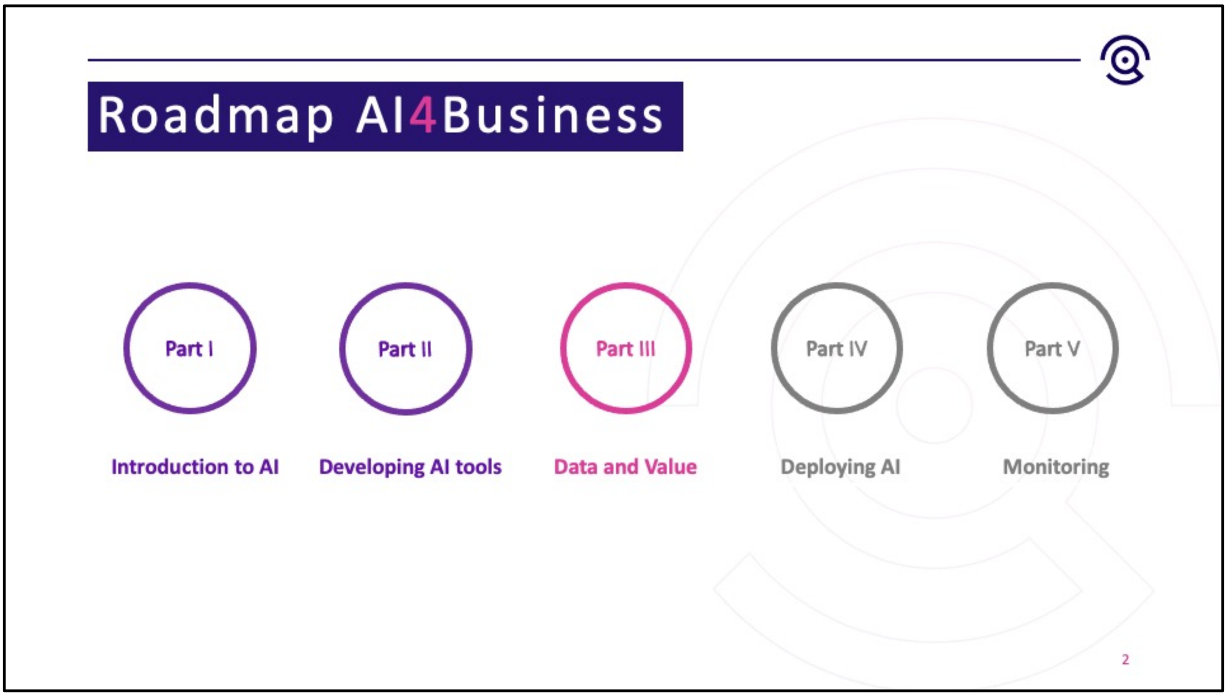
Welcome to the third module of the AI4Business course. In this module we highlight how you can actually start capturing business value with AI. In the end, this is of course the most important aspect for any company, namely to generate value for your customers. So let’s jump right into it.



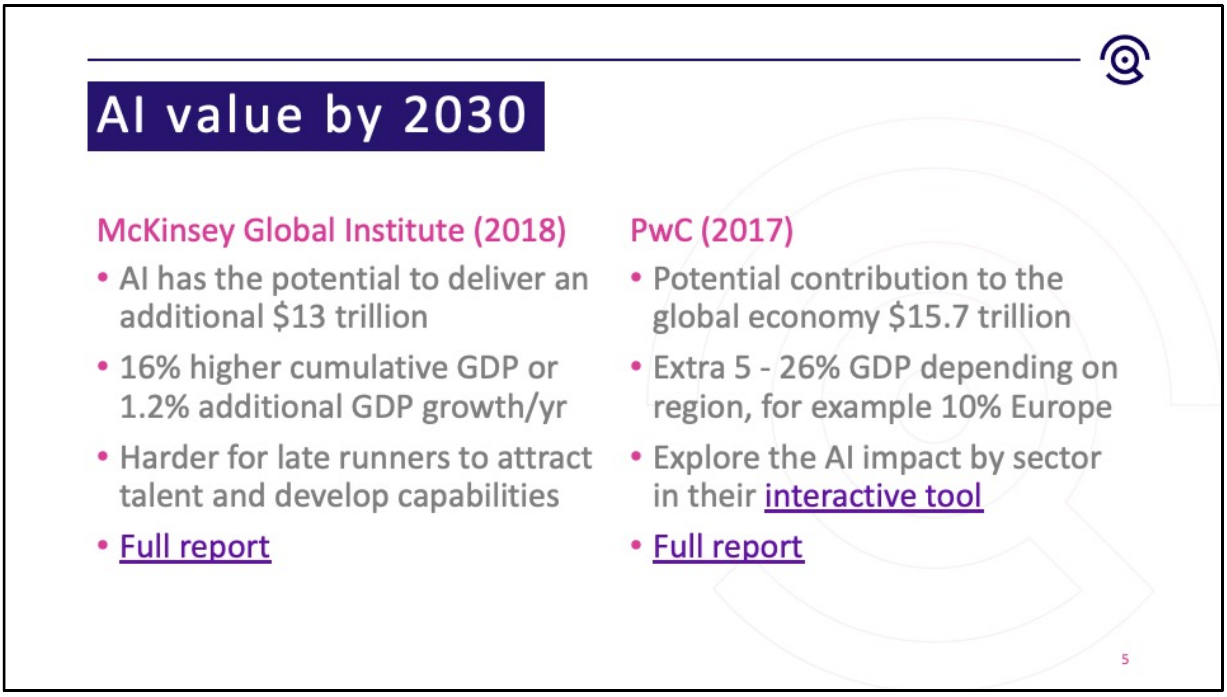
This module is the third out of five modules in the AI4Business course. We already covered the most important AI concepts and how to develop AI solutions from a high-level view. We will use that knowledge today and learn how to capute true business value with those solutions. The next modules build further on this and show how to bring that value to the real world and keep on capturing it over time.

Table of contents

3

We start this module by discussing how you can extract value from AI and from data. We then discuss the synergies at play between both data and AI. Because data is so important, we then go into the topic of DataOps, an agile approach that allows you to capture business value from your data. We finish this module with some tips and tricks on how to get value from the right business cases and how to become an AI company. I hope you will enjoy this!

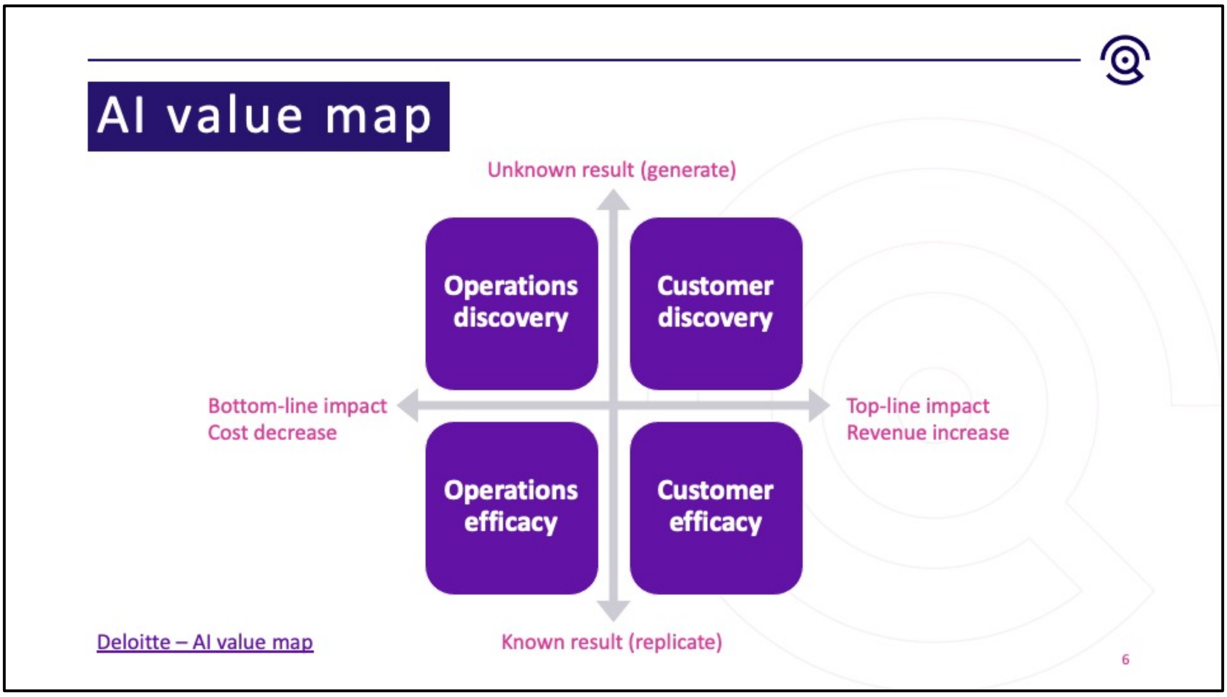
Let’s see how we can capture value with AI.



Two studies by leading consulting firms estimated the impact of AI on value creation in the global economy by the year 2030. Both full reports are linked on this slide.

In 2018, the McKinsey Global Institute looked at the economic potential for five broad categories of AI: computer vision, natural language, virtual assistants, robotic process automation, and advanced machine learning. AI has the potential to deliver an additional global economic activity of around $13 trillion by 2030, or about 16 percent higher cumulative GDP compared with today. This amounts to 1.2 percent additional GDP growth per year. They also note that late adopters might find it difficult to generate impact from AI, because front-runners have already captured AI opportunities and late adopters thereby lag in developing capabilities and attracting talent.

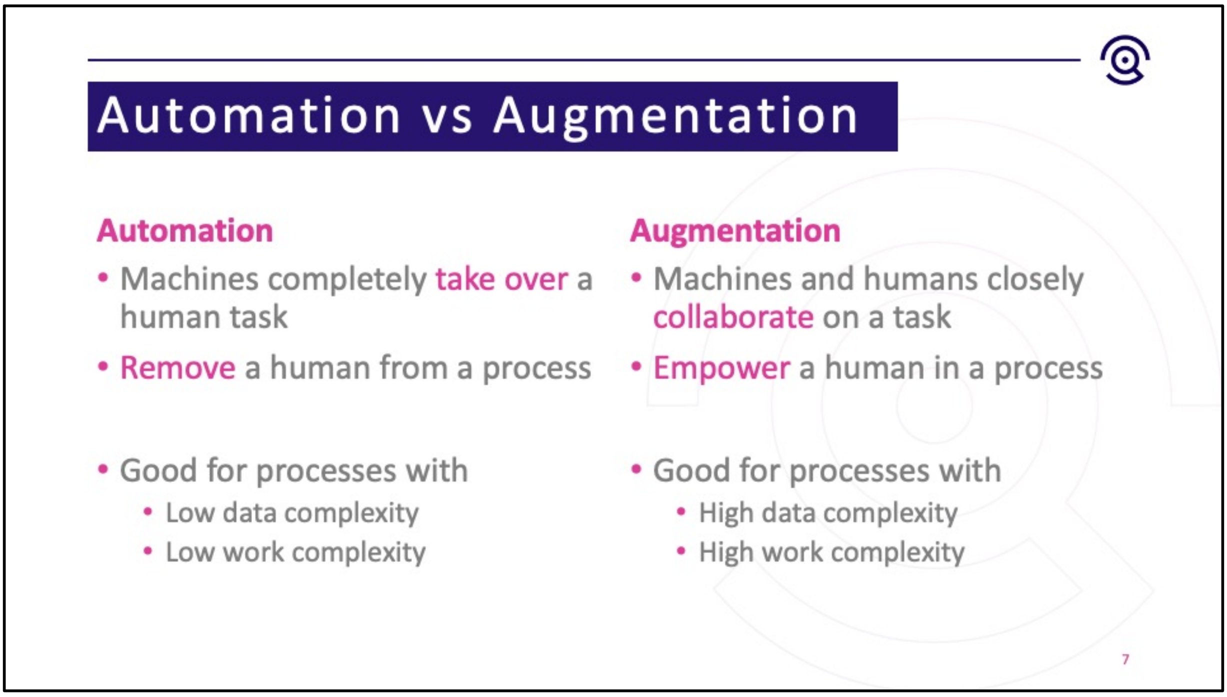
In 2017, PwC estimated the potential contribution from AI to the global economy by 2030 at $15.7 trillion. They do note that strategic investment in different types of AI technology is needed to make that happen. Different regions will experience different gains from AI, with for example a GDP growth of 10% for Europe, 15% for North-America and 26% for China. They published an interesting interactive tool where you can explore the AI impact by sector if you would be interested in this.



Deloitte defines four models of value creation for AI, where they map AI use cases in a 2x2 matrix according to their type of result and type of impact.

