



This work is licensed under the Creative Commons Attribution 4.0 International License.

To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.



You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:



Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.



Table of Contents

Overview	4
Buzzwords are everywhere	6
Sensing and understanding the real world	6
From capital assets to services	7
Automating decisions	8
Data Value Framework	9
Understanding work required to get value from data	9
Always start with value	9
From data to value, the core process	10
Different value, work and technology at various speeds	13
Scaling for value	14
The job is never done, improvement is essential	15
Learn first, deliver along the way	18
Miss-match between what we invest in and what the work is	18
Save yourself much money by designing great experiments	20
What you are piloting is a new way of thinking and working	21
Technology will work, if not now then soon	22
It is not that simple, data investments are complex	23
Software is everything	23
Difference between solutions and solution building blocks	24
Data security, if you have not been hacked, you will be	26
How to establish a healthy portfolio of data investments	27
Understanding the full solution lifecycle	28
Managing the portfolio and letting individual initiatives learn	29
Platforms enable scale for speed and profit	30
A few simple tools to help you	32
Investing in Data Overview Poster	32
Data Investment Canvas	33
Pilot Assessment Tool	33
Portfolio Visualisation Tool	34
References and Further Reading	35



Overview

At some point over the last few months, you would have been in a meeting discussing technology and its impact on your business. If you have not, you should have. It does not matter what kind of business you are in; technology has always re-defined industries, markets and our lives in general. The 21st century is well underway, and the impact of technology on everything is not only stronger but accelerating. What we knew to be impossible two years ago is now possible, and strategies we were sure will work, no longer work. There is much talk about automation, new mobile-driven services, the predictive capability of technology and everything in between. We see whole industries re-shaped by technology. To thrive in this fast-changing, incredibly exciting future, we need to not only change what we buy or do but how we think. If you are thinking and talking about responding to the disruption technology can have on your business you are already behind. Others are looking for technology to grow and differentiate their activities, not merely survive. It is time to embrace technology and re-imagine your business.

At the core of this technology revolution is data. It is the currency, the core asset and the unit of value that makes all modern technology magic work. Automated cars, automated customer insights, predictive machine maintenance and all other technology advances only work because there is data to power them. Investing in data is one of the most important investments you will make.

Many organisations are already investing in data and responding to the incredible amount of buzz and hype

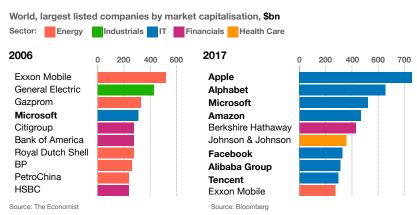


Figure 1 - Technology Transforming Markets

generated by new technology. These investments are mostly not living up to the hype, not because technology does not work, but that most organisations do not have the culture, structure or know-how to get the most out of technology. Figure 1. shows the change technology has brought to our world in just the last decade. IT companies are now dominating the markets as world's most valuable companies.

These companies invest heavily in data and

how to monetise it.

Investing in data is not just about buying and implementing technology. Throughout this paper, we will explore concepts critical to achieving value from data. This whitepaper is not a general collection of leader interviews and data about their opinions and projects. We will focus instead on foundational concepts that will enable you to understand how data investments work. The content and language herein is not technical and is designed to lead you through a structured way of thinking about data investments.

If you are investing in Analytics, Big Data, IoT, Data Platforms, Business Intelligence or any other data related technology, please read on. All of these techniques and trends are related, powerful and magical, but complex and hard to wield.



In this whitepaper you will learn that:

- Buzzwords that dominate conversations about investing in data are not the full story, but describe underlying
 patterns that are transforming our world.
- Investing in data is more than implementing technology or doing analytics. We will explore a Data Value Framework that shows how to design data investments focused on value.
- Critical learning approach as most organisations established ways of thinking and working are not going to
 achieve the breakthrough results data investments can achieve. Furthermore, using traditional project
 approaches to learn new skills, methods and technology will merely waste money without much result.
- Data investments are complex and that most organisations neglect this complexity to the detriment of their investments and value.
- Managing the data investment portfolio requires an understanding of the entire data solution lifecycle and is more than just a collection of projects.
- A few simple tools that can help you design your data investment strategy, pilots that help you learn and manage a portfolio of investments that transform your business.

If you are interested in understanding how to get value out of data, this whitepaper is for you. The language used in this paper is not a technology language on purpose. As you will learn the challenges and opportunities inherent in data investments are made possible by technology but require much more than technology to achieve results we want.

As with anything ActionTwelve, we are specifically targeting the following two groups of people:

- Leaders and managers governing, leading and managing data investments.
- Consultants who advise the leaders and managers.

If you are looking for quick answers, this document is not for you. If however, you want to deeply engage with how to invest in data and transform your organisation, read on.



organisations like yours.



Buzzwords are everywhere

"You shouldn't do things differently just because they're different. They need to be... better." - Elon Musk

Every company worth its stock price is talking, investing, building or buying Big Data, Predictive Analytics, Machine Learning based systems. Add Internet of Things, Edge Computing and Cloud to your investment portfolio, and you are left with huge budgets, failed projects and very little understanding why your investments are not getting the results. Technology literature is full of buzzwords and every year the repertoire increases, both to signal new and to re-brand older technologies. The trick with buzzwords is to separate them from technology trends. Buzzwords may lack concrete evidence, are typically vague and usually a product of marketing, but they contain seeds of truth. These grains of truth are technology trends. Technology trends are useful in understanding how your own business can transform and improve, given new capabilities made possible by technology. Let us explore these data related technology trends and how these can impact

Sensing and understanding the real world

Much of the content describing Big Data, Internet of Things (IoT), Digital Twins, Data Management and Analytics revolves around sensing and understanding the real world. It does not matter if you are trying to understand your customers better or your equipment, the modern technology available can cater for a vast diversity of use cases. The fact these techniques are generic and can be applied to a such a diverse set of challenges makes them unbelievable and risky at the same time. To better understand the implications, let us explore how organisations today sense and understand the world around them. Businesses sense their environment in the following ways:

- **People** most of the sensing within institutions is done by people, by observing, experiencing and interacting with customers, machines and one another. Capturing this data is laborious and time-consuming.
- **Software** the majority of work within organisations is driven by software. Whether it is your ERP, CRM, sales, procurement or management systems, many transactions are operated based on software-driven sensing.
- Hardware industrial and manufacturing organisations have a range of physical sensors that understand machines, physical processes or environmental conditions.
- Combination of all three where hardware provides a sensor signal that is interpreted by software and finally
 understood by people.

New technical capabilities allow for collection and contextualisation of data about people (our employees, partners and customers) and physical objects (machines, devices and the environment). The real trend we are experiencing is the technology-driven ability to collect data

What could we know about our reality we don't currently know?

about the world around us that has potential to drive new sources of value or optimise existing value. The nirvana is collecting data about our reality, making sense of it and making decisions that transform our profit and improve our impact on the society. Our critical investment is in data, as a core unit of value, not technology. Investing in data is about making our organisation more effective. Investing in technology is about efficiency.



Asking what IOT or Big Data technology should we invest in is not the most relevant question businesses should ask, but instead, the focus should be on one of the following:

- What could we know about our customers, processes or equipment so we may provide better or new products and services?
- What do we know within one part of our organisation that would be valuable to other parts that makes us aware of our value chains holistically, so we are optimising the results, not parts?
- How do we collect the data required to "codify" our knowledge so we can transform through automation?

Big Data, IoT and related technologies are spread across the maturity spectrum. Some are ready for out-of-the-box prime time, while others require deep technical expertise to wield their full power. While much fan-fare goes into the

development of technology and capability maturity models and assessments, the focus should be on your business maturity. The way you do business will probably change much more slowly than how fast technology matures. Being keenly aware of how you run your business, how could you run your business (#1) if you could and how will you run your business in the future is the best way to get most out of your data



investments. Data related technologies are numerous, at different stages of maturity and continuously changing. There are no "right technologies", the primary challenge is profoundly understanding how you run your business and how you can incorporate technology to drive transformation. The incredible promise of new data technologies is the ability to know more than ever about your customers and your business. What you would do with that knowledge is more important and how you get it.

From capital assets to services

Continuing our exploration of technology buzzwords and accurate technology trends, we cannot go past Uberification of everything. Uber has been called the most significant taxi company in the world, just like Airbnb is the largest accommodation company in the world, and so on. There are Uber of X companies springing up all over the world. The problem is that business models within these companies are still far from a stable profit design (#2). However, as with every buzzword, there is a kernel of a transformational technology fuelled trend. The trend is away from offering and consuming capital assets to providing services. Uber and Airbnb are platform businesses, that deal with connecting consumers and producers. Uber is not a taxi company; it connects customers with taxi companies (individual drivers who run their businesses). Classical enterprises, especially OEMs and product companies are noticing this trend towards services. For example, many

people are not buying cars, but are opting for on-demand services such as Uber. Tesla, one of the most innovative manufacturers today, is moving their strategy to include service offerings (#3), such as offering their customers ability to make money when their cars are not in use. The reason for this trend is that selling products creates an interaction between companies and customers at the point of transaction, the

How can we increase number of touchpoint with out customers and introduce better or new

services?

purchase. Creating more interactions with customers provides additional opportunity to offer value and therefore generate new offerings to those customers. Historically, service and product businesses have been viewed as different (#4). What we are witnessing is the merging of these business types. Services are entirely reliant on access to data, about the machines, customers, processes and everything in between.



