PALO IT

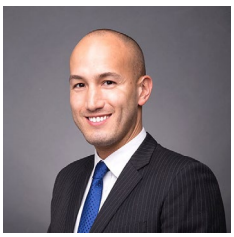
QUINLAN &ASSOCIATES

THE REAL LINGUA FRANCA (PART 2/2)

UNLOCKING THE POWER OF DATA

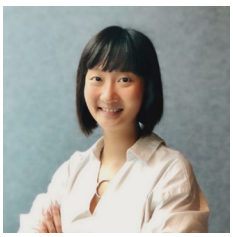
THE AUTHORS

QUINLAN &ASSOCIATES



BENJAMIN QUINLAN
CEO & MANAGING PARTNER
BCom (Hons 1) LLB (Hons), Macquarie University

PALO IT



JING LEI
MANAGING DIRECTOR
MSc in Management, EM LYON Business School

JUSTIN CHUNG
ASSOCIATE
BBA (Hons 1), HKUST

JUNG HONG KIM
DATA LEAD
BEng & BBA (Hons), HKUST

CHARLES YAU
SENIOR CONSULTANT
BBA (Hons), HKUST

CONTENTS

EXECUTIVE SUMMARY 3

SECTION 1: FOR DATA'S SAKE 4

SECTION 2: DATA STRATEGY OVERVIEW 13

SECTION 3: DATA STRATEGY – BAIT (BUSINESS) 20

SECTION 4: DATA STRATEGY – BAIT (APPLICATION) 30

SECTION 5: DATA STRATEGY – BAIT (INFORMATION) 38

SECTION 6: DATA STRATEGY – BAIT (TECHNOLOGY) 54

SECTION 7: CHANGE MANAGEMENT STRATEGY 68

SECTION 8: REAL-WORLD CASE STUDY 71

SECTION 9: HOW CAN WE HELP? 77

EXECUTIVE SUMMARY

For many years, companies across the globe have sought to capitalise on the wonderous possibilities of data. Underpinned by technological innovation and rising digital adoption, businesses are actively exploring the boundaries of what their data can do, both for their customers and internal operations.

As mentioned in our previous report, today's data explosion has created significant fears of missing out ("FOMO") amongst many senior executives. In response, corporations' data investments have surged, with global data-specific technology spending reaching USD 474.7 billion in 2020 and projected to grow to USD 639 billion by 2024. However, we believe that many firms have gotten lost in the hype, neglecting the fundamentals of a data project in exchange for empty promises and wasted investments.

The significant wastage in data-centric projects reflects two key problems: (1) most data is left unused, untapped, or unknown to a business (i.e. "dark data"); and (2) the cost of recovering and making flawed data (i.e. "imperfect data") ready for use is extremely high. In fact, dark data accounts for 55% of data expenditure wastage, costing companies USD 193 billion in 2020. However, this pales in comparison to the recovery costs of imperfect data.

The fact is, only 3% of data within organisations are "perfect" and ready for processing. The remaining 97% requires some form of manipulation to make it ready for use, which expends 10x more effort (and cost) as opposed to vanilla maintenance costs. We estimate that the costs of making imperfect data usable amounted to USD 1.5 trillion in 2020 alone.

Wastage stems from multiple areas, including a lack of understanding of a company's data value chain, mismatched technology and business needs, inadequate integration of siloed data systems, and failed transformation planning. However, the vast majority of this wastage is attributed to one basic theme: the lack of a defined organisational data strategy. At its core, this consists of two complementary components: the (1) Business, Application, Information & Technology ("BAIT") framework and; (2) a robust change management strategy.

The BAIT framework outlines steps from project formulation to execution, together with key action items. The change management strategy outlines appropriate policies and incentives to develop a data-centric culture that unifies and maximises data investments throughout the lifecycle of a data strategy project. Both components can be scaled up or down, depending on the type of project: namely, strategic, tactical, or operational. The type of project is determined by considerations such as resources, time, or infrastructure required. These components rightly consume considerable airtime and focus. But there is one centrepiece that holds an entire data strategy project together that is often neglected. And that is a fundamental business strategy.