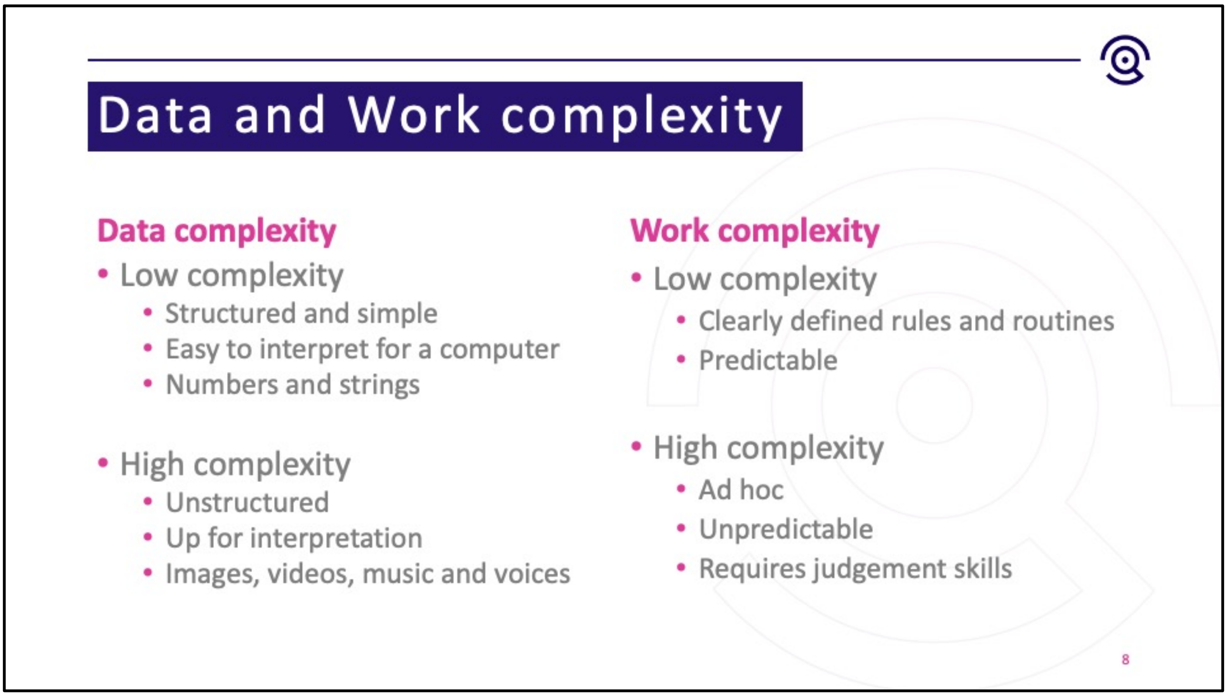
The vertical “type of result” axis distinguishes whether an AI application produces unknown or known results. The replication of known results matches input with a predetermined set of answers or actions, for example a chatbot with an underlying database. On the other hand, the generation of previously unknown results creates new insights and answers, for example new customer segments or new product designs. The horizontal “type of impact” axis distinguishes whether an AI a plication delivers financial impact on the top or bottom line. Bottom-line improvements are achieved through reduced costs while top-line improvements are realized through increases in revenue. Four areas for value creation with AI emerge when combining the two dimensions.

AI applications reproducing known results with the primary intention to create bottom- line impact enable operations efficacy, for example by automating repetitive tasks or making knowledge accessible in real-time. Top-line impact can be delivered through AI applications targeted at customer efficacy by employing AI to perform actions that are known to increase customer satisfaction and consequently conversion and loyalty, for example sales, marketing, and pricing automation. AI applications for the generation of unknown results with the primary intention to deliver bottom-line impact enable operations discovery. Such applications may for example discover new insights about process inefficiencies or cost drivers and serve to increase quality or cost effectiveness. Alternatively, top-line impact can be delivered through AI applications targeted at customer discovery. This aims at discovering new insights, actions or products that serve to strengthen the relationship with the existing customer or the acquisition of new customers. Such tasks can include for example the discovery of novel criteria for successful customer segmentation or data-driven product development.



AI can improve any given process through either automation or augmentation, but what is the difference? Automation implies that machines completely take over a human task, thereby removing the human from the process. Augmentation means that humans collaborate closely with machines to perform a task, thereby empowering a human in the process.

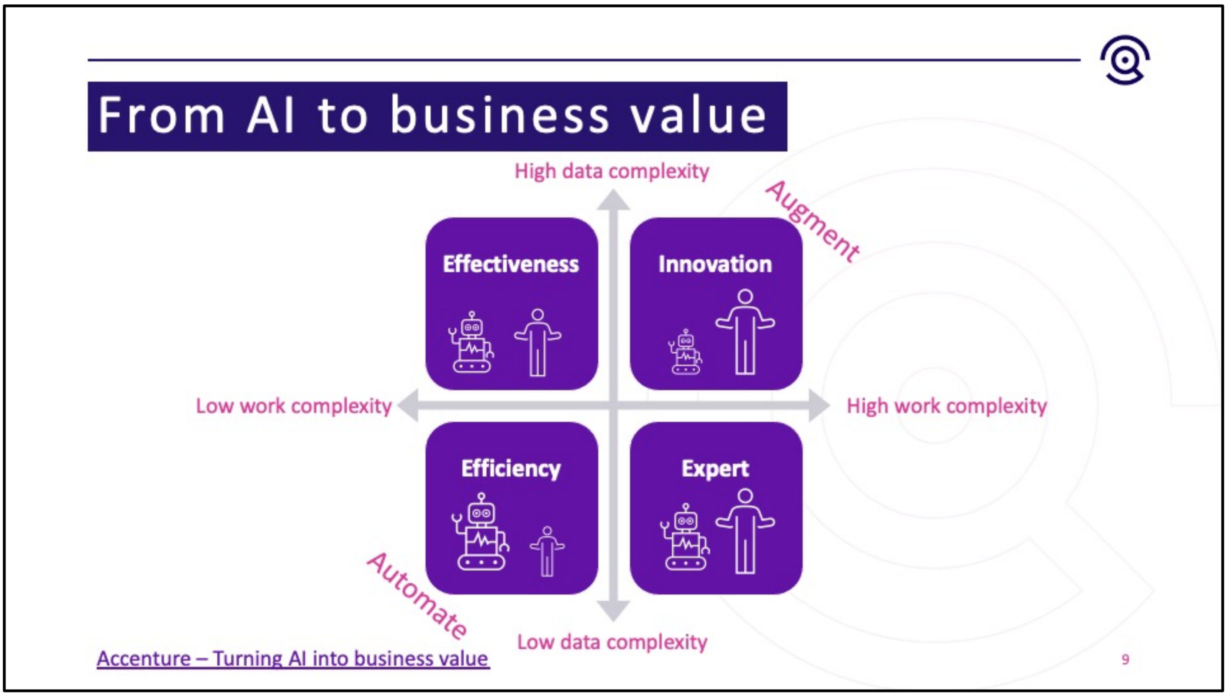
Humans are more than happy to delegate tasks to an AI if they are more robotic, focused on speed, accuracy or repetition. Nevertheless, there will also be tasks that people want to control themselves without relying on automation. Or sometimes the system is so complex that automation just isn't possible. For this reason, repetitive task where speed or no special cognitive skills are required are good candidates for automation However, not all processes can be automated, or they can only be partly automated. Automation is good for processes with low data and work complexity, while augmentation is good for processes with high data and work complexity. Now what do we mean by this?



Data complexity is the easiest of the two variables to explain. Data with low complexity is typically structured and simple, usually strings of text or numbers. The data is easy to interpret for a computer. High data complexity is often unstructured and up for interpretation. Images, videos, music, and voices are examples of complex data.

Work complexity is mostly about routines and rules. If a work has clearly defined rules and routines, it becomes predictable, and has low complexity. However, if work is unpredictable and ad hoc, then it requires judgmental skills, which would result in high work complexity. Computers are excellent at predicting, but not so good at judging.

So now we repeat what we just said: If a process has low data and work complexity, it is suited to be automated because computers can deal with this. If a process has high data and work complexity, it is suited to be augmented because human judgement is required here.



Accenture published a white paper “Turning artificial intelligence into business value” where they use the axes of work and data complexity to describe processes. The further to the bottom-left a process is, the more it can be automated. On the other hand, the more to the top-right it is, the more it can be augmented. In between, they propose four models for process strategies with AI: efficiency, effectiveness, expert and innovation.

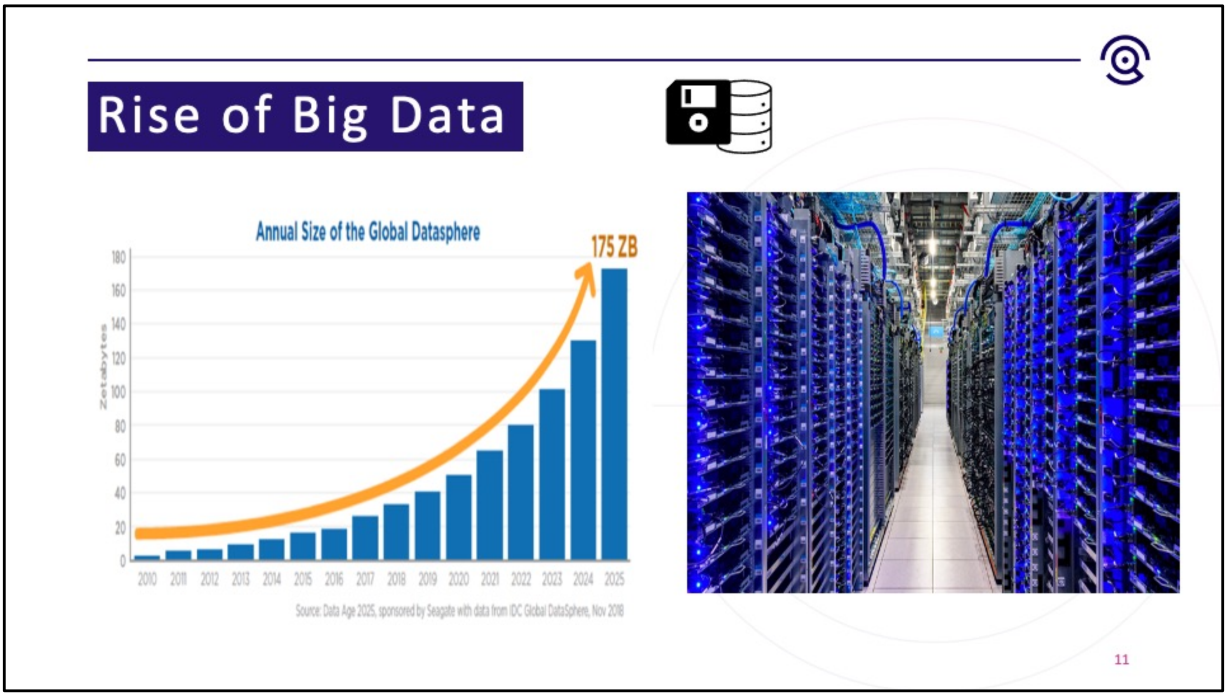
If a process has low data and work complexity, it can most certainly be automated. This is the simplest model, and the one that most companies are trying to adopt today. It is a clear first step for any company wishing to start their AI journey. Efficiency is all about optimizing your business, generally reducing costs. Processes that are placed here have very clearly defined rules and routines. Examples are automated credit decisions, purchases, recruitment or delivery.

If the work is simple, but the data is complicated, you will want to focus more on effectiveness rather than optimization. This model revolves around the communication and coordination of workers, with AI serving the role of an assistant. Processes placed in this model are used to make the work of employees more effective, by either eliminating or simplifying the act of scheduling, communicating, or monitoring. Examples are automated scheduling, customer support or task assignments.

When the data is simple, but the work complexity is high, AI can be used to leverage expertise. Processes placed in this model rely heavily on expertise and judgment. Therefore, unlike the left-most models, any decision made here is taken by humans. AI is offering a supportive role, opting to offer advice and insights. In this model, humans are always in control. Expert systems deal with anything from large amounts of money to human lives, meaning humans must always be responsible for decision-making. Examples are medical diagnosis or financial investments.

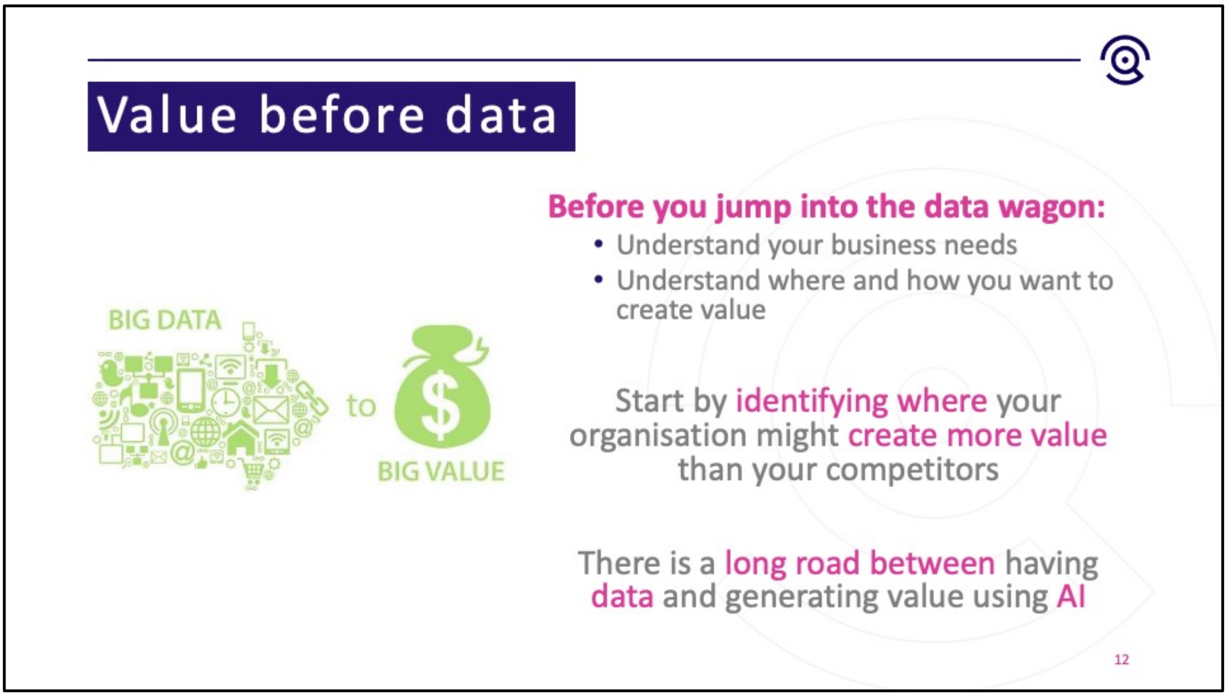
Processes that have both high data and work complexity can be used to enable creative work. Humans are in complete creative control, while AI is used to identify recommendations and alternatives. As you might imagine, this model is particularly interesting for jobs that are creative in nature, such as designing, researching, writing, cooking, composing music, drawing, filming and so on. Examples are recommendations while writing, composing, video-editing or performing a job interview.