So how should we think about this move from capital assets to services regarding our investment in data? Firstly, the concepts are different if you are exploring transforming your organisation towards offering services, or if you are a consumer of those services.? If you are an OEM or a product company, collecting data about your products while in

use will enable you to improve those products, offer valuable services to your customers like insight into the performance of the products you sell and expand your revenue streams towards service-based subscription models (#5). For customers, outsourcing monitoring and analysis of machines make it easy to get the benefits of new technology without having to invest in talent acquisition strategies and cultural transformation.



Are we retaining access to data, even if we are outsourcing to service providers.

However, customers of data-driven services should be aware that consuming these services is still outsourcing. Much focus should be placed on signing contracts where the customer always retains access to data themselves. Without this access, there is a high probability of significant value leakage and inability to implement data investment use cases beyond those outlined in the services agreement (#6). Besides, simple access to data is not enough; data must be accessible at different times depending on use case. For example, having access to sensor readings as a once-a-month download is of no use for automation, even as the data is technically accessible.

Automating decisions

The final trend we are going to explore is automation. We are all familiar with robots in factories, and incredible improvements automation has achieved in manufacturing. Automation delivered by software that runs our

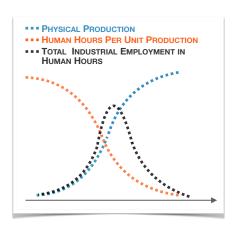


Figure 2 - Technology Transformation Model

businesses and economies has provided even more significant transformation, both in capability and efficiency. Automation is about transforming productivity by changing work away from being executed by humans to being run by machines and software. Figure 2 illustrates what typically happens in all industries. The production initially takes many human hours to achieve, but as technology enters a sector the physical output goes up, but so does the employment. Eventually, technology keeps increasing production while declining human employment (#7). Historically we have invented new work for people to do. People employed by farming and displaced by tractors got jobs in manufacturing. Once manufacturing started automating, people got jobs delivering services instead of making products. However, what is next? We are now not only automating

physical, dangerous or repetitive work but decisions made by humans. Autonomous vehicles, de-centralised renewable power grids, 3D printing, Internet of Things, name the buzzword, they are all driving towards collecting incredible amounts of data that is now being used to substitute not only human physical labour but decisions as well. The centre of this automation revolution is data. Without it, none of the magic is possible.

To close off this section on buzzwords, remember that technology trends are not just describing more gadgets and better entertainment, but a fundamental shift in how our economy and society operates (#8). The core technology trends we see today are focused on sensing and understanding the world around us, providing new and cheaper ondemand services, and not just automating physical work but decisions we make. These are by no means the only technology trends shaping our world, but a few related to data. Data is the lifeblood of the digital economy. Investing in data is not just a good business, it is a survival strategy.



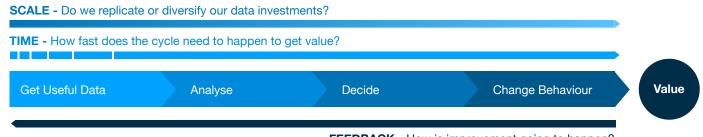
Data Value Framework

"He who loves practice without theory is like the sailor who boards a ship without a rudder and compass and never knows where he may cast." - Leonardo da Vinci

To further understand how to invest in data effectively, we need to understand the process by which we use data and get value. The framework described in this section is, of course, an over-simplification, but is useful in understanding broadly where our effort and investment needs to go to get maximum value.

Understanding work required to get value from data

The Data Value Framework depicted below, shows a central Data Value Chain and three critical considerations of how that value chain should be designed. The value chain consists of getting useful data, analysing it for understanding, deciding on actions to take and finally changing some behaviour given this new knowledge. How this value chain is designed, including technologies used, depends on how fast the value chain needs to run (Time), how many times it occurs (Scale) and how often does it need to change (Feedback).



FEEDBACK - How is improvement going to happen?

Figure 3 - Data Value Framework

Investing in data, without holistically understanding each of the above concepts is the number one reason why data investments fail or don't live up to their expectations.

Always start with value

Many data investments begin with data, with the idea that if we look at data long and hard enough, we will develop insights. Starting with data is a costly investment, one that makes us ask the wrong questions and formulate non-viable solutions. Our data investments must always begin with a clear objective and value-based questions. For example, collecting all the data for a piece of equipment might sound like a good idea, but it often leads to complex solutions, an overwhelming amount of data and unsuccessful implementations. The main reason for starting with the value first is the following:

- A specific set of data about a machine, or a customer, is rarely useful for valuable investments. This data
 must be integrated with other data to be useful. For example, getting sensor data for a truck without
 understanding what the truck is doing from the business perspective will make decisions hard to make.
- Depending on the value we are chasing we will need different data at different times. If we start with Getting
 Useful Data, we will continuously have to re-work our technical implementations. We need various data at a
 different speed to make decisions at different timeframes. Deciding what to do within next few minutes is
 very different from choosing what to do in the next month, especially from a technology point of view.

realistic and achievable.



• It is very rare that people making decisions are the same people implementing the technical and business change. Starting with value ensures we have a strong hypothesis about what it takes to make our investment work. A strong hypothesis can only be achieved by engaging stakeholders across the Data Value Chain.

Without starting with the value first, even the simplest of questions are hard to answer. What data should we collect, retain and use, where is the data going to come from, do we have the data already and so on? Every piece of data is useful for sure, but you are in general not going to need all of it. Start with particular questions like "How can we increase customer conversion from 15% to 20%?" or "How can we increase the time between maintenance outages for our laser cutters by 30%?". Starting with these types of questions anchors

our data investment in real value and allows us to make sound technical and change decisions. The most important thing we need to understand is "What does it take to get the value?". Starting with a strong business value hypothesis allows us to explore what we need to do, so our investments are

Is our investment anchored in a clear business value hypothesis?

Before we move on, there is a second type of business value we should mention, one that answers questions like "How can we reduce our cost of running data systems by 20%". The difference between this objective and others we have discussed is a difference between effectiveness and efficiency. When it comes to data, you should focus on pure business value cases first to get experience with investing in data (effectiveness). Only then should we focus on optimising that investment (efficiency). Good news is many companies have already invested in data and are potentially ready to invest in "Data Platform" initiatives, but you should not start there. We will discuss these types of actions later when we look at reusability across use cases.

From data to value, the core process

The core process (our Data Value Chain) mentioned below is deceptively simple. However, each step in the process has some critical considerations, when not taken into account can seriously undermine our investments.



Figure 4 - Data Value Chain (core process)

Get Useful Data Phase

Let us start with **getting useful data**. Firstly, not all data is equal, or valuable. Collecting data is expensive, not just concerning storage, but implementation, maintenance and infrastructure to keep the systems running. Data is also typically generated and managed outside our "data systems". There are several key considerations for ensuring we get the useful data to realise our investment.

Data ownership and access are possibly the most critical and misunderstood element in getting access to useful data. Much of the data originates within systems and solutions that were implemented before our data investment use cases were ever considered. The legal agreements and licensing contracts that govern the use of these systems in many cases preclude extraction of data out of those systems for analysis or automation. Most people are used to exporting data using spreadsheets and then applying analysis, but that is not ideal when automating data extraction. ERP systems (e.g. SAP) are especially prone to this problem (#9). Equipment manufacturers (OEMs) also



limit access to data their machines produce (#10). The reason for this data protectionism is simple; data is the lifeblood of all "digital transformation" endeavours. Just like you are reading this paper to invest in transforming your

?

Do the contracts and licence agreements allow us access to data?

organisation, software companies and equipment manufacturers are doing the same. Their business plans are to sell you insight thought services, not just software licences or equipment. Move from buying assets to buying services is one of the core technology trends we discussed earlier. These strategies can not work without access to data.

Data ownership and access considerations are mostly concerned with your relationship with third parties, but there is a critical internal consideration as well. Many organisations do not have **internal data contracts** between different parts of their team (#11). Having internal agreements may sound strange, but is essential in getting our data investments to work. In many cases, people responsible for producing data and people using data for analysis and decision making work in separate parts of the organisation. For example, an insurance company may be structured down different insurance product lines, like health, commercial or life insurance. These departments are responsible for serving customers who use those products. The data they collect about the customers is critical for the risk management department, even if the product line departments do not use the data directly. Likewise, an engineering department in an industrial organisation is responsible for designing and building control and monitoring

systems, while maintenance department uses data from those systems to optimise maintenance schedules. If people do not use the data, they will see little value in maintaining and improving systems that generate the data, unless objectives and KPIs are aligned across the organisation. To deal with multidisciplinary nature of modern business, some organisations are moving away from "functional" organisational models and are

? Are objectives and KPIs between data producers and consumers aligned?

implementing "product" based models (#12), where all the expertise for delivery of a product or service is in one place. This type of restructuring is outside the scope of this whitepaper, but the problem of aligning the producers and consumers of data is critical to the success of data investments.

Analyse Phase

The **analyse** part of our data value chain is where most of the insight work happens. Analyse phase is where we convert useful data and information into knowledge and insight. Even though the modern technology industry term for this work is data science, this is the realm of statistics, modelling, general mathematics and data mining (#13).

Much of the work during the analyse phase is highly technical. The success of this stage depends heavily on other parts of the value chain, including setting the clear business value hypothesis at the beginning. It is well documented that most analytics investments fail due to factors other than technology limitations (#14). Most failures are attributed to "dirty data" or hard to access data, lack of talent, lack of management support and so on.

