations looking to create new business models and improve products and services.

Leading-edge insurers are starting to (or are planning to) exploit big data in at least 12 different ways, each of which adds value to the organization in one or more of the ways described above (see chart on page 11).

1. **Making performance improvements in existing data warehouse environments.** IT has begun to adopt Hadoop-based architecture to speed up ETL in a data warehouse environment to meet reporting business SLAs.

2. **Detecting fraud.** Chief claims officers (CCOs) are adopting a multi-channel approach to fraud detection by looking at structured data in their claims and policy data warehouses and combining it with textual data in adjustor notes, police reports, and social media. Special Investigative Units (SIUs) are very interested in identifying suspicious claims or knowing which claims have subrogation or litigation potential. Text Analytics and NLP capabilities are enabling innovative solutions in detecting claims fraud, in addition to automated business rules, predictive analytics, social media analytics, and link analysis techniques.

3. **Combining customer channels.** Combining direct customer connections (email, call center, agent, portal, faxes, adjustor reports, etc.) with indirect customer connects such as social media, blogs, log files etc. provides a more holistic, 3600 view of each customer. This helps to create a personalized, unified communication response, enabling CMOs to achieve better brand value and gain competitive advantages while directly impacting the bottom line by reducing communication waste. Also, minimizing CapEx by using cloud technologies, along with mobile visualization techniques, enables executives to make decisions quickly and cost-effectively.

4. **Optimizing call center workload.** Analyzing telecom data from the switches (call detail records), combining it with claims, helps in understanding what activity was performed by whom, how efficiently, and is used to provide training guidelines for employees. Temporal call patterns analysis on voluminous and raw telecom data can help assist in staffing optimization as well.

5. **Using telematics data to derive prescriptive and predictive value.** CIOs are currently investigating how analytics can suggest better driving behavior by sensing telematics data and responding in near-real time. In addition to prescriptive analytics i.e. the devices alerting on risky behavior such as speeding or seat belt non-usage, they are also interested in predicting driving behavior risks, road delays, etc.

6. **Leveraging cross-sell and up-sell potential.** Similarly, by analyzing text and speech in a near-real time environment, organizations are presented with new opportunities to convert the call center from a cost center to an investment center by providing cross sell and up-sell capabilities.

7 **Using NLP, text analytics to improve sentiment analysis.** Using Natural Language Processing (NLP) and text analytics for social media as well as speech analytics for call center conversations, insurers are able to improve their sentiment analysis to better meet customer service improvement objectives.

8 **Utilizing social media to introduce new products and services.** Insurance CMOs are utilizing social media to introduce new products and services and target customers in specific regions in a cost-efficient and effective way compared to capital-intensive television and Internet promotions, innovatively changing the business model. Insurers can experiment with different segments, and then upgrade their strategies to a national level.

9 **Closing the loop between pricing risk and claims.** Underwriters are studying loss and fraud propensity of existing claimants in order to better price risk for new prospects, especially in the property casualty insurance business. This helps in minimizing risks and to a large extent, price the risk appropriately.

10 **Leveraging external data for more accurate pricing.** Using real-time location, traffic, and weather data can lead to more appropriate pricing on property casualty insurance based on how and where insureds actually drive their cars.

11 **Enhancing search capabilities.** Many reinsurers and insurance agents are using big data to discover innovative ways to search their intranet documents for providing fast search capabilities in unstructured PDFs and Word documents that were not possible previously. These are being used by their financial divisions, as well as in call center scenarios to provide real-time recommendations.

12 **Creating comprehensive customer satisfaction surveys.** Most insurance organizations perform customer surveys using a very small customer sample size. Big data enables the insurers to survey their entire customer base, processing the survey results in a fast and cost-effective way to paint a truer picture from their customer service responses.

	Use Cases	Increase Transparency	Improve performance, near-real time analysis	Improve customer segmentation	Provide actionable insights	Create new business models to improve products, services
1	ETL Performance Improvements		●			
2	Detecting Fraud	●	●	●	●	●
3	360° Customer View				●	●
4	Call Center Optimization			●	●	
5	Telematics	●		●	●	●
6	Cross-Sell and Up-Sell		●	●	●	●
7	Sentiment Analysis				●	
8	Social media for new products and services		●	●	●	●
9	Closing loop between pricing risk and claims			●	●	●
10	External data for pricing			●	●	
11	Search on unstructured documents				●	●
12	Better customer surveys	●			●	

The Opportunity

According to June 2012 research from analyst firm Novarica (2), in general, big data is not yet a big priority in most insurance organizations. In general, insurers are hampered by severely fragmented data environments and information silos, as well as insufficient investment in tools and technologies to support a big data strategy. Most carriers are still maturing and expanding their use of traditional data analytics and predictive models to improve processes, reduce losses and generally improve their book of business. Insurers that aren't exploring and embracing big data and developing a big data strategy will find that not only are they losing their competitive advantage and potential to achieve efficiencies, but they are not able to glean actionable insights from the mountains of data flooding into their organizations to achieve their desired business objectives. While a vast majority of insurers are using analytics for actuarial (95%) and pricing (83%), fewer than half of insurers are using analytics to improve operational areas like sales, marketing or optimized work assignment for underwriters or claims adjusters.

