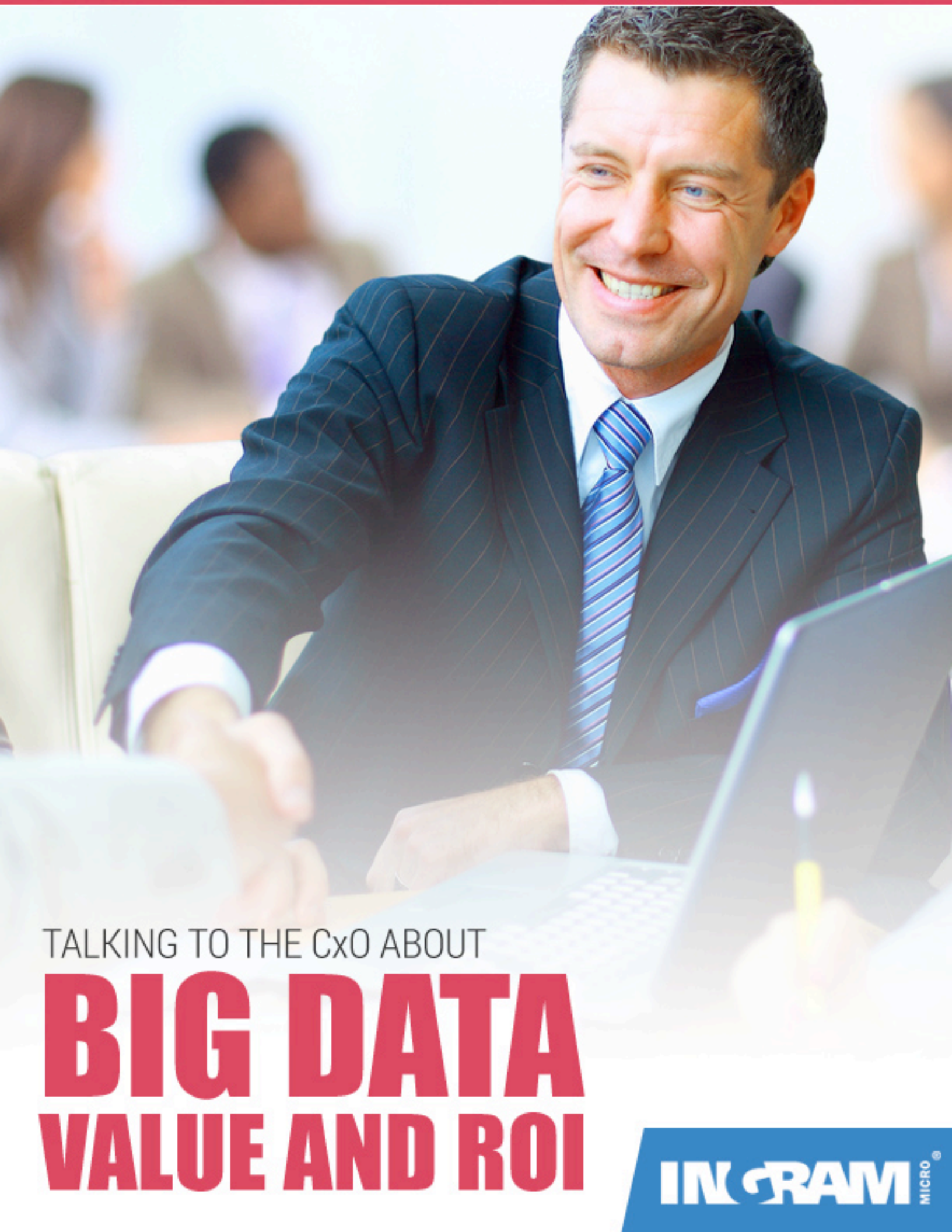


TALKING TO THE CxO ABOUT

BIG DATA VALUE AND ROI

INGRAM MICRO®

How To Speak Big Data

Table of Contents

Executive Overview

Organizations of all sizes in all markets are turning to big data to get a better handle on customer buying habits and go-to-market strategies. However, big data is a big investment. For resellers and integrators offering big data services, that means selling big data to the CEO, CFO, and CMO, as well as the CIO and IT department.

