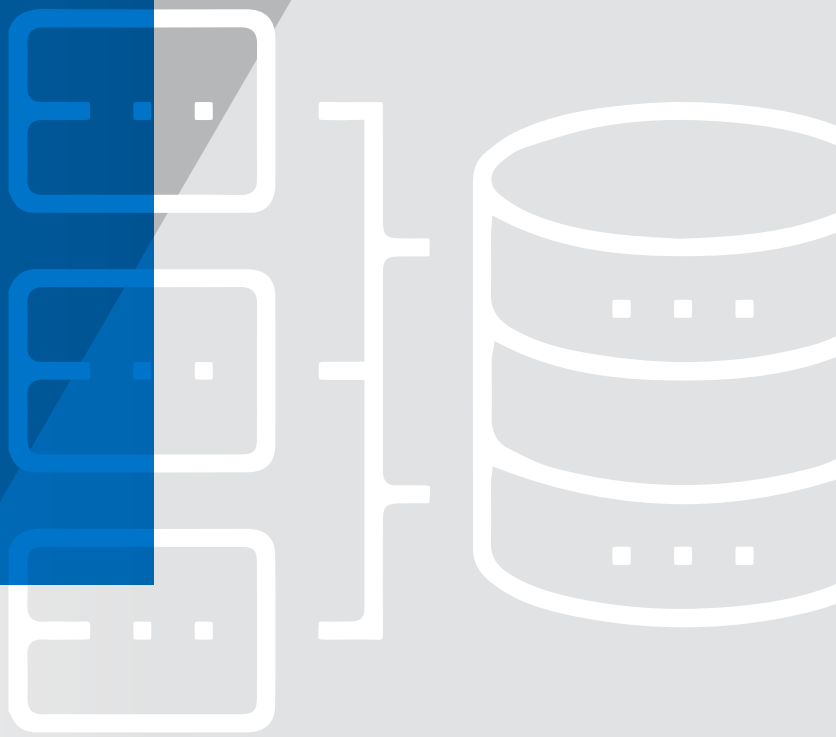
WHITE PAPER

BIG DATA FOR BUSINESS: HOW TO USE IT TO YOUR ADVANTAGE



BY: FRANÇOIS VIENNEAU BINETTE
Designer of Big Data Solutions

NOVIPRO

TABLE OF CONTENTS

Introduction 4

What is a big data project? 5

Why take the plunge into big data? 6

Impact on businesses 8

Higher revenue 8

Lower costs 10

Forget big data unless you've defined your business needs 11

What is a business need? 11

How do you start defining business needs? 12

Identifying business needs 13

Development methods 13

Understanding the business 14

Developing ideas 14

Summarizing needs 16

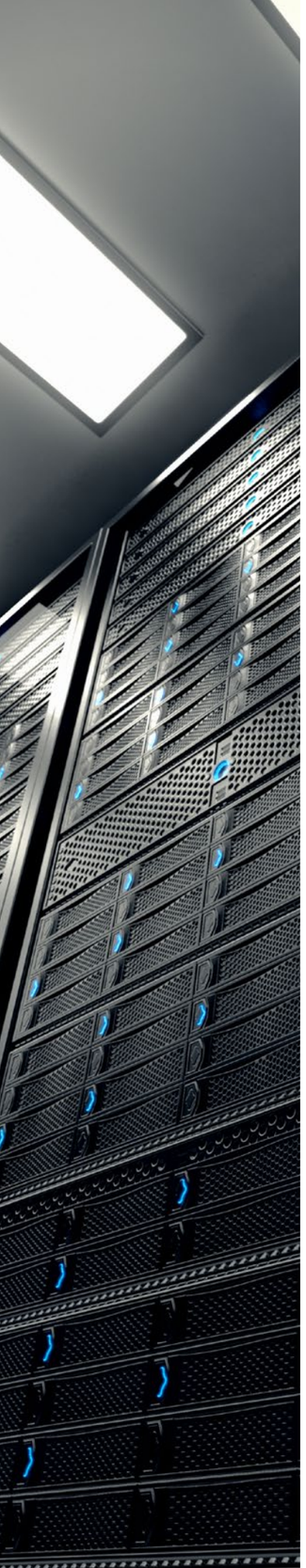
Reference use cases 17

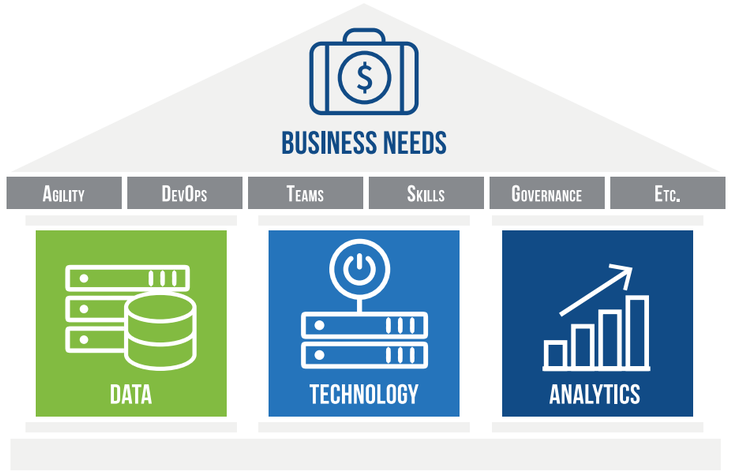
Modernizing architecture 18

Oversight and understanding 19

Optimization and prediction 20

Conclusion 21





INTRODUCTION

Big Data has generated a lot of buzz in the past few years. In fact, some people are convinced it's the Holy Grail for businesses. While that isn't entirely false, big data is an often-misunderstood field. This could be due to its tendency to break down data silos, or more simply because businesses still don't know what to do with so much information. And the confusion has only gotten worse since artificial intelligence, machine learning and the Internet of Things (IoT) have become such hot topics. All these practices are rooted in the depths of big data.

The purpose of this paper is to present a methodology for identifying needs and developing ideas on how big data technologies can be used to respond to them. We will lay out a technique to quickly determine whether these new technologies can be beneficial to companies interested in exploring them. One of the most important aspects of any big data project is the definition of business needs i.e., "What do I want to do with my data?" And yet this aspect is often taken for granted and left undeveloped.

We cannot, however, present a methodology without first providing the context for it. For that reason, we will look at several topics, including why businesses initiate big data projects and what they can expect from them. To better understand the background, it makes sense to review the reference use cases i.e., what businesses are currently doing with big data. And finally, we will broaden the discussion to include business innovation.

WHAT IS A BIG DATA PROJECT?