The skills shortage is a critical consideration for data investments. Data Science is applicable across the Data Value Chain but is especially crucial in the Analyse Phase. Figure 5 illustrates vital skills required for Data Science. What this means

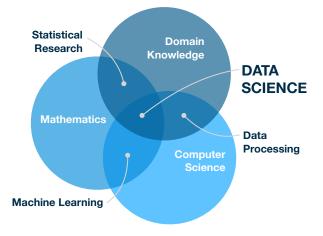


Figure 5 - Data Science Skills Model



for data investments is articulating clearly where the talent is going to come from, either through partnering, acquisitions, outsourcing or bringing expertise in-house. The in-house strategy is the hardest as competition for talent means that most companies will be competing with technology firms, where culture and benefits are more aligned with expectations of experienced data science capable people.

Phow are we going to get the skills we need to get the most out of our data investments?

Being keenly aware of why analytics investments are hard is essential, but reasons, why they are hard, are found elsewhere within our Data Value Framework. The analysis is only as good as the data we use. Likewise, analysing anything does not add value unless we are willing to act on the conclusions of the analysis, providing the findings are correct.

Decide Phase

The **decide** phase of the data value chain takes insights from the analyse phase and articulates steps that need to be considered. How these actions are articulated, depends on who or what is going to execute the actions. If people are going to perform the actions, complex insights need to be understood by people who do not necessarily have the background in data science or mathematics. This is why **visualisation** is a critical component of the Decide Phase. The hard thing about visualisation is that it needs to convey ideas, not in the language of data science or

Are insights communicated in the language of the people who will act on the insights?

mathematics but in the language of people who are going to carry out the actions. Many visualisation approaches merely present the data in a novel interactive way. While this approach has value, we should also visualise insights in a way that is biased for action. For example, visualising a fault is useful, but visualising a fault with suggested actions to remedy the defect is far more valuable.

There is, however, a cultural barrier that prevents the work in the Decide Phase to succeed, and that is that we are introducing data-driven decision making into our organisation. For many organisations, especially larger enterprises, making decisions based on evidence is hardly a typical day at the office.

Politics, power struggles, committees and endless bureaucracy are how most organisations make decisions. Ignoring the culture of the organisation when making data investments is one sure way not to realise the value we are seeking. If your organisation is not used to making evidence and data-based decisions you have to invest in educating your leaders and staff on this new way of working.

? Is data or evidence based decision making a part of our corporate culture?

There is a second class of decisions, those that are not going to be executed by people but will be performed by software. This is the role of automation. Decide phase of the Data Value Chain should take a significantly different approach. Visualisations such as dashboards and reports are of little use to software. If software is going to action

? Are the decisions going to be executed by people or software?

the insights from the Analyse phase, we need to express those insights in the language software can understand. This means having the capability to integrate systems, standardising how those systems communicate and how those decisions are monitored and followed by people in charge. Automating decision making is not an easy task.



Change Behaviour Phase

The final part of our Data Value Chain is to **change behaviour**. Changing behaviour is by far the hardest of the phases and is rarely executed well. Let's be clear; we are not merely talking about change management. Typical data investment plans will have a "change management" task running throughout the plan. This is equivalent to saying "magic happens here". Let us explore an example. Imagine we are trying to predict a failure of a laser cutter in our steel fabrication process based on the sensor data attached to a laser cutter. Let's say we can anticipate an overheating event 3 hours in advance with reasonable confidence.

Maintenance supervisor in charge of maintenance operates on a four-week shutdown plan and is measured based on her adherence to that plan. Telling her to ignore the plan and perform maintenance in next 3 hours on the laser cutter that has not failed goes against her KPIs. These types of examples exist across the organisation. Insights and decisions



we are getting from our data investments are typically not enough to get value; we will need to change behaviour to achieve the benefit. Data investments that do not include the Change Behaviour scope are doomed to fail.

We will learn more about how to design our Data Value Chain in subsequent sections. The ideas so far offer a view on several core issues faced by data investments. Next, we will explore three core considerations when designing the Data Value Chain for our data investments, how fast the value chain needs to run (Time), how many times it occurs (Scale) and how often does it need to change (Feedback).

Different value, work and technology at various speeds

The technology we select for our Data Value Chain depends heavily on how fast we want the value chain to run. This speed is dependent on nature of value we are chasing and how often we need to make decisions. For example, changing content on our website based on user behaviour will require very different approach than making decisions performed for an annual budget cycle. One of the primary considerations is selecting between streaming (continuous data and decisions) and batch (periodic data acquisition and decisions). Some use cases may require both approaches. Batch processing requires data to be collected over time and sent for analysis and decision making once enough is collected. Stream processing transmits the data for analysis as soon as it is available. Technology, skillset, levels of automation required are all dependant on how often decisions need to be made.

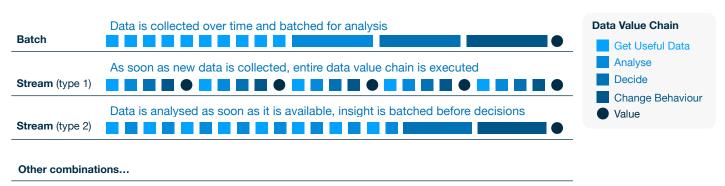


Figure 6 - Data Value Chain Speed Driven Implementation Types

There are many architecture approaches for designing technical solutions for data investments. With names such as Lambda (#16), Kappa (#16) or Zeta (#17) data technology architecture is a continually evolving field. There are no correct answers, but the clarity of what we are trying to achieve with our data investments makes technology easier to navigate.



Below are few core implications of different speed of decision making on each phase of the Data Value Chain.

Speed	Get Useful Data	Analyse	Decide	Change Behaviour
Slow	 Cleansing data at the source where possible. Temptation to collect all the data and copy it across different data management environments. 	 Can be outsourced more easily. Diverse data is harder to integrate. 	Visualisation of insight, hopefully biased for action and built with users.	Centralisation is more appropriate.
Fast	 Deploying and integrating sensing and monitoring technology. Impact on related technologies such as communication networks and computing infrastructure. 	 Separate environment for algorithm development (offline environment). Harder to outsource and integrate back into the existing systems and processes. More uniform data, but higher volume. Integration with slow data. 	 Providing human oversight on automated decisions and how to implement human intervention. Decisions are made closer to the action, more systems to maintain. Visualisation of actions, not just insight. 	 Heavily automated execution and notification systems. Communication with systems executing decisions require APIs. Distributed systems are more appropriate.
Both	 Data access rights. Integration with other systems, some of which will be legacy systems. Data cleansing and associated business process change for sustainability. Clear accountability for data quality, ideally by data producers. 	 Translation of data into a common data structure. Critical skills required and internal culture required to attract and retain the talent. 	Creating a separate team, decisions should be integrated with existing processes and producers and consumers of data.	 Data driven / evidence based decision culture supportive of executing on data insights. Underestimating the change on existing business and operating models. Temptation to implement one data management system.

Figure 7 - Data Value Chain Speed Considerations

Scaling for value

Investing in data is firstly about learning what it takes to make data investments a reality in your business. Once you establish some working knowledge and have delivered few high-value use cases, you can consider scaling your investment across the organisation or market. Scaling data investments can be made in two dimensions. One way is to apply same use case to a different part of your organisation. For example, if you have multiple factories with similar equipment, after the first successful investment in a single plant, you may consider implementing the design

for your Data Value Chain to other facilities. The second way to scale is to use the expertise, knowledge and technology and apply it to new use cases. You should not attempt both at the same time as once you try and scale there are a whole set of learnings that you still need to understand. During initial data investments before we attempt to scale, we should understand the following:



- **Teams that work** Data investments require multidisciplinary teams as illustrated in Figure 5. Also, a capable leadership team is essential, not just sponsors for data investments, but senior leaders who are educated in what it takes to drive value from data (18#). Talent acquisition, partnering and implementation strategies need to be robust and critical uncertainties tested and understood.
- Toolkit and Infrastructure Investing in data will require a whole new set of tools. This includes technical
 tools, management and leadership tools and both technical and organisational infrastructure to make the
 investments work.



Model shown below outlines a few considerations when scaling your data investments along the replication or diversification dimensions.

HIG

Replication (# of times we deploy a single use case)

HIGH

- Individual implementations do not add substantial value, but replication can potentially add up to a significant benefit.
- May require standardisation of processes, standards and technology, but should leave space for diversity and innovation, don't standardise everything.
- Automation of how solutions and change are deployed, use as much self-service as possible.
- Replicating and Diversifying requires a cultural transformation.
- High level of technology investment, data investment becomes core business, limit outsourcing.
- Focus on platform strategies, enable producers and consumers of data to create their solutions, careful use of technology patterns.
- · Never start here.
- Start here, select a few high-value use cases in limited parts of the business.
- Focus starting initiatives on learning what it takes to make data investments a reality in your business or market.
- Explore partnerships, talent acquisition strategies and organisational models to build a capable team.
- Focus on developing your data investment toolkit.
- Focus on high-value use cases; scale value is in re-usability of expertise and technology for different use cases.
- Diverse data will make data integration critical for faster scaling to different use cases.
- If use cases do not rely on same data, consider outsourcing.

LOW

....