Let’s start with a definition of AI.



The exponential growth of data has overwhelmed the capabilities of traditional approaches and tools. Technology is growing at a faster pace than ever before, all to keep up with the needs of organizations. This technological and methodological boom has allowed Big Data to reach almost every single aspect of our lives. The world is becoming a data-driven place. The ever-growing amount of data provides a vast number of opportunities and challenges in terms of data capture, storage, manipulation, management, analysis, knowledge extraction, security, privacy and visualisation.

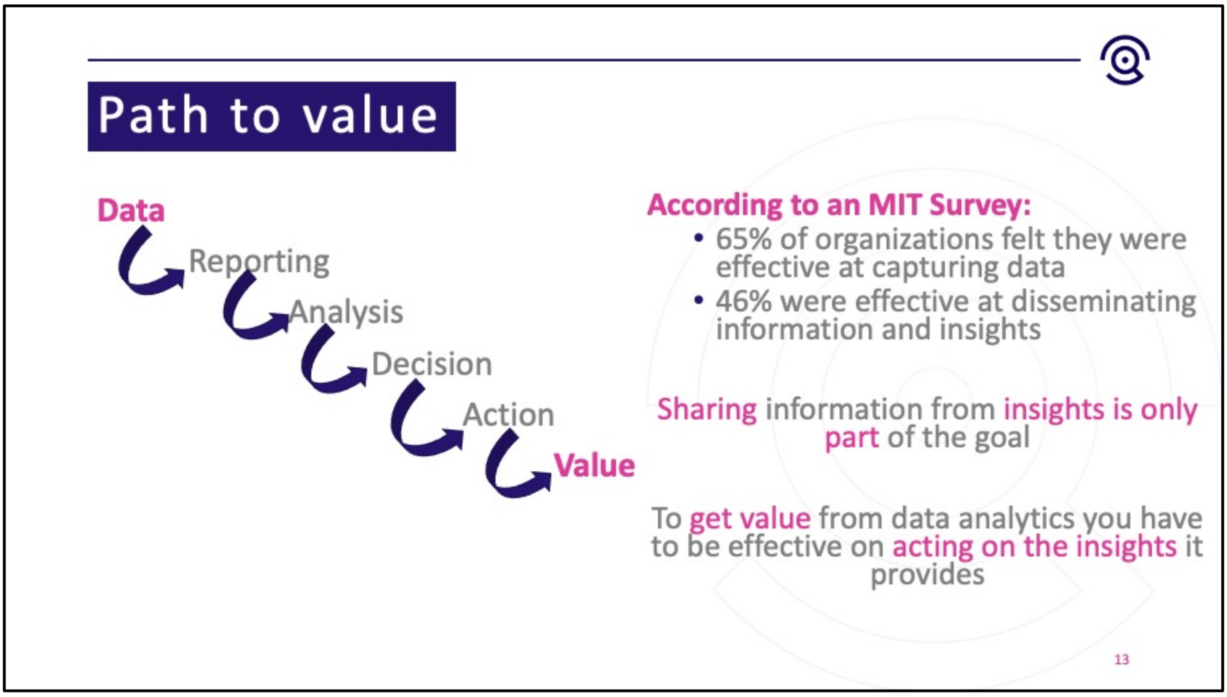
Though the potential of big data is unquestionable, there is still a wide gap between potential and realization. The rise of super computers with tremendous power and Big Data technologies appear to have empowered AI in recent years. However, in order to take advantage of what AI has to offer, it is not just necessary to be able to access and handle data. It is also important that we understand the path to value.



There is a visible common trend in the most advanced economies on how they are transitioning more and more form physical production to services and intangible assets. We could say that the information era is officially here, and at the center of it is Data. Companies are increasingly looking for innovative ways on how to harness the power Big Data, AI and data analytics in general. They also put focus on how to make the right investment to monetize their data.

The paradoxical things is that you should not start with data, but first understand your business. This means understanding what your value proposition is. Start first with where and how you want to create value and then move to data. Not the other way around. It is true that there are many opportunities in the data domain. Given the fast pace at which technology advances these days, our main limits are those of imagination. But your first step should not be spending large amounts of time thinking how much data you have, where it comes from or how you can use it. Start by identifying where your organisation might create more value than your competitors. Then try to understand the data requirements you would need to move forward. Only after going through that process you will really understand the scope of the investment you will need to harness the power of your data.

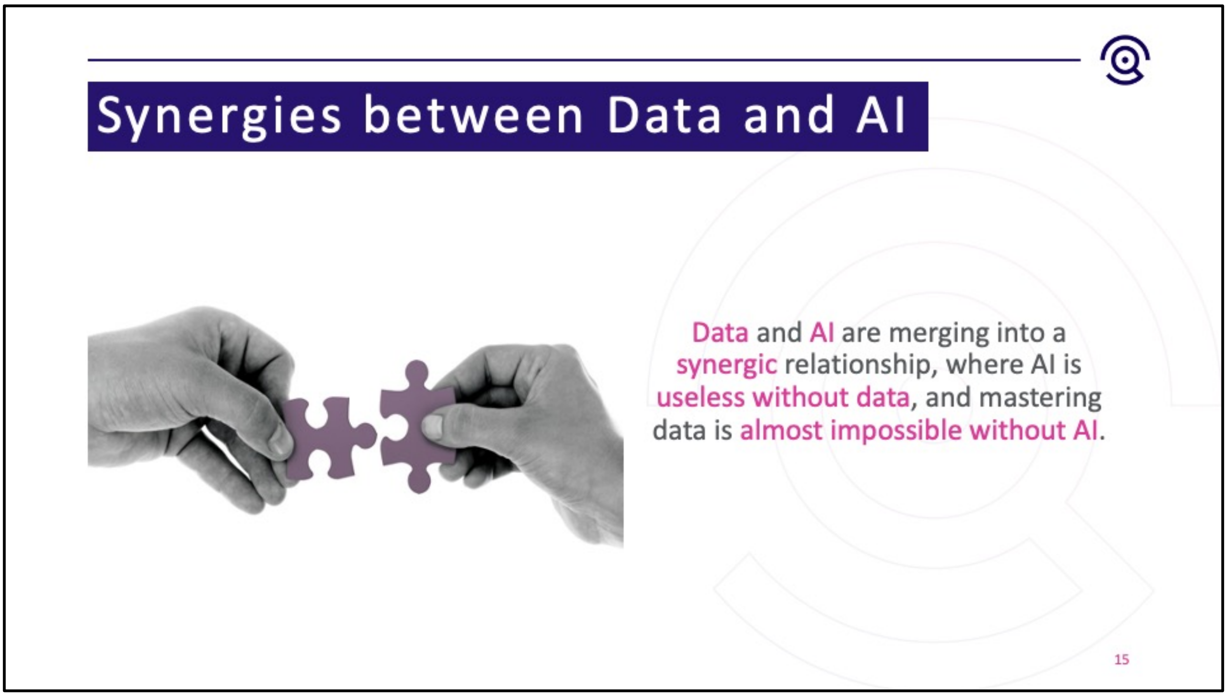
And before you start even thinking about AI, you should have your data game on. There are a lot of “boring” but important utility processes that you will need to invest in before AI is a possibility. Do not underestimate how much time and money it will take to have the data in good enough shape to move forward. Since everybody seems to be already in the data wagon, it is tempting to jump in just for the fear of missing out. If you decide to do, first adopt a strategic and systematic approach to extracting value from data, only then you will be able to reap the rewards of your investments.



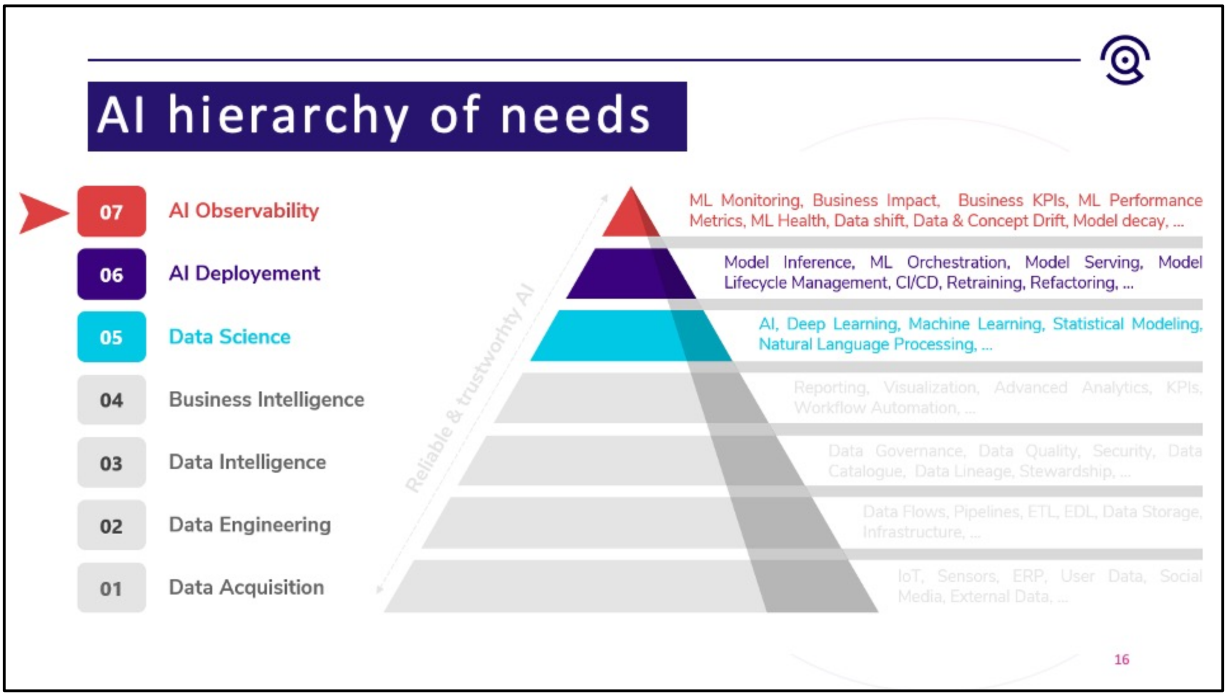
According to an MIT Survey, 65% of organizations felt they were effective at capturing data but only 46% were effective at disseminating information and insights. This is already a sharp decrease, but sharing information from insights is only part of the goal. To get value from data analytics, you have to be effective on acting on the insights it provides. And that is already a much more difficult task. So just collecting data will only bring you a step forward in the journey for value, but not so close to the finish line yet. There is a path that most organizations follow to generate value, and if one of the steps is misaligned it can impede an organization's ability to generate value from its data.

- As we can imagine, collecting and maintaining data is the first step of the journey. Without useful and trustworthy data there is no point to this whole exercise.
- After your organization has collected its data, this raw material usually needs to be cleaned, converted, combined or organized before it is truly useful and accessible to business users. Once that's done business reporting transforms data into information, through different means such as scheduled reports, scorecards and dashboards.
- This information can be used for analysis. While reporting is useful, analysis is the way in which we can turn information into actual actionable insights. This unlocks the potential of monetization. The big problem with many organizations is that they confuse reporting with analysis.
- Before insights can be translated into action, a decision maker needs to decide whether or not to act on a particular insight or recommendation. This can be a human decision maker or a computer in case of a fully automated process. No decision means no resources nor mandate to act on the insights.
- Finally, after a decision has been made, the next step is to plan an appropriate course of action and execute on it. It is possible to act badly even when receiving the right insights. One should have a strategy and be ready to learn from the mistakes of past decision. And this can also be done through data of course.
- Only following this approach, then we will be able to generate value from our data.

Let’s see how data an AI are working so nicely together.

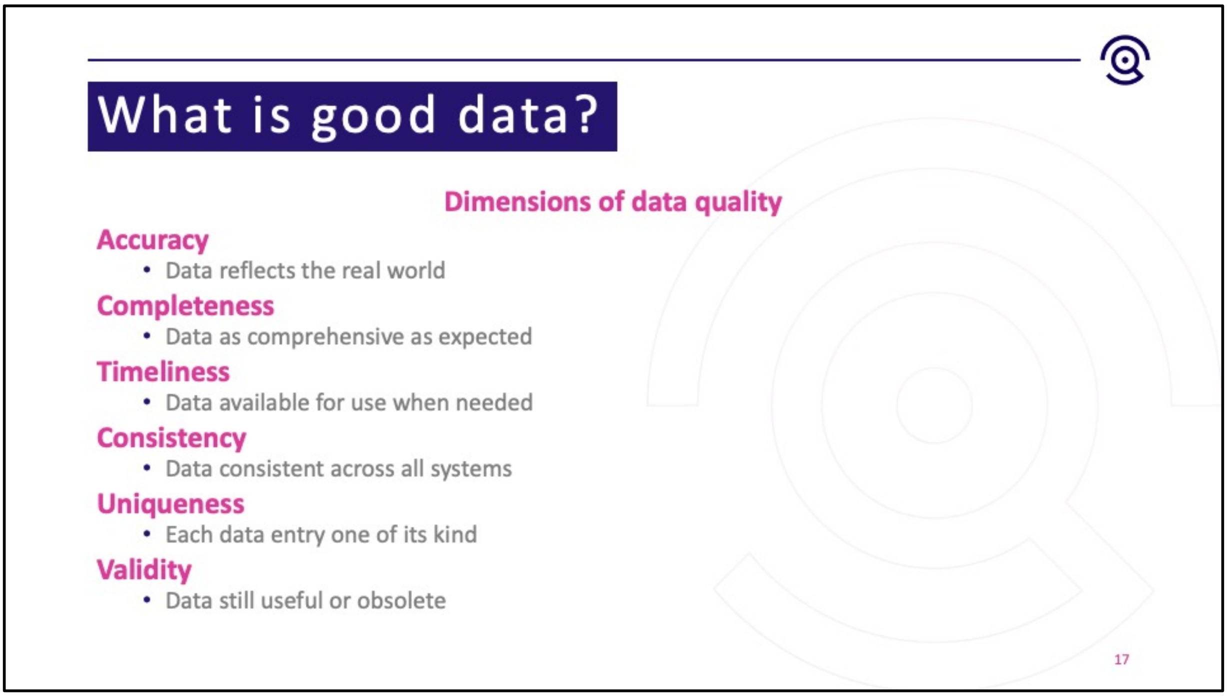


Big Data is everywhere nowadays, and by know it is expected to be here to stay. Similarly, it is also expected that AI and the demand for AI will only grow in the foreseeable future. More importantly, Data and AI are merging into a synergistic relationship, where AI is useless without data, and mastering data is almost impossible without AI. By combining the two disciplines, we can begin to see and predict upcoming trends in many different sectors, like business, technology, commerce, entertainment, and everything in between. It is undeniable that Big Data and AI have a synergic relations. Big data will continue to grow larger as AI becomes a more viable option for automating more tasks. On the other hand, AI will become bigger as more data is available for learning and analysis. Some fields of AI, like Natural Language Processing or Computer Vision, wouldn't even be possible without millions of samples, recorded and broken down into a format that AI engines can more easily process.