Before we go any further, let's take a moment to define what a "big data project" is. A big data project or initiative rests on three pillars that are simple yet essential and work together to respond to business needs.

A big data project obviously involves **data**. An abundance of it. But the format and production frequency of this data can be hard for traditional information systems (IS) to control. [I]

There are also **technologies** to capture, store, process and transform significant volumes of data in a way that is effective, efficient and timely. Indeed, there is a wide range of big data technology on the market. The assessment criteria include ease of use, compatibility and integration with other systems, licencing models (open or proprietary), and available support. Big data is often associated with NoSQL technology designed to store semi-structured or unstructured data. [II]

There are also **statistical** and **algorithmic methods** that allow you to convert data into value-added information. Often called "data science" in the big data context, these methods are found under other techniques or appellations, such as machine learning, artificial intelligence, deep learning, data visualization, etc.

Raw data doesn't have much value on its own; but after analyzing and transforming it, data can be used to help with decision-making. [III]

A big data initiative relies on these three pillars to respond to **needs, issues** and **business problems** determined in accordance with the use case.

The key takeaway is that big data projects don't just involve technology, they are broader and encompass data, practices, project organization and management, change management, data governance, etc.

[I] Note concerning data: We often hear about the 3 Vs or 5 Vs of big data, of which the most important are the velocity that data is produced, integrated, transformed and used; the volume of data to be stored and used; and the variety of data. Information from an operations database is structured, while an audio or video clip is unstructured. These varied factors prove complicated for legacy systems used by businesses.

[II] Note concerning NoSQL: Not only SQL databases were designed to store and retrieve data that can't be modelled in tabular relations. There are several NoSQL formats, with the most popular being key-value stores, column stores, document or object databases, and graph databases. The key takeaway is that, in most cases, the right technology exists for each specific use case and data type.