It is our belief that most data strategy projects fail because data's place as an integral part of the larger business is often overlooked. So many executives have been caught up in the hype behind data that they shoehorn it into a company's operations without detailed analysis as to "why" they are doing so; or worse, do so simply to appease shareholders.

At the end of the day, technology is guided by the business's objectives and likewise limited by it as well. A solid foundation in basic data operations must exist before data can be allowed to guide any business. Much like a language's primary purpose is to convey ideas and facilitate communication, it must first have a solid foundation of rules and structure before being able to evolve and guide the language as a whole over time. After all, data is the real lingua franca of business.

SECTION 1
FOR DATA'S SAKE

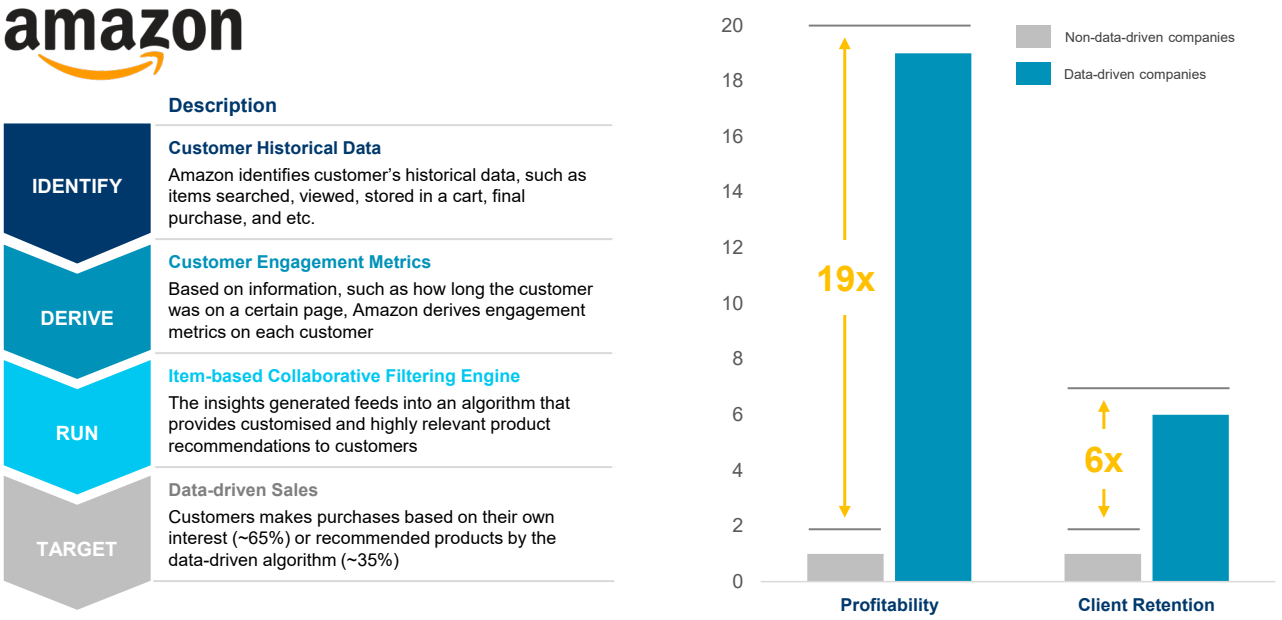
INTRODUCTION

As data rapidly evolves from a competitive advantage to a necessary cornerstone of the business world, it is becoming critical for executives to understand how to leverage its full potential. From advertising insights to decision-making to customer relationship management, more than half of all companies use data for their business processes.

There is growing evidence that drives corporate bottom-lines. In fact, data-driven organisations are 19 times more likely to be profitable (and 6 times more likely to retain customers) than their traditional counterparts.[1]

For example, Amazon's explosive journey to number one in the e-commerce space is largely attributed to its quality customer experience; using key engagement metrics and customer history, Amazon integrates customised recommendations into every step of the purchasing process. This data system, called item-based collaborative filtering, is directly correlated with – and generates 35% of – the company's global sales (see Figure 1).

