WHY DO ANALYTICS AND AI PROJECTS FAIL?

Whitepaper Series
August 2024

CENTRE FOR BUSINESS ANALYTICS

THE UNIVERSITY OF MELBOURNE

MELBOURNE BUSINESS SCHOOL

EXECUTIVE SUMMARY

Despite the hype, interest and AI frameworks, research shows that over 80% of all data science projects still fail[1]. Everyone has their opinion as to what causes failure but a detailed study analysing the real causes and then providing practical, field-tested solutions has been missing – until now.

There are hundreds of reasons as to why data science, AI and data analytics projects fail – knowing where to start can be overwhelming. However, by adhering to the 80/20 rule and focusing on the primary issues causing most failures, this white paper distils the problems down into four main themes and provides recommendations on how to overcome the endemic failure in the industry and begin succeeding with AI. The key recommendations for each thematic area are:

- **Strategy:** *Find those strategically important projects for which data science and business analytics can help accelerate the strategy.*
- **Process:** *Make sure you've got the right data to solve the problem at hand and don't expect any miracles if your data is no good.*
- **People:** *If you don't have the right people or people don't support an analytics approach, then don't even try.*
- **Technology:** *Invest in the right types of tools to solve problems that are known to have solutions you can implement.*

Each of these recommendations seems obvious, so why do the problems keep occurring? In this whitepaper we'll explore the research on this topic, present a case study illustrating the challenges and provide a more detailed explanation of these solutions.

The lessons learned here will assist you in boosting your success rate and confidence in using data to make decisions.

1 See the forthcoming book, Why Data Science Projects Fail by Doug Gray and Dr Evan Shellshear

The information presented here is extracted from the forthcoming book *Why Data Science Projects Fail* by the Global Director of Supply Chain Analytics at Walmart, Doug Gray, and Managing Director and CEO of Ubidy, Dr Evan Shellshear.

About the Author:

DR EVAN SHELLSHEAR

Dr. Evan Shellshear is the Managing Director and Group CEO of Ubidy, an innovative global recruitment marketplace that leverages AI to connect employers with specialist agencies. He holds a Bachelor of Arts and a Bachelor of Science (single and double major in mathematics) from the University of Queensland. Evan earned his PhD in Mathematical Economics (Game Theory) at the Nobel Prize-winning Institute of Mathematical Economics at the University of Bielefeld.

With over fifteen years of experience, Evan has built global AI and statistical algorithms for a wide range of industries, including manufacturing, retail, healthcare, supply chain, oil and gas, and energy. He has authored or co-authored four books and has served on multiple advisory boards at state, national, and international levels. Evan holds multiple accreditations across various digital platforms and is a recognized thought leader in AI and innovation, having published nearly 100 articles in leading blogs, magazines, and news outlets. He is currently an Adjunct Professor at the University of Queensland and Queensland University of Technology, where he teaches business analytics and AI strategy.

About THE CENTRE FOR BUSINESS ANALYTICS

CENTRE FOR BUSINESS ANALYTICS

Melbourne Business School has a proud history of advancing quality business education in Australia. The school is home to Australia's first MBA program, launched in 1963 and also the Master of Business Analytics, which is ranked 15th in the world by QS and recognised as the top program in Asia and Oceania.

The Centre for Business Analytics is at the forefront of data-informed decision making, uniting scholars, practitioners, students, and organisations driven by the challenge of leveraging data for organisational success. Since our inception in 2014, we have helped many organisations solve business challenges using data and furthered the data culture and maturity of many businesses. We have transformed the data culture and maturity of organisations through executive education, student practicums and talent acquisition, research and thought-leadership events.

Connect with the Centre for Business Analytics

cfba@mbs.edu

CONTENTS

WHY ARE DATA SCIENCE AND AI PROJECT FAILURES SO HIGH?

The rise of data science marks a significant shift in how organisations operate and make decisions. As businesses and institutions accumulate vast amounts of data, the ability to analyse and interpret this information has become crucial to gain a competitive advantage.

AI projects, which encompass data collection and storage, processing and analysis, and then tool development and deployment, enable organisations to uncover insights that drive strategic decisions, improve operational efficiency, and enhance customer experience.

While advancements in big data technologies, machine learning, AI and data science are leading some organisations to make better, data driven decisions, at the same time, the incredible hype surrounding it is leading many more to make poor investment decisions. Why is this the case?

To delve into the reasons behind these failures, we collected the necessary data in three ways:

1. We looked at what fellow practitioners wrote in blog posts, white papers, podcasts, videos and similar outlets, reviewing more than 100 pieces of content.
2. We analysed more than 2000 peer-reviewed articles published in scientific journals and conference proceedings related to the topic.
3. We interviewed and talked to dozens of world leading practitioners to hear their stories.

Although the research is very broad, upon analysing the comments and experiences of others, certain themes clearly begin repeating themselves with a long tail of other sporadic issues. One of the biggest differentiating factors from the data is the performance difference between analytically mature organisations that are experienced in delivering data science projects, and analytically immature organisations. Mature organisations were observed to produce double the successful outcomes than their immature counterparts.

Analytically mature vs analytically immature organisations

There are organisations that are thriving in these data driven times and others that are failing miserably. So, what are some of the characteristics of the analytically mature AI overachievers compared to their analytically immature peers?

To answer this question, we borrow ideas from the many, well-known frameworks for data and AI maturity which provide benchmarks for where an organisation sits on a scale of AI capabilities. They can each be summarised as follows:

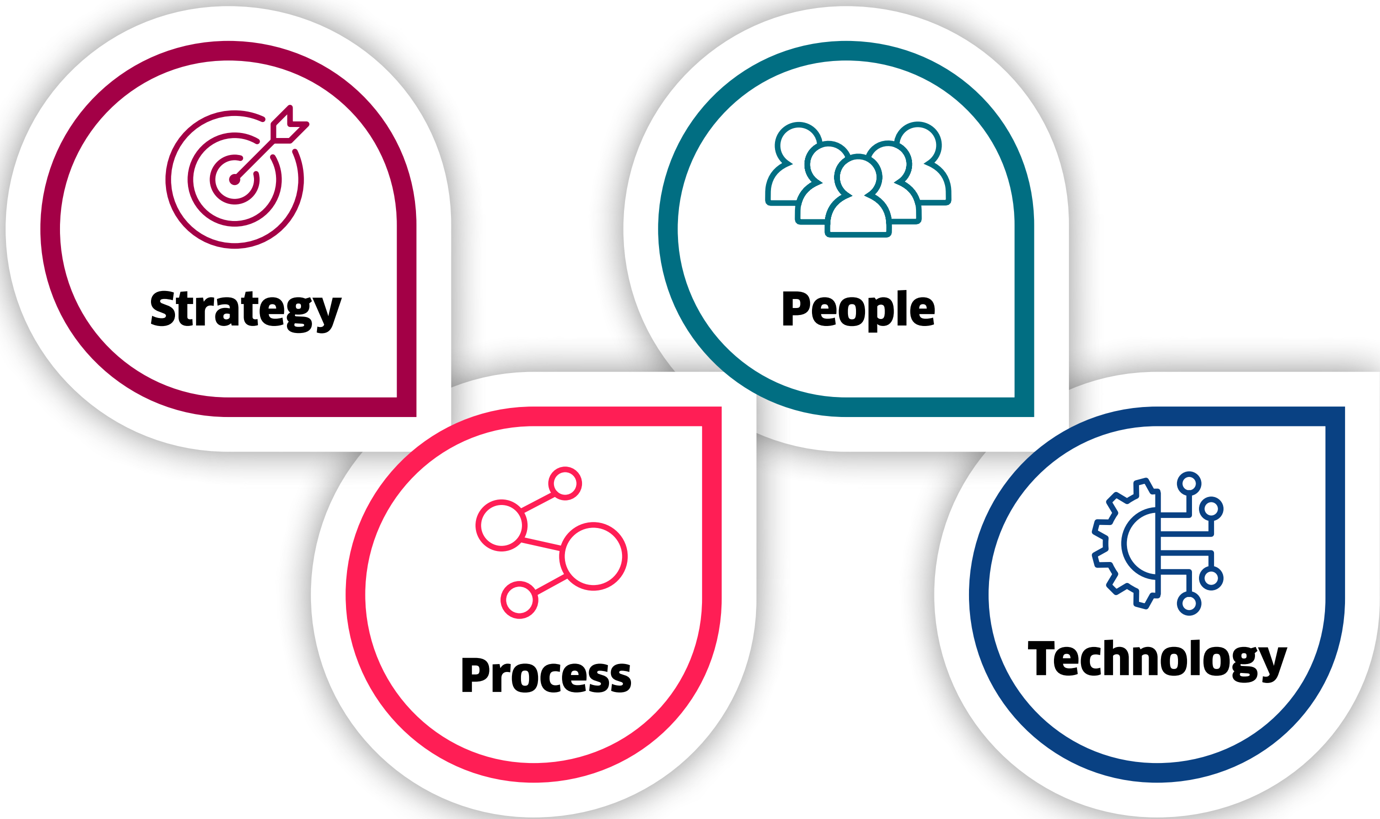
- An analytically *immature* organisation is typically one that lacks most of the requirements to deliver on analytics projects, such as high-quality data, in-house analytics capability and senior stakeholder buy-in. They typically will try and tackle AI, data science and data analytics projects as isolated activities that are not connected to the organisation's strategy. They often will encounter the issues we describe in this whitepaper.
- Analytically *mature* organisations are ones that don't typically suffer from the failures we'll discuss in this whitepaper. Usually, there will be senior stakeholder buy-in and a desire from the top, as well as a strategy to leverage analytics to improve decision-making. They generally have the required human and technological resources and have experience applying analytics and realising at least some commercial benefits from these activities.

As we progress through the whitepaper, it will become apparent how these characteristics lead to many of the common data project failures (or successes). Later, we provide recommendations on how to address some of these issues to help progress your organisation from being analytically immature to analytically mature.

What does the data show?

> **Previous research shows that around 80% of all data science and AI projects fail to achieve their stated goals or come close to them. However, further analysis reveals a more concerning issue – the actual failure rate for analytically immature organisations exceeds 90%. How was this finding made?**

To begin our analysis, we first grouped our research data into categories based on the well-known People, Process, Technology (PPT) framework. We also add strategy to the PPT categories, presenting them in the following order:



To come up with this order of importance, we categorised each failure reason from our research to one of the above four themes and plotted them on a yearly basis as to how often each occurred. Our goal was to understand the importance of each theme over time, so to avoid skewing the data, we counted how often each one arose each year and divided that by the total number of items from that year. This gives the relative importance of each theme (see Figure 1, the scale in the figure is just to be used as a reference for magnitude).

Looking at the data over the years, we see that process issues seem to be the most cited cause of failure, then strategy, and finally people and technology.

Theme By Year

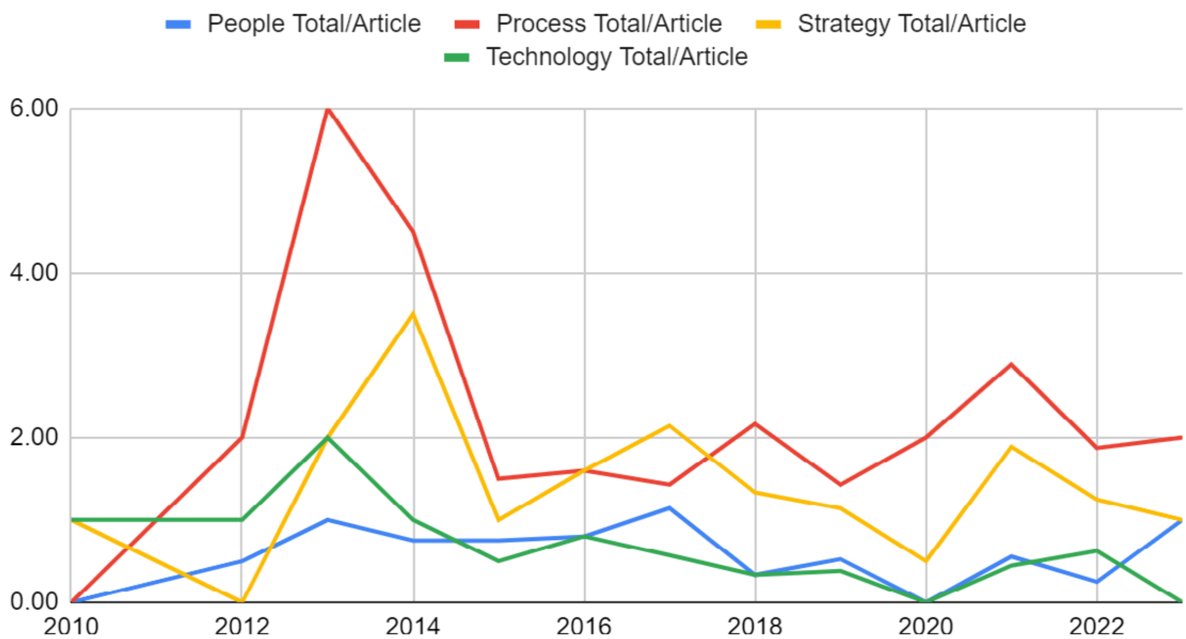


Figure 1.1. A graph of the weighted failure mentions in the literature by year and type.