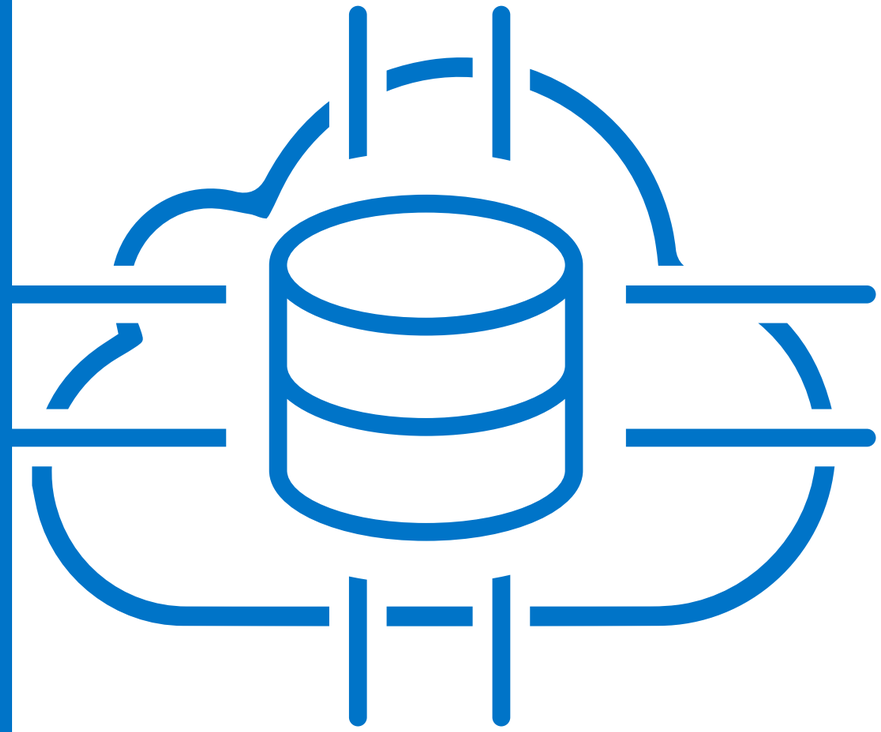
[III] Without going into the technical details, these algorithms generally require the type of computational power that big data technologies can support. Often, these algorithms are old. What's new about them is that calculations can be performed on petabytes of data. Previously, statisticians were limited because only supercomputers could process that much data. Using such large volumes of information is valuable because it generates more precise results. When you take into account all observations of a phenomenon, you are sure to get more exacting results than if you used just a sample of this information. With big data, you don't need to rely on a mere sample.

WHY TAKE THE PLUNGE INTO BIG DATA?

Just because your company produces a lot of data doesn't mean you should suddenly start capturing and storing it. It can be tempting to jump into a big data project. After all, it has become trendy. Everyone is talking about this seemingly innovative topic, and upper management may ask the IT team to kick off a big data initiative out of a desire to appear current. There are lots of reasons. But the truth is that IT projects cost money, and investments need to generate returns. As obvious as this may seem, this rule seems to be overlooked when it comes to headline-grabbing practices such as big data.

However, a big data initiative makes sense in some conditions. Specifically, when the company comes up against technical, operational or strategic roadblocks, such as:

- **An inability to respond to data analysis requests coming from the business units, either because the data is hard to access or because the lead time is too long;**
- **Trouble acquiring, storing or processing certain data produced by the company or generated externally;**
- **Poor performance using current databases and data warehouses;**
- **A change in the competitive landscape and trouble keeping up with industry developments.**



THERE'S SOMETHING I CAN'T DO

"Apparently our logs are interesting to analyze, but because of the format, we can't put them in our SQL databases."

WE CAN'T RESPOND TO NEEDS

"Our divisions want to use machine learning, but we don't have the tools to do it."

SOMETHING ISN'T WORKING PROPERLY

"The IT team wants to produce reports faster, but the processing takes several days."

HOW DO THEY DO IT?

"Our competitor is making more sales than us, but we don't understand how they manage it."

In other words:

These situations occur because the very nature of data has changed over time. Legacy systems were set up to meet specific goals and needs, and they're not suitable for today's challenges.

The impact of these restrictions is often seen as a small problem or even a necessary evil. But in reality, businesses that are constrained by their IT systems lose out on opportunities that could considerably change their operations and financial situation. Especially where data is concerned, the advantage goes to businesses that exploit information. And in competitive markets, there can be serious repercussions for companies that don't take an interest in the subject.

Naturally, big data technologies aren't the miracle answer to all your IT needs, but they can often be the solution to challenges related to data or to the nature of your data.

Start small, think big!

IMPACT ON BUSINESSES

What is the real impact on businesses? In big data projects, the return on investment can be seen in different ways, including higher revenue, lower operating costs, or an improvement in the way some divisions work, which can be hard to quantify.

SALES

Improve operations

Create new products or services

COSTS

Reduce IT costs

Enhance security

HIGHER REVENUE

When we talk about higher revenue, we're referring to the impact of a business decision based on data analysis that generated a positive financial result. It seems straightforward, doesn't it?

Cases of big data leading to increased revenue can be split into two categories:

1. Improved operations

There are two ways that operations can be improved thanks to big data: **decision support,** which is in the same vein as more traditional business intelligence; and **automated decision-making,** which involves direct interactions with clients and requires no human intervention.

Decision support isn't new, but the ability to cross-reference different sources of data and process all the information (whether it's a gigabyte or 100 terabytes) is. There are practically no limits left for data analysis.