HIGH

Diversification (# of different types of use cases)

Figure 8 - Data Value Chain Design Scaling Considerations

The job is never done, improvement is essential

Most organisations have improvement systems within, tasked with improving productivity. Industrial and manufacturing organisations have for the most part embraced continuous improvement or are implementing continuous improvement. Plan-do-check-act cycle can be found in various industrial and manufacturing organisations. Likewise, service organisations are implementing agile practices, which have roots in continuous improvement efforts in manufacturing. The relationship between continuous improvement and data is essential. Data is vital as it allows us to confirm beliefs about which practices and strategies are working and which are not. Many investments in data originate from continuous improvement strategies within organisations. People in charge of bringing forward the next wave of savings, throughput, customers and so on are continually looking to measure their process, equipment or relationships better. However, the correlation between data and improvement is complicated, and with new investments in data, it is essential we understand this complexity.

Data is a proxy for results of the system

Many organisations collect data and set targets for some future improvement, forgetting that we are not here to improve the number, but the situation we are facing. Numbers are proxies for reality and can distort our view of facts for the following reasons:

Are we improving the

 The measurement might not be an excellent proxy for what is happening because it distorts our view of the system or the numbers themselves are distorted. Are we improving the actual situation within our business or just numbers?

- Tenancy to set targets that are merely numbers improve one part of the system at the expense of the overall performance.
- Data itself is not useful without a model and models are approximations that are useful only for some situations, not universally (#19).



• The proxy nature of the measurement can diminish over time. What was once a useful measure is no longer, due to the fact we are improving and changing.

What this means is that even if our data investments are showing promise and value when we start, that value will diminish over time and might not be universally translatable when we scale the investment, either by diversifying or replicating.

There is a difference between continuous and discontinuous improvement

Value of data investments is typically expressed regarding improving some capability within our organisation (#20). It does not matter if we are improving our customer acquisition rates, the performance of our machines or quality of our products, all these improvements have to translate into a change to the way we work. However, designing an effective Data Value Chain depends on what kind of improvement we are seeking. Figure 9 illustrates a typical

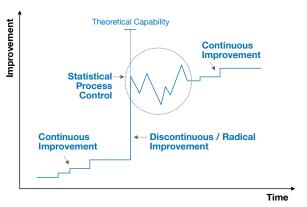


Figure 9 - Continuous and discontinuous improvement

improvement trajectory and different phases and situations we will find ourselves in as we improve. Continuous improvement requires data about our current process and behaviour. Discontinuous improvement tries to create a significant step change, which requires an understanding of the theoretical capability we are building. Data investments for discontinuous improvement require a different approach, as we may need to simulate future conditions that don't exist. Once we start implementing our step change, the process will exhibit variation that will require statistical process control, to stabilise the process and reduce variation. Data requirements during

this phase are very different again. During this period, we need to stabilise or productise the process. This requires close integration between our theoretical or simulation data and our real-world experience. Once we have stabilised the process, we can implement continuous improvement and start closing the gap between theoretical capability and our new reality. Being keenly aware of what kind of change we are achieving will significantly impact the design of the Data Value Chain and the value we get from our data investments. Figure 10 outlines some of the considerations when designing the Data Value Chain depending on the type of improvement we are seeking.

Improvement Type	Get Useful Data	Analyse	Decide	Change Behaviour
Continuous	Instrument and integrate current equipment, processes and systems.	Analysis of actual running processes and work.	Distributed decision making constantly acting on insights.	Improvement culture drives improvements bottom-up.
Discontinuous	Baseline current performance data Capture data from simulated and prototype processes	Simulation technology Ability to analyse simulation output	Ability to decide when to move from simulation, to prototype to live implementation.	Typically a larger transformation driven top- down.
Statistical Process Control	 Data captured from existing processes and simulations at the same time. Implement new instrumentation. 	Analysis of process design to stabilise output and enable continuous improvement.	Changing design parameters for the improvement and improving the theoretical models.	Close working relationship between transformation and on-the-ground team.

Figure 10 - Continuous and discontinuous improvement Data Value Chain considerations



Instrumenting and optimising the Data Value Chain

We have spoken about a relationship between data investments and improvement. However, how do we improve the Data Value Chain itself? Firstly let us examine a few scenarios where we will need to change the Data Value Chain:

- When things go wrong, and failure occurs.
- When our results and value are not aligned with design or expectations, especially when it changes over time for the worst or degrades.
- When our results and value change as we scale our solutions as we replicate or diversify.

Being aware of why our Data Value Chain is not performing is essential for its sustainability and continuous value delivery. Monitoring and testing of Data Value Chain must be designed and built into our solutions.

Let us examine how we design our Data Value Chain for stability and maintainability. We can either optimise for how fast we can recover when failure offers (typically measured as Mean Time To Recovery or MTTR) or optimised to remove failures altogether (usually measured as Mean Time Between Failure or MTBF). Optimising for both is expensive, we should

Are we designing and building our data value chain for maintainability?

do both, but optimise for one. Continuous improvement requires monitoring while discontinuous improvement requires extensive testing. We will, of course, have to switch from one or the other, but at any given time it is prudent to optimise for one, not both. Table below outlines some considerations for both of the scenarios.

Scenario	Example Optimised For	Typical Approach	When is it appropriate.	Outcome
Mean Time To Recovery (MTTR)	ZOUSSIDE TO SERVICE TO	 Monitoring for technology, process and data quality. Don't forget testing, but optimise solution for monitoring. 	 When replication scaling. During continuous improvement phase 	 Data Value Chain that is easier to change and continuously improve. Monitoring "big data" systems becomes a "big data" problem in itself.
Mean Time Between Failure (MTBF)		 Testing for technology, process and data quality. Don't forget monitoring, but optimise for testing. 	When diversification scaling. During discontinuous improvement phase and stabilising.	Data Value Chain that is transformative and offers step change.

Figure 11 - Data Value Chain maintainability considerations

To achieve the maintainability of our new Data Value Chain, we have to collect data about the Data Value Chain itself. Think of this diagnostic data or "data about data". Data value chain is just another process, of course, a new

? Are we collecting data about the performance of our Data Value Chain?

process, but one that requires improvement and care, just as much as the process we are trying to transform or improve with our data investments. Making data investment sustainable is only possible if we are aware of how technology, processes and data we are introducing performs against design and exceptions.



Learn first, deliver along the way

"If you think adventure is dangerous, try routine, it is lethal." - Paulo Coelho de Souza

Many companies are getting started with data investments. A typical way to get started is to run one or more proof of concept (POC) or pilot projects. Pilots should be designed to teach the organisation about what it takes to invest in data. Unfortunately, many data investment pilots are just classic projects rebranded as pilots. The inconvenient truth is that most organisations are not set up to get value out of their data investments. The following are few signals that indicate that pilots are focusing on the wrong things:

- Focusing investments on technology implementation without a clear understanding of value.
- Pilots are subject to the same governance, control and approach as all the other projects in the organisation.
- The scope of the pilots is focused on delivery of value, not learning how to deliver value.

In this section, we will explore how to get started with data investments, how to establish effective pilots and proof of concepts and their scope.

Miss-match between what we invest in and what the work is

If you look at data investments in your organisation majority will be branded with phrases like Analytics, Data Management, Visualisation, Machine Learning, Prediction and so on. Looking closely at their scope, we will find that majority will invest in analytics and visualisation as illustrated in Figure 12. The most laborious and most time-consuming work is in getting useful data and changing organisational behaviour. Investments have little chance of realising value even if their scope is mostly in Analyse and Decide phases of the Data Value Chain, without valuable data and change in behaviour. This is why data investments, cost more than initially thought and take longer than initially planned.



Figure 12 - Where we invest vs what the work is

The reason for this mismatch between what we invest in and where the work is, comes down to several organisational culture problems. These are:

- Organisational pressure to deliver real value, even when the way forward is unclear. Executive pressure
 to justify value delivery, forces project teams to focus too much on technology, which is typically under their
 control. An organisational culture that does not value learning, just delivery, will continuously misappropriate
 data investments.
- Lack of skills in analysis and data science forces executives to outsource data investments. Since both
 Getting Useful Data and Change of Behaviour phases require significant internal change, outsource providers
 want to prove value and deliver projects and therefore stay away from complex cultural problems.



What is the inherit risk

associated with our data

investments and how do

we manage that risk?

- Going too big too quickly with typical Data Management Platform initiatives. This occurs when technology planners look at the capability already in the organisation and identify a lack of modern data infrastructure. As we have learned previously, this is focusing on efficiency when we are not effective.
- Broken or legacy technology landscape makes data hard to access. Many people across the organisation will claim "if only I had access to data we could do amazing things." and so we focus our efforts on endless dashboards and reports. We do not ask why haven't we done this already? The answer to that question will bring us back to data is not valued, and our organisation resists change. Response to those problems can only be found by implementing leadership and cultural change, not technology.
- Plans, when there is lots of uncertainty, is a fail before you begin. Many organisations require initiatives to provide detailed plans before they get funding. This is fine for change we know how to do, but when we do not, having a plan that can't change is a disaster before you start. Failure is an option when you do not know what to do, organisational cultures that do not tolerate failure or see learning as valuable will struggle.

When we do not know how to invest in data, we should establish proof of concept or pilot initiatives. We should care about the following three types of uncertainty.

