FROM BIG DATA TO BETTER DECISIONS

The ultimate guide to business intelligence today.

THE BI GUIDE: WHAT YOU'RE GOING TO LEARN.

From sales opportunities to supply chain logistics, and from accounting software to social media stats, your organization is bursting at the seams with data. Business intelligence (BI) is the combination of tools, processes, and skills that help turn that vast amount of data into digestible information.

With information coming from every part of your organization, everyone needs better access to data to do their job well. Chances are, you need it, too. It's why you're reading this guide.

Browse our list of chapters to go straight to information specific to your needs, or feel free to read it cover to cover for a holistic view at how you can use BI to shape your work.

BI HAS OUTGROWN SPREADSHEETS AND DATA WAREHOUSES.

Before business intelligence was "business intelligence," it was nothing but numbers written on spreadsheets (the actual paper variety). But as technology grew, little changed with how business leaders consumed information—they moved from paper to digital, and when the volume got to be big enough, they moved the data from desktop spreadsheets to a massive table known as a database. In the end, the results were the same: static information presented in a document, maybe with a few graphs thrown in for good measure.

But you don't need new ways to replicate antiquated business practices. If you can get real-time updates on obscure college friends' lives through social media, then you should be able to access information from your business anytime, anywhere.

Now it's time to learn how to access the right data at the right time. Keep reading—you're in the right place.

Five reasons to read this guide.

You see BI as important and want more info as part of your professional development.

You'd like to pursue a career in BI.

You've got the gist of BI and want to brush up on some details.

You want to know what the BI team in your organization really deals with day to day.

You've heard so many buzzwords—"big data," "data science," "business analytics," "predictive analytics," "BI," etc.—and want to know what all the fuss is about.

WHY YOU SHOULD READ THIS GUIDE.

Data is on everybody's minds—from executives pushing their teams to take advantage of all the data the business collects, to consumers worrying about sharing too much of their personal lives. This guide cuts through the buzzwords and the technical jargon to give you an overview of business intelligence—the tools, processes and skills that help us harness the data explosion to make better and faster decisions. A state-of-the-art BI environment ensures the shortest and most reliable path from data to decisions that make your business more successful.

CHAPTER ONE

A flood of data, and how BI addresses it.

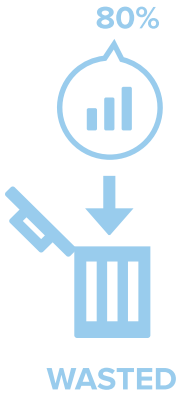
A FLOOD OF DATA.

We are living in a data deluge. The amount of new data created annually will grow ten-fold between 2013 and 2020, according to IDC, from 4.4 trillion gigabytes to 44 trillion gigabytes.

If you can swim in this flood of data, you win. According to MIT researchers, companies that excel in data-driven decision-making are 5% more productive and 6% more profitable than their competitors, on average. A study by IDC found that users of big data and analytics that use diverse data sources, diverse analytical tools, and diverse metrics were five times more likely to exceed expectations for their projects than those who don't.

DATA WITH NO ANALYSIS HAS NO VALUE.

Navigating the flood of data is much easier said than done. IDC predicts that companies will continue to waste 80% of customer data they have collected. More broadly, IDC estimates that in 2013 only 22% of all data in the world was useful (i.e., could be analyzed) and less than 5% of that was actually analyzed.



A University of Texas at Austin study put these general estimates in a business context: it found that for the median Fortune 1000 company, a 10% increase in the usability of its data translates to an increase of $2.01 billion in annual revenues and a 10% increase in remote accessibility to data translates into an additional $65.67 million in net income per year.

$2 BILLION ON THE LINE, BUT NOTHING NEW FROM BI.

With $2 billion on the line, CIOs have reported in Gartner's surveys that business intelligence has been a top priority for the last nine years. CEOs are also getting on the bandwagon, demanding more and more access to more and more data.

While the need for timely, accurate, and accessible business intelligence is greater than ever, the use of business intelligence tools has plateaued at about 20%-25% of business users in a typical organization over the past few years.

As Gartner recently observed, "despite the strong interest in BI and analytics, confusion around big data is inhibiting spending on BI and analytics software."

The frustration is widespread, according to surveys conducted by businessintelligence.com:

• Only 25% of CEOs say their reports contain the information they need and want.
• 44% of executives say that too many of their critical decisions were based on incomplete or inaccurate data.
• 75% of vice presidents surveyed said that they were dissatisfied with their access to the data they need, and 69% were not happy with the speed of information delivery.

These data management challenges are compounded by bloated solutions, complex deployments, and overly complicated user interfaces. The emergence of new tools and technologies for harnessing the data deluge, aimed at solving these issues, **may actually slow down the adoption and widespread use of business intelligence.**

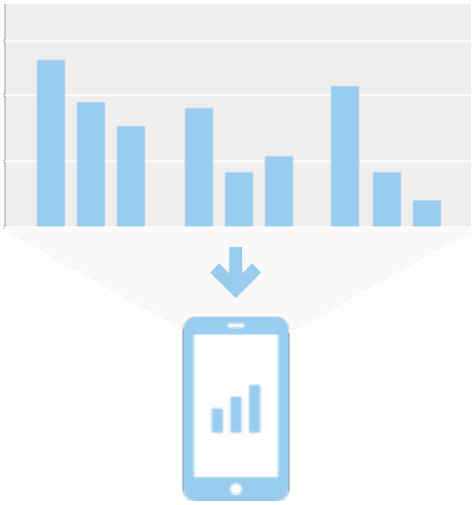
Want to learn more? Read the executive brief, "The big BI disappointment: Troubling gaps between BI expectations and reality."

