Gartner

Big Data Strategy Components: Business Intelligence Essentials

Analyst(s): Douglas Laney

Overview

Key Challenges

- Most enterprises are planning to embark on big data-related initiatives, or are doing so already, yet almost no organizations have an articulated strategy for big data.
- Big data initiatives are unique, not only in terms of technology, but also from a business and organizational perspective.
- Many big data initiatives originate with business units leaving IT in their wake or ill-prepared to adequately support them.

Recommendations

- Be prepared that big data initiatives will bring about significant business, organizational, technological and, in some instances, industry transformations.
- Start scouring for relevant big data sources from "dark" underutilized data, but also consider external commercial and public data service providers and social media.
- The biggest ideas for big data will likely come from outside your own industry. Adopt and adapt those that can help run, grow and transform your business.
- Develop a culture and leaders that value facts and sound analyses to overcome human limitations in decision making, insight discovery and process optimization.
- Become ever more pragmatic about information-related investments, considering the formal costs of acquiring, administering and applying information assets versus their economic benefits.

Introduction

The Global Language Monitor (GLM) uses text analytics of online content including press releases, social media chatter and the blogosphere to determine which buzzwords are currently baffling the world. According to GLM, the most confounding term of 2012 is "big data."

A burgeoning interest in big data coupled with this lack of understanding causes an information strategy gap that introduces tremendous risk for enterprises. On one hand, companies are at risk of overinvesting in big data before they are prepared to execute on it. On the other hand, they are at risk of underinvesting in it and ceding competitive advantage. The opportunities of big data are only beginning to be realized. Regardless of the absence of an overarching strategy for big data (let alone enterprise information management) many early adopters have still achieved one-off, high-value successes but scant few CIOs (our estimates put the figure at less than 1%) have helped lay out for business executives how big data can be transformative or disruptive on an enterprise or industry scale. Nor have they laid out or fully considered the many critical elements required to coalesce a big data IT and business strategy (see Table 1). Many big data initiatives

are driven and funded directly by business functions (marketing, for example) with IT not invited to the party until it can better demonstrate the value it offers.

Table 1. Strategy Essentials for Business and IT

Business Strategy Essentials	IT Strategy Essentials
• Acknowledge how big data initiatives are unique • Generate big ideas for big data • Identify potentially valuable data sources • Build business leadership belief in data • Become even more pragmatic about investments	• Ensure infrastructure adequacy • Consider alternate information architectures • Anticipate and govern risks • Expand your analytic capabilities • Assemble necessary skills • Alter IT organization structures

Source: Gartner (October 2012)

Analysis

Acknowledge How Big Data Initiatives Are Unique

The first component in any big data strategy is the recognition that big data initiatives are different from other business and IT initiatives in a variety of ways. Among some progressive, nimble and risk-embracing organizations, big data initiatives are already paying handsome rewards. From early predictors that enable unprecedented agility, to targeted personalization that generates goodwill and revenue boosts, to previously impossible product and process innovations, big data has proven it can be a game changer for enterprises and industries. Regardless, a Gartner Research Circle study conducted in July 2012 shows that 31% of businesses are holding off on big data technology investments, 16% are waiting up to two years and 15% expect to start investing within 12 months. Perhaps rightly so as big data initiatives are unlike other traditional data management, analytics or application development projects.

Big data initiatives are all about change — changing business processes, data sources, infrastructure, architecture, skills, organizational structures and economics. They often result, not in incremental improvements to existing business processes, but in radical changes to existing processes or even their outright displacement. We have also witnessed them transforming entire businesses and disrupting industries as well. Traditionally, the majority of business intelligence (BI) or analytics initiatives deal with transactional data, whereas big data projects more frequently involve the accumulation of external or sensor/monitor-generated sources and their integration with internal operational data. More frequently, sources of big data are unstructured when it comes to text or multimedia, for example. Compared to other projects big data projects tend to concentrate on acquiring, integrating and preparing information rather than the data's functionality which may be as straightforward as identifying correlations, anomalies or patterns. This shift in focus can strain traditional approaches to enterprise architecture, project management and role definition.

The greatest perceptible difference with big data projects, and the one we believe is given a disproportionate amount of press as a result, is the underlying technology. Traditional, even state-of-the-art, hardware, database management systems and analytics capabilities are often dispensed with in favor of technologies specific to accommodating massive, swift and diversified data and analysis (Hadoop, for example). For organizations and IT professionals indoctrinated in the traditional ways of data warehousing and BI, these changes can be arduous.

Finally, big data initiatives require a degree of financial rumination and discipline focused on the question, *"What value can we generate from this data, and is it more than it costs us to accumulate, administer and apply it?"* The outcome of big data projects can be uncertain. Even more uncertain is the ability of many businesses to act on what they find in the data. With time being money, how quickly can you get from focused experimentation that yield insights or innovations to their implementation and institutionalization?