What makes big data different is its value proposition. Big data delivers big insights. For the first time organizations can gather data from all aspects of their operation as well as from external market and machine sources and create algorithms that give them answers to questions that were previously unanswerable before. Big data can tell you about new market opportunities. Big data can tell you where to develop new products. Big data can tell you more about customer attitudes and buying patterns. Big data can tell you about your competitive landscape. And big data can tell you how to optimize operations.

What big data offers is a means to link unconnected datasets in a way that will generate new insights. For example, in addition to structured data such as sales figures and inventory tracking, big data can integrate unstructured data such as social media conversations or other content that doesn't neatly fit in a spreadsheet. Assimilating structured, semi-structured, and unstructured and using sophisticated algorithms reveals new patterns that make for faster and better business decisions.

How you arrive at big data insight can take you down various paths and may have various infrastructure requirements. You have to determine what question you are seeking to answer and how large a dataset you need to find a meaningful result. You have to assess your storage requirements to handle all that data, and what kind of analytics and computing power you will need to process the data. These are the technical considerations, but what is driving the big data sale is the need to answer the big business questions.

Selling big data is about selling actionable insight. That's why an effective big data sales pitch is made to the business visionaries and decision-makers of the company, not IT management. The chief executives are the ones struggling with the big questions about how to generate new revenue, increase productivity, cut overhead, and open new markets. Those are big data questions. Once you have sold them on the big data approach to answer these questions you can talk to the CIO and IT management about what technology will be required for execution.

Unlike traditional technology sales, selling big data is not about speeds and feeds. The performance and scalability of the infrastructure only has value if it supports the outcome. Since insight is more important than infrastructure, the VAR's role becomes more that of a business consultant than an IT professional. The reseller has to be able to talk about burning business issues in P&L terms that the CEO can appreciate. As a business consultant, you are building a business case where you can demonstrate how big data delivers concrete returns. It's all about turning information insight into revenue or substantial operational savings.

If you can bring a deeper knowledge of the customer's industry to the problem it will go a long way toward aiding your big data consulting efforts. Much of the success of big data analytics hinges on being able to formulate the right question and assimilate the appropriate data streams and variables. If your customer is in retail, for example, knowledge of stocking strategies and supply chain issues could be invaluable.

This guide has been developed to help you provide big data counsel to CEOs, CFO, CMOs, and others in the executive suite struggling with the big business questions. The objective of this guide is to help you adapt your IT expertise and knowledge of databases, enterprise infrastructure, data storage, and analytics to the businesses challenges your customers are facing. Selling big data requires a different approach, starting with the business value and insight big data delivers before you talk about the technology to deliver that insight. We hope to show you how to "talk the talk" so you can work with your customers as clients, helping them define their business issues, define the scope of their big data initiatives, and offer solutions that turn data into insight that is actionable and delivers real results.

Understanding Big Data Insight

Let's start by considering what executives know about big data and what they think they know.

Big data is still one of the most ill-defined and poorly understood terms in technology. In essence, big data means any dataset that is so large and so complex that it cannot be captured, curated, and processed using existing computer technology within a reasonable time. Big data datasets often run into petabytes (1000 terabytes of 1015 bytes); datasets too large to readily manage using existing enterprise systems.

IBM estimates that 2.5 quintillion bytes of data are generated each day, and that 90 percent of the world's data is less than two years old. The data explosion opens new opportunities for identifying trends and extracting meaningful information from the ever-rising tide of data, assuming you have the capacity and know how to sift through it.

In 2001, Doug Laney of Gartner defined the problem of data growth and dealing with ever-increasing volumes of available information. In defining the challenge of managing the increasing flood of data, he offered a three-dimensional viewpoint to define the big data problem:

1. **Volume** – Coping with the shear amount of data being generated.
2. **Velocity** – Handling data at high processing speeds, in and out (I/O).
3. **Variety** – The types of data being generated and the number of data sources are increasing and need to be incorporated into the analytical mix.