Here are some examples:

- Cold or historic data is typically stored outside decision-support systems due to their age and volume. This data can be used in its entirety to refine operational analyses and to better understand the company's evolution over the long term.
- Inventory management can be improved by crossing historic sales data with distribution data, and even meteorological data. We could assume that the number of air conditioners sold is increasing along with the rise in temperatures, but has this really happened? And how can we strike the right balance between being short on stock and overstocked?
- By analyzing the various steps in a manufacturing process, you can understand it better and see where improvements can be made to reduce waste, improve lead times, better synchronize with other operations, etc. The idea is to continually improve while adjusting to circumstances.

Automated decision-making is a useful innovation that makes it easier to respond to customers and provide them with relevant services quickly, easily and with no human intervention.

There are all kinds of examples, but the most common ones are:

- Virtual assistants (chat bots) that can answer customer questions, complete transactions, carry out operations for users, etc. These assistants can do their jobs because they've learned from large volumes of data and because, when customers ask them questions, they perform calculations based on the user's background information.
- Multimedia data recognition is both popular and useful for things such as computer recognition and computer-assisted driving. These techniques require large volumes of data from photos and videos to do things like recognize faces, roads, road signage, etc.

2. New product or service creation

Businesses can analyze internal and external data to help them create or refine products and services that will better meet customer needs and improve the company's chances for success. This is a way for businesses to be proactive and innovative, instead of waiting for the next trend. In addition, it's easier to justify the cost for developing the product or service.

For example:

- We are seeing a new business model in the auto insurance field, in which the premium is calculated based on how frequently the insured uses their vehicle. This wouldn't be possible without the use of decision-making technologies and processes to collect and use massive volumes of data.
- A business can use key attributes of successful products to create new ones. All kinds of data sources must be taken into account (shopping cart, distribution, marketing initiatives, etc.) to understand the influencing factors, enrich the analysis and ensure that the results are statistically accurate.
- Acquiring and processing data from social media can help companies understand what their target audience likes about their products and those of their competitors. This gold mine of data can be used to create new products or services.

Most major companies with an online presence and start-ups base their operations and service offer on data they have processed and analyzed.

LOWER COSTS

As with higher revenue, there are generally two reasons for lower costs:

1. Lower IT costs

This involves reducing operational expenditures, and particularly IT costs. In other words, it's about paying less for better results. Without going into the technical details, this is because several big data technologies are open source and don't require specialized equipment. What this means is that you can replace a tool that requires a paid licence with a free, open-source program.

2. Enhanced security

There are several aspects to enhanced security, but the essence is simple: How can the company avoid losses by proactively analyzing the data generated by its information systems?

For example:

- Analyzing traffic and consumer habits on the company's transactional website to detect the risks of fraud such as identity theft.
- Analyzing logs to detect potential intruders in the company's IT system.
- Extracting weak signals from the company's business data to detect fraud or abnormal activities that could negatively impact the company.

All of these examples—whether they involve higher revenue or lower costs—are situations in which data analysis (and not necessarily big data analysis) can lead to positive returns for the company. This may involve a simple event that, when it happens, triggers an action when the data is created.

When it comes to repercussions, some cases are less obvious than others. But the thing to keep in mind is that when a decision is based on data, it will probably result in a positive consequence that should be identified and repeated over time.

FORGET BIG DATA UNLESS YOU'VE DEFINED YOUR BUSINESS NEEDS

Until now, we've focused on the reasons for starting a big data initiative. It may seem simple, but to be honest this kind of initiative isn't always a sure thing. Every manager wants the resources to boost their company's revenue and cut costs. But how can you do it without getting into the technological details? The only way you can is by defining your company's business needs before doing anything else. In other words, what, specifically, does your company want to achieve by acquiring, storing and processing big data?

WHAT IS A BUSINESS NEED?

In this discussion, the terms "business need," "business requirement" and "business challenge" are all synonyms for "use case." In other words, a big data use case is a series of data processes and events that have at least one specific objective and have positive repercussions for the business, i.e., higher revenue or lower costs.

According to IBM, a use case helps a business solve a specific business challenge by using patterns or examples of proven technological solutions. This is why they should be properly defined and documented to reveal any individual or group concern that is irrelevant to the success of the project. In other words, you should make sure your initiative is truly profitable.

Businesses usually have some challenges to meet, problems to solve and sub-par processes to improve. These challenges often form the business needs on which you should base your initiative. The question you need to ask is: how can I respond to these needs and use cases with the right technology? The concept isn't new, but it's still appropriate in a big data context: "Tell me what you want to do, and I'll tell you what you need to do it."