This is the AI hierarchy of needs. The general idea of the AI hierarchy of needs is that you should not move up the hierarchy until you've done the basics in the prior step. Many organizations are eager to start their data science journey, and with good reason. AI is present everywhere and there is proven evidence that it can generate value. However, some companies tend to fail in their quest to produce AI from data, and the reason is that that they tend to jump some steps in this pyramid.

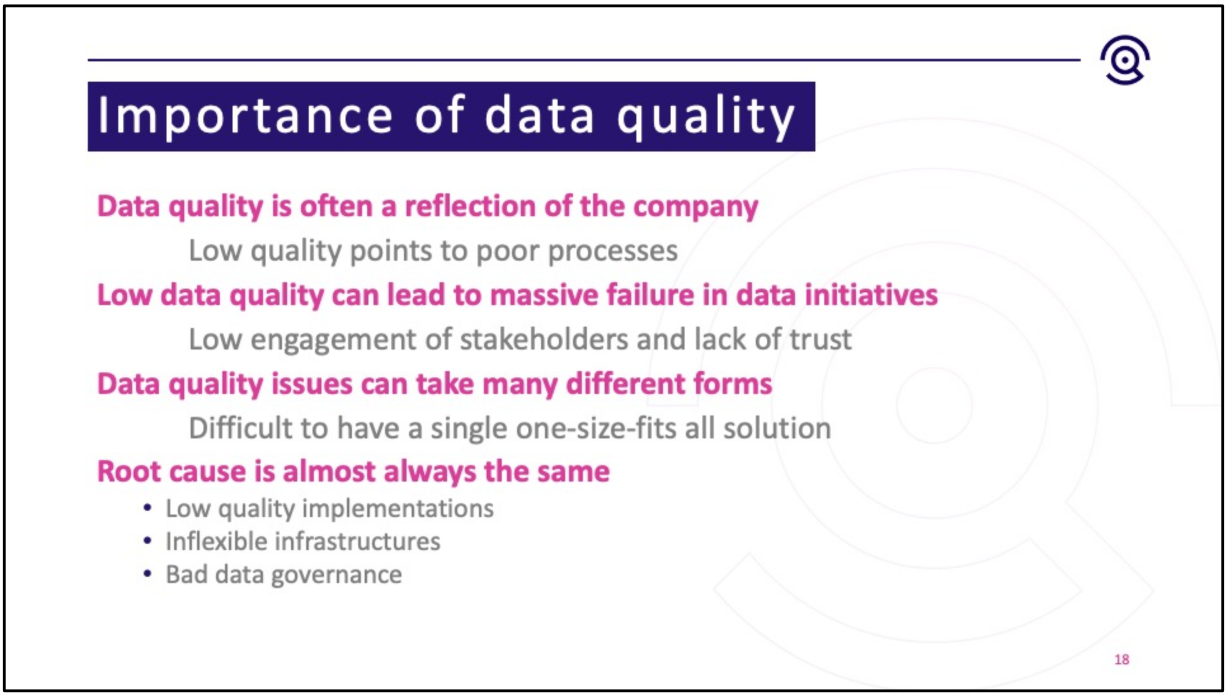
It is not the same to do a proof of concept with some static data, than trying to put a model into production. To succeed we need good data acquisition processes, robust and flexible infrastructure, well designed data pipelines, good data quality and appropriate data governance. It is true that there is no need to have everything figured out before starting with AI. However, the better the lower levels of the pyramid are implemented, the more likely we will be to produce reliable and trustworthy AI.



We know that AI needs high quality data in order to be able to succeed. That poses the question, what is good data for AI in the first place? We should be able to identify good data by looking at several dimensions of data quality:

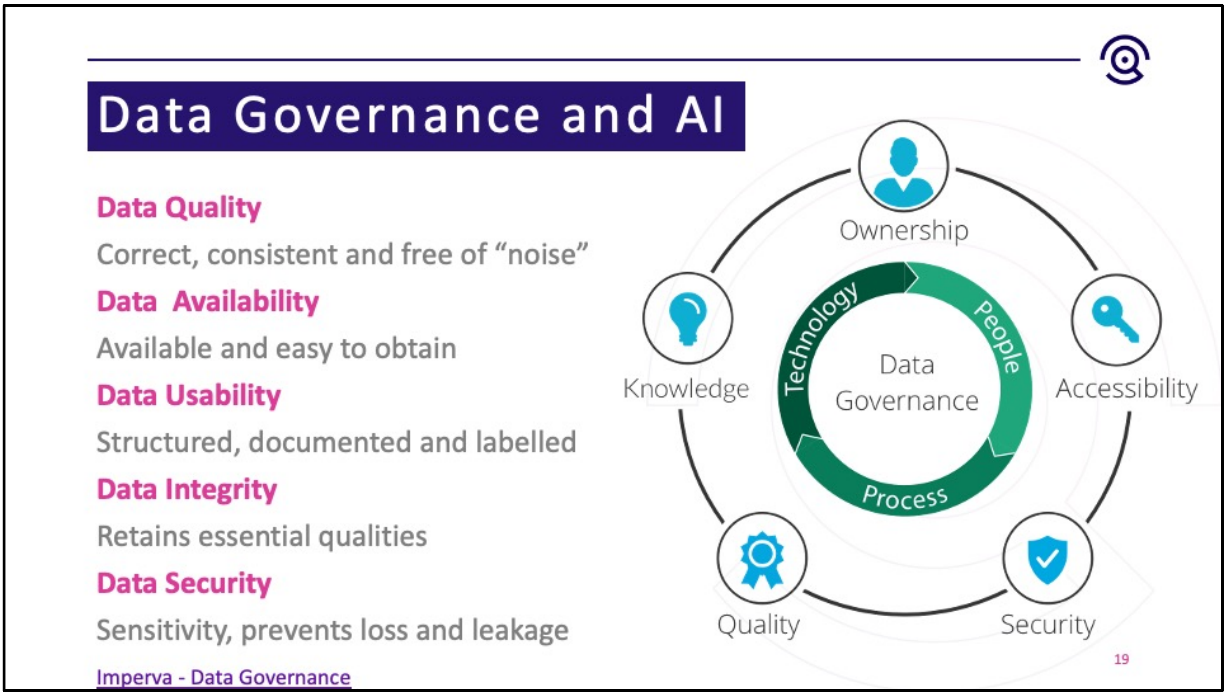
- Accuracy checks if the data reflects the real world. How accurately does the data reflect the event in question or the real-world object?
- Completeness checks if the data is as comprehensive as expected. For example, when someone gathers phone numbers, we expect it will include the area code.
- Timeliness checks if the data is available for use when needed. For instance, if we forecast on a daily basis, relevant new data should at least come at the same frequency.
- Consistency checks if the data is consistent across all systems. Do all systems across the data ecosystem contain the same information?
- Uniqueness checks if each data entry is one of its kind or if there are a lot of duplications.
- Validity checks if data is still useful or whether it has become obsolete over time.

- This shows that good data quality is a multi-faceted aspect with a lot of important dimensions to consider.



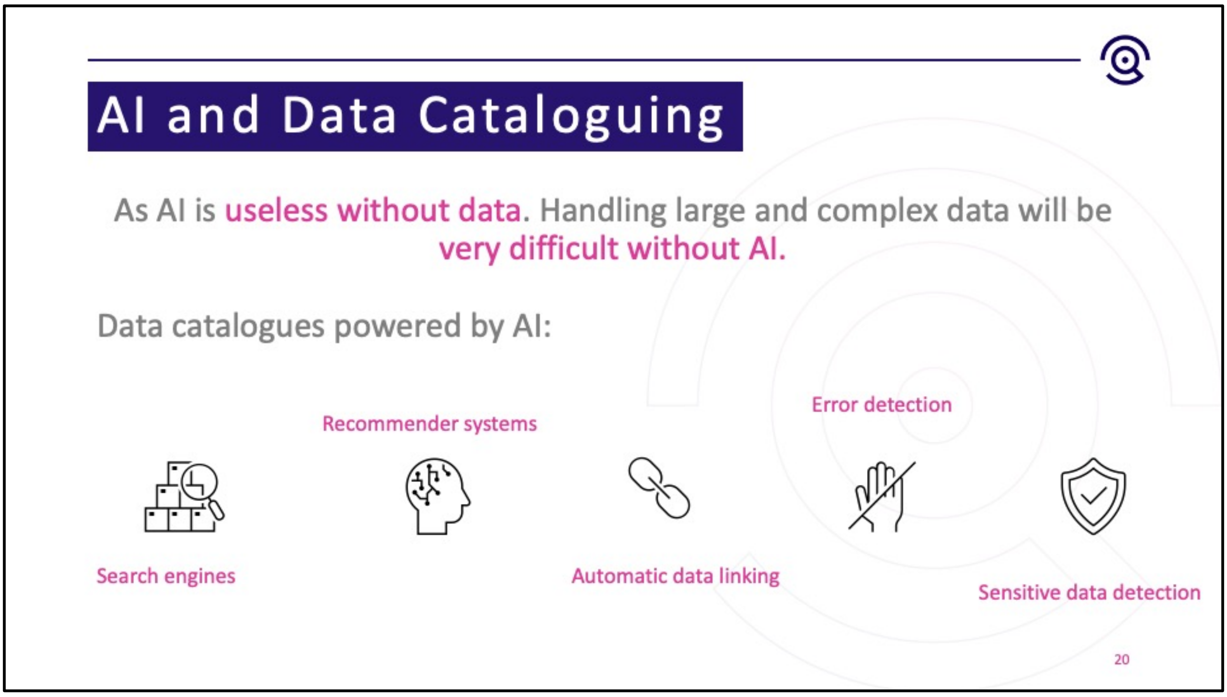
For a company to be data-driven and exploit the benefits of AI, data quality is a critical pre-condition.

- Data quality issues are known to explain limited trust in data from corporate users, waste of resources and poor decision-making. Often low data quality is nothing but a reflection of the company culture itself and it exposes the lack of good processes.
- Low data quality can lead to massive failures in data initiatives. In fact it is often the main reason behind such failures. A combination of low stakeholder engagement and lack of trust in data often leads to the abandonment of some systems, which in turn negatively impacts the original KPIs and success criteria.
- Data quality issues can take many different forms. It is therefore difficult to have a single one-size-fits all solution. Not having the right validation tools and processes in place for your data will make any data initiative you take way more difficult than it should be.
- The root cause for failure in data initiatives is almost always the same, namely low quality implementations, inflexible infrastructures and bad data governance.



Ensuring good data governance is the best (and maybe only) way to secure good data quality and to ensure that AI projects make it to production. Data governance is more than simply managing your data. It jointly manages the people involved, the technologies as well as the processes. It also ensures observability of the flow of data and enforces ownership over data, so there is always a responsible person at all times.

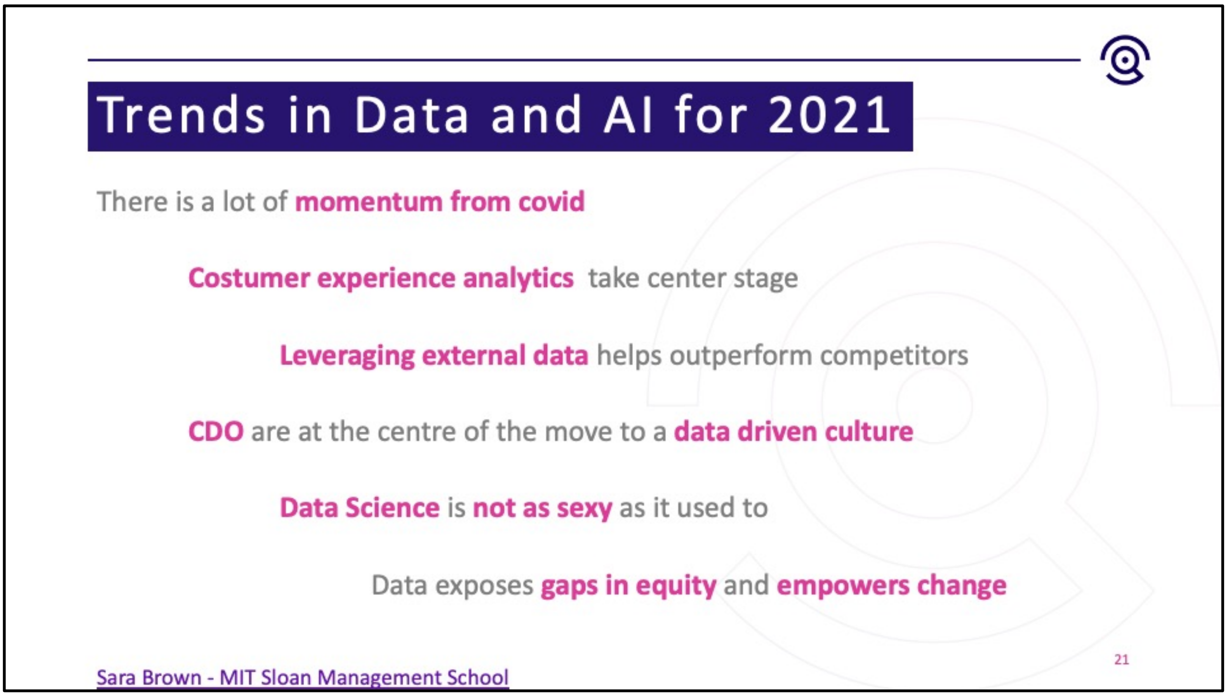
Governance guarantees the data accessibility, enforces security protocols to safeguard the data. More importantly, good data governance also ensures the data is of as high quality as it could possibly be. In short, data governance policies are guidelines that you can use to ensure your data and assets are used properly and managed consistently. Without good data governance it would become increasingly difficult, if not impossible, to deploy complex AI systems.