FIGURE 1: BUSINESS OUTCOMES FROM DATA

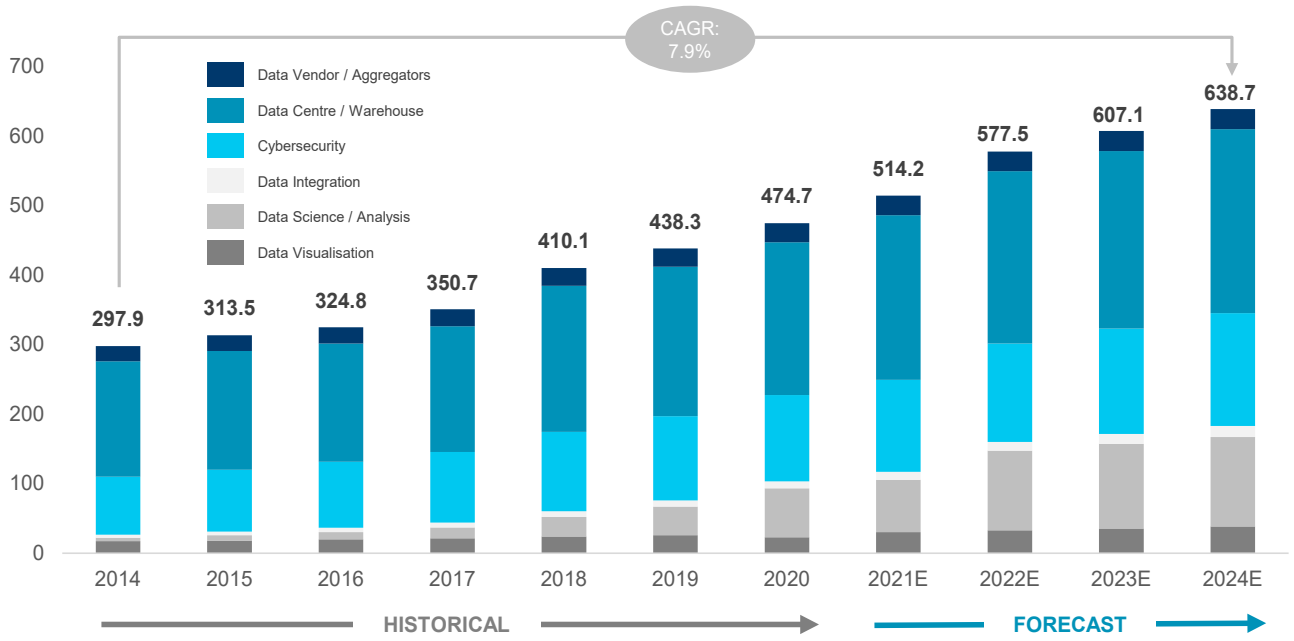


Source: Nanyang Technological University, Arizona State University, Millimetric.ai, Quinlan & Associates analysis

Given the vast monetisation opportunities available, companies worldwide have invested heavily in data analytics and other data-related solutions to support their data operations. In fact, we estimate that total data-related IT expenditure worldwide is expected to reach USD 639 billion by the end of 2024, growing at a CAGR of 7.9% since 2014 (see Figure 2).

[1] Millimetric.ai, ' What To Do When There's Too Much Data', available at: https://www.millimetric.ai/2020/08/10/data-driven-to-madness-what-to-do-when-theres-too-much-data/