Generate Big Ideas for Big Data

The major opportunities for big data are not just around insular decision making or incremental improvements to existing business processes, they are around ways to transform the business and disrupt the industry by:

- Asking and answering "chewy" questions that were never possible before
- Radically changing existing business processes or introducing new ones
- Introducing new products and services even down to market-of-one customer segments

Asking "Chewy" Questions

Chewy questions are ones that go beyond the mundane types of questions answered by basic BI tools such as, *"How much did our business grow in the past year?"* Instead they are questions that make full use of broader, deeper and more real-time data and, if answered and acted upon, could have profound effects. For example: *"How can we increase customer shopping basket value by 20% and loyalty by 33% by better understanding their individual interests and behavior, and considering a range of economic forecasts and competitor moves."*

Chewy questions typically have many of the following characteristics:

- Specific yet open-ended
- Relates to a business process and aligned with strategic goals
- Focuses on optimizing or innovating, not merely informing
- Considers change relative to other indicators or processes
- Leverages and integrates internal and external inputs
- Is forward looking rather than backward looking
- Is more about differentiation than just comparison
- Considers various scenarios
- Is actionable: More about "do it" than "prove it"
- Requires advanced tools and techniques to answer

Radically Changing Business Processes

Because of its high volume, velocity and/or variety, big data is an ideal resource for introducing groundbreaking ways of optimizing (or even eliminating) business processes. Large quantities of data can give deeper insights than ever before (using sentiment analysis or forensics, for example); faster or more frequently flowing data enables operational processes to respond in real time or near real time to closed-loop stimuli; and the wider range of data available for integration and correlation can generate understandings of causality that enable predictive and prescriptive analytics.

Introducing New Products and Services

Product management, marketing, sales and information managers should work together to determine how big data can lead to the development of new offerings in one or more of the following ways:

1. Identifying new markets
2. Identifying new feature needs for targeted submarkets
3. Identifying opportunities for completely new offerings
4. Personalizing and mass customization
5. Aggregating, packaging and selling information products

New markets may be geographies or customer segments determined by analyzing available public or commercial data for gaps among competitors or between distinct markets. Opportunities for new features can be discovered through the analysis of aggregated customer touchpoint data (transactions, warranty and call center data, for example) or broader social media streams. Similarly, new offering concepts can be created by performing predictive analytics against combinations of market and feature data or by identifying and "listening to" social media influencers. Finally, enterprises are starting to discover the economic value of productizing the data they generate and collect.

Identify Potentially Valuable Data Sources

Many great ideas for big data initiatives come from an understanding of the range of data sources available and what questions can be answered if they were integrated and correlated. Gartner has identified five distinct types of sources:

Operational Data

This is information about customers, suppliers, partners and employees that is readily accessible in online transaction processing and/or online analytical processing databases. It typically includes transactional data, contact data and master data. Frequently, enterprises have the opportunity to collect even more information during the course of business via sensors or process monitoring (smart meters, Internet-connected devices, voice and RFID, for example).

Dark Data

Dark data is information collected during the course of business that remains in archives, or is not generally accessible or structured sufficiently for analysis and could include emails, contracts, documents, multimedia, system logs or other intellectual property. Parsing, tagging, linking or otherwise structuring or extracting usable data from these sources is considered the greatest immediate opportunity by most businesses among all types of data.

Commercial Data

For many years industry-specific data aggregators (D&B, Nielsen and IRI, for example) have made available syndicated credit, real estate, postal, household and other data by subscription. Today, marketplaces are emerging for almost any variety of legally-available data from companies like Microsoft Azure or Infochimps. Even privately among business partners, information assets are being used to barter where once only financial and material assets were. CIOs need to be aware of those that relate to their market and assess their potential as well as working with business partners to encourage the availability of their data.

Public Data

Many governments have also begun opening their data coffers. Open government initiatives to support economic development, health, welfare and citizen services are in various stages of implementation throughout the world. This data can also have significant mercantile value, especially when mashed with other data sources, to understand and act on local/global market conditions, population trends and weather, for example.

Social Media Data

Participation by individuals and businesses in blogging, tweeting, yammering, Facebook and LinkedIn updating has created another fast-growing, invaluable source of data about preferences, trends, attitudes, behavior, products and companies. Posts, trends and even usage patterns themselves are increasingly used to identify and forecast target customers and segments, market opportunities, competitive threats, business risks and even in selecting ideal employment candidates.

IT must help business leaders understand the range of data available and business leaders must put this data into the context of organizational goals.