- Value uncertainty. We do not know how to extract value from our data investments. Should we chase continuous or discontinuous improvement and what kind of payoff do we get from either. We might also be unsure where to begin, which process or function in our organisation would benefit most from our investment.
- Scale uncertainty. Scaling investments beyond initial trials bring with it more considerable value and more substantial uncertainty. Is our initial hypothesis going to be right elsewhere in our organisation or market?
- **Technical uncertainty.** How are we going to implement solutions required? Technical in this sense is not just about data related technology, but technical considerations such as contracts, talent acquisition and retention, operating model design and so on.

Designing initial data investments to remove types of uncertainty mentioned above is essential for the long-term returns (#21). Approaches such as Lean Startup (#22) are focused purely on removing uncertainty.

What we are talking about is what is the inherent risk associated with our

data investments and how do we manage that risk. Figure 13 below illustrates a simple risk management process and related considerations for each type of uncertainty.

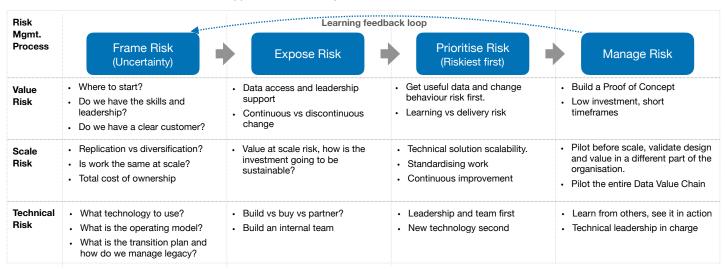


Figure 13 - Uncertainty risk reduction and management considerations



Save yourself much money by designing great experiments

Starting data investments in your organisation will require re-thinking many things you take for granted as well as learning new things. There are no playbooks or 12 step programmes to guide you. At best, there are frameworks (such as ideas in this whitepaper) to help you think about your data investments. Lack of empirical methods means you are going to have to learn what works and what doesn't. If we consider the Data Value Chain, we have been discussing we will find that much of the success depends on how well we understand the overall system within which data investments will exist. It spans IT, vendors, our organisational culture and so on. Learning about how to invest in data is just as much about learning how your organisation and markets work as it is about learning how the new technology works. Most literature suggests taking an agile or startup approach to data investments (#23). There is much uncertainty when we begin our data investments, and the only sure way to deal with uncertainty is to experiment and learn. Designing experiments are not trivial, especially not in organisations where conventional project approach is a standard way to achieve change. We have also mentioned earlier that data investments require a cultural shift in how we make decisions, towards data or evidence-based decision making. We should also seek and use evidence while implementing data investments. Designing experiments is about using evidence to guide our data investment decisions, basically a scientific method way of bringing change. We will not dwell on these methods, like Design Thinking, Lean Startup, Agile and so on (#24), as there is much literature around to explain these. Here we will focus on a few principles of effective experiment design and a few practical ideas to illustrate these principles. Firstly, what does a good experiment design look like:

- It has a valid hypothesis. Customer-problem hypothesis is designed to test if your users or customers experience the problem you think they have. Problem-solution hypothesis examines if your solution achieves the value outcome you are proposing. Having a strong hypothesis allows you to learn and adjust both how you think about the idea and the end solution.
- It addresses most significant risks first. This is the trickiest part of experiment design. Since we are
 proposing a scientific approach, we are looking for actions, behaviour, mental model, technology or anything
 that you have very little data to prove or disprove your assumptions. These are not project risks like schedule
 or budget overruns; these are risks associated with investing in something you know little about.
- It articulates state of buy-in and accountability. Not having buy-in is a massive risk to the outcome of any investment. Even in medium size organisations, who pays, who does and who benefits are typically different people. Clear accountability for leadership, learning and value are critical.
- An iterative approach to learning with quick cycles. Quick data investments?
 learning cycles translate to small initial investments, with evidence of how to improve the learning cycle over time. That requires not just awareness of customer-problem-solution fit but a keen awareness of how we are approaching the investments.
- It considers all aspects of Data Value Framework. Experiments are designed to learn about uncertainty based on riskiest first, but overall finds wholistic elements we have been discussing in this paper. If the trial is not testing particular uncertainty, it is because it is not the riskiest, but it is still on the list.

benefits for each of our



Next, we will look at a few practical ways to remove uncertainty that are typically overlooked for more expensive methods and approaches. Such as:

- Thought experiments are cheap. Too many organisations have pilots that build technical solutions to prove if they would work. In many cases, all that is required is a few stakeholders and a thought experiment to answer the question, will it work? There are teams out there building technical solutions under "pilot" or "POC" banner that are removing political uncertainty where there is little value, technical or scale uncertainty. What is worse, they are not capturing or removing political uncertainty, just spending money.
- Sometimes, you don't need to build anything. Pilot programmes are typically used to introduce new technology into an organisation, but when asked how many customers of this solution have you spoken to, the answer is none. You can remove much uncertainty by looking at other's experience doing types of investments you are making. Find others with experience and ask, instead of building new solutions on your own. Sometimes, a slide deck is just as valid at learning something about your customers as a fully working prototype. Find a simple way to get the data points to learn from; this is what experiments are all about.
- Copy experiment results from others. You would be amazed the
 answers you would get to questions to "have we done this before
 and what were the learnings?". You could remove the uncertainty
 of data investments by learning from your own experience.

?
Have we, or others, done this before and what were the learnings?

Starting with technology implementation or significant transformation programme as the execution vehicles for data investments is fraught with risk, especially if there is much uncertainty. Learning is critical for effective data investments and requires a different approach from traditional project-based methods to change.

What you are piloting is a new way of thinking and working

Data investments are rooted in technology, which in turn changes how we get work done. Most of the time, when piloting data investments, we are discovering new ways of working. Here are some significant ways data investments will change how we work:

- Having better market and customer insights. So now we have better ideas, now what? We need to change how we market our products and services and so on. It changes work, or it has no value.
- · Understanding what makes our processes, equipment and functions work and how to make them better.
- Data will become an asset, an asset we manage, improve and use every day.
- Data based decision making is the most prominent change, one that will have a profound impact on our culture.

The majority of effort, when piloting data investments should be placed on changing how the organisation thinks, feels and behaves. The core way you are organised, the work you do and how you do it is described in your operating model design. Operating model design needs to evolve based on the learnings from the early data investments. It is essential to repeat; data investments should not be treated as technology investments, or even worse, technology projects.

It may sound cliche but changing how work gets done is a people problem (#26). Even if we automate every single process in our organisation, we still have people around, defining and creating automated systems. This section of



this whitepaper is short, but a critical one. The central message is that if you are not changing how you work no amount of data will produce value. Work is always about people and how they think, feel and behave, changing those is not easy.

Technology will work, if not now then soon

Technology is, and the rate at which it is adopted is changing rapidly. Figure 14 below illustrates this acceleration, which has profound implications on how we invest in data and change in general. There are three fundamental principles we need to understand:

- The technology works, if it does not, it will soon. Digital technology especially is subject to rapid improvement and transformation. Unless you are a technology leader, you are unlikely to be innovating digital technology.
- How fast you adopt new technology is **critical.** The game for most organisations is the establishment of culture and ways of working that allow for rapid learning and integration of new technology into their day to day life.

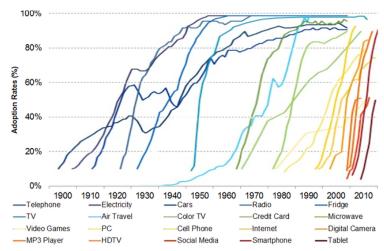


Figure 14 - Adoption of Technology in the US (1900 to the Present)

Your customers and employees are faster at adapting to new technology then your organisation. Because of the decreasing costs and improvement in user experience, we typically adopt new technology faster in our personal lives than at work. Most business and operating models are not designed for this rapid change.

Are we focusing on how to adopt new technology, not iust whether it works.

New products and services, new processes, new skill sets are all part of a very challenging problem. The winners are those with the agility that allows them to respond to the technology trends and abilities faster than their

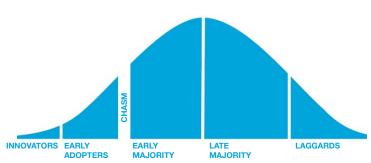


Figure 15 - Technology adoption curve

competitors. While technology is a driver of these new possibilities, it is not the primary challenge itself. The fundamental problem is the way we work every day.

Early investments in data should not place significant focus on technology, and whether it works, but instead on what possibilities does the new technology and data create; and how can you

bring those opportunities to life. For most organisations, the core question is not whether the technology works, but do you know what to do with it. The truth is in most cases organisations do not know what to do with technology. Figure 15 illustrates the curve showing different technology adoption approaches (#27). Understanding your current approach to technology adoption and where you need to be, will help explain what kind of pilots you need to design and run and the scope of those pilots.



It is not that simple, data investments are complex

"...the only simplicity to be trusted is the simplicity to be found on the far side of complexity." - Alfred North Whitehead

Outside the considerations described so far, based on our Data Value Framework, data investments are complicated due to the nature of the technology used to implement them. In this section, we will explore several considerations that make data investments complex that is typically neglected in the standard literature. We will explore the following:

- **Software is everything**(#28). Looking at our Data Value Chain, we can imagine that every step of the chain is moving towards automation. We create automation systems by building software.
- Data security is critical(#29). As data investments result in new technical solutions and user base grows our risk of exposing these systems to external threats increases as well.
- You will need the talent to make it work (#30). There is a minimal number of real solutions on the market.

 What exists is a collection of advanced solution building blocks that need specialist expertise and new ways of thinking to turn into solutions.