The synergy between AI and data is most often thought about in the direction from the data towards AI. However, the synergy also flows the other way around. AI has become key in the handling of complex data models. Enhancing data catalogues is one very good example. Modern data catalogues do more than just describe data, they are complex pieces of software powered by AI systems in order to enhance the user experience. Some examples are the following ones:

- Search engines, similar to a google search. AI powered search engines facilitate the data research. Tanks to metadata, we can potentially filter through thousands of data sources by just using a couple of filters and keywords.
- Recommender systems are also useful in this context. Having data sets recommended based on previous searches might speed up the search process. Moreover, a more intelligent catalogue should be able to suggest what kind of documentation a dataset should have based on tags and other metadata.
- Another use case is automatically recognizing and linking similar data sets which helps users visualize their data lifecycle.
- Using AI, it is also possible to automatically search for errors. Not only on the data's quality, but also the quality of its documentation.
- Finally, another place where AI has proven to be very is useful is helping data

professionals adapting to GDPR. Some smart catalogues are able to identify and flag sensitive data automatically.

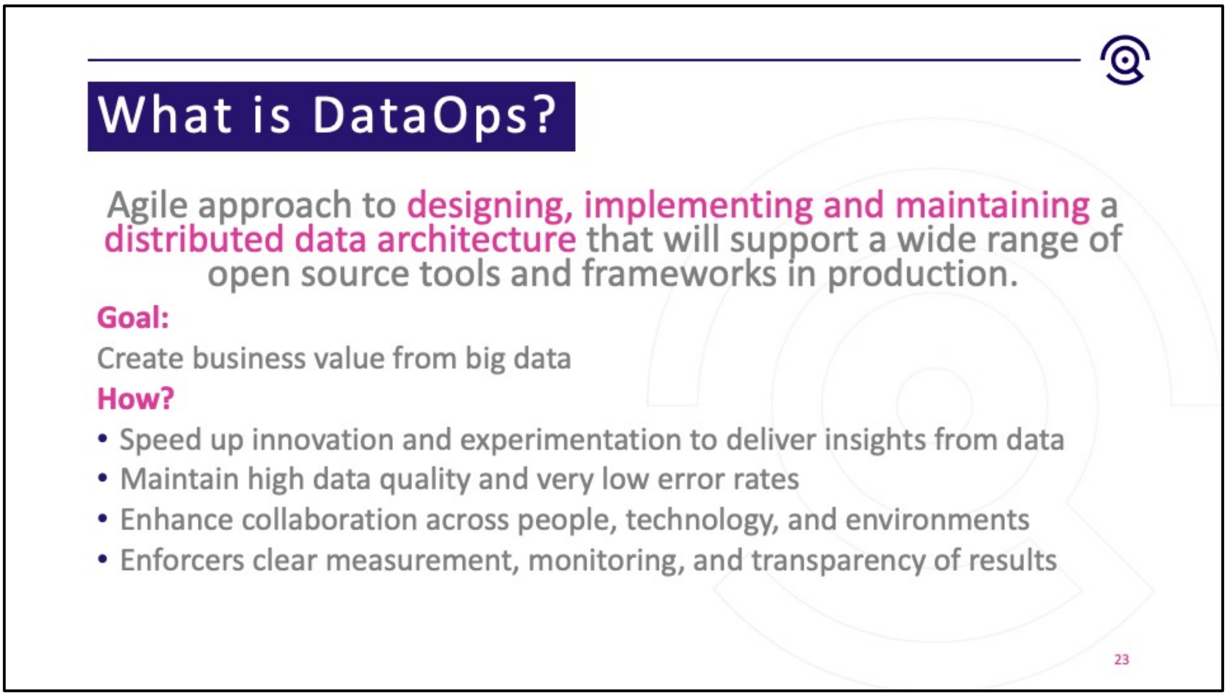


Cindi Howson, chief data strategy officer at analytics platform provider ThoughtSpot, puts forward the following six trends in data analytics. These should help us understand better how the relationship between AI and Data is evolving.

- Covid 19 forced companies to react and pivot in order to stay afloat. This helped to accelerate the adoption of a more data driven culture and opened up many organizations to embrace new things. Trying to keep that momentum is good and we need to keep three things in mind: Accelerate cloud migration plans. Use this disruptive time to challenge the status quo. And finally, focus on the first 1% of change. It is good to make long and detailed plans, but the most important part is to take the first steps towards it.
- Nowadays, businesses have many touch points with costumers where data collection is possible. Bringing all this data together so you have a holistic view of the customer is important. This is where new data technologies and principles can be leveraged by AI systems.
- Those who leverage external data outperform competitors by double digits. The range of data sources has exploded, combining internal knowledge and data with what's available out there can give you the edge you need. Imagination is the ceiling.

- One of Howson concerns is “How can we create a data-fluent world, or a data-driven organization, if people, business leaders, everyday decision-makers, are afraid of data?” The proliferation of Chief Data Officers (or CDOs) and their growing importance in the organization has helped mitigate these concerns. They are helping staff to upskill in data and are showing the business the potential value good data can have. CDOs lead the charge towards a data driven culture within their organizations.
- Some companies are finding that data and machine learning projects haven’t yet delivered returns and aren’t as impactful as they’d hoped. This does not mean it is not useful, but expectations need to be recalibrated. Especially more emphasis should be given to, not just the models, but to all the processes needed to support them.
- Finally, data exposes wide gaps in equity and also empowers change. So there is even hope that data will make this world a better and more fair place.

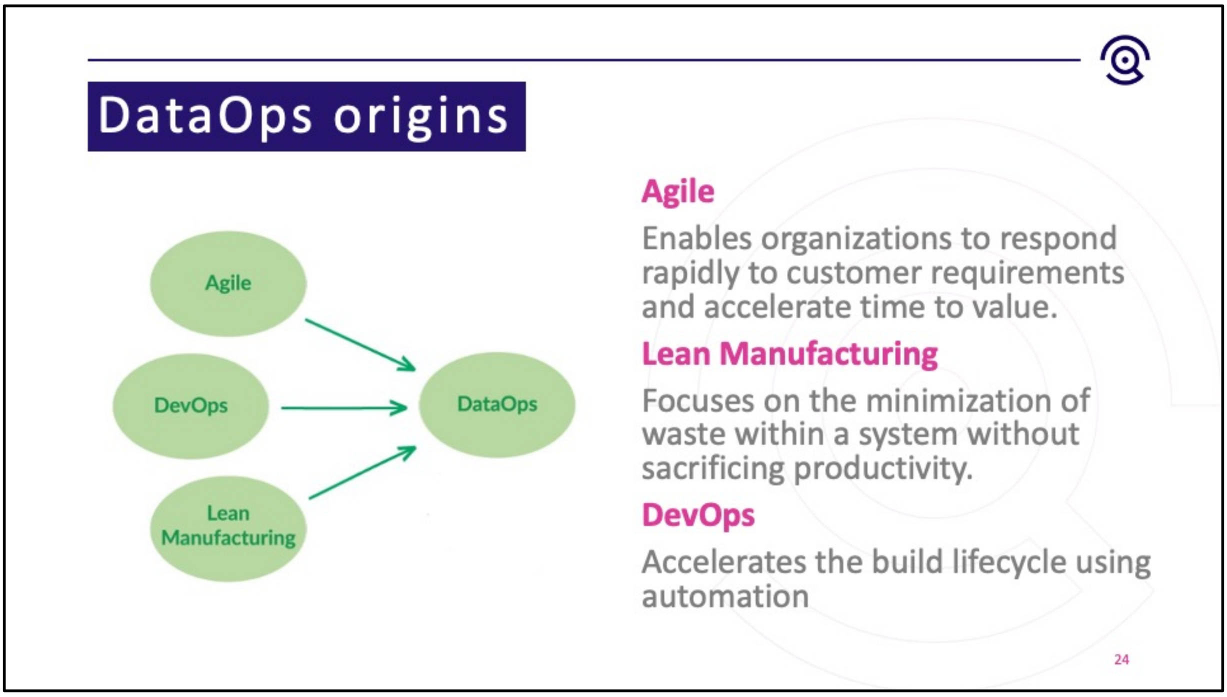
Let’s discuss DataOps, an approach to data that allows you to generate business value from it.



In this course, we have already discussed how data can help you realize immense value for your organization, for example through AI or BI. However, handling data to enable this is not a trivial matter. It is for this reason that having a solid data strategy is very important. But even more important is to have a methodology and the right set of tools to help you realize that strategy. This is where DataOps enters the picture. DataOps is nn Agile approach to designing, implementing and maintaining a distributed data architecture, that will support a wide range of open source tools and frameworks in production.

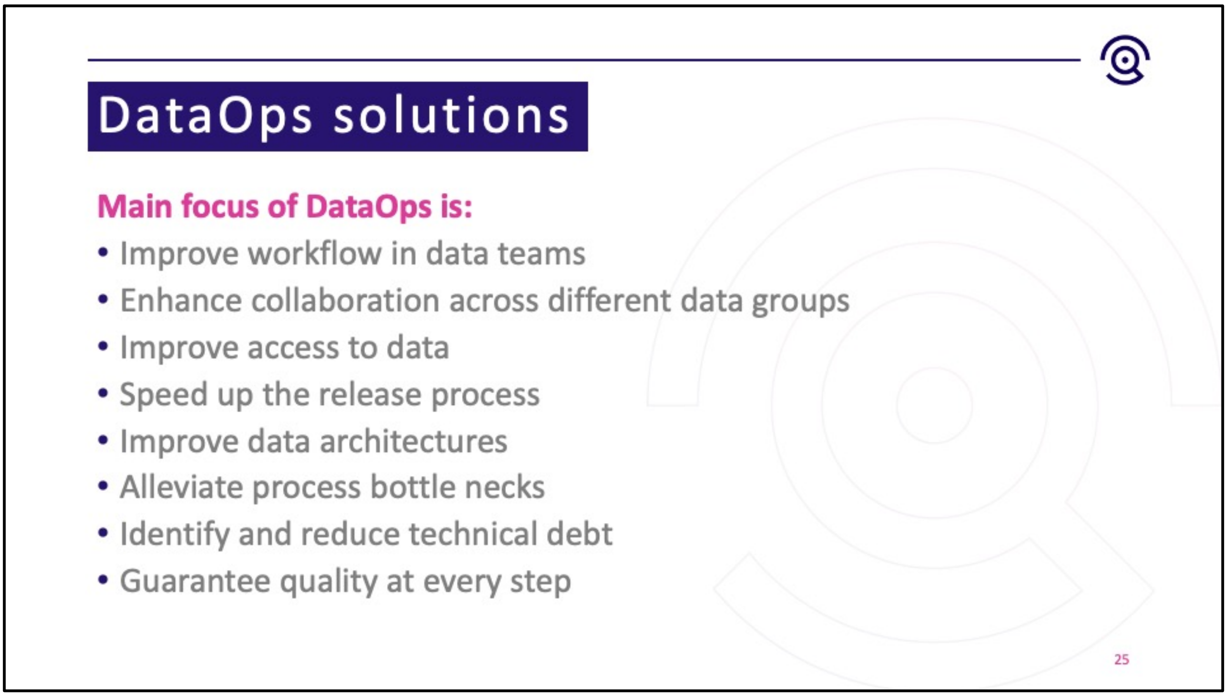
The main goal of DataOps is to help organizations generate value from data by helping it in the following ways:

- To speed up innovation and experimentation to deliver insights from data.
- Help to maintain high data quality and very low error rates.
- Enhance the collaboration across complex arrays of people, technology, and environments.
- And finally, enforce clear measurement, monitoring, and transparency of results.



Although DataOps is a very new concept, it borrows from some tools that have been around since World War II. Especifically from the manufacturing methodologies that helped revive the Japanese economy. These manufacturing techniques, known as lean manufacturing, are now being widely applied to software development and IT. DataOps brings them further into the data domain. Apart from lean manufacturing, DataOps also borrows heavily from both DevOps and agile frameworks and applies them to data analytics development and operations.