FIGURE 2: DATA-SPECIFIC TECHNOLOGY SPENDING

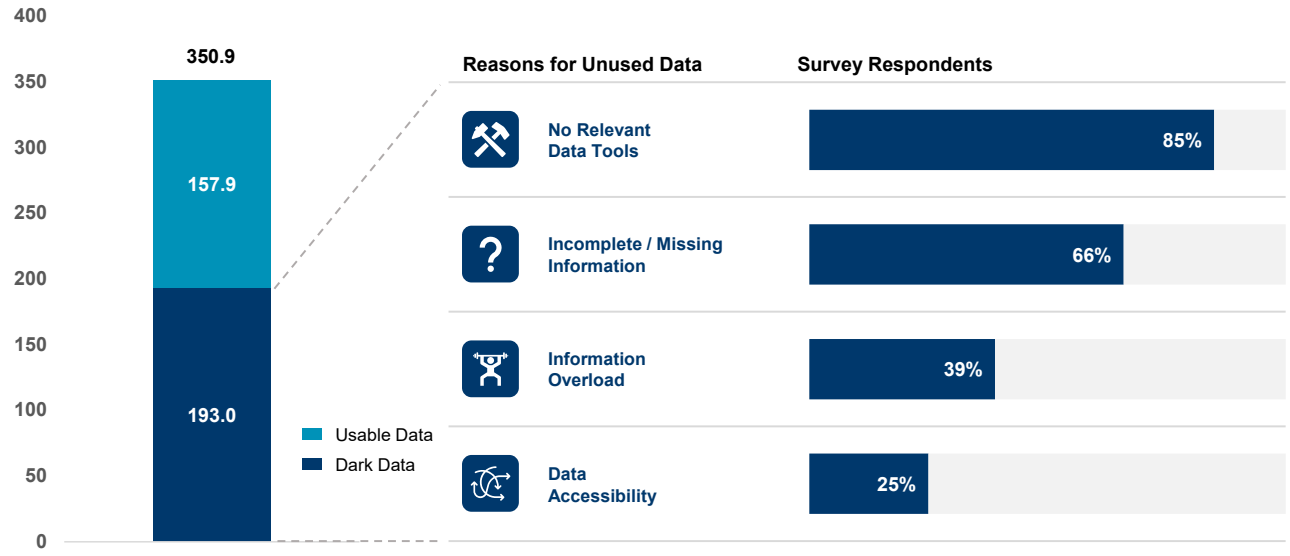


Source: Fortune Business Insights, Gartner, Market and Markets, Quinlan & Associates analysis

Excluding spend on cybersecurity, global data spending reached USD 350.9 billion by the end of 2020. In an ideal world, every dollar of this investment should have delivered tangible value to the companies making these investments. However, a lack of strategic metrics and granular planning of data utilisation has seen much of this spend go to waste.

Unused, unknown, and untapped data is commonly referred as “dark data”, which accounts for more than half of collected data by organisations. According to a survey of 1,300 companies conducted by TRUE Global Intelligence, 85% of the respondents stated that a lack of relevant tools to reap the full potential of the data is the primary driver of the existence of dark data, followed by incomplete information, information overload, and data accessibility. We estimate this dark data cost organisations USD 193 billion in 2020 alone, representing 55% of the total data expenditure, excluding cybersecurity spending (see Figure 3).

FIGURE 3: COST OF DARK DATA – INVESTMENT WASTAGE



Source: HIS Markit, IDC, Seagate, CISCO, Quinlan & Associates estimates

As highlighted in the figure above, the issues surrounding dark data are not solely attributed to data infrastructure; many also stem from suboptimal employee data practices (e.g. poor behaviour, lack of governance, etc.), where data is collected, stored, and managed without consistent standards. These inconsistencies (and a lack of completeness) results in considerable deadweight loss, as critical business decisions can be delayed (or even completely misguided) due to inaccurate, unreliable, irrelevant, and untimely data insights.