The crucial thing is to think before you act. Don't start with the technology. As a matter of fact, ever since businesses computerized, their challenge has been to connect technology with their business needs and processes. Big data is a hot topic, and because of this and many other reasons, when big data technology is discussed, managers often look at the issue backwards: "What am I going to do with my newly installed big data technology?"

It's been proven many times that this kind of approach doesn't give the results you need because it's often the case that the needs haven't been properly defined (so you don't know whether the investment will provide enough of a return). Even if you get around to identifying your needs after the technology has been installed, it may not even have what it takes to meet them. In short, your big data initiative must start with your business needs, and not with the technology. Whether or not a business need has already been defined, it is easily developed and refined with the appropriate support, expertise and methods.

Here's an example of the value chain for a manufacturer:

HOW DO YOU START DEFINING BUSINESS NEEDS?

When you start looking into big data, what you want more than anything else is to positively impact the company's main business units through data analysis. As mentioned, you also want to have a positive effect on its IT, even though IT is usually viewed as a support division. Don't limit yourself. If there's a way to use the data and profit from it, it doesn't matter where the idea comes from. Support divisions often have data analysis initiatives that help the company cut costs, while in the main business units, big data analysis often results in potentially higher revenue. However, these are not mutually exclusive.

The goal is to monitor and analyze any data-based (or non-data-based) business processes and see how you could improve decision-making or even automate it and make it more effective. In other words, "What decisions are being made in my company and how can I improve decision-making using hard data?"

In the end, the business need doesn't have to be perfectly clear. All you need is an idea, a big or small issue that could be improved. Once you've got that, everything falls into place.

Here are some examples of business processes that could be improved through data analysis:

- A business conducts maintenance on its trucks at 200,000 kilometres. Would analyzing the logbooks and repair records help prevent breakdowns and improve vehicle maintenance?
- A company contacts clients who make more than 100 purchases a year, offering them special deals in an attempt to boost sales. Would analyzing the transactions of all its clients help it target potential buyers that it may not have identified from their purchasing volume?
- A department store's parent company wants to learn more about its clients' socioeconomic characteristics and better understand buying trends, both in-store and online, so that it can improve its marketing campaign.

IDENTIFYING BUSINESS NEEDS

As you already know, use cases or business needs form the foundation of any big data initiative. Our experience shows that just because a business has installed big data technologies doesn't mean that it will properly use the data it collects. It's an old story: technology that holds a lot of promise for a company is ultimately not always used. You've seen it before, haven't you? An IT project has to offer a certain rate of return over the medium or long term, otherwise it's a net loss for the business. Big data projects are no exception to the rule.

So, what does a business need, before even getting into the technology, to confidently (and successfully) launch a big data initiative and never look back?

What is important is defining your own use cases and the needs that are specific to your company (or business unit), and evaluating any potential return. A complementary approach is to simply adopt reference use cases that have already been proven. However, this is a less attractive option because people are more likely to resist a change if they are forced to adopt something that they haven't necessarily chosen or developed. We will look at these reference cases in the next chapter.

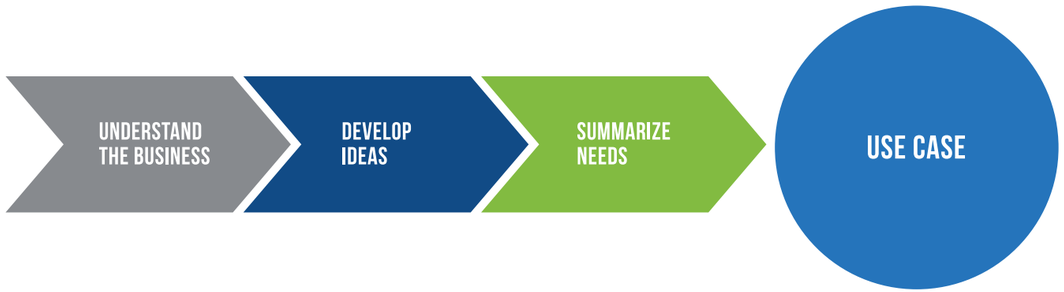
DEVELOPMENT METHODS

Determining which steps to follow isn't easy. The best way to do it is to use a structured, accessible and non-restrictive approach that leaves room for creativity and innovation and allows ideas to be developed.

Most of the steps involved in developing a big data project aren't new: they are rather conventional planning and organization processes. Something that is new, but isn't always respected, is really taking the time to develop the right use cases.

NOVIPRO offers a simple, proven methodology for developing use cases with the support of a large number of players using a structured method. This methodology is accessible to all, but we recommend choosing one person to lead the initiative, and who will be responsible for the whole thing. This leader must be a big data expert and know how to lead interviews and group debates.

The methodology has three major stages: understanding the business, developing ideas and summarizing needs.