As Gartner defines the principle of the three V's: "Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization."

For those familiar with the concept of business intelligence, big data goes one step further. Where business intelligence uses descriptive analytics to measure things and predict trends, big data uses inductive statistics to infer truths from large datasets (e.g. using regressions, non-linear relationships, and causal effects). Where business intelligence tells you what is happening, big data also tells you why it is happening. Big data also uses large volumes of unstructured data that cannot be processed by DBMS, RDBMS, or other conventional database techniques.

What Big Data Can and Can't Tell You – The Value from Big Data Projects

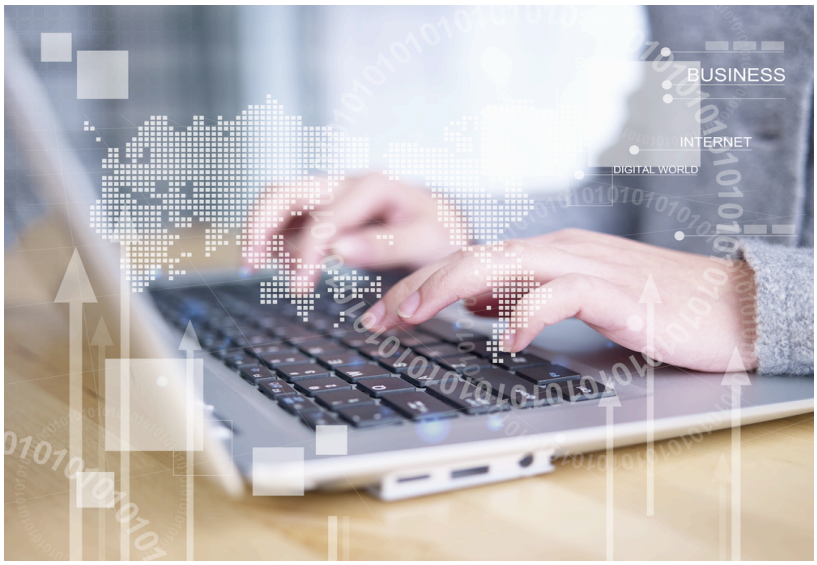
One of the other differences between business intelligence and big data is the application of the findings. One way to think of business intelligence is in terms of using static information to uncover business trends. The data used for business intelligence analytics is "steady state," such as the previous year's sales figures, purchasing history, or inventory statistics. These data points will tell you where you have been, and you can only extrapolate future trends based on past performance.

With big data you get insight from changing metrics or real-time data. Big data is designed to help business leaders make immediate decisions based on immediate data. For example, in addition to using historical mobile phone records and GPS archival data to assess consumers' mobile shopping habits, big data lets you integrate real-time cell phone activity to deliver immediate marketing sales messages to consumers based on their location or activity. Companies like Facebook and Amazon use big data analytics every day to match consumers' online behavior to real-time offers and advertising.

When approached properly, big data will deliver insight that can be used immediately to implement a strategy or make a tactical decision or course correction. To be effective, you have to start with the right question before embarking on a big data project.

According to Forrester research, the three key issues that lead to big data failure are:

- Underestimating the technical skills and expertise required;
- Creating a data silo with a big data project; and
- Not starting the project with a question.



This last point may seem counterintuitive to most senior executives. Isn't big data supposed to reveal what you don't know? Doesn't it provide new insights so you can find the needle in the data haystack? You can have data scientists plowing through streams of data looking for new insights, but without context there is no way to tell whether you are getting real insight or just interesting noise.

Starting with a focused question doesn't limit the potential scope of your findings. Consider how web search operates. The first question yields results that allow you to refine the question and the process repeats itself until you get the insight you want. With big data, it's a matter of having a context to identify the right datasets and create the right algorithms to reveal meaningful results. You have to start with a hypothesis so you know when you get the results you are seeking.