Next, we grouped the themes by whether they were issues facing mature or immature organisations. As expected, for analytically immature organisations, the issues written about appear more often than those facing more mature ones.

Company Maturity Level Failure Count By Year

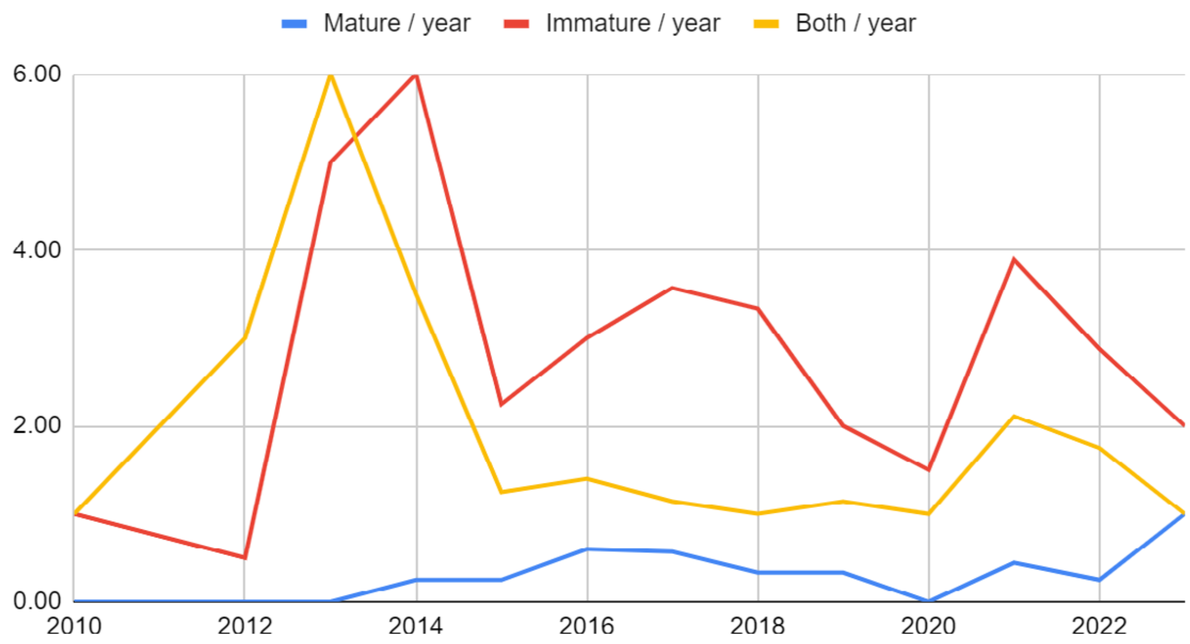


Figure 1.2. A graph of number of failure mentions in the literature by year and company maturity.

The numbers in Figure 1.2 corroborate the idea that as a company becomes more mature, it can reduce its failure rate.

If we take these time series of failures as a proxy for organisations' success rates, then calculations will show that approximately 95% of all weighted mentions point to a failure attributable to an immature organisation and 45% of all weighted mentions attribute the failure to a mature organisation (or both). However, this does not address the question of what percentage of AI projects fail in analytically mature organisations compared to their analytically immature counterparts[2]. To answer this question, we need more data. This additional data comes from a Massachusetts Institute of Technology (MIT) whitepaper covering how organisations are closing the gap between ambition and action in AI[3], as well as other project failure rate data[4]. The report showed that analytically mature organisations make up around 23% of all companies, meaning approximately 77% of all companies are analytically immature.

By combining the fact that 80% of all AI projects fail (with these failures occurring in both analytically mature and immature organisations) with the data from Figure 1.2, we can conclude that twice as many failures are due to analytically immature organisations compared to analytically mature ones. Given that 77% of the companies are analytically immature, we can estimate that the actual project failure rate of the analytically immature organisations is over 90%, with analytically mature organisations failing in "only" 40% of their attempts.

2 These statistics only provide data on the overall failures attributable to each type of organisation but given that there are different numbers of analytically immature and analytically mature organisations, we need to factor that into account as well. We do this in the following analysis.

3 Ransbotham, S. (2017, September 6). Reshaping business with artificial intelligence. MIT Sloan Management Review. https://sloanreview.mit.edu/projects/reshaping-business-with-artificial-intelligence/

4 O'Neill, B. T. (2020, October 28). Failure rates for analytics, AI, and big data projects = 85% – yikes! Designing for Analytics. https://designingforanalytics.com/resources/failure-rates-for-analytics-bi-iot-and-big-data-projects-85-yikes/

Why does a high failure rate matter?

If this high failure rate persists, it will significantly damage the analytics industry. Therefore, it is essential that the groundwork be laid in a way that is palatable and maximises the chances for organisations to improve their experiences with data science. Otherwise, companies will dismiss data science as "snake oil" and a false promise and fail to understand the effort required to achieve success on such initiatives.

Our primary goal is to **temper the AI hype with a dose of realism**. While we firmly believe in the power of mathematics, computing and analytics, if false expectations are set and practitioners and leaders don't fully understand the complexities of AI projects, then a stunning **80% (or more) of analytics projects** will continue to **fail**.

There is also the financial cost of failure to consider. The cost of developing a robust data science system depends on the scope of the problem and the scale of the company, but a *properly developed AI system* typically requires an investment of between $200,000 and up to $50 million. If, for each successful project, four are unsuccessful, then we are looking at between $800,000 to $200 million wasted ***per successful project***—this is a critical problem that we need to address immediately.

It is estimated that by the year 2030, the market for AI will have grown twenty-fold and be valued at close to two trillion US dollars[5].

If 80% of this ends up being wasted, it is a monumental misallocation of limited resources. We need to do something about this before we adversely affect people's attitudes toward one of the most powerful approaches to assisting decision-making: the data- and model-driven approach.

5 Thormundsson, B. (2023, October 6). Artificial intelligence (AI) market size worldwide in 2021 with a forecast until 2030. Statista. https://www.statista.com/statistics/1365145/artificial-intelligence-market-size/
Next Move Strategy Consulting (NMSC). (2023, January 1). Artificial Intelligence Market Size and Share | Analysis - 2030. Next Move Strategy Consulting. https://www.nextmsc.com/report/artificial-intelligence-market

Why analytics projects fail: An academic perspective

Research partnerships between industry and academics in business analytics seem like a no-brainer. Academic experts bring deep area knowledge, backed up by international networks, that allow them to identify and implement cutting-edge data analytics tools and innovations to solve challenging problems. Yet they are rarely connected to the coalface of practice and can lack contextual insights built up by organisations over years. On the other side, many businesses are actively pursuing strategies to use high quality data analytics to improve their operations and decision-making. But in doing so, they can face challenges that they are not well placed to address, including technical hurdles and a lack of experience in framing complex analytics problems.