UNDERSTANDING THE BUSINESS

First, you must study the company and the business environment in which it operates, at least to a certain extent. It's important to properly understand the company's various internal and external environments, just like you would if you were defining its business strategy.

	STEPS	DESCRIPTIONS
1	**Understand the business (and support) processes**	Study processes that add value for the business, along with the data used (or not used) by these processes. Over the short term, these are easily identifiable processes that could benefit from a big data initiative.
2	**Take an inventory of data produced**	It is important to be aware of all the data produced by the company, the context in which it is produced, who uses it, what are the rules for managing it, etc. The point is to have a good comprehensive view to evaluate the company's strengths and weaknesses with respect to data access.
3	**Identify challenges in terms of internal and external operations**	The idea is to get a basic idea of the company's problems. Don't limit yourself to processes that use data. Instead, look at all the bottlenecks and challenges because they may benefit from data analysis.
4	**Collect some preliminary improvement ideas**	Through interviews with different stakeholders, you can quickly home in on ideas for data analysis. At this point, you're focusing on the quantity rather than the quality of improvement ideas. Once again, they aren't necessarily based on data usage.

Whether you're talking about processes, data, challenges or ideas, anyone can participate and express their opinion on these subjects. Anyone, from the head of the company to a warehouse employee, has the potential to solve problems. This is why interviews can and should be conducted as needed with any stakeholder.

In short, it's important not to limit your examination of the business and to avoid any preconceived notions. Furthermore, the size of the company may be a factor. For example, in a SME, it may be realistic to look at all business units, while in a big business, the only viable method would be to start with a very specific business unit.

DEVELOPING IDEAS

This means putting the company's various stakeholders on an equal footing and then asking them to think about problems, challenges, solutions or needs that they've encountered in their day-to-day work.

Experience has shown that people don't necessarily lack for inspiration when developing big data use cases. However, if they aren't aware of the full scope of possibilities offered by big data, it is harder for them to present ideas on the subject. When expressing their needs, the various business units tend to limit themselves by presenting ideas that match what they believe to be possible, without being aware of the full range of technological possibilities.

This session is therefore an excellent opportunity to present what's being done in the industry with respect to big data use cases. This should always be done with the goal of opening people's minds and helping them better reflect on what would work for their own business.

	STEPS	DESCRIPTIONS
5	Educate stakeholders (and even more people) about the **state of the art** and **big data concepts**	Whether we're talking about IT, business units, support divisions, management or analysts, people may not necessarily know as much as they should about big data. So it's absolutely essential to provide basic training for all stakeholders to foster development and innovation with respect to business needs. As we already mentioned, business units tend to limit themselves when expressing their needs. Exposing them to the reality of big data, including the possibilities offered by its three pillars, will help spark their imaginations. These educational sessions are also an excellent opportunity to update the various observations and discoveries made in the first stage. You're basically equipping participants with all the tools they need to properly prepare for the next step.
6	Lead **brainstorming** sessions to express business needs	In practice, this means forming a variety of groups and discussing ideas on improving (or even creating) business processes through data analysis, decision support, automation, etc. No one is better able to reinvent their job than the experts in the field who must themselves adopt new ways of doing business. Everyone can be invited to this kind of activity. Naturally, some people will probably have better ideas than others, but you never know who in the company may have thought of a solution, a process that could be improved, or even uses data that could assist them in their work. The idea is to not limit yourself. A well-organized and directed brainstorming session whose participants possess authentic expertise in their areas of business and have been educated about big data concepts (step 5) could very well produce use cases and business needs that have real added value.

We are suggesting a brainstorming activity, but many other techniques for developing ideas can be implemented, such as design thinking, storyboarding, reverse thinking, etc. There are no good or bad ways to develop ideas. All you have to do is bring the stakeholders together and take the time to do it well.

There are many benefits to doing this exercise one or more times with various mixed-stakeholder groups. First, the company won't end up doing what everyone else does just because it's what other companies do. The ideas and needs you come up with will be based on your company's circumstances. Second, if people identify their needs themselves and claim ownership of them when implementing new technology, the risks of failure due to resistance to change are reduced because new methods are not being imposed on people. The ideas come from the users.

The second stage could take a long time, meaning it often lasts several weeks because you have to bring the stakeholders together. But it is still essential. Companies that move too quickly and adopt technology before thoroughly reviewing the issue and needs run the risk of its users refusing to implement the technology.

SUMMARIZING NEEDS

This involves consolidating all information collected and translating it into business needs or practical, achievable use cases.