Where to Seek Value from Big Data Projects

When helping customers formulate the right big data question, consider the areas where big data analytics have proven most insightful:

4. **Supporting business decision-making with analytical proofs** – Storing transactional data allows companies to collect both real-time and near-real-time performance data about everything. Being able to process all that stored data and add in real-time data streams makes it easier to perform “what if” analyses. Big data insights come from setting up controlled experiences to analyze performance variables, identify underlying causes, and ultimately manage performance.
5. **Improve business performance by creating transparency** – Making data more accessible creates value in itself. The big data process has to break down data siloes and makes information more readily available and searchable. The result is that big data can improve performance because transparency highlights inefficiencies and bottlenecks in business processes.
6. **Identify new business opportunities, new products, and new markets** – Big data can be extremely useful in developing growth strategies. Big data analytics can assimilate existing product data, including anecdotal customer data from social media and other sources, and aid modeling of the next generation of products. Similarly, big data can identify new niche markets or areas of opportunity.
7. **Supporting customer and market segmentation** – Big data can help tailor products and services to meet the needs of a niche market or specific customer. Using big data insights for microsegmentation makes it possible to target narrower customer groups in a way that yields more sales.
8. **Competitive analysis** – By using what you know about your competition and bringing in additional data streams reflecting market conditions, customer attitudes, raw material availability, and other trends you can create an accurate map of the competitive landscape. Big data makes it easy to identify strengths and vulnerabilities in target markets.
9. **Replacing human decision-making** – Risk engines driven by big data algorithms can automate routine decision processes and minimize human error. For example, inventory and pricing controls can be automated using real-time sales data.

Big data is not the product itself, but the insight it can deliver can be invaluable. By couching big data results in practical terms, you can demonstrate the concrete value of big data initiatives.

An effective argument that sways many senior executives is how you can use big data to harness insight from social media. Since big data can assimilate both structured and unstructured data, you can use Facebook posts and Twitter streams as unstructured datasets to learn more about consumer attitudes and behavior to learn more about customers. Properly applied with the right questions, a big data analysis of customer buying habits fueled by social media can reveal more about customer behavior than was ever possible before. It's best not to oversell big data from the outset. The most effective approach is to define a use case as a test. Once you have proven the value of the initial project you can expand. Applying big data is a lot like panning for gold. You have some indication of where to start looking, and when you strike a rich vein of information you can invest more to mine it for more insight.

Start with a Big Data Question

To improve the chances of big data success, you have to define the scope of the big data project so it delivers insight that has ongoing value and is immediately actionable. You have to start with the right question.

A study by Infochimps reveals that 55 percent of big data projects are never completed or fall short of their objectives. The two primary reasons for failure are a lack of business context around the data (51 percent) and a lack of expertise to connect the dots (51 percent). By properly defining your big data question (creating context) and formulating a hypothesis that will derive value from the data (connecting the dots), you will have a much higher chance of success.

Build on What You Know

The best place to start is by looking at existing business problems that the company is striving to solve using their existing database and data infrastructure. Whether the challenge is risk analysis, lead generation, cost containment, marketing spend optimization, cross-sell, upsell, supply chain analysis, pricing, or any other business problem, ask yourself if adding more data to the mix will give you a better analysis that is more statistically accurate.

When determining the type of big data question you want to address, consider these three criteria:

1. **Does the answer to the question require multiple data sources?** If the business challenge cannot be addressed by data already available, then the problem may be a good candidate for big data. Assess what specific types of data will be required to address the question, and be sure to inventory data silos within the organization. If you need to add external data streams or assimilate structured and unstructured data then you have a need for a big data project.

2. **Will the answer to the question yield actionable results?** Most big data projects yield insights that have ongoing value. For example, if you try to determine the size of a potential market that is a one-time question which means you can invest a lot to build a big data infrastructure to get the answer. If, however, you want a better approach to retail pricing using multiple data sources, that question has ongoing value; once you assemble the data streams and build the right algorithm you can use that infrastructure again and again to guide pricing decisions.