Thus, there is real potential for academic experts to work hand-in-hand with industry partners as co-researchers to the benefit of both parties. This has long been recognised by the Australian Research Council (ARC), which has a special "Linkage Scheme" to partially fund such co-research. Yet, it is frequently the case that such projects – whether funded by the ARC or not – fail to deliver on their identified promise. While every project is different, there are some common causes of failure. What are these and how can they best be avoided?

The first is that the partners can differ in goals and objectives. Academics typically focus on publishing research, while industry partners aim for practical applications and commercial gains. Yet these differences are rarely incompatible. For example, publication of the results may be fine if an agreed level of anonymisation is adopted by the academic when writing a paper. Or the commercialisation of the project's findings may be an objective as long as it is clear where IP lies. Each case is different, but a clear discussion and formal agreement at the outset is almost always a necessary requirement.

A second is a lack of cultural understanding between individuals involved in the project. Academics are surrounded their whole working life by other academics, while staff at an industry partner are unlikely to have worked with academics before. This is fertile ground for misunderstandings, negative stereotyping and resentment to occur. Yet most projects start with a lot of good will, and a clear recognition of each other's strengths at the outset helps build trust. The academic really is an expert, otherwise they would not be there. Staff at the industry partners really are problem-solvers, otherwise they would not be there. A clear understanding should be developed that this is a division of labour after all!


Business Model
Market Research
Total Costs
Volatility
April
July
August
Sustainabilit

A third reason that projects can fail is the setting of inadequate milestones. Of course, this problem affects many projects, not just those involving co-research between academics and industry partners. But this issue is greatly compounded here by the different timeframes that both parties face. Academia can be painfully slow, with the whole sector being designed to undertake the careful and balanced assessment of evidence before coming to conclusions. In contrast, the competitive imperative in industry – especially in fast moving areas like analytics – can act to promote speed of solution over the quality of solution. But any co-research that requires academic and industry collaboration is likely to be important enough to find a middle ground, and marking it out with clear staged milestones is often essential to success.

A final reason that projects can fail is that the project itself is a poor candidate for co-research between academic and industry partners. Good candidates usually involve solving substantive and reoccurring problems that face a business or sector. Smaller or one-off problems are not suitable for co-research and it might be better to engage in a pure consultancy arrangement. For example, I was involved in co-research with a major airline where the objective was to develop advanced modelling and forecasting of individual-level customer demand for passenger flights using big data from their booking and customer databases. This was an ongoing problem that was fundamental to the airline's operations and it was willing to invest in continuous improvement of their solution. They funded several PhD research students who worked inside the airline but were supervised and guided by the academic team. This proved remarkably successful, as the research students had the time to bridge both worlds in order to solve problems effectively, and build trust and understanding between the academic and industry team.

This list makes clear that co-research between industry and academic partners has a risk of failure. However, forewarning of the specific types of risk in such projects goes a long way to managing it and fulfilling the project objectives.

About the Author:

PROFESSOR MICHAEL SMITH

Chair of Management (Econometrics)

Michael Smith has held the Chair of Management in Econometrics at MBS since 2007. He is a leading researcher in Bayesian statistics and business analytics.

Michael completed his PhD at the Australian Graduate School of Management at the University of New South Wales. Prior to joining MBS, he held positions at Monash University and the University of Sydney. He has also held visiting positions at Ludwig Maximilians University in Munich, the Wharton School at the University of Pennsylvania, McCombs School of Business at the University of Texas, London Business School and UCL.

Past major awards include an Alexander von Humboldt fellowship and an Australian Research Council Future Fellowship. In 2021 he was awarded the University of Melbourne Faculty of Business and Economics Deans' Award for Research Excellence.

Michael's research focuses on developing methods for the analysis of large and complex datasets that arise in business, economics and elsewhere. On the methodological side, he has worked on Bayesian algorithms, spatial and time series analysis and multivariate modelling. On the applied side, he has worked on marketing models for advertising effectiveness and consumer response, neuroimaging, and macroeconomic and business forecasting. He has a long-standing interest in the electricity markets, including the modelling and forecasting of demand and spot prices.

Michael's research has been published widely in the leading academic journals in statistics, econometrics, marketing and forecasting. He is regularly invited to speak at international conferences and workshops, and is involved with a number of prominent international academic societies.

Michael has taught courses in econometrics, statistics, decision sciences and business analytics at all levels – from undergraduate to PhD level. At MBS he currently teaches Data Analysis on the part-time MBA and Risk Analytics in the Master of Business Analytics.

WHAT CAUSES FAILURE?

AI and data science projects are hard. They add mathematical sophistication and additional change management to already challenging IT projects. AI practitioners are so-called knowledge workers and due to this, even before a single line of code is written, they are up against a formidable challenge.

"In knowledge work… the task is not given; it has to be determined. 'What are the expected results from this work?' is… the key question in making knowledge workers productive. And it is a question that demands risky decisions. There is usually no right answer; there are choices instead." Peter Drucker – Management Consultant.[6]

Let us begin to understand what causes failure by presenting a detailed summary of the research that went into the foundations of the book *Why Data Science Projects Fail*. Let's start with an experience deficit revealed by Kaggle's annual data science survey.[7]

Kaggle, founded in 2010, is one of the world's largest online communities for data scientists. Its annual survey consistently shows that roughly 60% of its members have less than two years of work experience. Putting aside the response bias, it gives us insight into the level of experience of those currently undertaking data science work, highlighting a lack of seniority in the field and therefore, a capability gap when it comes to knowing how to strategically run a data science function from the top of the organisation, down.

The problem with missing skills at the top when something is popular (e.g., the hype surrounding AI) is that we can sometimes see a deference to the technical gurus who themselves may lack the knowledge or experience. In the classic book, *To Engineer is Human*, Henry Petroski states:

"[I]nexperienced engineers are tempted to work beyond their competence because of the availability of powerful software, and once the numbers are crunched, engineers tend to rely on the results rather than their own judgment, e.g., the roof of the Hartford Civic Center, which was designed using a computer model to analyse the stresses. So confident were the designers that they brushed aside the questions of workmen who had noticed a large sag in the roof well before it collapsed under the snow and ice of a January 1978 storm."