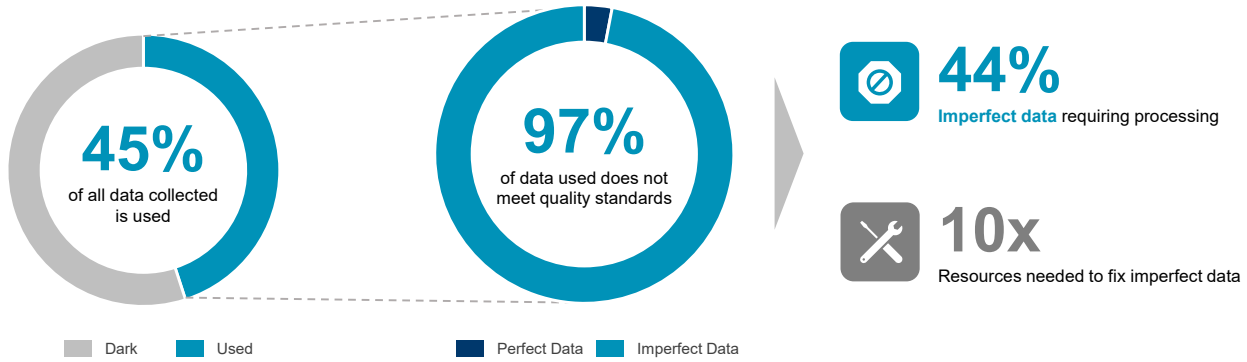
COST OF IMPERFECT DATA

A study was conducted by Harvard Business Review ("HBR") with the objective of measuring the data quality score of multiple firms by counting the number of error-free records at a company. HBR found that only 3% of data met the acceptable quality standards; [2] in other words, 97% of data collected needed to be recovered by revisiting, consolidating, cleaning, and correcting them to become business relevant and ready for use. The cost of was found to be approximately "ten times as much to complete a unit of work when the data is flawed in any way compared to when the data is perfect." [3]

With 97% of records subject to some form of data recovery, and maintenance costs for this data being ~10 times more than a vanilla data maintenance costs, we estimate organisations worldwide wasted USD 1.5 trillion in managing imperfect data in 2020 alone (see Figure 4).

[2] Harvard Business Review, 'Only 3% of Companies' Data Meets Basic Quality Standards', available at: https://hbr.org/2017/09/only-3-of-companies-data-meets-basic-quality-standards
[3] Harvard Business Review, 'Assess Whether You Have a Data Quality Problem', available at: https://hbr.org/2016/07/assess-whether-you-have-a-data-quality-problem

FIGURE 4: COST OF IMPERFECT DATA (1/2) – RECOVERY



	Imperfect Data	Loss (USD)	Multiplier	Total (USD)
Recovery Costs from Imperfect Data (2020)	**44%**	**153bn**	**10x**	**1.5 Trillion**

Source: HBR, Quinlan & Associates analysis

As highlighted in Part 1 of this two-part report series, the primary driver of these losses is a fundamental lack of understanding of the data value chain, with most companies focusing on: (1) collecting masses of information without a structured approach; and (2) processing it with patchwork enhancements to fit a specific algorithm. Additionally, to fully understand data, having a good grasp of underlying technologies (e.g. data architecture, infrastructure, etc.) and applications (e.g. visualisation tools, decision-making tools, etc.) is important as data should not be considered in a silo, but as part of a full value chain.

FOUR BUSINESS GOALS & DATA

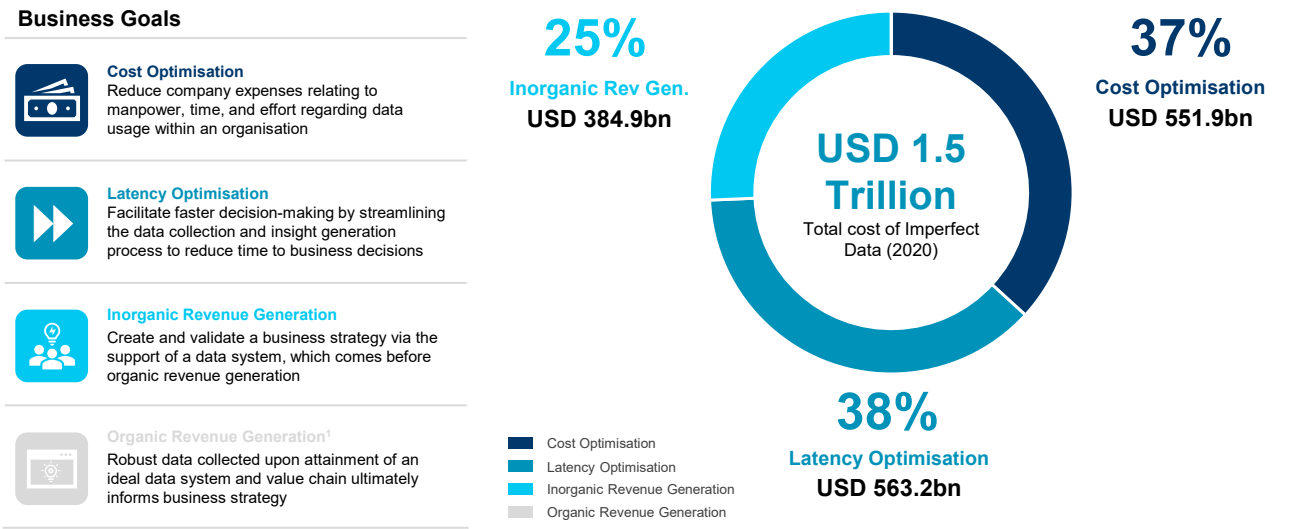
As the saying goes, "The road to hell is paved with good intentions". And it is undoubtedly true that most companies start out with the best intentions when it comes to utilising data.