WHAT BUSINESS LEADERS NEED FROM BI.

Today's leaders no longer make decisions based primarily on intuition. Instead, making decisions today is a team sport, involving all the relevant people in the organization, and taking advantage of new technologies to collect and analyze all the relevant data. In this all-hands-on-data environment, decision makers expect:

- Data from all relevant sources in one place.
- Real-time data, not having to wait for an analyst to deliver it or wait for IT to respond.
- Data that is accessible anytime and anywhere, on any mobile device.
- Data that represents one version of the truth
- Self-service data and analysis, reducing the reliance on experts.
- Data and analytics that help predict what's coming.

Today's business intelligence is embedded in all levels of the organization, allowing anyone that needs to make a decision—operational, tactical, or strategic decision—to make it based on the best data available. Business Intelligence is the combination of tools, processes and skills that help us turn the data deluge into better and faster decisions.



Want to learn more? Read the analyst report, "7 Steps to Making Big Data Accessible to Executives."

CHAPTER TWO

The business intelligence market.

HOW BIG IS THE BI MARKET?

Gartner estimates that the worldwide business intelligence and analytics market was $14.4 billion in 2013, growing at 8% annually. Assessing the larger market for business analytics, IDC estimates it had reached $104.1 billion in 2013, at a growth rate of 10.8%. The big data segment of this market was $12.6 billion in 2013, with a growth rate of 27%.

The business intelligence market is dominated by a few large players—SAP, Oracle, IBM, SAS, Microsoft, Teradata—accounting for about 70% of worldwide revenues. The balance of the market is accounted for by hundreds of small players, including numerous new startups, most of them focused on one or two segments of the market. Established business intelligence-focused companies include Actuate, Information Builders, Panorama, MicroStrategy, QlikTech, Tableau Software, and Tibco Software. New startups include Alteryx, Birst, Domo, Good Data, and SiSense.

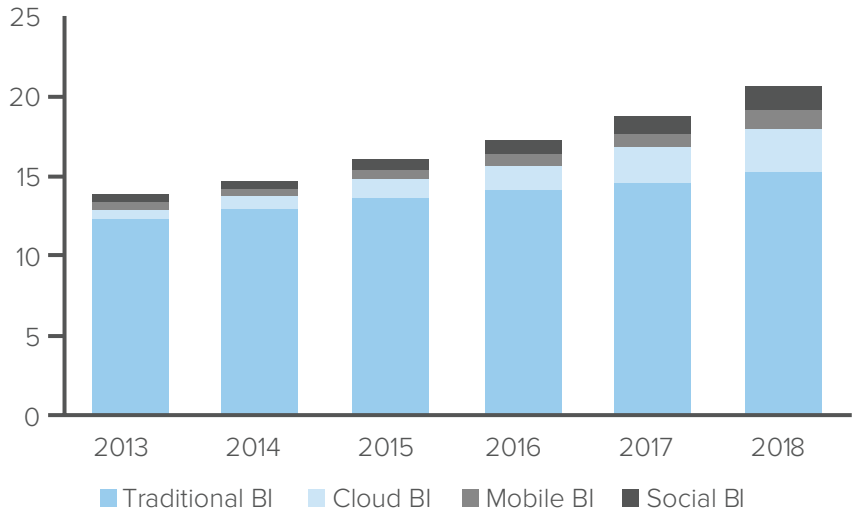
HOW IS THE BI MARKET CHANGING?

In older tools—and even in most current solutions—BI tells you what happened in a specific segment of your business. With how quickly business is moving today, that kind of BI is as problematic as driving down the freeway by looking only in your rear-view mirror.

With new technology and new expectations, BI is moving toward a more predictive model that shows you what will happen. New BI systems are now beginning to show how all the various parts of your organization work together to produce an outcome, and business leaders can finally see the big picture and make faster, better-informed decisions.

This transformation started over a decade ago as more and more firms started to compete on the basis of statistical analysis and data management prowess. It's what drove today's online giants like Netflix, Google, and Amazon—each with a reputation for mastering data, measurement, testing, and analysis—to be what seem like unstoppable forces.

In response to these giants' success—who barely existed 20 years ago—many established companies now invest in statisticians and operations research personnel, build business analytics departments, weave modeling, prediction, and forecasting into their processes, and acquire new hardware and software tools to support these activities.



Global intelligence market size, by technologies, 2013–2018 ($ billion)

Sources: Gartner, Redwood Capital

HOW ARE ORGANIZATIONS MEETING THE DEMAND?

More recently, another new layer of the business intelligence market has emerged and become known by the somewhat misleading name of big data. Again the main culprits were online firms such as Google, Yahoo, and LinkedIn but this time the new layer of the market was created around the new technologies (e.g., Hadoop), and the new roles (e.g., data scientists) that were invented by these companies to support data-driven decision-making and turn their data into revenue streams. Now, every organization has to reconcile itself to the rapid growth of available data, the competitive pressures to excel in data mining and analysis, and the increasing need to bring these capabilities to all levels of the organization.