3. **Will the answer to the question deliver sufficient return to justify the cost of the big data initiative?** The big data question can't be too big or too small but has to be just right – the insight has to have lasting value to justify ROI, as with our pricing example, above, and it needs to address a business problem with sufficient scope and returns to justify the big data expense.

Use an Interdisciplinary Team

One of the byproducts of big data is that it breaks down data silos within the organization. Big data is as much a cultural initiative as it is a technical initiative. You want to break down departmental walls to pool all the company's data for more accurate results.

You also want to make sure that all the stakeholders are included. Be sure to include sales, marketing, manufacturing, accounting, any department or group who can benefit from the big data findings or who will be affected by the outcome of the big data project. Gather together

representatives from all affected departments and get their input on formulating the scope of the big data project. Are you asking the right question? What data sources are available that will add insight to answer the question? How would they like to use those findings to do their jobs better?

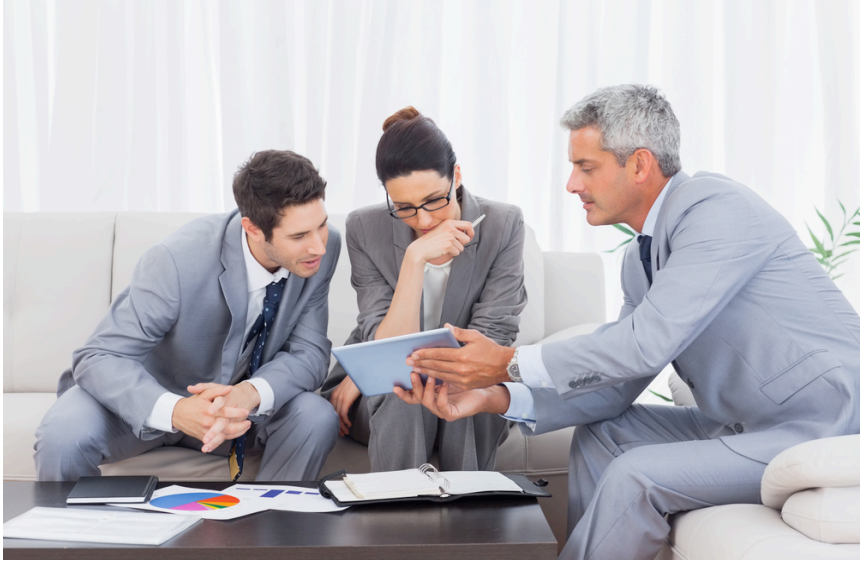
Embracing big data often requires a corporate cultural shift to pool rather than horde data for mutual benefit. Promoting data unification requires executive commitment, but the outcome is that everyone in the organization gets better information to support decision-making and improve revenue and operations.

In making your pitch to the executive team, make sure they appreciate that big data isn't an end in itself but a means to an end. To achieve that objective, i.e. to answer the burning questions that will lead to success, they will have to adopt a holistic approach that will affect the entire organization. Once they embrace big data there is no going back.

Build a Use Case

In addition to developing the right question you have to set the appropriate scope for the project. The same Infochimps survey we cited earlier revealed that 58 percent of the IT professionals who responded said their big data projects failed because of inaccurate scope. Setting the scope of the project so that it is revealing but not all-consuming will promote big data success.

To make sure that the scope of the big data project is achievable you have to plan in advance. Work with interdisciplinary team to agree on the objectives of the project in order to define its scope. You want to establish expectations in advance so the customer understands what the project will yield in terms of actionable insight that will deliver ROI.



The best strategy is to start with a use case. A use case is defined to address a specific business challenge by analyzing patterns that can be customized for the specific big data question at hand. If you can develop a smaller, more manageable use case it will be more useful as a proof of concept. Once you have proven the value of big data you can develop a broader use case and justify a bigger big data infrastructure.