Businesses typically look to utilise data in an effort to accomplish four main business goals, namely:

1. Cost Optimisation;
2. Latency Optimisation;
3. Inorganic Revenue Generation; and
4. Organic Revenue Generation.

In the process of accomplishing these goals, inadequate strategic planning and execution can lead to subpar results in the form of substantial financial losses. We have found that when these four business goals and respective journeys go awry, they contribute significantly to the cost of imperfect data in organisations (see Figure 5).

FIGURE 5: COST OF IMPERFECT DATA (2/2) – BREAKDOWN BY BUSINESS GOALS

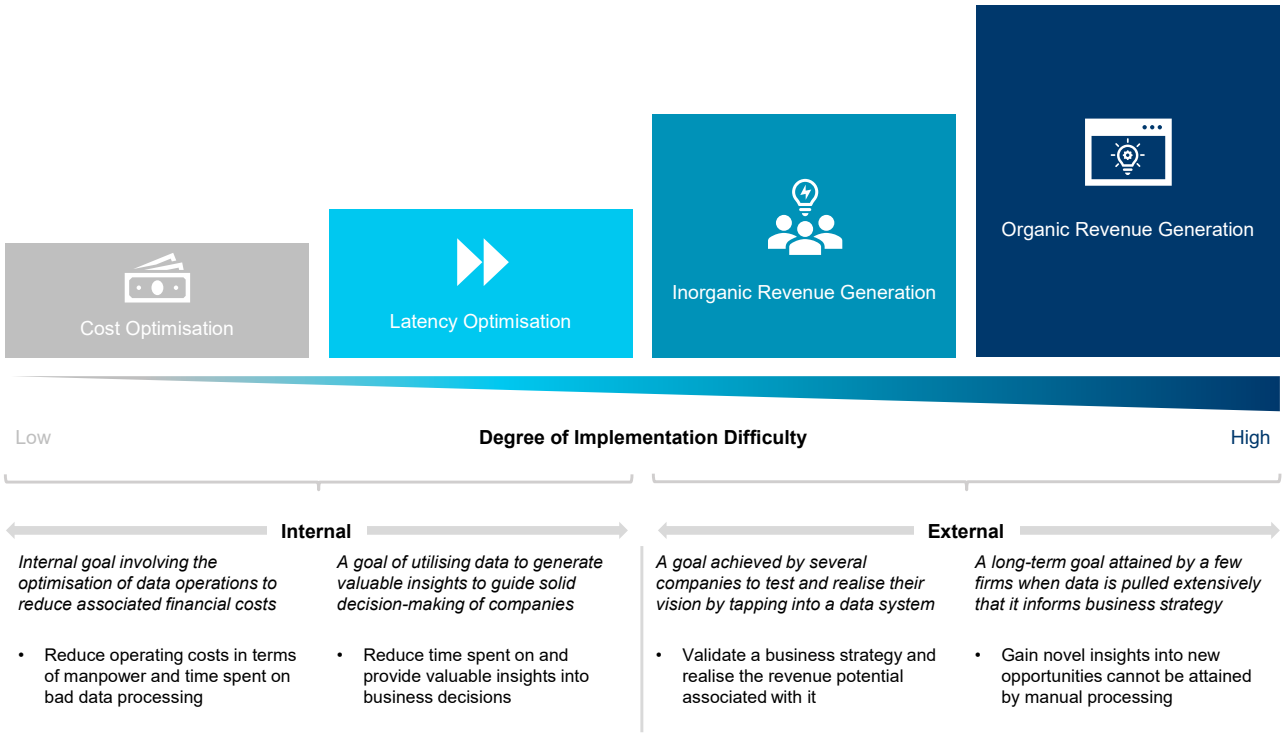


Note: Companies with organic revenue generation are unlikely to have significant wastage from investments in data projects as it requires tremendous attention to the data value chain and a robust organisational data strategy to achieve this goal

Source: NewVantage Partners, Capgemini, Harvey Nash, Quinlan & Associates analysis

These business goals can be further separated into two main categories: (1) internal goals; and (2) external goals (see Figure 6).

FIGURE 6: BUSINESS GOALS DEFINITION



Source: Quinlan & Associates analysis

1. INTERNAL: COST OPTIMISATION

As mentioned earlier, many businesses spend considerable time recovering or making data ready for downstream usage. The hourly cost of manpower skyrockets when staff are forced to waste additional time and effort collecting disparate data points from a variety of sources (or painstakingly creating additional processes to parse data when it could have been collected and stored properly to begin with). This represents the first internal business goal of optimising for costs, based on time or money spent on bad data processing.

2. INTERNAL: LATENCY OPTIMISATION

Another key internal goal of many businesses is reducing latency in decision-making. This is not to be confused with cost optimisation, which can include employee time spent on data recovery. Instead, this goal looks at utilising data to generate fast business decisions on the fly, from the point of data ingestion within an organisation to insight delivery at the end of the process.

Some companies may already have a data process in place, where critical information is passed to management for executive decision-making. Any latency generated in this process (from collection to insight generation) could seriously hamper quick decision making and, by extension, blunt a firm's competitive edge. Unless this is resolved, the problem typically compounds over time, leaving many organisations trailing their competition in an increasingly digitally focused world.