But no economy had enough trained talent—data scientists, analysts, systems managers, etc.—to meet the sudden demand, prompting a burst of new technologies meant to fill the void. Thus, investment in BI tools and technologies is primarily driven today by the trend towards wider adoption of BI, giving end-users easy-to-use tools for accessing, viewing, analyzing and manipulating data. This "democratization of business intelligence" or "self-service BI" is accompanied by growing investments in embedding BI capabilities in various business processes and applications. These new applications, leveraging new data types and new types of analysis, are increasingly installed on mobile devices, drawing on data that resides in the cloud, supporting users anywhere, anytime.

"Major changes are imminent to the world of BI and analytics including the dominance of data discovery techniques, wider use of real-time streaming event data and the eventual acceleration in BI and analytics spending when big data finally matures."—Gartner

In developing their business intelligence capabilities, organizations have always had the option to buy outside services to supplement their own in-house activities. They could buy specialized skills, consulting, or even specific data from data aggregators. This segment of the market, now called "data-as-a-service," has recently grown rapidly with the emergence of new players providing data services with embedded BI and analytic capabilities. To alleviate the analytics and data science talent shortage, some vendors focus on providing the required skills on a project-by-project basis.

WHAT ARE THE KEY COMPONENTS OF THE BI MARKET?

The BI market is typically segmented according to product functionality such as "query and reporting," "online analytical processing (OLAP)," and "dashboards." It is easier, however, to understand the BI market if we look at the process of business intelligence or the steps required to get from data to decisions. In a nutshell, the process of business intelligence has three steps: Ingestion, Analysis, and Delivery.

"Experts often possess more data than judgment." —Colin Powell

It starts with the ingestion of data—identifying the right data sources and preparing the data for analysis; continuing through the analysis stage, including processing the data and applying analytical models to the data; and concluding with delivery—presenting the results of the analysis in easy-to-consume manner and at the most convenient point of consumption for the user.

These three steps will be covered in detail in the following three chapters.

See the infographic: "The World Needs Data Scientists."

CHAPTER THREE: THE BI PROCESS

Step 1 - Ingestion

INGESTION:

The process of business intelligence starts with identifying the data sources and the type of data that can support specific decisions and business objectives. Once you have the data, you need to make sure it is ready for processing and analysis.

HOW IS THE BI MARKET CHANGING?

Before an organization can take in data, business leaders need to understand where it's coming from, what format it's in, and how to turn raw data into something useful.

Here are some of the basics:

The data for business intelligence comes from a variety of sources, internal and external to your company. Internal sources include engineering and manufacturing processes, Enterprise Resource Planning (ERP) systems, sales force automation and customer relationship management (CRM) software, and financial and accounting activities. External sources include supply-chain and logistics systems, business and distribution partners, social networks, websites, location/GPS systems, mobile and stationary sensors, and click streams. There are also many "open data" sources on the Web that make data collected by government agencies, non-profits and businesses available at no charge.

DEALING WITH DATA STRUCTURE.

The data coming from these disparate sources is in many types and formats, including rows and columns in traditional databases, images, text documents, video, PowerPoint and HTML files, email

messages, sensor data, web-based transactions, and IT systems logs. These data types are usually classified into three broad categories: Structured, semi-structured, and unstructured data:

Structured Data

(e.g., the numbers in a customer invoice) *can be easily ordered in the rows and columns of a traditional database table (e.g., customer account number, invoiced amount) or some other type of database with a defined structure.*

Semi-Structured Data

(e.g., HTML or email files) *conforms to a partial structure or a standard format and contains specific markers that give it some type of organization.*

Unstructured Data

(e.g., an image) *is not organized in any pre-defined manner.*

THE STRUCTURING OF DATA: A HISTORY.

"Structured" and "unstructured" are somewhat misleading terms. All forms of human communications have some structure (e.g., language), and machine-generated data typically has a structure because it is designed to have one. What we have is a continuum that extends from a highly rigid structure, which is defined before the processing and mining of the data to highly flexible structure that is defined after the processing and mining of the data.

The "highly rigid" end of the continuum gave rise in the 1970s to technologies such as relational databases that exploited the structure imposed on the data. The focus on "structured" data, (i.e., data with predefined structure), continued until the 2000s. At that point, online search and web analytics companies started digging into "unstructured" data, (i.e., data without a predefined structure). New techniques are now available that take in data that has loose structure (e.g., log files) or implicit structure (e.g., natural language) and extract that structure rapidly and at scale, making it available for analysis in a time frame where it is still useful.

DATA PREPARATION.

Given the variety of sources and types of data, a lot of work needs to go into preparing the data before it is stored and analyzed. The data could be of varying quality (e.g., an address may be missing a ZIP code or may contain a spelling mistake), may not be consistently recorded in the same manner in different sources, and may have a different format. All of these issues are dealt with and the inconsistencies and imperfections of the data reconciled in the process of data preparation. It is usually referred to as the Extract-Transform-Load (ETL) process, where the data is taken from its source, changed to fit certain rules or standards, and then moved to where it is stored, typically a data warehouse.

Following the Garbage In, Garbage Out (GIGO) principle, the "cleansing" of the data turns out to be one of the most crucial steps in the BI process and requires careful attention. This has become very important recently with the rise in the quantity and variety of data sources and it is often said that 80% of a data scientist's time is spent on cleaning the data. Being the new kids on the data mining block, data scientists have recently invented new terms for it such as "data munging" and "data wrangling."