	STEPS	DESCRIPTIONS
7	**Collect the various needs and ideas expressed and transform them into practical use cases**	In concrete terms, what has been expressed, what does that imply and how can you respond using the three big data pillars? Is there any underlying data? How can you respond to this need?
8	**Evaluate the potential return in these use cases**	A use case must have a positive financial impact. Can you evaluate it before taking the plunge? If so, is the financial contribution it makes big enough to be at least one of the reasons to invest in new technology?
9	**Select some promising use cases**	In general, a variety of use cases could spur you to invest in big data technology. These use cases offer sufficient value for the business units in terms of decision support, innovation, improved operations, etc.
10	**Optional: share use cases**	Once you've decided on your use cases, it could be worthwhile, albeit somewhat redundant, to discuss the selected cases with the people involved in the development process to both invite further clarification regarding the need and ensure it has been properly understood.

From this point forward, the company can evaluate whether launching a big data initiative offers any financial benefit, as clear, achievable needs have been identified and, if they are met, could boost the company's performance. The solutions to these needs may not necessarily mean using big data technologies or capturing additional data, but in any case this exercise will put the company on the right path.

The third stage may be completed within a reasonable timeframe, based on the scope of the reconciliation work.

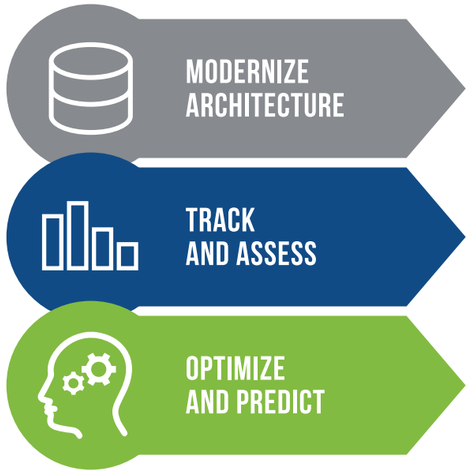
REFERENCE USE CASES

Up to now, we've shown that big data technology must basically meet a well-defined business need.

We explored a relatively simple and structured methodology to identify these business needs. But what about general industry practices? Are there cases that are so frequent that a company cannot ignore them, even if it initially hadn't identified them as part of its needs? The answer is yes.

These cases have already proven their worth in the industry. The developers of big data technologies quickly recognized them, first for their communications and marketing needs, but also because the companies that adopted them saw a real return on their investment.

It is hard to fit all big data projects into very specific categories, but we have identified three that might overlap but ultimately help us properly understand the breadth of these cases.



The following tables group the most common use cases in companies that launched big data initiatives. Of course, the simple fact that a competitor has implemented a use case does not mean it is right for the company, but it is a good indicator of what should appeal to some business units.

MODERNIZING ARCHITECTURE

USE CASE	BUSINESS NEED	↑$	↓$
Data warehouse optimization	Access all historical data from several sources to produce more complete analyses.	×	
Active archiving	Access all historical data from several sources to enhance analyses. Data warehouses rarely hold more than a few years of data, which makes historical analyses more difficult.	×	
Built-in data loading	Replace or improve tools for loading/extracting data (ETL) through better performance and the ability to manage data volume, variety and speed.	×	
Disparate data collection	The ability to acquire information from inside or outside the company, regardless of its form: data from the Internet of Things (IoT), logs, public data, data in audio, video or text format, etc.	×	
Reduction of operating costs and total cost of ownership	Reduce fees, such as licencing fees for a broad range of technologies: databases (DBMS), data collection (ETL) tools, statistical data processing tools, data visualization tools, connectors and more. Established statistical software developers are often replaced by more effective open technologies.		×
Mainframe load reduction	Reduce the amount of historical data stored in your mainframe computer and make it easier for analysts to access data.		×
Indexing and research	Allows you to search all words in a variety of file formats: CSV, PDF, Word, etc.	×	
Data lake (data services)	Break open the company's various data silos by storing their data in a single place. This makes it possible to perform enhanced data analysis.	×	

OVERSIGHT AND UNDERSTANDING

USE CASE	BUSINESS NEED	↑$	↓$
Activity report	Improve business intelligence through the ability to process a nearly unlimited amount of data faster and more efficiently.	×	
Discovery and exploration	Analyze raw or processed data without preconceptions to discover trends and make the most of the data. This is hard to do with traditional data warehouses or data marts, which are built using specific rules, as they provide little scope for exploration.	×	
Operations analysis	Examine operations once again to maximize results and cut costs.	×	
Data enhancement	Improve analysis with data joins.	×	×
Disparate data analysis	Helps you take advantage of data that isn't always analyzed such as logs, telephone calls (audio), textual data, sensor data (IoT), etc.	×	
Social mapping	Makes it possible to analyze the relationships between data. Graph-type analyses are often associated with graph databases, helping you understand the interactions and relationships between participants, such as individuals.	×	
Payment monitoring	Speed up analysis of sales channels to prepare for and possibly reduce cases of fraud, whether or not they recur.	×	
Shopping cart analysis	Be capable of analyzing all shopping carts so you can identify the relationships between purchased products, thereby helping with cross-selling.	×	
Journey mapping	Track customer journeys throughout all of the company's points of contact: website, branches, call centre, etc. This will help you understand and improve processes.	×	