---

6 This quote appears in the best selling book, *Getting Things Done* by David Allen

7 See for example some of the following sources:
Kaggle. (n.d.) 2018 Kaggle Machine Learning & Data Science Survey. https://www.kaggle.com/datasets/kaggle/kaggle-survey-2018;
Mooney, P.T. (2018, November 12). 2018 Kaggle Machine Learning & Data Science Survey. https://www.kaggle.com/code/paultimothymooney/2018-kaggle-machine-learning-data-science-survey
Mooney, P.T. (2020, November 19). 2020 Kaggle Data Science & Machine Learning Survey. https://www.kaggle.com/code/paultimothymooney/2020-kaggle-data-science-machine-learning-survey
Mooney, P.T. (2021, October 14). 2021 Kaggle Data Science & Machine Learning Survey. https://www.kaggle.com/code/paultimothymooney/2021-kaggle-data-science-machine-learning-survey

This capability gap is causing a large schism between the analytically mature and the analytically immature organisations.

Inexperienced companies are merely following the hype without understanding the complexity of data science projects or developing a compelling vision or strategy. Expectations are high and executives, who often lack a deep understanding of data science, view analytics initiatives as disruptive, rather than incremental. This is because when leaders envisage a project, they aim for ambitious goals and target their most significant problems, instead of starting with something commensurate with their experience and current analytics maturity.

Our research shows that these issues then cascade as follows:

Data issues

When an analytically immature organisation begins a project, the first and main factor they underestimate is data issues, including a lack of one or more of the data sources in a centralised location for analysis (in a data lake, data lakehouse, or other platform/framework) as well as problems with data quality, data access, data relevance, metadata, data lineage, and data management/processing tools. This underestimation is one of the greatest causes of failure and is a challenge for all companies, every time.

Lack of strategic planning

Less mature organisations often lack a well-defined use case that aligns with the company's strategy. They want to use AI and leave it to the wrong people (e.g., the data scientists) to figure out where to use it. Unfortunately, data scientists typically aren't trained in management or commerce, so they invariably fail in trying to solve both the commercial and technical side of the business challenge. Many data scientists also aren't trained in the proper way to engage end users and stakeholders in a business, nor how to extract the requirements needed to deliver a project that adds business value. This means the wrong project is selected, one that is unlikely to produce a good ROI or align with the business strategy. In addition, many data scientists often haven't no experience in the domain of the company with which they have just started working for and therefore struggle to interpret their results in light of the business environment.

No buy-in from leadership

Because companies are reacting to hype and hastily launching into AI development by delivering projects without strategic planning, the organisation's leadership team most likely doesn't properly understand the basis for the initiative. Therefore, there is a failure to support the project in its entirety. This leads to a lack of resources, ill-defined processes and often, unrealistic expectations on what the project can deliver. In addition, they will often fail to set business targets with clear commercial goals, the very thing needed to help obtain proper buy-in from end users, superiors and peers.

Poor resourcing

Many companies starting out don't properly resource their teams due to lack of experience, knowledge, and access to experienced data scientists.

Treating AI projects like IT

Due to a lack of understanding, analytically immature companies may treat AI projects like IT initiatives because it is the closest context they know. They therefore focus on the technical side of the project and give it to their IT teams. However, data science projects are significantly more complex than standard IT projects, incorporating crucial, new human elements. They have the potential to completely change or, on rare occasion, eliminate job functions and decision-making processes within a business.

The above, and many more reasons, lead to a lack of organisational support throughout the project, poor utilisation of the final product (if there is one), and unengaged end users.

Finally, once the project is completed from the technical side, change management is not carried out because the data science team doesn't know how to, and in an immature organisation, resources aren't allocated for it. Because the data scientists are inexperienced, they don't understand all the effort required to get the model to production therefore projects take longer than the original optimistic estimates, with stakeholders either losing interest, or the project failing altogether.

Although AI projects are often led by technical individuals with Masters and PhDs, we can see that the above causes of failure are rarely due to technical issues, but mainly a lack of "soft" skills.

This finding is not new, in fact it is something that has been known for over a century now, with Dale Carnegie summarising the state of knowledge in his famous book *How to Win Friends & Influence People*:

> **"Investigation and research done a few years ago under the auspices of the Carnegie Foundation for the Advancement of Teaching uncovered a most important and significant fact – a fact later confirmed by additional studies made at the Carnegie Institute of Technology. These investigations revealed that even in such technical lines as engineering, about 15 percent of one's financial success is due to one's technical knowledge and about 85 percent is due to skill in human engineering – to personality and the ability to lead people."**

WHAT IS THE SOLUTION?

Any solution to this problem will be challenging and multifaceted because, in reality, these problems are relevant to any innovation, not just AI. Turning back to the research, we can draw some specific recommendations as to what we can do to reduce our chance of AI project failure in each of the four areas of Strategy, Process, People and Technology.

There were numerous reasons for AI project failure related to strategy (as presented in the table) with the most common reason for failure being the simplest:

Failing to build the need in the organisation (poor use case, no clear business value, no actionable insights, solution looking for problem)

The next most common was:

Lack of leadership/upper management buy in

It was fascinating to see these two reasons occurring most frequently as they seem to be the two most basic elements that any organisation should ensure are in place before beginning a data science project.

Based on these findings our recommendation for overcoming the strategic challenges is:

Find those strategically important projects for which data science and business analytics can help accelerate the strategy.

This starts at the top of the organisation and requires the whole business to be aligned, know and work towards the strategy, even in the AI team.

Reasons
Competing investment priorities
Failing to build the need in the organisation (poor use case, no clear business value, no actionable insights, solution looking for problem)
Lack of leadership/upper management buy in
Lack of vision or strategy (or alignment to, poor data maturity)
Lacking data governance upfront (including security, ethics, etc)
Lacking security & compliance (managing data, regulatory, etc)
Limited focus rather than company wide
Not clearly measuring success (unclear deliverables, missing actionable insights)
Not enough effort in vendor analysis
Not identifying business parts with quick wins
Treat as once off rather than on going

Strategy

The table outlines the set of reasons that our research revealed as the causes of process failures. The most common reason for failure occurring was the most expected one:

Data quality and reliability related issues

The next most common was:

Not setting clear or reasonable expectations

Just like the common causes of failures identified for strategy, these seem like basic elements an organisation should already have in place. The clear recommendation from these findings is:

Make sure you've got the right data to solve the problem at hand and don't expect any miracles if your data is no good.

By following this advice, likely many project failures could be avoided by simply not starting projects that are set up for failure.