"It is a very sad thing that nowadays there is so little useless information."
—Oscar Wilde

But cleaning the data is a small part of a very large process. The proliferation of data sources requires that data scientists find ways to reconcile all those data sources to each other in a process called "data integration." Data integration refers to tools that are part of the ETL process and help combine data from different sources to ensure a single, unified representation of the data.

MANAGEMENT, GOVERNANCE, AND VALUE.

The rules and standards for cleaning, transforming, and integrating the data are defined in what is called Master Data Management (MDM). Master data is the standard description of people, things, places or concepts that are important to the business, (e.g., customers, products, or sales regions). Master Data Management is the combination of tools and processes that create and maintain consistent, accurate and comprehensive lists of master data.

Data Governance

People
» Master Data

Management Processes
» Cleaning
» Transforming
» Integrating

Tools
» ETL
» ERP

Master Data Management is an important component and one of the key deliverables of a much larger, often enterprise-wide activity, called Data Governance. It is an umbrella term which includes all the people, processes, and tools required to create a consistent and appropriate handling and management of an organization's data. You will find Data Governance in action especially in business activities that require compliance with government regulation (e.g., financial services). More attention is being paid today to Data Governance in a variety of industries, however, with the increased privacy and security concerns regarding consumer data (and the mishandling of it in many cases).

THE BUSINESS IMPACT.

Finally, the increase in the quantity and variety of data sources has been linked to an increased need to support BI tasks in near real time or real time, leading to faster decisions. The goal of what is known as Complex Event Processing (CEP) is to identify meaningful events that may serve as opportunities or threats to the organization and respond to them as quickly as possible. CEP represents a unique data preparation challenge in that it is based on real-time data and as such is not part of the established ETL process.

Businesses today track and process streams of data about events that may impact their fortunes in the near- or long-term. These may be internal events such as sales leads, customer orders or customer service calls; external events such as news items, text messages, social media posts, stock market feeds, traffic reports, or weather reports; and the events may signal a change of state, when a measurement exceeds a predefined threshold of time, temperature, or other value. Streamlining all the data streams by integrating them into one coherent and manageable body of data is key to streamlined data processing and analysis which is the next step in the process of business intelligence.

CHAPTER FOUR: THE BI PROCESS

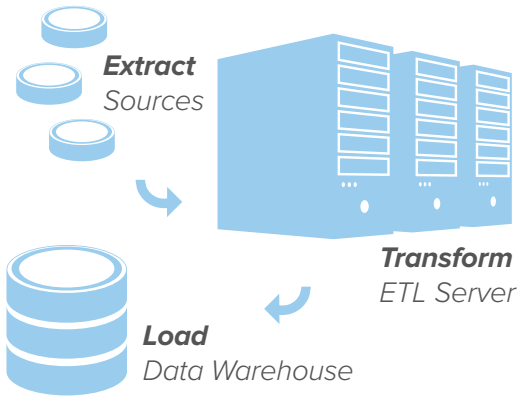
Step 2 - Analysis

ANALYSIS:

Integrated, standardized and "clean" data is stored and processed in databases or other specialized data management systems and analyzed by applying statistical models and methods to the data.

DATA STORAGE AND PROCESSING.

The Extract-Transform-Load (ETL) process typically loads the data into a data warehouse which is a specialized database used for data storage, reporting, and analysis. Traditionally, the database of choice for these tasks has been of the relational database management systems (RDBMS) variety with a popular query language called SQL (for structured query language).



Relational databases use tables to store information. The data is represented as columns (fields) and rows (records) in a table. With a relational database, the user can easily find specific information (e.g., a customer's address), sort the data based on any field (e.g., customer's name, address, type of purchase, etc.) and generate reports that contain only certain fields from each record (e.g., a record may contain all the data for a specific customer).

With a relational database, the user can quickly compare information because of the arrangement of data in columns. The relational database model takes advantage of this uniformity to build completely new tables out of required information from existing tables. In other words, it uses the relationship of similar data to increase the speed and versatility of the database.

Want to learn more? Check out the executive brief, "The Data Warehouse Dilemma."

ONLINE ANALYTICAL PROCESSING (OLAP)

A more specialized type of databases or data storage and processing systems is Online Analytical Processing (OLAP) tools. They expose the multidimensional view of data to applications and enable BI operations such as consolidation, drill-down, filtering, and slicing and dicing. Databases configured for OLAP use a multidimensional data model, allowing for complex analytical and ad hoc queries with rapid execution time.

"Information that is imperfectly acquired, is generally as imperfectly retained," —William Playfair, inventor of the pie and bar charts, 1786

OLAP databases have typically run on disk-based storage. Recently, however, as the cost of computer memory continues to decrease, analytics processing is more and more performed in-memory, (i.e., over data) that resides in computer memory rather than on a hard drive. This results in faster analysis and greater flexibility in using data from a variety of sources.

NoSQL AS A STOP-GAP.

In the early 2000s, a new type of database has started to gain popularity as it facilitated the storage and retrieval of data that is not organized the in tables used by relational databases. Collectively called NoSQL, the new databases of this non-relational type successfully managed "unstructured" data such as documents and graphs.