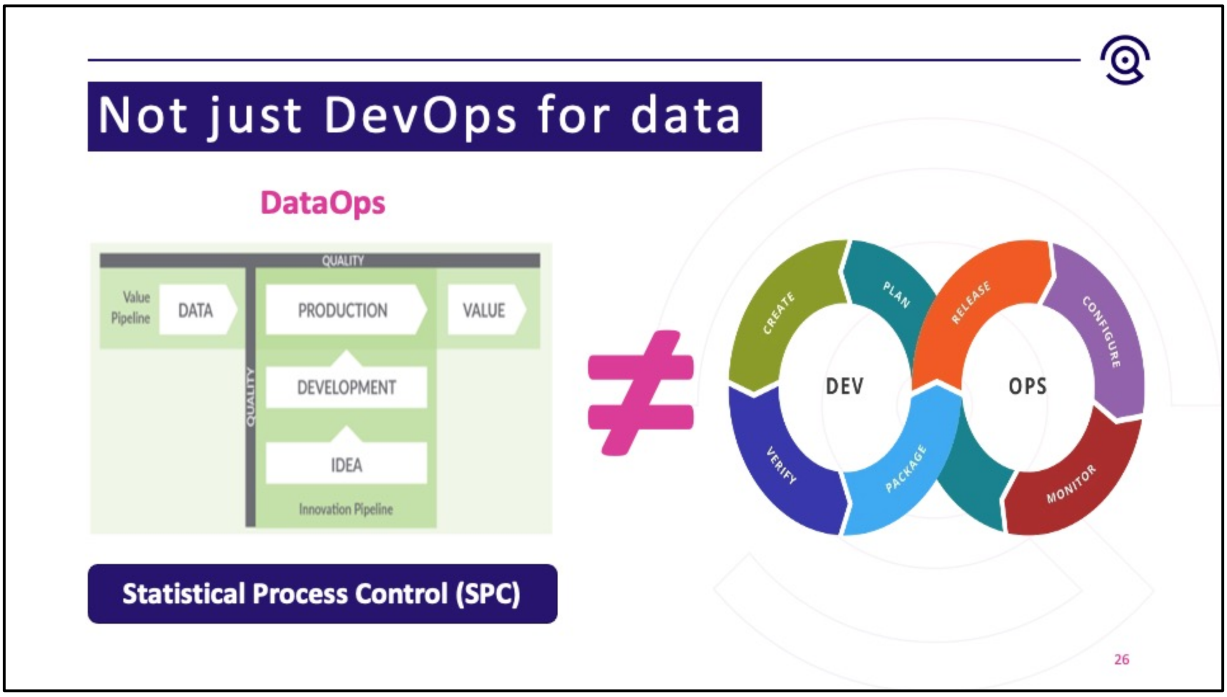
Being agile enables organizations to respond rapidly to customer requirements and accelerate time to value.
Lean Manufacturing puts focus on the minimization of waste within a system without sacrificing productivity. Moreover, lean manufacturing also brings focus on quality to data analytics, using tools such as statistical process control.
Last but not least, DevOps accelerates the build lifecycle, formerly known as release engineering, using tools for automation.



Before we go more in detail, perhaps it is good to understand what the type of problems are that DataOps tries to solve.

The main focus of DataOps lies on the following type of solutions:

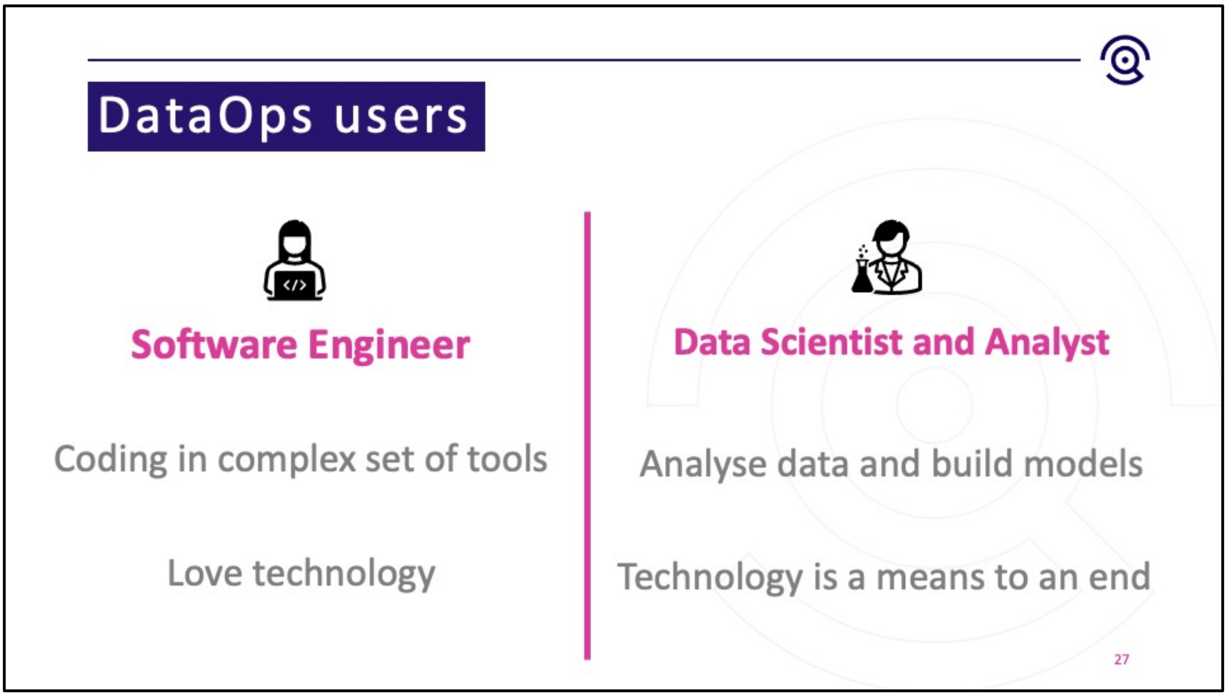
- Improving the workflow in data teams, as collaboration is not always easy.
- Enhance collaboration across different data groups, not only in the different organizational silos.
- Improving access to data, meaning not only timely access but also security.
- Speed up the release process through DevOps. Since data is delicate, many teams sacrifice quality over time when there is no DataOps.
- Improve data architectures, as inflexible data architectures are perhaps one of the most common problems in some large organizations.
- DataOps also aims to alleviate process bottle necks, identify and reduce technical debt and guarantee quality at every step by means of observability.



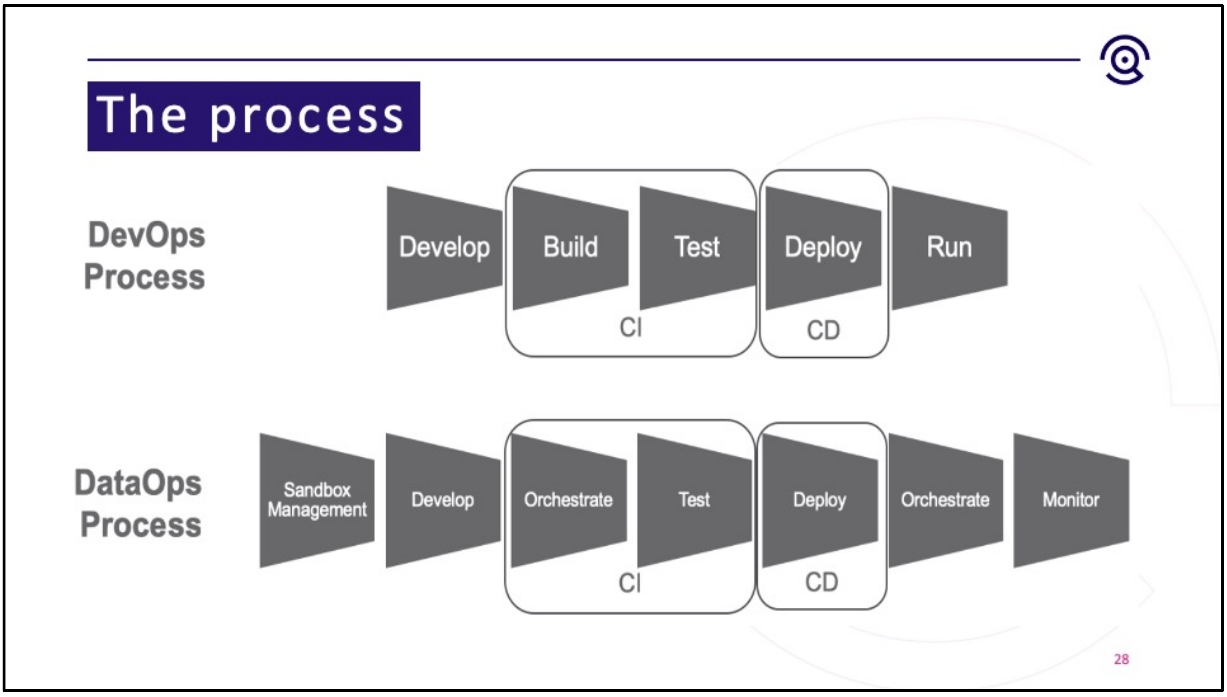
One of the most common misconceptions about DataOps, is that it is simply DevOps for data. In al fairness, it is true that having good DevOps processes and tools is fundamental for DataOps to succeed. But DataOps is more than that.

DevOps cycles are generally depicted as infinite loops, in which each step feeds a following step and that iteration keeps on going indefinitely. This is also true in DataOps, but this time we have two active pipelines that intersect each other, and each of them is also continuously iterating. On the one hand we have the Value Pipeline. This takes raw data sources as input, and through a series of orchestrated steps, produces analytical insights that create value for the organization. On the other hand we have the Innovation Pipeline. This is the process by which analytical ideas are introduced into the value pipeline. This pipeline's lifecycle resembles that of DevOps, but we will see later how the orchestration process adds a layer of complexity.

One tool coming form lean management that is worth having a look at is statistical process control or SPC. SPC measures and monitors data and operational characteristics of the data pipeline, ensuring that statistics remain within acceptable ranges. With SPC in place, the data flowing through the operational system is verified to be working at every step. If an anomaly occurs, the data analytics team will be the first to know, which enhances efficiency and improves the quality. By enforcing SPC, DataOps monitors the quality of data flowing through the pipelines



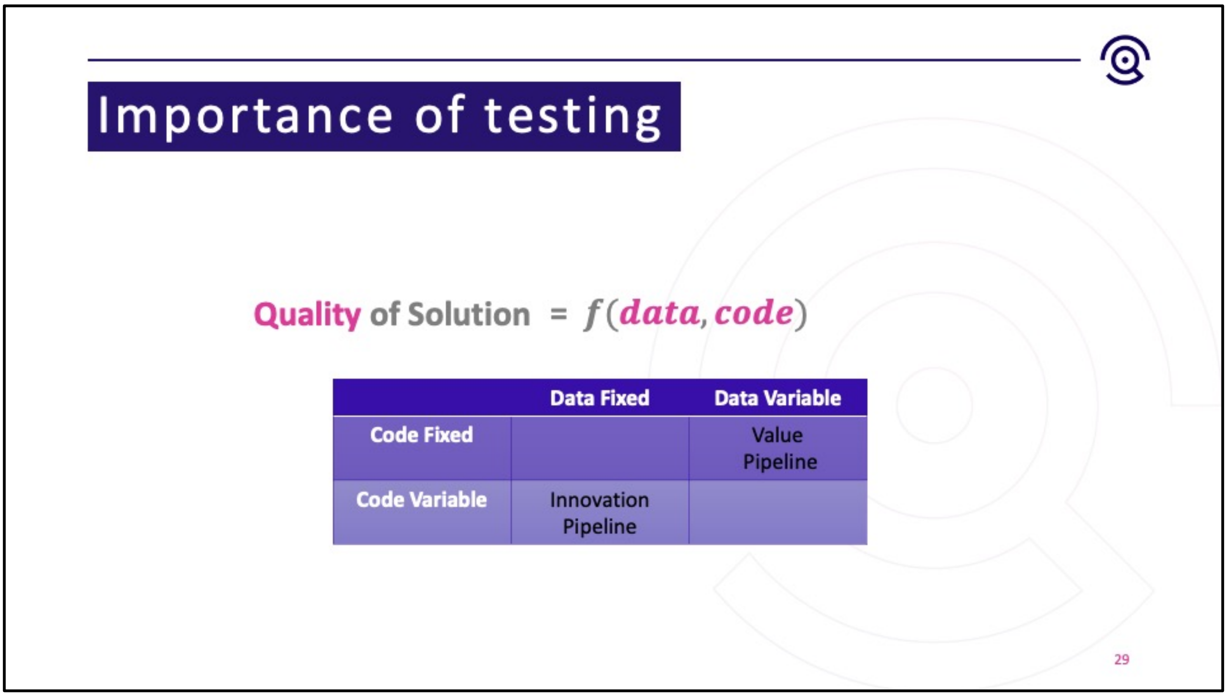
Another big difference between DevOps and DataOps are the expectations, which are set by none other than the users. While DevOps was created to meet the needs of software developers, DataOps is targeted to a different and more diverse group. Software engineers are comfortable with coding and the complexity of multiple languages, tools and hardware/software systems. They love coding and embrace technology. The requirement to learn a new language or deploy a new tool is often seen as an opportunity, not a hassle. Meanwhile, DataOps users are often the opposite of that. They are data scientists or analysts, who are focused on building and deploying models and visualizations. Scientists and analysts are typically not as technically savvy as engineers. The technology used to create these models and visualizations is just a means to an end. So for many of them, having to deal with more than two technologies can already feel like a big hassle. This adds an extra level of complexity, as DataOps accepts that data professionals live in a heterogeneous multi-tool world, and it seeks to make that world more manageable for them.