OPTIMIZATION AND PREDICTION

USE CASE	BUSINESS NEED	↑ $	↓ $
Single view (360° view)	Improve tracking and engagement by collecting all available data (including scorecards) on a client or supplier in a single place. This helps you better understand them, fosters strong relationships and enhances decision-making.	×	×
Customer service	Improve customer service through automation: virtual assistants (chat bots), artificial intelligence, sensor data (IoT), etc.	×	×
Sentiment analysis	Find out more about the company's reputation by capturing and analyzing both internal (customer service) and external (social network) data, usually through text mining.	×	
Customer retention	Be able to identify whether a customer is likely to leave the company and go to a competitor by analyzing all of their behaviour.	×	
Segmentation	Better understand and target the company's client categories to tweak marketing strategies and enhance sales, loyalty and satisfaction.	×	
Recommendation engine	Offer customers products that they are likely to buy based on their purchasing history.	×	
Product and service design	Use data analysis to design a new product or service that successfully targets customers. In other words, can we add value to our products and services through data analysis?	×	
Fraud/weak signal detection	Use specialized algorithms to detect activities or transactions that are unusual but not necessarily flagrantly so. The ultimate goal is to reduce fraud by detecting new fraud methods.	×	×
Predictive maintenance	Predict breakdowns or accidents before they happen: for devices, vehicles, appliances, buildings, etc.	×	×

It is almost impossible to categorize all use cases. Sometimes you have to force them to fit. So, when you look carefully, it's no surprise to see some repetitions and certain cases that could very well belong in another category. Despite this, the goal is to show that most cases on the market ultimately have similar objectives.

Furthermore, some of these cases apply to specific business units, but others are broader and apply to any business that produces and manages data and wants to make use of it, depending on the maturity of the business. You have to learn how to walk before you can run. Even though some of these cases may be highly interesting, they are not always within the reach of all businesses (or budgets). Some more complicated cases require a company to have a certain amount of experience with these kinds of projects. This is partly because of how complex they are to implement, but also because of technical complexity, IT equipment, acquisition of data through silos, etc. Feasibility must be assessed within the context of the business. [v]

[v]Data exploration: This use case is not a true example. When a big data initiative matures and the investments in technology and talent have proven beneficial for the company, some businesses seize the opportunity to explore the varied data available to them, just for the sake of it. Thus, without any prior indication that the data or data processing is relevant, the company will ask its analysts to explore the data at their disposal. The idea is to detect trends, patterns, correlations, etc. There's always a chance that within that data, there's a nugget of information that could help management make an important decision or find a competitive advantage that hadn't occurred to them before.

CONCLUSION

This white paper was intended to present readers with a suitable method for launching a big data initiative. Big data investments must provide a return, whether by increasing revenue or cutting operating costs. This return results from responding to at least one business need that the company will be able to meet using big data technology.

Obviously, there are many other aspects that must be considered for a big data project. For example: How do you complete a big data proof-of-concept? What skills are needed? Do you need a data scientist? How do you organize and manage a big data team, along with agility, data exploration, data governance, communications, resistance to change, the appropriate technological layers, and so on?

All of these questions come next. They must be addressed once the company has defined its needs and use cases. Business analysts have spent decades saying that technology must serve clearly defined business needs, and it is no less true today, no matter how useful big data is supposed to be.

Finally, taking the plunge, or at least starting to examine whether it's worth taking, is not a complicated issue. It's a development process like any other. However, support from experts in the field at all stages of project development must be integrated into the plan so that you can predict the time and investment needed, quickly identify potential benefits and maximize returns.

FRANÇOIS VIENNEAU BINETTE

DESIGNER OF BIG DATA SOLUTIONS

With several years of experience on big data projects in Europe, François specializes in technical and functional support for the big data value chain; from the strategic vision of an initiative to the installation and management of a Hadoop implementation project. Addressing both business line managers and IT executives, François aims to consolidate big data efforts and make them profitable in order to promote the return on investment long awaited by open-source technologies. He is also a certified administrator and developer for the Apache Hadoop ecosystem.

NOVIPRO

www.NOVIPRO.com | www.e-spacesystem.com | france.e-shs.com

1-866-726-5353 | 1-844-377-7978

DELL EMC PARTNER PLATINUM