SQL	NoSQL
Programming language that forms the basis for relational database solutions.	Broad class of data management.
Data is stored in a single structure for consistnecy in operations.	A distributed file system stores objects across a pool of commodity resources.
Specific instructions are used to query and manipulate data in a defined table.	Different algorithms are used to query and manipulate data based on the solution.

Source: CompTIA, "Big Data Insights & Opportunities," Sept. 2013.

But even this new type of databases could not deal effectively with the rapidly growing Web and the requirements of search engines. Google, the company at the forefront of indexing and analyzing the Web, invented a completely new approach to storing and processing unstructured data.

Want to learn more? Check out the analyst report, "What Business Leaders Hate about Big Data."

MAPREDUCE

The third approach, called MapReduce, solved the problem of waiting a long time to read lots of data (later to be called "big data") from disk drives. It did it by distributing the data over many commodity servers and their disk drives and then reading and writing the data in parallel. This new approach (often described as a "framework") to storing and processing data was developed further as the open-source project Hadoop and has become the foundational technology for managing big data.

MapReduce is a batch query processor, (i.e., it runs over the entire dataset) and it does so at reasonable speeds. As such, MapReduce is a good fit for applications where the data, typically unstructured data (i.e., it does not conform to a predefined schema or structure), is written once and read many times. In contrast, relational databases are good for structured data that is continuously updated. Today, the differences between relational databases and MapReduce/Hadoop are blurring as many vendors bring to the market data management solutions that combine attributes of both approaches.

DATA ANALYSIS – DESCRIPTIVE.

Once the data is stored and processed in an optimal data management solution based on the business need and the type of decisions supported by the data, it is ready for analysis. Analytics (or Business Analytics) is the application of descriptive, diagnostic, predictive and prescriptive models to data in order to answer specific questions or discover new insights. Analysis techniques range from historical reporting telling the decision maker what happened recently to looking at the future, predicting what is going to happen and recommending the best course of action.

An example of descriptive and diagnostic modeling in widespread use is the concept of Key Performance Indicators (KPIs). KPIs define a set of values against which the performance of the entire organization, a business unit or function, or specific projects or employees is measured on a regular basis. By establishing KPIs, the business defines for its various constituencies what "success" means and a set of clear priorities. The periodical assessments of the performance of the business against its performance indicators often lead to identifying potential problems and areas for improvement.

> ***"Big data is nothing without its little brother—traditional KPIs."***
> ***—Bernard Marr***

DATA ANALYSIS – PREDICTIVE.

Predictive and prescriptive modeling makes use of statistical methods that identify trends and recurring patterns in a set of data. Largely known today as predictive analytics, these methods can be applied to any type of unknown whether it is in the past, present or future. Predictive analytics uncovers the relationships between what explains the situation we are trying to understand and a similar outcome, based on observed past occurrences.

Lately there has been more emphasis on analyzing the future than the past and the new tools and techniques supporting this shift are sometimes referred to as advanced analytics. This future-orientation and the growing use of new tools for optimization and simulation have been spurred by the arrival of big data and its practitioners—data scientists.

Watch the webinar, "Choosing the right BI solution: Overcoming 5 common concerns."

DATA ANALYSIS – MACHINE LEARNING.

The new discipline of data science, touted by the Harvard Business Review as the "sexiest job of the 21st century," is analytics on steroids. It combines statistics with computer science and knowledge of a specific business domain. Data scientists typically approach data without a pre-conceived notion of what could be found in it and analyze it to discover the "unknown unknowns" (as opposed to the "known unknowns")—what we don't know we don't know.

Another important aspect of data scientists' work is to automate the analysis of data, using computer technology. They do so primarily by using machine learning techniques.

Machine learning is a branch of artificial intelligence and is best thought of as the application of computer technology to learning. Similar to our basic learning process, the computer is "trained" by data which is labeled or classified based on previous outcomes, and its software algorithms "learn" how to predict the classification of new data that is not labeled or classified. For example, after a period of training in which the computer is presented with spam and non-spam email messages, a good machine learning program will successfully identify and predict which email message is spam and which is not without human intervention.

Even when machine learning methods are used, humans still call the shots regarding what they want the machine to learn and what data should be used. In the next and final step in the business intelligence process, humans play an even bigger role as the recipients of the analysis on which they base their decisions.

Definition of Machine Learning:

"A scientific discipline that deals with the construction and study of algorithms that can learn from data. Such algorithms operate by building a model based on inputs and using that to make predictions or decisions, rather than following only explicitly programmed instructions."

CHAPTER FIVE: THE BI PROCESS

Step 3 - Delivery

DELIVERY:

Getting the analysis to the right people at the right time and the right place in a format that is easy to consume is the essence of the third and final stage in the business intelligence process.

REPORTING

There are a variety of ways to present the results of data analysis to the decision maker. These tools range from static displays of data summaries and data graphics to more dynamic output that allows the user to interact with and manipulate further the analysis.

Reporting is a basic BI capability that provides highly formatted, print-ready and interactive reports, typically based on pre-defined triggers or conditions. Some reporting tools have ad hoc report/ query functionality which allows users to ask their own questions of the data, eliminating the need for the creation of another report, typically by the IT department or the BI group. Similarly, interactive visualization tools allow users to analyze the data by interacting directly with its visual representation. This allows users to rapidly explore patterns and find outliers in the data. A relatively new development in this space are search-based data discovery tools which allow users to employ search terms and use a search interface to explore and refine different views of the data.