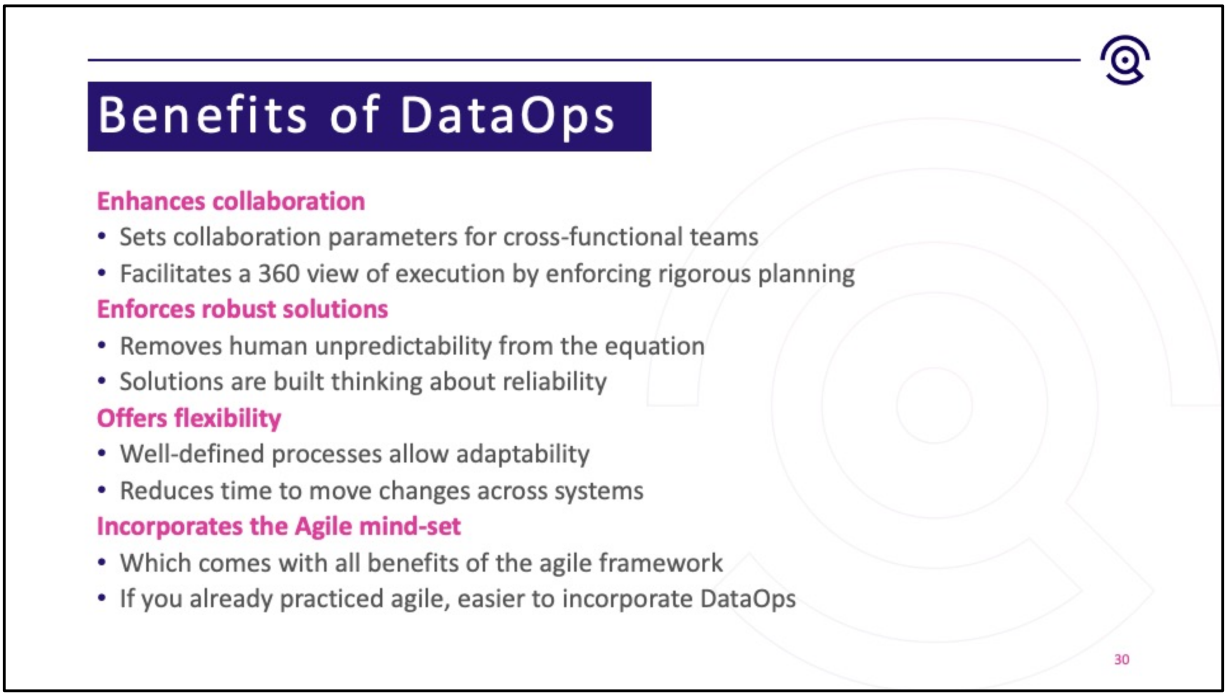
DataOps also has a slightly more complex process that DevOps.
In DevOps we have the following stages:

- Develop which creates/modifies an application.
- Build, which assembles application components.
- Test, which verifies the application in a test environment.
- Deploy, which transitions code into production.
- Run, which executes the application.

However, in DataOps there is also the need for orchestration. For example, the data factory might consist of a pipeline process with many steps. The "orchestrator" could be a software entity which controls the execution of the steps and handles the exceptions. These steps can be very complex in themselves and can have dependencies. Notice that orchestration happens two times, as orchestration of the data factory is also needed after deployment.

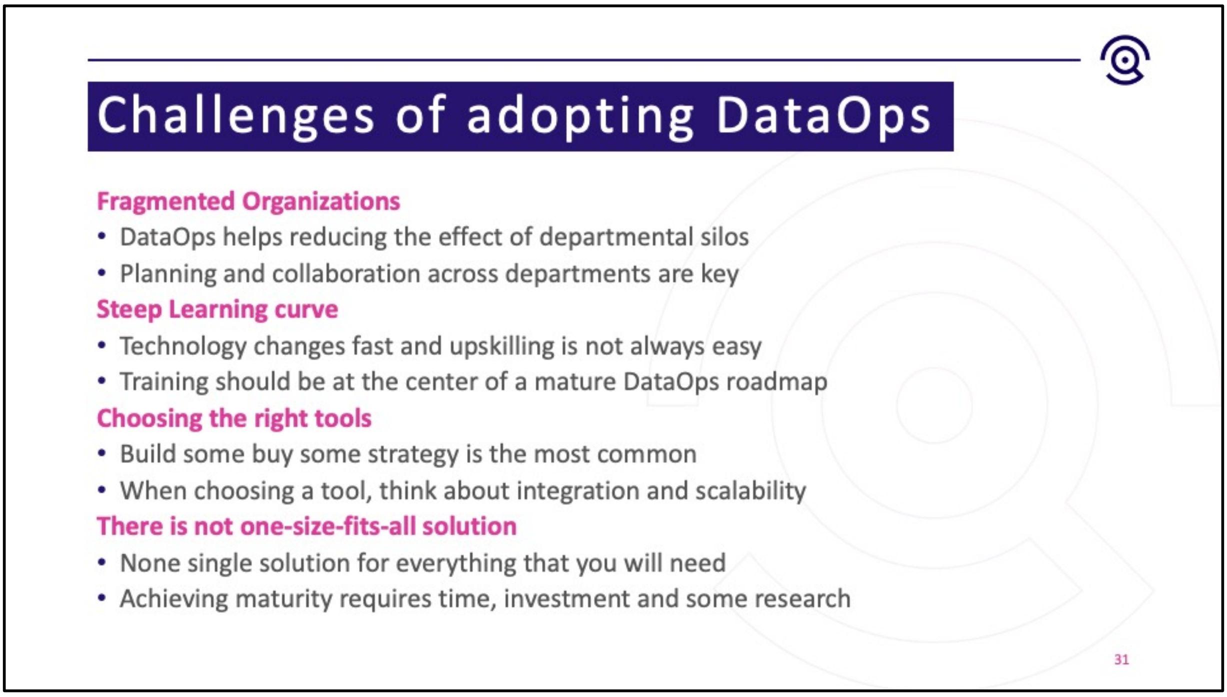


In DataOps, tests have dual value. In the Value Pipeline, tests monitor the data values flowing through the data factory to catch anomalies or flag values that are outside of the expected parameters. In the Innovation Pipeline however, tests validate new analytics before deploying them. This exposes how having these two separate pipelines adds an extra layer of complication. The final quality that the customer receives depends on both code and data. We should therefore test not only the joint behaviour, but also the behaviour of holding one fixed while letting the other one vary. In the Innovation Pipeline, code is variable and data is fixed, as these tests target the code and not the data. On the other hand, in the Value Pipeline, code is fixed and the data is variable. Data that flows through the Value Pipeline is variable and subject to statistical process control and monitoring. Therefore, these tests target the data which is continuously changing, and not the code.



DataOps offers many benefits, at many different levels, but there are four high level benefits that stand out.

- First, it enhances collaboration. DataOps sets collaboration parameters for cross-functional teams. We know that a data strategy incorporates many different profiles across the organization, so collaboration is key. This facilitates a 360 view of execution by enforcing rigorous planning, because it will be impossible to move forward otherwise.
- Second, it enforces robust solutions. DataOps removes human unpredictability from the equation and puts focus on building automated and reliable processes. There is always uncertainty remaining of course, but DataOps helps mitigate them as much as possible.
- Third, DataOps offers flexibility and the capacity to quickly adapt to changes. Well-defined processes imply that changes require fewer hands and less time, with a lower probability of introducing errors. This reduces the time needed to move changes across systems, from development to production.
- Fourth and final, DataOps incorporates the Agile mind-set. Which comes with all benefits that the agile framework has to offer. As an added bonus, if you already practiced agile before, then it becomes easier to incorporate DataOps as well.



There are many challenges when trying to adopt a DataOps strategy. Managing data is a very complex task, and small errors can propagate across the whole organization. Here we list four bottlenecks that can pose a big challenge for an organization in the road to DataOps maturity.

- First we have fragmented organizations. While DataOps aims to reduce the impact of departmental silos, their very existing poses a huge barrier to maturing its processes. Planning should not be underestimated and discussions should be open to allow input from contributors. There lies difficulty in organizing such a diverse group of parties with different end goals and backgrounds, making processes take more time than expected.
- Second, there is a steep learning curve and often the lack of required skills. Lack of skills can present a barrier to implementing a proficient DataOps strategy. As technology keeps changing at a fast pace, many data professional are required to adapt as they go. Since most of these professionals tend to work in high-stress environments, taking time to proactively build skills isn't always an option. For this reason, training should be a key component when drafting the roadmap to DataOps maturity.
- Third, choosing the right tools is not easy. Most DataOps solutions do not rely on a

single tool, nor on a build everything in-house strategy. It is more of a build some and buy some situation actually. The important thing to keep in mind when choosing the right tools are the aspects of integration and scaling.

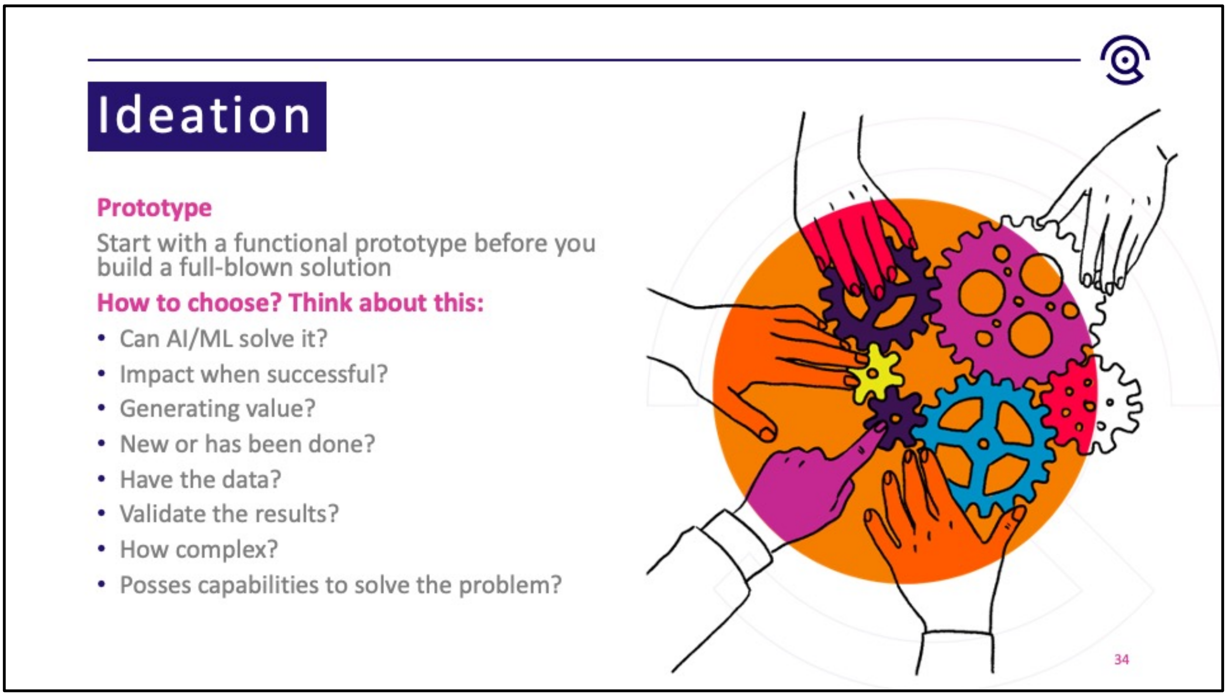
- Fourth and final, there is not one-size-fits-all solution. Even if you stick to the same vendor, there is still no one single solution for everything that you need to tackle when building your DataOps strategy. Achieving maturity requires time, investment and research. And even some finesse.

Let’s end with some tips and tricks on how to capture value with AI.

We have talked about how useful AI and high quality data can be for modern organizations. However, if your goal is to generate value from these resources and tools, probably the most important step is to select the right use cases. It sounds like a simple endeavour, from having a great idea, to collecting feedback, prototyping and validating the concept. Nonetheless, it is common for organizations and AI experts to fail while trying to develop AI systems. Apart from the technical challenges, it is also possible that the AI systems fail to meet expectations, like for example not meeting original KPIs or not increasing the revenue.

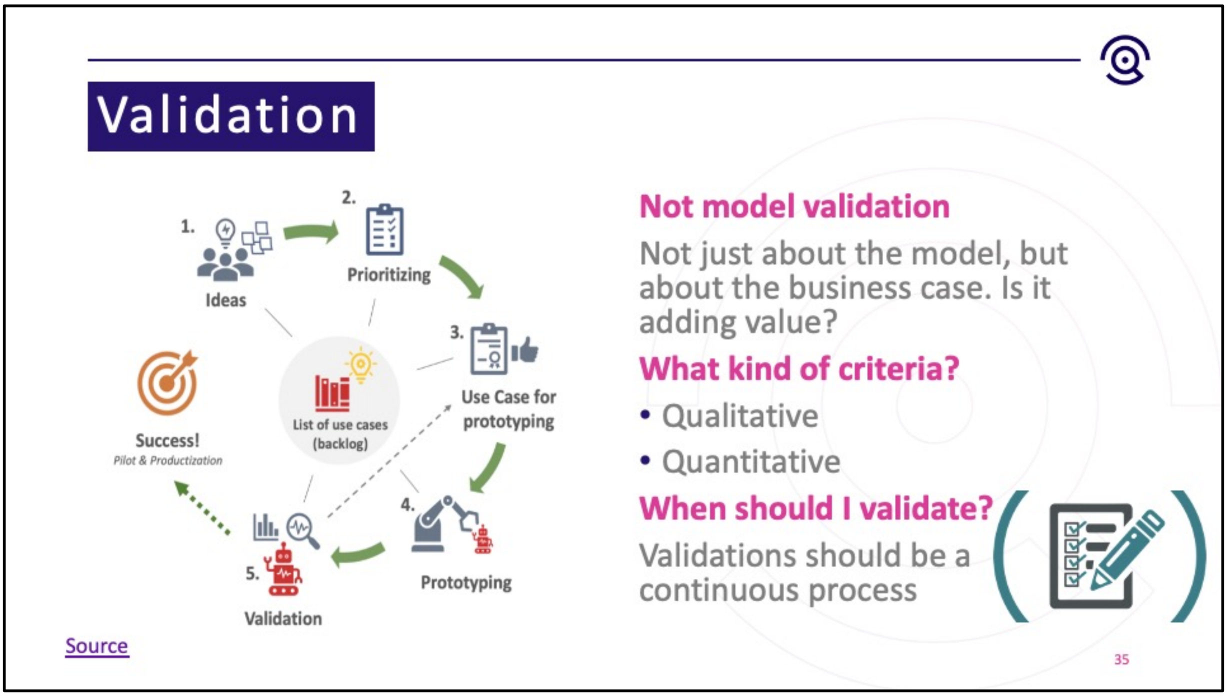
It is important that, before you can adopt a machine learning solution, you must learn how to identify the right machine learning use case for your business. Sometimes, ML might not even be the right solution for you. To truly benefit from machine learning, you need a clear understanding of your business problems. So what are your needs, as well as your resources. These problems should be actual business problems, with the possibility to generate value. Machine learning can solve very interesting problems, but one thing you should always consider is: Are these problems relevant for my business?