Apply a Universal Use Case

Different types of use cases will be valuable to companies in different markets, but there are two use case concepts that seem to be universal:

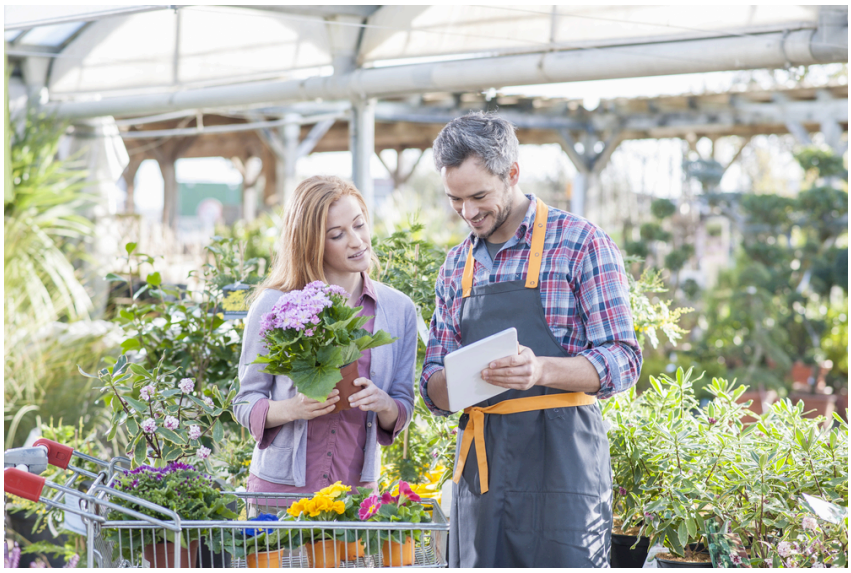
1. **Get an Enhanced View of Your Customer**
   Big data is ideal for learning more about consumers or customers and their needs and buying habits. Using data already stored within the enterprise and additional data sources such as social media streams you can create a 360-degree view of customers, revealing what they buy, how they shop, what they will buy next, and what leads them to recommend your company or products. Big data also can reveal better ways to engage with customers.

2. **Perform an Operations Analysis**
   Machine and operational data can be incorporated into big data analytics to assess operations. Output from sensors, meters, and other devices can be used in complex analyses to gain real-time or near real-time visibility into operations and transactions. The “Internet of Things” is generating data from everyday devices such as cell phones, cars, watches, televisions, even smart diapers. Any new input into the data stream can be mined for big data analytics.

Consider Vertical Market Use Cases

Big data also lends itself to addressing problems unique to specific industries and markets. Depending on the customer, consider creating a use case that is unique to their industry but that will deliver big returns. Here are a few examples:

Health Care – Using big data, a pharmacy benefits management company can assess patient behavior in order to improve process. For example, the company can use predictive analytics to identify patients at risk of missing their prescribed dosage, or send customized messages to patients to help them save on medicine. Or consider a hospital that uses big data to integrate financials, operational, and clinical analytics systems so physicians can enter queries at any time and access real-time information to review patient records.



Retail – Retailers are using big data analytics for stocking strategies, pricing, store design, and a variety of applications. Consider, for example, that a retailer is stocking the latest hot video game title. Using big data analytics they can assimilate historical data, such as selling season and past sales performance, with other information such as game advertising, social media buzz, movie tie-ins, etc., to predict demand. Using geographic modeling they can even predict demand by retail market. They can even create dynamic pricing models and targeted promotional campaigns.

Manufacturing – Manufacturers in all areas can benefit from big data. InformationWeek recently published a profile explaining how Intel used predictive analytics to save $3 million in manufacturing costs. The chip maker was able to implement an end-to-end big data process analyzing machine data from all the equipment on the factory floor. As a result, they were able to dramatically reduce product testing. Intel executives predict that ultimately they could save well over $100 million in cost savings and cost avoidance using big data analytics.

These are just a few of the use cases by industry. There are similar use cases to be found for telecommunications, financial services, and insurance – any industry that can benefit from analyzing trends from large datasets. By talking with customers you will be able to find a use case that addresses an immediate business problem where the findings will have immediate value, including a large return on the pilot project.

Assess the Big Data Infrastructure

Once you have defined your big data problem and built a working use case it's time to assess available resources to determine what you have already and what you need. There will be technology gaps, but you can fill them or find workarounds.