DASHBOARDING

Dashboards are visual environments that consolidate critical performance metrics and integrate monitoring, analysis, and prediction capabilities. Dashboards often employ visualization components such as charts, gauges, sliders, checkboxes and maps, and are often used to show the actual value of the metric compared to a goal or target value.

"Good displays of data help to reveal knowledge relevant to understanding mechanism, process and dynamics, cause and effect"—Edward Tufte

There are three primary types of dashboards:

» Operational dashboards *are used to monitor processes, events, and specific activities as they occur.*

» Tactical dashboards *measure and analyze the performance of specific projects, processes, and activities.*

» Strategic dashboards *(also sometimes known as scorecards) are typically deployed in a top-down fashion to review the progress towards achieving the strategic objectives of the business.*

DITCHING THE DESKTOP SOLUTIONS.

Traditionally, getting the data and analysis to the user was done via internal networks, from a central BI platform to a PC or a laptop. In recent years, however, the landscape of distributing business intelligence to decision-makers has changed considerably with the addition of cloud-based solutions and the explosion of smartphones and tablets.

As more and more organizations embrace cloud computing and more data is available in the cloud, BI-as-a-service or cloud BI is gaining in popularity. It promises rapid deployment and increased flexibility, and reduces the strain on already-overburdened IT resources. It is particularly valuable when the business collaborates with partners and customers, sharing data and analysis to achieve specific business objectives or in support of a joint project. For many small- to midsize businesses, as well as departments within a large organization, the cloud is an optimal solution for business intelligence.

Check out the executive brief, "The New School of Dashboards: 5 Reasons to Reconsider Your Approach."

EXPLORING THE CLOUD.

Delivering business intelligence from the cloud is especially valuable when the decision-maker is accessing the data and analysis using a smartphone or a tablet. Mobile BI takes advantage of mobile devices' native capabilities, such as touchscreen, camera, location awareness and natural-language query to provide data and analysis in static reporting or interactive modes. Mobile BI features include multiple visual query methods, use of GPS data for geolocation and geospatial analytics, animated displays, sensor-based queries, and integration with other enterprise mobile applications. It allows for targeting customers with offers and opportunities related to their current activities and location, matching marketing messages to most receptive market segments and individuals at any given point in time, and leveraging mobile as part of a broader interactive marketing campaign.

CHAPTER SIX

The benefits of the BI

THE BENEFITS OF BI.

Though the barriers to entry to turning data into usable information are many and significant, the proper use of business intelligence has made a positive impact on virtually all aspects of business and in all sectors of the economy.

Some of BI's most successful uses are to:

- Improve processes
- Increase the performance and productivity of enterprises and employees
- Provide better customer service and to delight customers
- Streamline work with partners and suppliers
- Make the work better through the activities of government agencies and non-profits

Want to learn more? Read the customer story, "RelayRides Customer Success."

In many situations, getting better data and analysis to decision makers clearly improves what they do and how their organizations function, even if there is no quantifiable and easy-to-measure impact. There are also situations where the impact of BI can be quantified—here are 10 such cases:

HOW UPS USED DATA.

Route optimization analysis by UPS led it to come up with a simple rule for its drivers: Minimize or even eliminate left-hand turns. As a result, between 2004 and 2012 it saved 10 million gallons of gas and carbon emissions were reduced by 100,000 metric tons—or the equivalent of pulling 5,300 cars off the road annually. Route-optimization analysis also saved UPS $98 million idle minutes or about $25 million worth of labor cost each year.

Read the article ›

DATA IN THE HEALTHCARE SYSTEM.

By sharing patient data among emergency departments, the Washington State Health Care Authority achieved the following results:

- Emergency department visits declined by 9.9%
- Visits by frequent clients (who visited five or more times annually) decreased by 10.7%
- Visits resulting in a scheduled drug prescription decreased by 24%
- Visits with a low acuity diagnosis decreased by 14.2%.

Read the article ›

HOW WEATHER DATA PREDICTS RETAIL SALES.

Tesco, the largest retailer in the UK, combines data from weather records with detailed sales data, broken down by store and products, to build computer models that predict future demand for product lines according to weather forecasts. This and similar types of analysis give Tesco a more accurate picture of demand, leading to savings of £100 million a year through a reduction in wasted inventory and to a 30% reduction in the number of instances of products on promotion being out of stock.

Read the article ›

BIG DATA VS. BIG CRIME.

The Los Angeles Police Department (LAPD), with assistance from researchers at the University of California, analyzed 13 million crimes recorded over 80 years so it can predict where a crime will occur in the future. The results of the analysis, focused on one LAPD precinct, led to a 12% decrease in property crime and a 26% decrease in burglary.

Read the article ›

LOVE “HOUSE OF CARDS”? THANK DATA.

Read the article ›

Netflix decided to outbid established TV networks and invest $100 million in two seasons of “House of Cards” based on its extensive analysis of its customers’ viewing habits and preferences. “House of Cards” brought in 2 million new U.S. subscribers in the first quarter of 2013, which was a 7% increase over the previous quarter. It also brought in 1 million new subscribers from elsewhere in the world. These 3 million subscribers almost paid Netflix back for the cost of “House of Cards” within a single quarter.

DATA MODEL INCREASES CUSTOMER RETENTION.