While relatively few insurers are fully immersed in a comprehensive big data strategy and reaping its benefits, the good news is that most insurers are planning their big data approach. Even fewer insurers capture, persist, and analyze big data within their computing environment today, but those that do typically leverage traditional computing, storage, database and analytics, in addition to newer platforms such as the Hadoop eco-system. In fact, larger property casualty insurers universally plan to embrace big data and analytics across all financial and risk management areas (as well as most operational areas). Only about half of smaller insurers are planning the same actions.

Getting Started

Big data solutions encompass a new generation of software and architectures designed to economically extract value from enormous volumes and variety of structured and unstructured data by enabling rapid data capture, discovery, and/or analysis. According to Novarica, the insurers that will be best positioned to profit from the potential value of big data will be those who have created a culture where business leaders trust analytics and act on the insights provided. All insurers should take steps to create that culture today if it doesn't already exist in their companies.

The key is to start small with a Proof of Concept (PoC). Here is an example of how insurers can leverage a big data platform and some key considerations to keep in mind when embarking on a PoC. In this

example, IT is interested in using a big data environment to speed up long-running ETL processes in a traditional data warehouse environment using structured data, which is leading the organization to miss meeting reporting SLAs for business.

PoC Objective: Increase Business Agility		
Big	Business Use Case/Hypothesis	Speed up ETL so IT can meet reporting SLAs for business
Small	Investment	< 120K$
Big	Executive Support	Cx 0 level support
Small	Scope/Design/Implementation	Few longest-running ETL scripts
Big	Architecture (Data Platform)	Hadoop eco-system — scalable for growth
Small	Team	5 FTE
Big	Collaboration — IT & Business	Needed for Harnessing & Harvesting respectively
Small	Duration	< 12 calendar weeks
Big	Business Value	Business is able to take timely decisions/actions
Small	Incremental Success	Provide funding for next phase

It is important for insurers to develop a good business use case for meeting the strategic objectives of that line of business. In addition, solid backing from a CxO/VP level executive is absolutely required not only for funding, but to evangelize and communicate the objectives and need to the larger organization, including partners and vendors. Although the scope and investment in terms of people (~ 5 FTEs), tools (e.g. open source Hadoop eco-system), technologies and infrastructure (e.g. commodity hardware or cloud) might be small, the architecture should keep the long term view in mind. For the right harnessing and harvesting, good collaboration between IT and business is imperative to iteratively experiment and drive actionable insights. Insurers can then use this incremental success is to get increased funding for next phases and/or use cases.

As insurers identify and understand the scenarios for applying big data within their businesses, they will need to tweak their existing processes to be able to ingest the data "variety," having good "veracity," increasing "volume" or growing need of real-time "velocity," to derive objective-driven actionable "value." Insurance organizations that are able to develop an analytics-driven culture and that learn how to harness the power of big data and harvest the valuable information and insight big data provides can create competitive advantage and positively impact their brand and their top and bottom line.

Big Data Challenges: At-A-Glance

Insurers are faced with a number of factors that combine to make big data a big challenge:

- The explosion of data and proliferation of channels;

- An increasingly competitive landscape, especially in the P&C and life sectors;

- The financial tsunami of the past several years, as well as the resulting increasingly demanding regulatory requirements in both North America and Europe;

- An unusually high number of catastrophic losses caused by brush fires, hurricanes, earthquakes, and other natural disasters in recent years;

- The constant stream of innovations that have made it essential for insurers to deal with big data at the same time they are managing the challenges of their existing small data environments.

References

(1) McKinsey Global Institute report, "Big Data: the Next Frontier for Innovation, Competition and Productivity," May 2011.

(2) Novarica, Insurance Technology Research Council, Analytics and Big Data at Insurers: Current State and Expectations, June 2012.

TATA CONSULTANCY SERVICES

Experience certainty.

Contact

For more information about TCS' consulting services, contact **global.insurance@tcs.com**

Subscribe to TCS White Papers

TCS.com RSS: http://www.tcs.com/rss_feeds/Pages/feed.aspx?f=w
Feedburner: http://feeds2.feedburner.com/tcswhitepapers

About Tata Consultancy Services Ltd (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model™, recognized as the benchmark of excellence in software development. A part of the Tata Group, India's largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at **www.tcs.com**

IT Services
Business Solutions
Outsourcing