Reasons	
Data quality and reliability related issues	Not planning for many iterations and ongoing
Lack of change management	Not setting clear or reasonable expectations
Lack of upfront or ongoing planning (RACI, lack of PM, etc)	Not starting small and simple
Missed timelines	Not supporting rapid growth (scale big quickly)
No clear organisational communications plan or poor communications (also not selling the benefits, poor report/dashboard)	Not using agile processes for solution delivery
No customer focus and value co-creation (wrong questions, project doesn't match the user's workflow, etc.)	Silo thinking rather than big picture
No formal training	Time wasted on extensive data exploration (or led by)
Not having ownership move from IT to business unit	Treating it like an IT or software project instead of ML/AI
Not managing open-source risks	Vendor issues like: Lack of vendor support, vendor hype, bad vendor fit, etc.

People

The table outlines the set of reasons that our research revealed as the causes of people failure. The most common reason for failure occurring was as expected:

Lack of the right resources

The next most common was:

Poor company analytics culture

There is a clear pattern occurring here, as once more, these are two seemingly fundamental elements which should be in place. The recommendation based on these findings is:

If you don't have the right people, or leadership disapproves of an analytics approach, then don't even try.

These two elements are easy to test. Firstly, verify if there are individuals who can deliver on the planned projects and secondly, will you (the leadership) support them even if it conflicts with the way you currently do business?

Reasons
Lack of the right resources
Leadership has a lack of sufficient knowledge of AI and its applications
Misaligned or conflicting interests
Poor company analytics culture (e.g. islands of analytics with "Excel" culture)
The Old School Mindset - not used to analytics, can't describe the problem in the right level of detail

The table outlines the set of reasons that our research revealed as the causes of technology failure. The most common reason for failure was not unexpected:

Lacking high quality data- and tool infrastructure

The next most common was:

Technical reasons (e.g. inaccurate predictions, wrong or poor models)

The first reason is closely aligned with the process-related issue of lack of quality and easily accessible data, indicating a lack of investment. The second is not a surprise and is probably what many people think of when a model fails. Our recommendation to avoid these issues is to:

Invest in the right types of tools to solve problems that are known to have solutions you can implement.

This likely comes back to people having the right level of analytical knowledge to know what they can solve (or is even solvable) and then being able to make the business case to invest in the right tooling needed to complete the project.

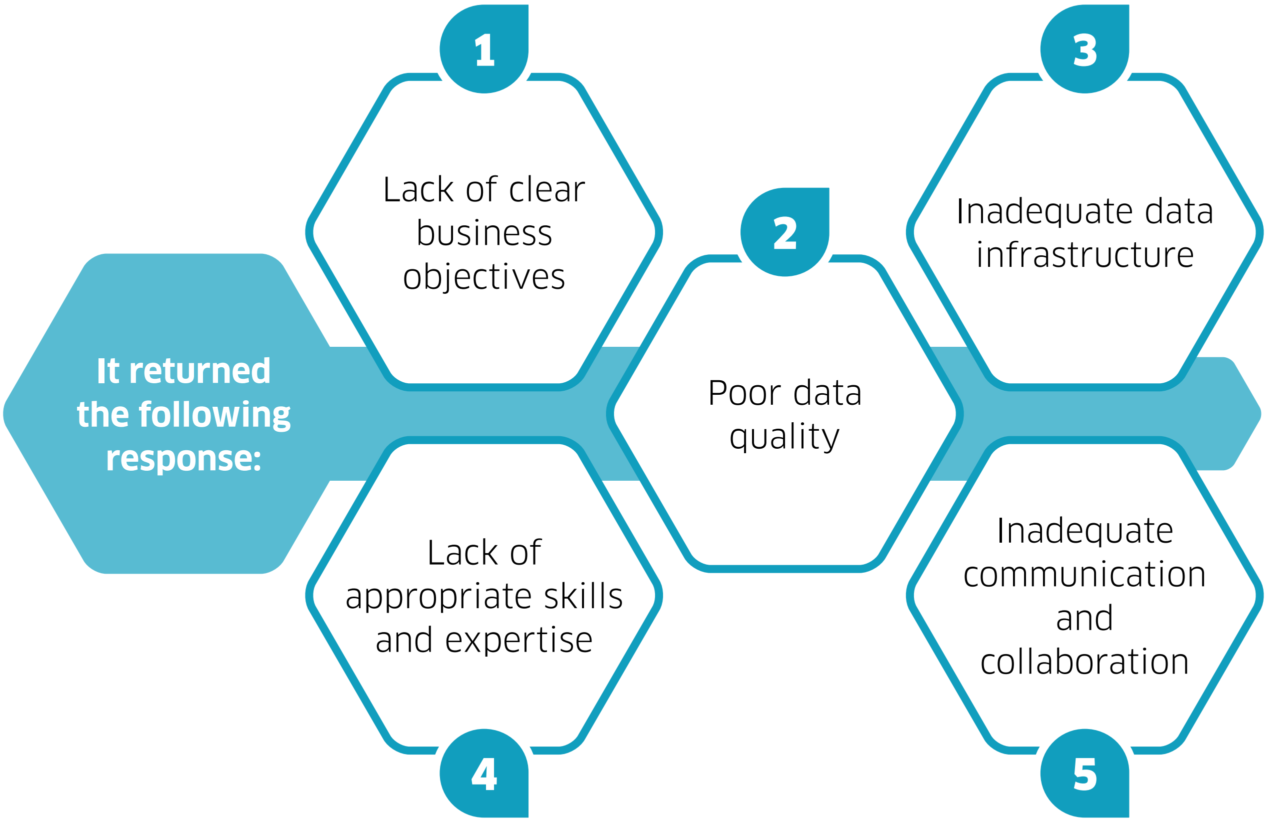
Reasons
Data projects aren't discoverable (i.e. the results aren't easily findable by users)
Failure to deploy (e.g. underestimate effort, etc)
Lacking high quality data- and tool infrastructure
No established company ontology or single versions of truth (also missing systems and standards)
Not sharing data easily (data silos, etc)
Technical reasons (e.g. inaccurate predictions, wrong or poor models)

Technology

HOW CONFIDENT ARE WE?

In a final step to validate our research we asked one of the greatest repositories of human knowledge ever created, ChatGPT (GPT-3.5). Using the following prompt, we asked this astounding knowledge base why it thinks data science projects fail with the following verbatim prompt:

The failure rates in data science projects are so extremely high in companies (greater than 85% according to reports) and we have many frameworks to help do projects successfully, so why do we still see so many failures? People know the reasons for failure so why do we still fail?