Read the article ›

Paychex provides payroll, human resources, and employee benefits services, primarily to small businesses. As it loses about 20% of its customer base each year, Paychex developed a model that predicts high-risk customers and can track what the Paychex branches are doing (or not doing) in terms of increasing customer retention. Based on the model, some branches developed a year-end retention program, targeting clients most likely to leave by providing free payrolls and loyalty discounts. When the retention strategy was applied, the customer loss rate was 6.7%, as opposed to 25.2% loss rate when nothing was done. The analysis also helped significantly the bottom line by helping the branches overcome their eagerness to touch all customers by offering discounts to customers likely to stay with Paychex, rather than targeting only those predicted by the model to be the most likely to leave.

REDUCE INEFFICIENCIES TO SAVE $80 MILLION.

Read the article ›

Rio Tinto, a British-Australian multinational metals and mining corporation, has reduced costs by $80 million by eliminating processing and logistics inefficiencies based on its continuous monitoring and analysis of operational data. Rio Tinto’s Process Excellence Centre is staffed by 12 mineral experts who analyze data from five of the company’s coal sites in Australia, and operations in Mongolia and the US. A large interactive monitor displays technical data in real time with the center receiving data 100ms after it is produced at the site. This is examined by 20 different analytical systems in order to allow processing improvements to be immediately introduced.

THE RIGHT DATA AT THE RIGHT TIME SAVES.

Read the article ›

German online and catalogue retail giant Otto Group, has used analytics to improve its demand forecasting, leading to annual savings of tens of millions of euros. The use of predictive analytics has led to a significant reduction in rates of return on key fashion items, saving about 10 million to 15 million euros. It has also improved gross profitability on men's fashion items by introducing dynamic pricing, changing prices based on demand, and forecasting what prices customers will accept on a particular day.

CHAPTER SEVEN

The challenges of BI

ROADBLOCKS TO GROWTH.

Successful use of data requires successful access to the right information at the right time—an aspect of BI that is easier said than done.

The challenges of developing and implementing BI solutions that deliver the desired results have slowed down the BI market. BI Scorecard's annual survey of BI users, administrators, and directors found that BI adoption as a percentage of employees has remained flat at around 22% for the last several years. Moreover, only 28% of survey respondents in 2013 indicated BI has delivered significant business impact—six percentage points lower than the year before and the lowest since the survey began in 2006.

A Domo survey of more than 1,000 business executives found that only 43% have access to the information they need, just 20% feel that the business information they receive adequately answers their questions, and 68% regularly have difficulty making sense of their business information.

Getting right the implementation of any information technology solution has always been a challenge, but the recent explosion of new BI tools and modes of delivery has certainly exacerbated the situation. Here are some major roadblocks to effectively using BI to deliver the right data and analysis to the right decision-maker at the right time and place.

1. YOU'VE GOT JUNK DATA.

Poor quality of the data used by your BI solution frequently leads to poor or no decisions. For example, missing data means customer profiles contain data on some attributes for some customers but not for others, either because of poor data entry procedures or failure to collect the right data. At the same time, you may have lots of data but it is mostly meaningless or incorrect. In other situations, the data may be pure gold but if it arrives late it could completely miss the window of opportunity for making the right decision and having an impact on the business.

Paying close attention to what goes into the BI solution—to the comprehensiveness, accuracy, and timeliness of the data—goes a long way to ensuring the quality of what comes out.

2. THE IT-BUSINESS GAP.

20% authorized analysis

80% rogue analysis

The infamous or even non-existent "business and IT alignment" can spell trouble to the success of your BI solution, especially if it depends on the IT organization for implementation and on-going care. The requirements of the business often clash with IT's goals and metrics for stability, cost-effectiveness, and standardization. To make matters worse, business and IT speak different languages and sometimes have a hard time understanding each other. The result is that as much as 80% of all data and analysis produced by BI solutions in a typical organization come from what's called "shadow IT", (i.e., users doing BI work on their own, mostly with spreadsheets and other off-the-shelf tools). This may lead to low-quality BI and to the proliferation of "data silos." A related problem is not having "a single version of the truth" which means a business unit may base its decisions on data that is different from that used by another business unit or the entire organization.

Striking the right balance between enterprise-wide access to data and analysis and standardization on one hand, and flexibility, agility, and innovation on the other hand, is key to BI success.

3. YOUR CULTURE IS STOPPING YOU.

More often than not, what technology you buy is far less important than building and maintaining a culture of business intelligence. Having an analytics orientation "baked into the culture" and

insisting on making decisions based on data, facts, and rigorous analysis is where BI succeeds. Transparency, getting to root causes, generating clever hypotheses, understanding the quality and validity of the data used, seeking data to refute—not support—cherished assumptions and conventional wisdom, are all hallmarks of a great business intelligence culture. An important attribute is the ability to strike the right balance between machines and humans, between people's judgments based on experience and expertise and what the data says. It is also important to balance the time it takes to gather comprehensive and accurate data with the need for a timely decision based on the best data available at a given moment.

Culture matters, and it matters more than technology.