Software is everything

We can imagine the Data Value Chain implemented without software since Data Value Chains have existed for a long time. Imagine individuals walking around with clipboards, collecting data from analogue gauges attached to machines and infrastructure and manually writing readings onto a piece of paper. We probably know of many instances in our companies where we collect data manually like this. Same applies to service businesses, executing customer questionnaires with pen and paper. This type of data collection is error-prone, time-consuming, expensive and in many cases dangerous. Moving along the Data Value Chain, we could perform analysis manually and draw charts on graph paper. We can even then instruct people to adjust some lever or dial or also change the process if

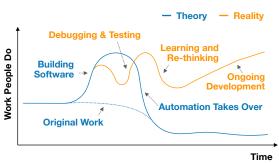


Figure 16 - Automation Theory vs Reality

required. However, that is not the type of data investments this paper is about. We are talking about using modern technology to automate the Data Value Chain. Moreover, automation of Data Value Chain can only be implemented by using the software. Here in lies a trap. Figure 16 shows theory of what we think software automation is like and the stark reality that finds many inexperienced organisations surprised. We believe automation once complete will remove work people do, but in fact, it replaces work with other work, one that is concerned with ongoing development and maintenance of technology. Sure, we can

outsource it and substitute our people, but the cost remains. Moreover, reasoning about work of software is much harder than reasoning about the work software automated. Of course, there are potentially enormous benefits, but only if we understand that software development is an ongoing process. Doing software development well is hard



and requires much cultural and technical infrastructure to work. The frequent failure of software projects is an ongoing challenge (#31), but our world operates on software, everything has software in it. The problem is that many organisations who are not software companies are investing in software. The general approach is to obsess with

project management methodologies and to exert control. There are no simple answers here, beyond the fact that investing in data requires an investment in software development capability. Once our data investments mature and start generating real value, software development automatically becomes a core competency for any business. Do not

Are we building organisational capability in software development?

underestimate the importance of being good at software, if you are not, find a great partner to help you build your software muscle, not just deliver your projects. Building internal capability is not a challenge that is met by only hiring software developers. It is about making your organisation capable of delivering software, continuously. It requires a change from leaders (#32) down to operational staff, and not just within your IT department, but organisation wide.

Difference between solutions and solution building blocks

There is a vibrant and ever-evolving landscape of data technologies driven by a significant amount of investment, see Figure 18. Typically these technologies are developed to solve particular technical challenges and are later packaged as solution building blocks. Solution building blocks are technology components used as part of the overall solution. For the most part, data technologies fall into this category. There is a small number of actual solutions on the market. Solution building blocks require technical expertise to extract value, see Figure 17. These are not install-and-use type technologies such as a word processor. What is more, solution building blocks typically require a number of them integrated together to deliver a successful overall solution.



Figure 17 - How to build data solutions.

Typically organisations misunderstand the sales pitch presented by vendors selling data solutions. The presentations usually cover previous use cases of what the vendor achieved and a collection of technologies used to produce the results. What is typically missing is what it takes to have viable solutions. Viable solutions are those that are sustainable (both technically and financially), technically feasible and achieve value. Most of the data

? Is this technology a complete solution, or a solution building block?

solutions do not just involve licensing costs (many are open source), but consulting, project and support fees to achieve them. Given a choice available on the market, we need to invest in strategy, selection and planning work. Data investments without the investment in technical

talent are not viable, given that most modern

data technologies do not achieve business value without the technical expertise to wield them. That does not mean just geek type of talent; it requires talent investment across the organisation. If your geeks are spending months convincing senior management of decisions needed and likewise, if senior

Is this investment sustainable and are we building internal expertise at all levels of our organisation?



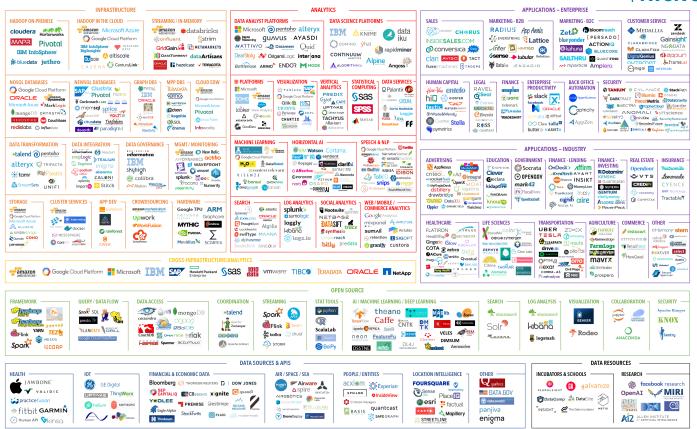


Figure 18 - Big Data Landscape (#33)

managers are not exploring new value potential of data, your data investments will not achieve value they could. The core message of this section is that amazing shiny technology you are being sold is probably a building block that requires specialist skills you probably don't have. Make it a habit of asking "how are we going to do this?" not just "what are we going to buy?" By studying the Figure 18 above, we realise the following challenges:

- There are no solutions that cover the entire Data Value Chain. Some technology focuses on data acquisition, others on data storage, some on analytics and visualisation and so on. There are a few vendors offering platform solutions that integrate these building blocks and provide single vendor for development and support. Technology integration is critical to achieving viable end solutions.
- Selection decisions are hard and temporary. Selecting
 appropriate technology from a fast-moving and complex market is
 a skill. Selecting technology does not just mean choosing from
 what is available but recognising what is possible.
- It takes time and money to create viable solutions; vendor lock-in is a real risk. Committing to a technical direction carries a risk of vendor or technology lock-in. The focus should be on the ability to modularise and integrate to reduce this risk. You do not want your data locked into a solution you cannot evolve at the pace you require, both to reduce costs and enhance capability as the new technology becomes available.
- Capability enhancement is not considered. Most organisations offering data solutions do not want you to
 get good at data investments as their primary revenue comes from strategic and project consulting. You
 need to be aware of the ongoing costs required to sustain data solutions and account for them appropriately
 in your investment calculations.

investment required to

at acceptable risk?

achieve and sustain value



Data security, if you have not been hacked, you will be

Data security is a critical risk for your organisation that will become more pronounced as your data investments mature. The reasons for this increase in risk footprint are the following:

- Data will become more critical to the decision making. Hopefully, data investments will mean you are relying on data to make decisions. Data security is crucial if you, your suppliers and customers are to trust those decisions.
- **Data volume will increase.** New insights will require more extensive and more diverse set of data.

 Acquisition and validation of that data will require management of a more significant technology footprint.
- Data will be widely distributed. Data is not very likely going to be all located in a large central system, but on mobile devices, edge devices, personal computers and so on. Securing this data will be critical without restricting access that can negate benefits we see from our investments.

Data breaches are increasing globally at an alarming rate. Organisations are typically investing in perimeter security or trying to stop violations from occurring, which while useful has done little to stem the tide of breaches.

A more useful way to think about data security is to move our mindset away from breach prevention towards breach acceptance. We need to think about what would happen when we get hacked, not if. We are not advocating not investing in breach prevention, but to invest in breach security at the same time. This means focusing on the following few critical concepts:

- Encrypt the data at every stage of the Data Value Chain. Encrypt the data as soon as you get it and thought it life, especially during transit and storage.
- Control user access. Most people are used to data being open by
 default in their daily lives. My facebook feed is open by default until I
 decide to close it. This is opposite of how we experience data at work.
 Controlling use access in a way that does not make data hard to access is
 critical to achieving data investment outcomes.

Are we securing data both in transit and when we store it?

• All this encryption needs to be protected. Data systems are typically sloppily managed, especially during early pilot phases. Protecting, storing and distributing encryption keys is a critical capability for any mature data enabled enterprise. Your data, even if encrypted is only as save as the keys used to manage encryption.

Data security and cybersecurity, in general, should be on a strategic risk register of any organisation investing in data. These concerns are now becoming core legislation in most countries and are tied to many industry-specific licences to operate, especially for industrial and consumer-centric organisations.

Most countries have cybersecurity strategies sponsored by the top levels of government (#34).

Are our cyber security practices evolving with our data investments?

One worrying trend we have noticed in many organisations is for data investment initiatives to get "exemptions" from cybersecurity protocols of their organisation

under the guise of "not going into production it is just a pilot". While this is acceptable, it typically signals deficiencies in how you are managing cybersecurity.



How to establish a healthy portfolio of data investments

"Risk comes from not knowing what you are doing" - Warren Buffett

Moving past individual pilots and proof of concepts, establishing a healthy portfolio of data investments requires concepts applicable to most technology-driven investments. Portfolio management is a topic for a separate whitepaper; here we will cover some key considerations and approaches that should turn your random list of data projects into a healthy data investment portfolio. Let us first recap a few principles we have encountered so far that relate to establishing our data investment portfolio:

- Start small, at one place in your organisation with an exploratory budget. Stay away from substantial investments initially, especially large platform programmes. Focus on learning about all parts of the Data Value Chain and what it takes to implement changes in your organisation. Remove the riskiest uncertainty first.
- Focus on capability build partnerships, not just delivery vendors. As we have seen so far, data investments require robust capability in software delivery to realise the benefits of modern technology. Your portfolio needs to have partnerships that help you build your internal capability.
- Invest in cultural change and leadership. Having a data investment portfolio that merely focuses on the
 delivery of solutions will not be able to sustain the proposed benefits. Moving the organisation towards datadriven decision making requires a change in organisation's culture.