This problem can be frequently seen in the retail industry. Company management typically relies on their distributors to collect and feed sales data back to headquarters. A significant amount of time is then spent collating and analysing this data before any meaningful insights can be generated. In fact, we typically see companies reviewing sales records quarterly (or monthly, at best).

Assuming perfect execution of data collation, after taking these quarterly insights and creating a tactical shift for a product, a few additional weeks has passed, such that management needs to make decisions based on only a handful of major data points each year. While better than semi-annually or yearly, it certainly does not hold up in today's fast-paced business world, where pivots may need to be made on a weekly, daily or even real-time basis.

3. EXTERNAL: INORGANIC REVENUE GENERATION

Executives often have hunches and a sense of how to steer their business in a rapidly changing landscape based on their experience. However, while undoubtedly valuable, these perspectives should be verified by data to either strip away personal biases or solidify instincts, forming part of a cohesive business strategy.

A good example of this is a company looking to offer new products through improving its use of data and digital solutions. Take Netflix: while it began as a DVD-by-mail service back in the late 1990s, it pivoted into streaming services in 2007. With this move, Netflix broadened its user base and began collecting a significant amount of information around its viewership habits and preferences.

Netflix subsequently utilised this data generated from its pre-existing business in two primary ways, namely: (1) the identification of new opportunities; and (2) as leverage for studio negotiations. Addressing the first point, Netflix had created customer metrics (i.e. when viewers play, pause, and rewind, when they watch content, device preferences, and location data, etc.)[4] and identified shows and genres that its viewers desired. This provided them with unique insights into targeted licensing from large studios, saving significant amounts of capital (as opposed to the shotgun approach of "license anything we can"). Secondly, its monopoly on viewership data has given Netflix an outsized advantage when negotiating with studios for subsequent licensing or production deals.[5]

In summary, Netflix created a new core product (streaming) and fundamentally changed its business model; creating efficiencies around capital allocation (targeted licensing) and yielded tremendous revenue growth. All of this was possible through the company's devotion to robust data collection and analysis. Additionally, this created a hoard of data for Netflix to utilise which transitioned the company into the summit of any company's data journey, complete organic revenue generation.