If you are starting your AI journey, be sure to find low effort high reward cases. Try identifying problems you can actually solve with your current capabilities, but that will also generate value of course. Remember that Machine Learning is expensive and complex. So for you to justify a machine learning investment, your business problems should be impactful. Planning is necessary, as it will help you course correct early if necessary and even realise if ML is necessary at all for the problem at hand. Finally, having a validation and ideation strategy is key. Let's discuss these a little bit deeper.



Data Science is a very creative discipline. Most times solutions are not straightforward and there are no quick recipes. However, if left unchecked, creativity can become a problem, because creativity is the death of productivity. Therefore, you should definitively invest in creativity and research when ideating your solutions, but you should also add a structure to it. To avoid big losses, you need to prototype before you start building a full blow solution. But more importantly, even before you think of prototyping, you should have a good sense of why this solution and not another idea. Also focus on the challenges that lie ahead. To help in this whole ideation process, you can think about the following questions:

How can AI solve my problem?
What is the impact that this solution will have on my users?
Is there any tangible or even intangible value?
Is my idea unique? How much research is involved? Is something similar available?
Is data available or is there the need to generate/collect it for prototyping?
Would it be easy to validate or can I even validate it?
How complex will it be? Do I posses the capabilities to solve such a problem?
How difficult will it be and how long will it take to release?



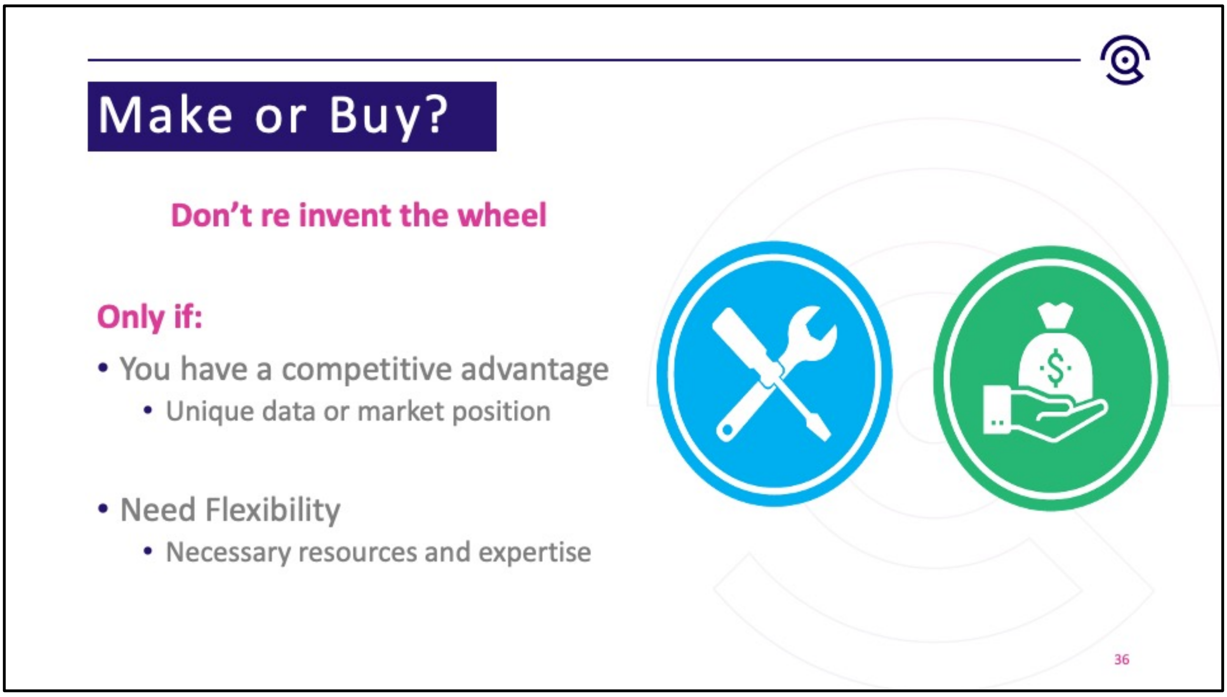
Validation is one of the more important, but often neglected parts, of selecting the proper use case. And when we talk about validation in this context, we do not mean model validation. That's just a part of the story, we mean validation of the whole business idea. Remember, at the end of the day, the most important question to answer is if this solutions is adding business value or not. In order to validate this question, we should have two types of criteria.

First, we need quantitative criteria. This should be a neat and clean way of measuring the performance of the ML system. And this should be generally easy, as often just the error threshold of the model is used. Even if it is a combination of performance metrics related to the prototype, this will not be particularly challenging to calculate.

On the other hand, qualitative criteria are certainly the most challenging. As there are representing the aspect that we just can't directly measure, like for example the value for the users. It is very important that these criteria are well thought out and representative of your end goal. However, remember that perceived added value is what in the end will matter the most.

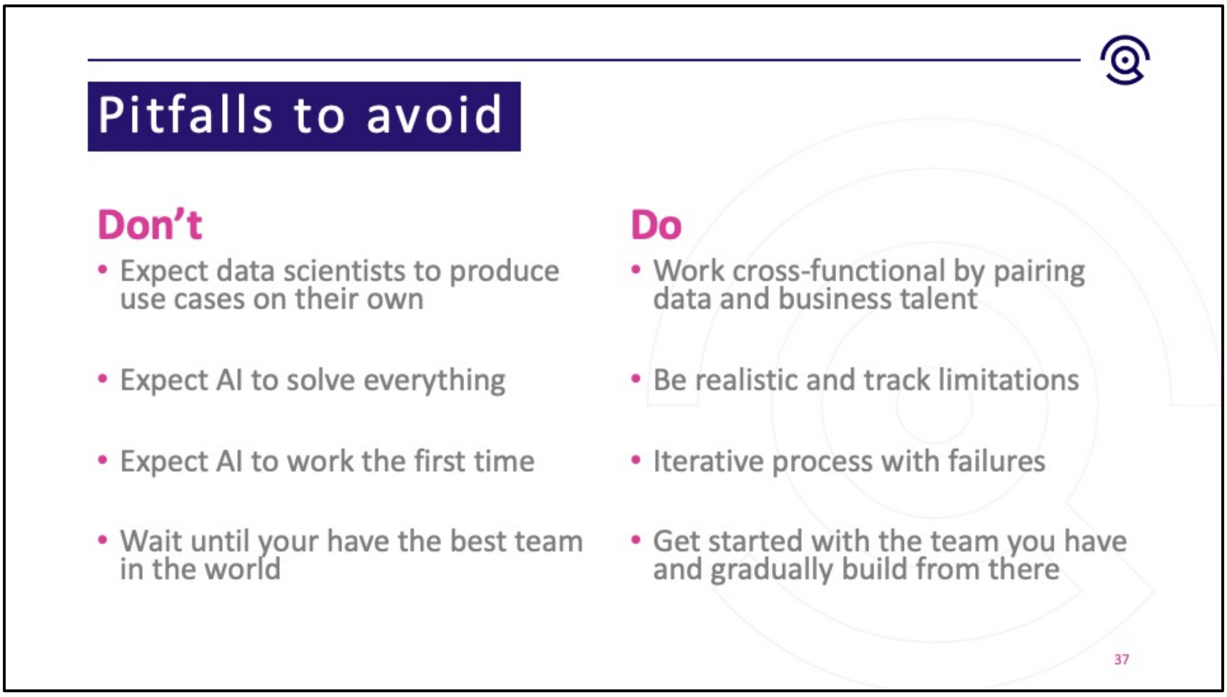
Another question that is often asked by the business is: when should we validate?

The simples answer is: as soon as you start prototyping. Validation should be a continuous and iterative process. If you wait too long before you validate, you run the risk of wasting resources that could have been useful somewhere else.



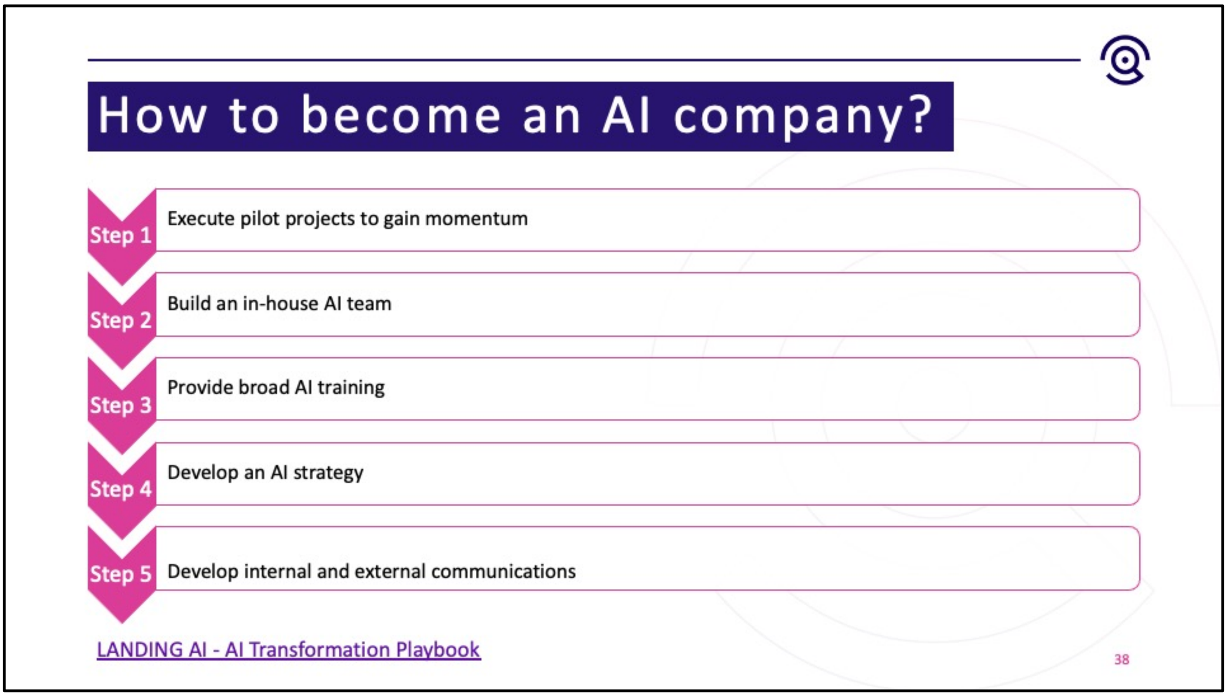
A recurrent questions when working in the field of AI and Data, is wheter to make or buy. Not necessarily about buying the whole tailored solution to the use case, but maybe only the technology necessary to solve it. Sometimes the choice is obvious. For example, a very small company will probably not want to build a full blown custom data platform from scratch. But the choice is of course not always so obvious. Sometimes, when too much customization is needed, making things from scratch might be the best solution. But most of the times, existing of-the-shelf solutions are going to be enough for most of the use cases. Or at least a hybrid situation in which we can smart copy some pre-existing solutions. So don’t spend your time reinventing the wheel.

Some exceptions exist of course when it comes to making your own solutions. For example, when you have a marketable competitive advantage, like unique data or a unique market position or expertise. Or sometimes existing solutions do not integrate well with your systems and it might prove easier to build from scratch than to customize something existing. However, this is only possible if there are the right expertise and resources available in-house. Which unfortunately in many cases is not true.



Here we list some AI pitfalls to avoid in order to find feasible and valuable business cases.

Do not expect your team of data scientists to come up with use cases on their own, but really work with cross-functional teams in the ideation phase. This allows to pair both data and business talent, resulting in much more realistic expectations on what the project can or should deliver.
Do not expect AI to solve all your problems, but be realistic about what AI can and can't do for you. Always keep in mind the limitations of data, technology, resources, engineering, and your own capabilities.
Do not expect AI to work smoothly from the very first time, but prepare for an iterative process towards the development of valuable applications. Multiple attempts will be needed and failure is a given.
And finally, do not wait until you have the best AI team in the world to start your AI journey. It is important to just get started with the team you have, and gradually build from there depending on the needs that develop over time.
Start small and keep on learning along the way.



LANDING AI published the AI transformation playbook on how to lead your company into the AI era. According to them, “The playbook draws on insights gleaned from leading the Google Brain team and the Baidu AI Group, which played leading roles in transforming both Google and Baidu into great AI companies.” Not too bad right? The play book outlines five steps to becoming an AI company that can capture value from it.

- Step 1 is the execution of pilot projects to gain momentum. Do a successful initial project, not necessary the most valuable one, but one with a high success chance. The goal here is to make the wheel spinning and generate quick results (observable in 6-12 months for example). These first projects can be possibly outsourced, this will still allow you to learn and convince the business of AI’s potential value.
- Step 2 is to build an in-house AI team. The best approach is to go for a centralized AI unit that is deployed in business units. Not only the people should be centralized, but also company-wide tools and platforms for software and data are recommended.
- Step 3 is to provide broad AI trainings. The content should differ depending on the profile you are teaching. Executives and senior business leaders should learn what

AI can do for your company, understand the basics of AI strategy and should be able to make resource allocation decisions. Business unit leaders working on AI projects should learn how to set project directions, both on technical and business diligence fronts, but also to track and monitor progress. Data scientists or ML engineers should be skilled to collect data, develop AI software and deploy these in production. There is a lot of good content online, so curate rather than create.

- Step 4 is to develop an AI strategy. Decide how to leverage AI in order to create a competitive advantage specific to your industry. It might seem counterintuitive that this in only step 4 out of 5, but you first need to understand AI and how it can help your business, before formalizing the strategy. At this step, also consider creating a data strategy. And align your AI and data strategy. Always remember that a better product results in more users, which gives more data, which creates better products again and so on.
- Step 5 is to develop internal and external communications. AI is likely to change your product or service, so communication about this is important. Clear communication should be done to all stakeholders, for example investors, government, consumer, user, talent, recruitment and employees.

- Following these steps gives you a high chance on capturing the value of AI within your business.

This is the end of the third module. Hopefully you now have a better view on the value that AI can bring to your business and how to continue to actually capture it. I hope to see you again in the next module, where we will bring this value to the real world, bye.