Let's start with the use case. Once you have defined the use case map out the data flows needed to address the big data question. This will highlight what technology is missing and what data capabilities are required. It also will allow you to determine what data should be included and what data is less useful; you only want to include datasets that have strategic value and will deliver insights. Peripheral data might be interesting but will also generate distracting noise that may not be meaningful. This step will also allow you to identify data silos. The interdisciplinary team should be part of this process so all potential data sources are considered. It's often best to start with "little data," i.e. data you have in hand, and then expand your project with additional data sources.

This is also a good time to develop business rules and assess the complexity of how the data interrelates. Once you have a better understanding of the data sources you can identify analytical queries and algorithms that will be required to generate the desired output.

As part of the assessment of big data capabilities, be sure to identify gaps in data handling. For example, what additional data quality requirements are needed? Do you have the tools to collect, cleanse, and aggregate data for analysis? What about data governance policies? You need methodologies to classify data, define its relevance, and then store, analyze, and access it.

Understanding the Enterprise Requirements

Naturally, the big data assessment also includes close scrutiny of enterprise resources. While the common myth is that big data projects use commodity hardware, the truth is that high-end hardware is often required. You need to have sufficient computing power to handle higher compute intensity (i.e. a high ratio of operations to I/O) and increased parallel processing capacity. Big data also uses a distributed architecture, with analytics processing being performed locally where the data is stored. That means virtualization and the ability to seamlessly integrate multiple computing resources both within the enterprise and in the cloud.

And then there's storage. Your big data project could demand petabytes of information. That means having the capacity to store all that data; some of it within the enterprise and much of it in the cloud. And the big data infrastructure has to have the elasticity to adapt as more data and more computing power are called on for analytics.

When assessing the enterprise infrastructure, scalability and ease of management are critical. Be sure your big data environments meet these three criteria:

1. You need a highly scalable system with extreme parallel processing and dense modular packaging. The system will likely need more memory, more bandwidth, and more throughout to handle multiple tasks simultaneously. It will have to respond to millions of events per second and have the processing capacity to handle advanced analytics algorithms within a reasonable timeframe.

2. The computing environment needs to be reliable and resilient. Demand will fluctuate and the system will need to be able to handle increases in workload without having to change the architecture. This will enhance workload performance with little or no downtime, and improve security since data processing is contained.

3. The system has to be built around open source. Whether you use Apache Hadoop or some other open source platform, the computing system has to be built with an open architecture to promote interoperability and offer flexibility. It also simplifies management of new workloads using virtualization and cloud resources.

Staffing Considerations

As part of the infrastructure assessment you need to make sure you have the right staff available to handle the project. Big data expertise basically falls into three areas: architecture, software, and analytics.

You probably already have enterprise architects on staff. These professionals understand enterprise networking design and how to design that flexible, scalable infrastructure you will need. The ideal big data enterprise experts will understand how to deal with high-speed I/O,

bandwidth, throughput, parallel processing, virtualization, cloud storage, and related issues.

The software experts are the database gurus who can help harness the data streams you will need. Ideally you want software experts with a database programming background who can work with the existing database systems such as DBMS and RDBMS. These professionals will be responsible for application development and scrubbing and cataloging data sources to support analytics.

Then there are the analytics experts. Ideally you want to recruit statisticians or big data scientists, but since that kind of expertise is rare you can recruit SAS programmers and in-house programming experts with a penchant for business problems. The role of the analytics programmer is to extract insight from big data, so they will need to be able to learn Hadoop or some other programming and create the algorithms to process the data while having an understanding of the tactical value of the results. They also will need the expertise to interpret the data findings and model the data in a graphic form that makes it understandable to senior management.

Chances are that most of the talent you need to launch a big data initiative is already on staff. Talk to the CIO and IT management to assess the team's skillset and identify areas where you may need external resources.

Measuring Big Data ROI

Now you are ready to launch your big data program and start extracting insight from your disparate big datasets. But how do you measure the return on investment? What metrics can you use to demonstrate that those big data insights are paying off?