In this section we will explore a few critical concepts that are beyond advice found in modern portfolio management approaches (#35) but are essential in getting an effective data investments portfolio:

- The total cost of owning data solutions throughout the solution lifecycle is vital. Implementing data investments across the Data Value Chain will change expenses and budgets across the organisation, not just IT or Technology function.
- Micromanaging the portfolio will stifle learning. Data investment portfolio model should manage high uncertainty initiatives differently then rollout and scale initiatives with high budgets and low uncertainty.
- At the right time, build the internal data platform. Many technologies and organisational building blocks
 can be re-used across the organisation and use cases. Investing in a platform capability will improve the
 efficiency of the data investment portfolio. Platform capability is a combination of skills, technology and
 approach investment that create reusable patterns that accelerate the path to value.

Data investments should form part of the broader technology investment portfolio. Portfolio management is essential even if you have a modest technology investment budget. The focus should be on managing and tracking our pathways to value not budget spent. Being great at managing technology investment portfolios is outside the scope of this paper but should be essential understanding for any leader investing in data.



Understanding the full solution lifecycle

Data investments will inevitably result in technical solutions and changes to work you do and how you do it.

Otherwise, there is no value in investing in data. Data investments will go through a lifecycle, that will change the focus and level of investment required to pursue value.

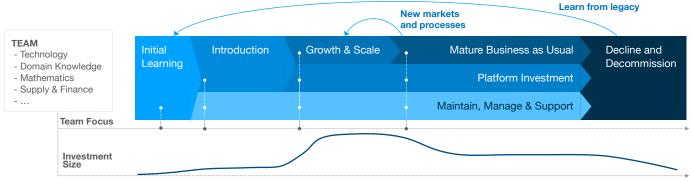


Figure 19 - Data Solution Lifecycle

Figure 19 illustrates a Data Solution Lifecycle. The model presented is self-explanatory. However, we need to call out a few critical concepts as these are seldom understood, especially at the leadership level:

- Cross-functional / multidisciplinary teams are crucial. One sure way to slow down and add risk to data
 investments is to engage in technology business partner/service provider debate. Regardless of the type of
 organisation you have, technology is core business. There is no scenario at which you are competitive in the
 market without technology. Data investments are organisation wide and require everyone's involvement, from
 end users to supply and finance to technical staff.
- Investment size will increase over time. Investment size will be variable. Make sure to start small and focused and only invest serious money during Growth & Scale phase. Spending serious money during Initial Learning and Introduction phases mean you are focusing on delivery, not learning. Invest big only when uncertainty presents a reasonable risk, and you know what you are doing.
- Platform Investments are critical for scaling and long term
 viability. Reusability of skillsets, technology and methods is
 essential for scaling data investments. Whats more, investments
 need to be structured to enable many teams to implement and use
 their solutions. Creating a central organisation that implements all
 data solutions is fine at the start, but as you scale, a platform mindset (#36) is critical.
 - Do we understand out path to production, that is viable, cost effective and manages technical debt?
- As soon as you have users you have maintenance, support, security and management concerns. It is
 effortless to label data investments as "innovation" initiatives given these are new to many organisations.
 Ensure your budget and team are ready to support and maintain the solutions as soon as they have users.
- Technical Debt and Legacy are our most significant technical risks. Sometimes you need to implement something quickly to respond to a market need but enforce messy processes or technology that make further changes harder. Sure, you could spend time designing and considering carefully what you do, but then you might miss market opportunity or value. Either approach is fine as long as you are aware you are creating technical or legacy debt and like any other debt it needs to be repaired, or you go bankrupt. Understanding the concepts of technical debt (#37) is critical for making sound investment portfolio decisions.



Managing the portfolio and letting individual initiatives learn

For data investments to expand beyond pilots and experiments into changing how the organisation works, you need to address the following two prevalent portfolio management (#38) sins:

- A collection of projects is not an investment portfolio. Asking all departments and parts of the
 organisation to report on their data projects is not a portfolio. Portfolios should start with a clear value
 hypothesis and a strategy describing how that value will be realised. Projects within a portfolio are designed
 to deliver value and validate and improve the strategy. Portfolios should be designed not collected.
- Focus on task costs, not task throughput. The constant focus on costs and not throughput will incentivise teams to sacrifice value for reporting requirements. Project teams should focus on task throughput and be improving their performance. This requires a keen focus on "How" we execute investments.
- Multitasking kills throughput. Inevitably data investments will have one persistent constraint, that being
 skilled resources to deliver solutions. Crucial people will be stretched between tasks and projects due to
 inevitable resource scarcity. Focusing the portfolio plan to take this fundamental constraint into account will
 enable better delivery. Limit multitasking as much as possible, let individuals concentrate on one thing at the
 time.

Most data investment portfolios are treated as any other portfolio of projects, even in medium size companies. The truth is that majority of initiatives fail, not because someone used a wrong project or portfolio management

framework, but because portfolio managers and decision makes did not understand what the work is. Decisions that are disconnected from work will always result in adverse outcomes. This entire paper is structured to explain the work that goes into investing in data effectively. If your portfolio and investment management systems do not reflect the reality of the work required, you will struggle. Getting a good grasp of how projects run

Do we understand the work we are doing and keeping decisions as close to work as possible?

is essential in making great decisions. Techniques such as Critical Chain Project Management (#39), Lean Startup (#40) and so on are a great place to start understanding systemic reasons why projects and portfolios fail and how to move past those challenges. You do not have to implement one approach or another, but you have to have a keen interest and systemic understanding of how to manage projects and investments.



Figure 20 - Effective Portfolio Understanding

The compelling data investment portfolio requires a deep understanding of what the work is. That work changes both based on the level of uncertainty and wherein the Data Solution Lifecycle the solutions are at. We have

? Are we understating the "how" uncertainty just as much as "what" uncertainty in our investment portfolio?

discussed at length that data investments must improve the work our organisation does or the value will not be realised. The complicated truth is that we may have to change the work our company does just to be able to execute the investments. This requires a deep understanding of how we structure our investments into portfolios and how we monitor and run those portfolios.



Platforms enable scale for speed and profit

The last section of this paper is dedicated to platforms. Most advice out there centres on platforms. Most investments are wasted on platforms, and most technical solutions are sold as platforms. Investing in a modern data platform is a critical part of investing in data. Most organisations invest in platforms either way too early and don't understand the reason for investment or how to use the platforms once constructed. Typically, platforms come to life due to the following reasons:

- Attempt to standardise the technology stack, support and spend. Even the most immature companies have databases and systems that manage data. It stands to reason that if you standardise these systems, we can reduce our operational budget, deal with a smaller number of vendors and enable integration.
- Attempt to centralise data expertise. Everyone in your organisation has to deal with data. If you ask each of your department heads for their two-year plan, there is a high chance each of them will have some data initiatives. This may sound like a waste of money, to let them all do their own thing, so you create a centralised team to develop solutions for each part of the organisation.
- Combination of both. One technology stack, one team, seems reasonable and cost-effective.

There are grains of value in all of the approaches above, but unfortunately, in reality, we get the opposite effect. Our delivery speed disappears, our budgets go out of control and accountability is nowhere to be seen. There is a fourth approach to building platforms that deal with core problems faced with the conventional wisdom on how to build and invest in platforms. Before we describe this platform approach, we need to understand why the above strategies do not work. The problem lies at the heart of standardisation. Standardisation and its cousin, centralisation, of

technology, talent or methods does not manage the variety of demand. If everyone in your organisation asked for the same thing or every customer wanted the same thing, standardisation is a right way forward. We know from experience that this is not true. Different departments want data presented differently, at varying timeframes to people with different skillsets

Are we designing the platform to satisfy real demand from customers?

and experience. Likewise, customers want slightly different things at slightly different times depending on their demographic or situation. Much of the hype with platforms comes from success stories of online platforms such as Uber, Youtube, Airbnb and so on. The reason these platforms are successful is that they do not provide end to end solutions, but attack a particular problem, like sharing videos in case of YouTube. The work and technology used to create and edit videos, is not part of the platform, same goes for Uber, Airbnb and others. However, most organisations that invest in their internal platforms try and provide the end-to-end solution in attempting to provide a single technology stack and a single team to deliver a variety of uses cases all implementing the Data Value Chain. This variety of demand is what makes typical platform initiatives very experience, complicated and eventually ineffective. Traditional platform designs respond to this variety of demand in the following ways:

- A complex technical landscape that is trying to satisfy everyone. This approach results in ever increasing scope and technical complexity. It becomes a competition of who shouts the loudest that gets their features built. Furthermore, this complexity rises costs that are hard to justify.
- An oversimplified technical landscape that does not satisfy anyone. A solution that does not enable the implementation of the full Data Value Chain. This typically manifests itself in the selection of a single vendor to deliver the solution, and as we have learned, there is no such vendor on the market.



• Push to "simplify" and "standardise" the business. Typically the excuse why the above approaches do not work is that the demand is broken or wrong. "If they only asked for the right thing this will work".

There is another way, one that takes into account the real power of platforms. Platforms should enable and make smooth the interaction between those who create value and those that need it. It provides methods, technology and expertise to make interactions easy. Platforms, therefore, have two types of customers, producers and consumers. Think of YouTube; video creators are producers, people who watch videos are consumers. Youtube makes the interaction between these two types of customers easy. Uber drivers are producers, while people wanting to go places are consumers, Uber makes this interaction easy. Platforms monetise this interaction. Uber does not own cars, yet cars are essential for the implementation of the end-to-end solution.