"The 'I' of BI—intelligence—can only be achieved by fully engaging the half of human-computer interaction that possesses intelligence: the human half. Business intelligence is only as effective as its ability to support human intelligence."—Stephen Few

4. MOBILE IS MISSING.

Struggling to equip always-on executives with fully functioning BI applications on their smartphones and tablets? Device-related challenges include lack of screen real estate and the limitation imposed by interactive gestures. Security issues, however, are the most important in the marriage of mobility and BI, and they extend beyond the device itself to how you protect data in general and to your BYOD policies. Seamless integration with existing applications, device independence, and giving users the flexibility to customize mobile BI applications for their specific needs are also issues that must be dealt with.

All of these challenges also represent an opportunity to excel in implementation of BI and achieve the desired business impact. For example, the same BI Scorecard survey that found BI adoption as a percentage of employees has remained flat at 22%, also found that companies that have successfully deployed mobile BI show the highest adoption rate at 42% of employees.

A lot of ground left to cover: An average of only 22% of employees in any given organization use BI. Even the "successful" mobile BI deployments show 42% as the highest rate of adoption.

CHAPTER EIGHT

The Future of BI

THE INTERNET OF THINGS.

The business intelligence market is teeming with innovation and new approaches to collecting, analyzing and delivering data. The growing realization by enterprises that high quality data and its efficient and effective analysis are the most important basis for competitive advantage—in all industries and sectors of the economy—will further fuel the explosion of innovative approaches to business intelligence.

In 2020, the world will create 44 trillion gigabytes of new data, a tenfold increase from 4.4 trillion gigabytes in 2013, according to IDC. Between now and then, the Internet of Things will be a major new source of data, presenting a new challenge and opportunity for business intelligence. IDC estimates that the number of "things" connected to the Internet such as sensors, microcontrollers, and wearables will grow from 14 billion in 2013 to 32 billion in 2020.

Many prognosticators predict that in the next few years enterprises will finally "get" social media and how to mine the wealth of data generated by them. But a far more significant challenge—and opportunity—will be presented by the data deluge coming from the Internet of Things. We already see this challenge-that-can-be-turned-into-opportunity in the emerging Internet of Things applications and market segments such as fitness bands, smart thermostats, and connected cars. The failures of some of these early Internet of Things consumer applications show that the key to success is doing business intelligence right:

Check out the infographic,
"Data Never Sleeps."

The data should be collected in non-intrusive ways:

» It should be collected from all relevant sources; the user should be involved in deciding what data is collected and how it is analyzed;

» The analysis should be provided to the user (or another "thing") in a timely fashion;

» The analysis should provide value (e.g., benchmarks, recommendations, predictions), delighting users and encouraging further use.

Check out the infographic,
"How Big Data Could Save the Government $500 Billion."

The Internet of Things promises to improve the internal operations of enterprises everywhere, cutting waste and improving productivity in numerous activities, from inventory management to supply chain logistics to customer relations. Just as with consumers, however, enterprises will not benefit from the true potential of increased connectivity if their workers are overwhelmed by data. If the data collected and transmitted by the Internet of Things is not relevant, if it is not provided in a timely fashion, if the analysis does not suggest ways to improve a work activity or process, then enterprises will not benefit from the Internet of Things. If employees, managers, and senior executives in enterprises big and small, private and public, don't have positive experiences with this abundance of data, they won't take advantage of it.

GREATER DEPENDENCE ON MACHINE LEARNING.

By 2020, we will see many new startups and established BI vendors addressing these issues with new tools and new approaches to collecting, processing, analyzing, and delivering data and insights. Spending a lot of time and resources on cleaning data will be a thing of the past as vendors offer solutions based on artificial intelligence and machine learning. Another set of tools will help organization measure the performance of data, helping them understand better what to keep and what to delete and what type of data yields valuable insights.

To find more and more information "needles" in big data "haystacks," enterprises will increasingly use solutions that provide data in context, based on the analysis of metadata—the who, what, and

Check out the white paper,
"How to Please Your Data Lover."

where of data and its links to other data. This metadata will give enterprises a much more granular view of their operations, allowing them to move from tracking assets to servicing assets to managing assets—even to giving assets more autonomy to manage themselves. The marriage of embedded BI with enhanced data from embedded devices will provide a new view and understanding of the internal workings of the enterprise.

WANTED: MORE DATA SCIENTISTS. A LOT MORE.

The new BI solutions will not, however, alleviate the dearth of analytics professionals, at least in the near future. The increased attention paid to data and its analysis will make the talent shortage over the next few years even more acute than it is today. But by 2020 it will disappear thanks to numerous new data science and business analytics programs and companies developing their own in-house data analysis training programs.

In addition, the democratization of BI will become a reality in most organizations with BI participation rates reaching over 50% of the employee population. As data becomes the key to all business activities and the essence of collaboration, employees will spend more time with dashboards on their mobile devices than they spend using email.

See the infographic:
"The World Needs Data Scientists."

ALL BUSINESS WILL BE DIGITAL.

Last but not least, the digitization of everything will make data—and business intelligence—the business of enterprises in all industries and sectors of the economy. The biggest business trend by 2020 will be the creation of new businesses based on collecting, analyzing, and delivering data. The business that up until now has been left to BI vendors, will become a new source of revenues for companies from health care to agriculture to banks to manufactures to retailers. Virtually any firm in any industry will be able to participate in the data-driven economy and many will do just that by providing new products and services based on their unique business intelligence skills.

By 2020, all businesses will be digital businesses. As ones and zeros consume the world, data will become the new product and business intelligence—finding the needle in a haystack—will be the new process of innovation.


DOMO