Wikibon conducted a survey of companies using big data and discovered that the expectation is to get between $3 and $4 return for every $1 spent. The reality is that the rate of return is more like $0.55 on every $1 spent. This is largely because of poor planning and implementation. The reasons most often cited for a less-than-expected outcome are: 1) lack of experienced big data experts with the skills to analyze the results; 2) immature technology and a lack of big data tools capable of delivering results; and 3) a lack of a compelling need for big data, which means insufficient advance planning and failure to tie the big data project to a measurable and actionable outcome.

Big data opportunities generally fall into two broad categories: game changers or business extensions. You have to determine how to measure the results based on how you categorize the insight you are seeking. Will the results be a big “aha” that will change the direction of the company, or will it be a discovery that will improve current operations?

While the insight should lead to better decision-making and positive action, too often the results are hard to tie directly to revenue. Another survey conducted by IDG Research of more than 200 IT and business leaders revealed mixed reviews of big data outcomes. More than 85 percent agree that big data enabled better informed business decisions, but 23 percent said their big data was a success while 52 percent said their big data projects were “somewhat successful.” Again, the disappointment stems from a lack of the right insight and an inability to extract laser-focused data points without expensive analytics consultants. Companies also report that the cost and time required to deliver the results were greater than anticipated.

Choose the Right Metrics for Return on Investment

The challenge with measuring big data ROI is that senior executives often have to initially rely on a perception of success rather than metrics, which is ironic since big data is designed to provide hard data to aid decision-making. To help executives gauge the returns on big data, you have to set expectations in advance during the planning stages.

Too many executives expect their big data investment to deliver a big payoff, such as the 300 percent return cited by the Wikibon survey. Big data is not an all-seeing oracle that can answer every business question. However, if you set the scope of the project appropriately and ask the right questions you will get results that will yield measurable benefits.

Instead of trying to measure direct returns from big data output, consider using different ROI metrics:

1. **The cost of an alternative approach:** What is the cost of the big data project compared to another route to get the same information? Can you expand the data warehouse to accommodate more data and different types of data? Of course, you could use OLTP technology and add database servers and software, but the cost will probably run into millions where a Hadoop framework can deliver the same results for about $200,000. You still extract more value from your existing datasets and the Hadoop infrastructure relieves the strain of adding data processing capacity.

2. **The overall value of big data insight:** If you limit the scope of the big data project, the value of the results will likely be greater by comparison. If you start sampling data looking for answers to narrowly defined questions it will reveal how rich the available data is, and what other insights you might extract. As the value of smaller big data projects becomes more obvious, you can expand the big data infrastructure to get more value. Using cost-saving strategies like cloud storage will help promote a greater ROI.

If you start with smaller big data projects you also can reduce "time-to-insight." Limiting the scope of the queries will make it easier to deliver easy to grasp, actionable results that will prove the value of big data for better business decision-making. Being able to deliver real insight quickly can offer customers a real and dramatic competitive advantage.

Conclusion

Once you have the proof of big data ROI you can discuss next steps. As big data questions become more intricate and complex, the types of data streams will grow and vary even more and projects will call for more computing power, more data storage, more analytics, and better backup and security procedures.

So here are some basic rules when engaging with senior executives about the value of big data:

1. Avoid techno speak and use concrete business examples.
2. Limit expectations and make sure the customer understands what to expect.
3. Include all stakeholders; big data is an interdisciplinary initiative.
4. Keep the scope of the use case focused and manageable, identifying a big data question that will yield an outcome that will prove the value of big data.
5. Gather the right metrics to tell a compelling business story.
6. Present the findings not the data; most executives want a graphic representation, not raw data.

Part of the measure of success is the impact that big data has on the corporate culture. Embracing big data is an all-or-nothing commitment. Once the executive team understands the value of big data they will start to ask bigger questions, get new answers, and generate new ideas that lead to new questions. There's no returning to the silos of business intelligence. Big data opens up a new world of possibilities that will continue to yield bigger and better insights and bigger and better profits for big data resellers.