A way to think about platforms when investing in data is to consider the following:

- Implementing various Data Value Chains is about the effectiveness or doing the right things.
- Implementing a platform is about efficiency by encoding re-usability and making interactions required to build Data Value Chains easier, or doing things right.

Platform investments centralise the tools, technology, methods and skills required to accelerate, de-risk or improve how Data Value Chains are constructed. Platform investments should not focus on building solutions, but building blocks that are used by people developing solutions. This re-framing of who are the customers of the platforms is the core mistake most companies make. Platforms deliver value, not by impacting the data investment value proposition directly, but by optimising data investments through:

- Acceleration codifying and packaging knowledge, expertise, tools and technology that makes the next data investments faster to realise.
- Cost efficiency in organisations with a most diverse set of technologies that are hard to operate it may
 make sense to standardise thought platform investments, but only if it genuinely reduces learning and
 organisational effort required to build full Data Value Chains. Focusing on costs is a one way to increase
 them, the focus should always be on value.

Reflecting back on Figure 19 and the data solution lifecycle, data platform investments are used for scaling the shared technology, skill and method patterns across the organisation. That is why you should never start with the platform investments, but focused pilots that require low levels of investment. Once you have implemented few full Data Value Chains and are ready to scale, either by replicating or diversifying, you are prepared to identify common patterns and invest in their optimisation. Without this experience, platforms will represent optimisation of process you know little about.



A few simple tools to help you

"The productivity of work is not the responsibility of the worker but of the manager."
- Peter Drucker

This paper offers an unconventional set of concepts describing how to invest in data effectively. We have compiled a collection of easy to learn and use tools to help you navigate your data investments. We use these tools when helping our clients with data investments, and we continuously update them as we learn more about what works and what doesn't. Please give us feedback about how the tools are helping you and how we can improve them.

The tools available in this section are:

- Investing in Data Overview Poster outlines critical concepts in this whitepaper.
- Data Investment Canvas a strategic tool for capturing data investment benefits, scope and approach.
- Pilot Assessment Tool understanding if your pilots are removing right kind of uncertainty.
- Portfolio Visualisation Tool a tool to help you visualise your data investment portfolio.

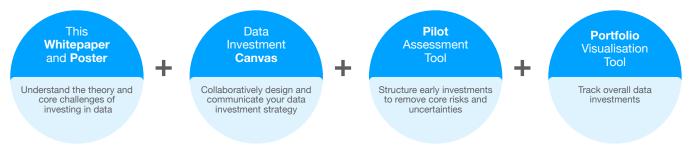


Figure 21 - Investing in Data Toolkit Overview

Investing in Data Overview Poster

A poster that outlines all the critical questions posed in this whitepaper. It also includes the following models all on a single A3 page:

- Data Value Framework
- Multidisciplinary view of data investments skills
- Transformation vs Improvement Model
- · Data Solution Life Cycle



When to use this tool?

- Great for recalling core concepts from the Investing in Data whitepaper.
- Great to take with you when explaining the concepts to others, it has useful diagrams to go through.
- Print and post on a wall to share with a broader team.
- Works well with Data Investment Canvas, as a prompter for questions to ask.

Where do I get this tool from?

https://www.actiontwelve.com/assets/images/publications/investing_in_data_overview.pdf



Data Investment Canvas

Data Investment Canvas is a tool used to capture the core strategy guiding data investments collaboratively. It is primarily a workshop concept capture tool as well as a communication tool. The canvas is broken down into four different rows, each designed to capture the following strategic concepts:

- Where to play, by examining products and services, core value chains and customers targeted by data investments.
- Why we play, by looking at results being sought, stakeholders being affected and resources required.
- How we play describes the customer-problem-solution hypothesis.
- Actions and Skills, showing the actual activities required to achieve the data investment strategy.

When to use this tool?

- Strategy workshops with a diverse set of stakeholders.
- To communicate overall data investment strategy.
- To guide a holistic approach to capturing data investment strategy.

Where do I get this tool from?

https://www.actiontwelve.com/assets/images/publications/investing in data overview.pdf

Pilot Assessment Tool

The Pilot Assessment Tool is used to ensure pilots are working on the right things, focused on learning and managing their budget and process well.

The tool is structured in two parts.

- **Overview**, outlines questions that every pilot needs to answer that are focused on ensuring strategic and value alignment.
- Detail outlines vital hypothesis, assumptions and how those are being validated.

Testing of the control of the contro

When to use this tool?

- To design pilots, ensuring correct focus and alignment.
- · To track pilot progress and learning.

Where do I get this tool from?

https://www.actiontwelve.com/assets/images/publications/pilot_assessment_tool.pdf



Portfolio Visualisation Tool

The Portfolio Visualisation Tool provides an overview of the data investment portfolio. It asks you to think of your portfolio regarding the overall goals, not just a collection of projects. It outlines the following:

- Results Chain that links resources, activities and outputs to outcomes. It provides a basis for understanding how all the initiatives in the data investment portfolio contribute to the overall narrative of value delivery.
- **Customer-problem-solution** outlines critical hypothesis about how data investments will change your markets or organisations.
- Investment pivots track key learnings and how the data investment strategy and investment portfolio are responding to the learnings.
- Project portfolio section tracks financial and delivery performance of all the initiatives in the portfolio.

Name Additional Control of Visualisation Tool Additional Control of Control

When to use this tool?

- To design, unite and track a set of data investment initiatives as a cohesive portfolio.
- When translating the data investment strategy into a set of actions.

Where do I get this tool from?

https://www.actiontwelve.com/assets/images/publications/portfolio visualisation tool.pdf



References and Further Reading

- 1. Pierson, Marc. (uploaded 2015). "Tape" of Ackoff's Bell Lab Lecture (video)
- 2. Newman, Daniel. (2017). Why The Uber Of Everything Is A Flawed Approach To Innovation
- 3. Musk, Elon. (2016). Master Plan, Part Deux
- 4. Thomas, Dan R. E. (1978). Strategy is Different in Service Businesses
- 5. BCG. (2017). Evolving Digital Offerings for Off-Highway OEMs
- 6. Bloomberg, Jason. (2017). John Deere's Digital Transformation Runs Afoul Of Right-To-Repair Movement
- 7. Hicks, Michael J; Devaraj, Srikant. (2017). The Myth and the Reality of Manufacturing in America
- 8. Rifkin, Jeremy. (2014). Capitalism is making way for the age of free
- 9. Howlett, Dan. (2017). SAP UK v Diageo an important ruling for customers with indirect access issues
- 10. Bloomberg, Jason. (2017). John Deere's Digital Transformation Runs Afoul Of Right-To-Repair Movement
- 11. Kramer, Staci D. (2011). The Biggest Thing Amazon Got Right: The Platform
- 12. Yeates, Clancy. (2017). ANZ Bank restructure to create '150 start-ups'
- 13. Hansen, Drew. (2016). How To Become A Data Scientist In 2017
- 14. Kaggle. (2017). 2017 The State of Data Science & Machine Learning
- 15. Palmer, Shelly. (2015). Are You Ready for Data Science?
- 16. Forgeat, Julien. (2015). Data processing architectures Lambda and Kappa
- 17. Scott, Jim. (2015). Zeta Architecture: Hexagon is the new circle
- 18. Weathington, John. (2017). Why CEOs must lead big data initiatives
- 19. Kramer, Aaron. (2017). <u>Understanding How and Why Your Model Works</u>
- 20. Ackoff, Russel L. (1994). If Russ Ackoff had given a TED Talk...
- 21. Blank, Steve. (2013). Why the Lean Start-Up Changes Everything
- 22. Ries, Eric. (2011). Lean Startup Methodology
- 23. Woods, Dan. (2017). Real-Time Feedback: Agile Product Development At Walmart
- 24. Schneider, Jonny. (2017). Understanding Design Thinking, Lean, and Agile
- 25. U.S. Department of Labor. (1999). Futurework: Trends and Challenges for work in the 21st Century
- 26. BCG. (2017). Twelve Forces That Will Radically Change How Organizations Work
- 27. Everett M. Rogers. (2010). Diffusion of Innovations, 4th Edition
- 28. Andreessen, Marc. (2011). Why Software Is Eating the World
- 29. Australian Cyber Security Centre. (2017). Australian Cyber Security Centre 2017 Threat Report
- 30. Oxford Economics. (2012). Global Talent 2021
- 31. Charette, Robert N. (2005). Why Software Fails
- 32. Xiao, Guo. (2017). The Next Big Disruption: Courageous Executives
- 33. Truck, Matt. (2017). Firing on All Cylinders: The 2017 Big Data Landscape
- 34. Australian Government. (2016). Australia's Cyber Security Strategy
- 35. Highsmith, Jim; Hollier, Natalie. (2016). Thoughtworks Insights: Portfolio Management in the Age of Digital Enterprises
- 36. Choudary, Sangeet Paul. (2015). Platform Thinking Labs
- 37. Kruchten, Philippe; Nord, Robert L; Ozkaya, Ipek. (2012). Technical Debt: From Metaphor to Theory and Practice
- 38. de Souza, Fernando B; de Moraes, Antonio AC. (2016). <u>Analysis of the application of critical chain project management in the product development process and portfolio management of an aircraft manufacturer</u>
- 39. Goldratt, Eliyahu M. (1997). Critical Chain
- 40. Ries, Eric. (2017). The Startup Way