These five main reasons directly align with our research. The order of priority even largely lines up. So, with the stored knowledge of billions of parameters trained on trillions of text tokens seeming to agree with our findings, not much is left to doubt.

CONCLUSION

There are many books available that praise the benefits of AI and pressure leaders to adopt it lest they be left behind, however, too much of this discourse seems misguided. Many analytical disasters use the techniques of successful companies and with so much focus on what these companies do, we need to be sure that these strategies are really what differentiate the mature from the immature.

Some inexperienced companies may be doing the same things as excellent companies, so how do we know what is truly critical to success? Without a proper review of the less than successful projects, how do we really know what works? We must face the failures to make sure we aren't deceiving ourselves.

Henry Petroski, mentioned earlier as the author of *To Engineer is Human*, captured our perspective perfectly when he wrote the following 40 years ago:

"I believe that the concept of failure…is central to understanding engineering, for engineering design has as its first and foremost objective the obviation of failure. Thus, the colossal failures that do occur are ultimately failures of design, but the lessons learned from these disasters can do more to advance engineering knowledge than all the successful machines and structures in the world.".

There is further support of this opinion. In a study published in 2013 in *Management Science* researchers showed that how we learn appears more complicated than you may think and failures play a key role in this. Examining more than 6,000 cardiac surgery procedures that used a new technology, the researchers found that "individuals learn more from their own successes than from their own failures, but they learn more from the failures of others than from others' successes."

This whitepaper aims to help further enable AI and data science to solve complex strategic, tactical and operational problems, and support and better enable data- and model-driven decision-making. It is time that organisations stop entertaining fantasies about why every company must forget what they are doing and suddenly become "AI first". Instead, we must realise that data science and AI are just tools to deliver better organisational outcomes, just like a forklift.

Failure is an extraordinarily effective form of feedback. We hope that by sharing our lessons and stories in our book, individuals will learn a lot more and begin to understand the reality of AI failures, and their root cause.

The content of this whitepaper is based on the book *Why Data Science Projects Fail* by Doug Gray and Dr Evan Shellshear. To order a copy go to:

https://www.routledge.com/Why-Data-Science-Projects-Fail-The-Harsh-Realities-of-Implementing-AI-and-Analytics-without-the-Hype/Gray-Shellshear/p/book/9781032660301

CASE: AN ANALYTICAL NIGHTMARE

You would not expect a top 200 publicly listed company to stumble on a well understood part of the data science deployment process. However, this is precisely what happened when a team of external accounting consultants thought their expertise with numbers provided them with sufficient skills to execute and deliver a data analytics project like a professional. Unfortunately, the outcome delivered failed to achieve the level of a beginner.

Our story began when a large retail chemical manufacturing conglomerate, engaged a local accounting company with a small analytics team to create a holistic supply chain model. The model was supposed to provide detailed costs for all items in their warehouse. The impetus for the engagement was to review the whole process—from sourcing to warehouse management—through an analytics lens. This would allow the manufacturer to take a product all the way to the retail shelf and calculate the profit margin based on the full cost of handling and delivery. The vision was to apply an optimisation engine to the whole process and use this AI tool to improve the cost performance of their supply chain.

To build the proof of concept, three months of data were collected from one of the small warehouses, which alone amounted to millions of rows of data. The manufacturer managed their company's data on SQL databases and the audit team's lack of data science experience led them to believe they could use SQL scripts to run the whole optimisation solution. So, they began exporting data in CSV files from the database and uploading them to a SQL one in their own environment to start developing and testing code.

The end algorithms were designed to pick up live, current data from the central database, review it for a particular time range, and optimise that operating period. However, along the development journey, the accounting team forgot to enter the time range of analysis and although the intention was to use this program for three months of data (a few million rows), it was actually configured to extract hundreds of millions of rows of data from over a decade stored in the database. If the accounting team were to deploy the algorithm to extract this data, it would cause a company-wide meltdown because this central storage is used for everything. The database would be overloaded with queries, rendering all areas reliant on real-time data useless, including point of sale terminals.

If this were to happen, it would also bring down the entire manufacturing production system. Colleagues working with dangerous compounds would not be able to identify the chemical properties of their inputs which could lead to adverse, and potentially lethal and uncontrollable molecular reactions. The impact of such an incident on the manufacturer's reputation would be disastrous and likely cause their stock price to enter free fall.

Fortunately, the week that the consultants were preparing to deploy the SQL script, the senior manager of the accounting firm had a discussion with another analytically interested director in their company and explained their project. Once the director heard what was planned, he asked some simple questions, like whether the team had run this in a test scenario using IT infrastructure similar to the planned production environment.

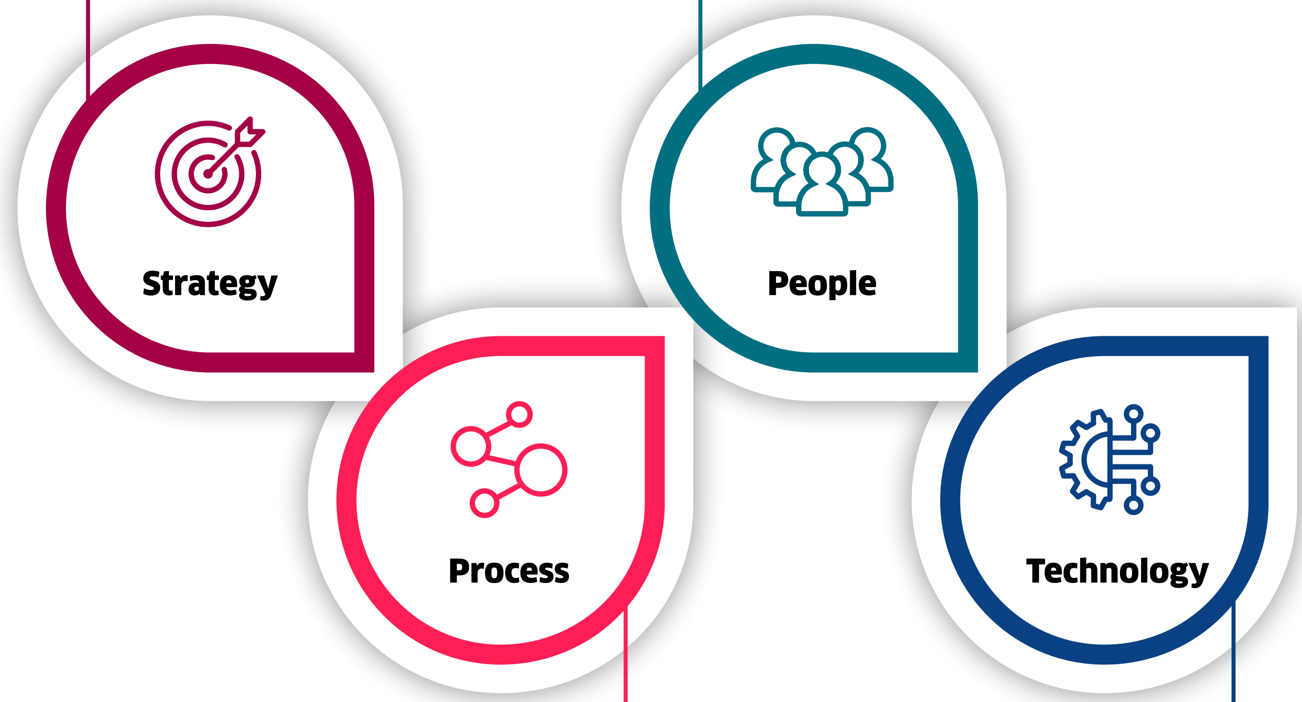
The answers to these questions shocked the director because not only had they not run the program in a test IT environment, but it also turns out they didn't have any testing protocols at all, and normally deployed new, untested changes directly to the manufacturer's live working database.