Build Business Leadership Belief in Data

Despite surveys indicating the contrary (see "CEO Survey 2012: Financial Services CEO Agenda") we find that many business leaders are still resistant to relying on data for decision making. Especially in matters of strategy, deep personal or professional experience, or multidimensional factors business leaders rely on intuition more often than benefits their organization. In strategic decision making leaders tend to overemphasize past individual experiences despite new or differing data indicating situational change. Even more common today, as information becomes more complex and analytic techniques become more sophisticated, is the inclination merely to discount data or formulae that one doesn't understand.

Even simple vignettes like the infamous Monte Hall paradox show that the math behind an apparently simplistic scenario (selecting the highest probability winning door after one of the three is eliminated) is eminently more complex than it seems and can lead to missed opportunity. In larger studies, Berkeley economist Terry Odean examined the behavior of stock traders who, to their great detriment, have a pre-disposition to "get even." What behavioral economists Meir Statman and Hersh Shefrin refer to as the "disposition effect" is so rooted in human wiring that consciously overcoming it is impossible. Another common fallacy that creeps into gut decision making is the belief that the law of averages will bring values back to an assumed center quickly. In a simple example, a coin flipped several times in a row revealing heads, is not "due" to come up tails the next time.

Limitations on human decision making are not just intellectual, they are also biological. What social psychologist Roy Baumeister calls "decision fatigue," demonstrates how people make inconsistent decisions under different physical conditions or even at different times of the day.

Remedies to this discounting of available data that CIOs should enact and business leaders should embrace include:

- Executive education in basic statistics, risk/scenario planning, "group think" avoidance and even decision theory
- Decision competitions among individuals or teams
- Communicating analytic insights and their transformative opportunity

- Pairing data scientists directly with executive teams
- Hiring and promotion of individuals who best leverage available information
- Brainstorming sessions to generate ideas for big data
- Dedicated project management offices to drive opportunities for implementing and deploying big data
- Introducing roles such as the chief data officer and/or chief analytics officer to organizationally emphasize the importance of information assets
- Data governance, data quality and metadata management to ensure confidence in and comprehension of data

Become Even More Pragmatic About Investments

Big data doesn't dramatically alter the economics of acquiring, administering and applying information assets but it does amplify it. No longer can organizations ignore the need to balance these information supply chain costs with the tangible value derived from information. As information becomes ever more recognized as a corporate asset, irrespective of its shocking oversight as a balance sheet asset by the accounting profession, CIOs and CFOs need to get in alignment with how information asset costs and benefits are measured (see "Maverick* Research: The Birth of Infonomics, the New Economics of Information.") Doing so will engender the pragmatism required to justify big data initiatives. Ultimately, analytics projects are a means of delivering aggregate and summarized information related to a particular business problem. With big data initiatives a greater component of that expense involves the data itself, so accounting for it financially is imperative.

One important scheme for tipping the balance of big data benefits to outweigh its cost is ensuring that the data serves multiple business purposes. Compiling, hosting and processing petabytes of data for a single business process rarely makes for sound financial fundamentals or good use of scarce skillsets. Neither does going through the expense and effort to integrate a multitude of highly complex data sources or streaming subsecond data.

Big data is generally derived from external sources or enhanced operations monitoring rather than as a byproduct of business transactions. This means that data acquisition costs are explicit and could arguably be capitalized as opposed to being obscured in operational expenses. Likewise, the cutting-edge integration, management, storage, processing and analytics technologies often demanded by big data initiatives need special investment consideration. Some innovative technologies targeting big data projects such as NoSQL database management systems (DBMSs) like Hadoop, cloud-based hosting and novel predictive analytics products can even offer economic advantages over traditional relational DBMSs, on-premises storage and enterprise BI solutions.

Additionally, organizations have range of tools and techniques at their disposal to control or balance expenses including:

- Adopting portfolio approaches to projects
- Quantifying the costs and potential versus realized value of information assets
- Limiting pre-justified implementations to sand boxes and experimentation
- Making big data "small" via data sampling, subselection, parsing, aggregation and tiered storage
- Leveraging outsourced or subscription analytics as available
- Hosting data in the cloud for flexibly provisioning storage and processing
- Broadly consumerizing data and analytics among employees and partners
- Productizing non-sensitive datasets as a revenue stream

About Gartner Business Intelligence & Analytics Summit
The **Gartner Business Intelligence & Analytics Summit 2013**, March 18 – 20, in Grapevine, TX, provides independent advice, thought leadership and insight you need to fast-forward with BI and analytics, discover how to build a strategy, architecture and team that delivers value now and in the future. Delegates will benefit from a three day learning experience where you can learn directly from Gartner analysts and expert BI practitioners, hear the latest research findings, take-away actionable insights and also share knowledge and experience with your peers. Hot topics include: Predictive analytics, mobile BI, big data analytics, logical data warehousing, and visual data discovery.

Additional information from the event will be shared at www.gartner.com/us/bi and on Twitter at http://twitter.com/Gartner_inc using #GartnerBI.

Save $300 on the standard registration rate with priority code GARTSBA.