The director straight away stopped the delivery of the code and let his superior, the accounting firm partner, know about the impending reputational disaster. As a result, the partner immediately discontinued the data science project, recognising that admitting defeat was a much better outcome than the reputational and market risk of deploying the SQL script. Whilst this was a painful blow to the partner's ego and bottom line, it was as much smaller failure than if the code were deployed.

Reviewing this disastrous implementation, if we examine the four areas of Strategy, Process, People and Technology, we can identify clear failures from each, especially the ones that we identified to be the most common issues.

There was a lack of effort put into building the strategic business case in the organisation. What's more, both organisations were missing proper data governance (including safety, security, ethics, etc).

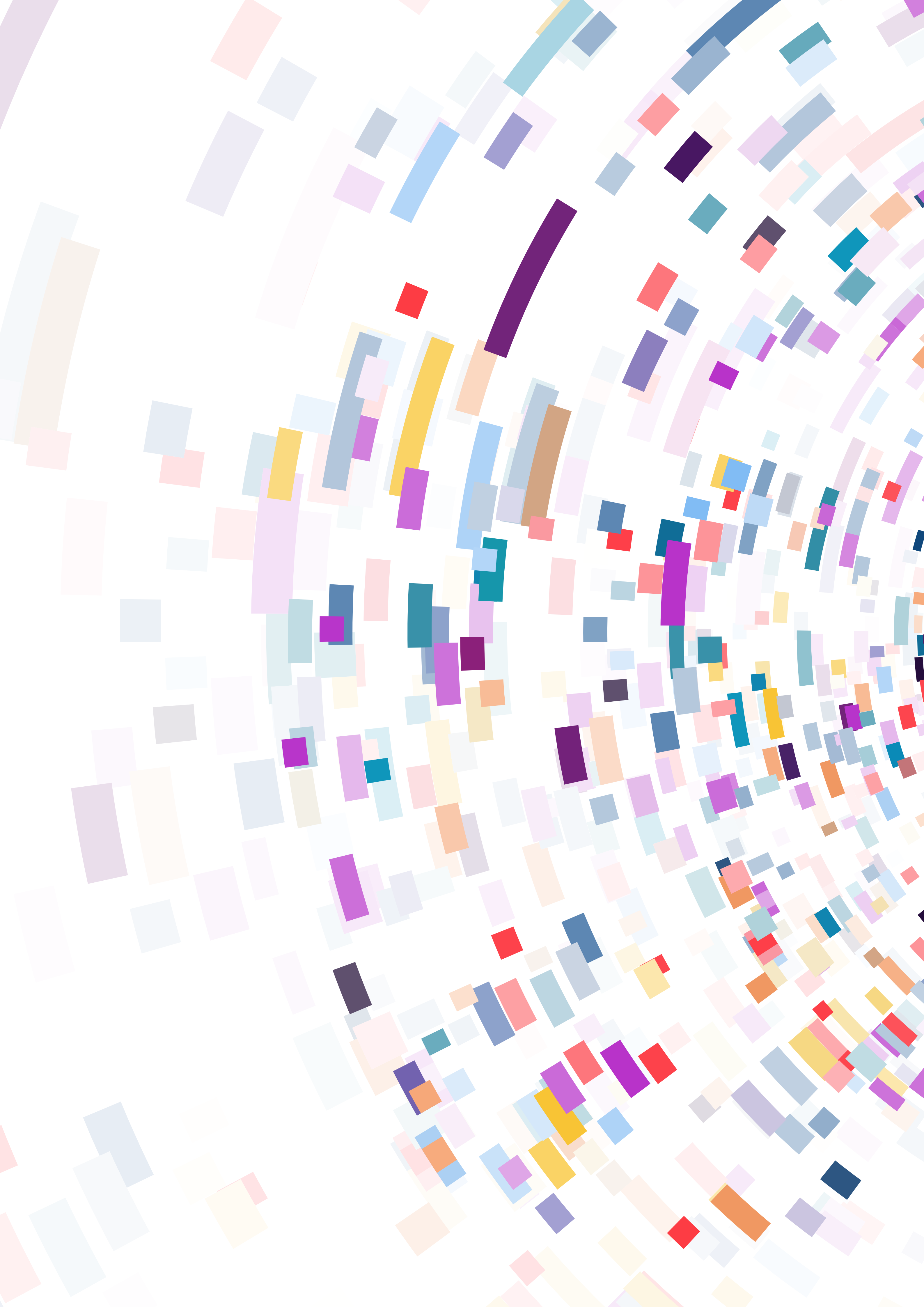
There were many failures that fell under the people category, but the main ones were a clear lack of the right data science skills and talent combined with leadership lacking sufficient knowledge of AI, its applications and deployment methodologies.



In this case, there was a lack of upfront or ongoing planning (RACI frameworks, lack of a project manager, etc). The project team also did not plan for multiple iterations nor think about the daily usage by the business.

Finally, they lacked high quality data and tool infrastructure and failed for technical reasons (e.g. inaccurate predictions, poorly coded or wrong models).

In summary, their failures were truly a masterclass in how not to execute data science projects.



Connect with the Centre for Business Analytics

cfba@mbs.edu

CONNECT WITH US

@MelbBSchool

/MelbourneBusinessSchool

melbourne-business-school

/MelbBSchool

@melbbschool

CENTRE FOR BUSINESS ANALYTICS

THE UNIVERSITY OF MELBOURNE

MELBOURNE BUSINESS SCHOOL