4. EXTERNAL: ORGANIC REVENUE GENERATION

Organic revenue generation occurs when big data begins to develop insights that can directly point a company towards the best way of making more money. In this case, datasets become so large – and analysis so robust – that it begins to develop perspectives from blind spots that manual processing is unable to achieve. As highlighted in Part 1 of our two-part

[4] The Netflix Tech Blog, 'Netflix Recommendations: Beyond the 5 stars (Part 2)', available at: https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-2-d9b96aa399f5
[5] New York Times, 'Sony Pictures enters a streaming deal with Netflix', available at: https://www.nytimes.com/2021/04/08/business/sony-netflix-streaming-deal.html

report series, this is quite a rare scenario, as businesses would need to have such an efficient (and effective) end-to-end data value chain that all data is automatically and readily accessible (and usable) to generate insights from novel questions that decision-makers may have.

In the case of Netflix, the company created a recommendation engine that dynamically alters the movies shown on its landing page, tailored to individual user behaviours. This manifested in an attractive user-centric functionality. While a fantastic feat of engineering, Netflix's recommendation engine is only a front to the true intricacy of the company's data prowess; the ability to create curated content based on viewership preferences. The sheer amount of data Netflix sits on means that it can generate a profile of a movie or TV show that has yet to be produced but has immense potential with its current audiences.

A prime example of this could be seen with its hit series, 'House of Cards', which gave Netflix management the direction they needed with the following viewer insights: (1) David Fincher's a 'Social Network' was frequently watched from start to finish; (2) the British version of 'House of Cards' was well received; and (3) viewers who enjoyed these two series also enjoyed Kevin Spacey and/or films directed by David Fincher.[6]

Netflix went on to invest USD100 million into the series, which has since been credited with cementing the company's dominance in the streaming space, with other hit series launched in the same manner.[7] As of the time of writing, Netflix currently trades at USD 664 per share as compared to USD 23 per share back in February 2013, when 'House of Cards' first launched.

Today, Netflix has a plethora of curated movies and TV series which are not available anywhere else, creating a significant strategic advantage for Netflix and a defensible revenue stream, a clear testament to the power of a tightly managed and holistic data strategy.

GOAL DIMENSIONS

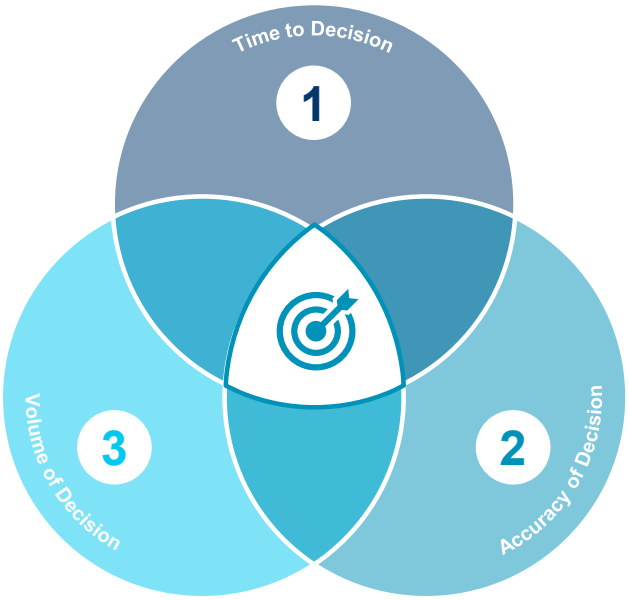
As highlighted earlier in this report, companies waste significant sums of money in data project investments, largely due to an incomplete understanding of the specific business problems they are trying to address.

There are three key dimensions that firms should consider when deciding what level of investment is necessary, centred around the business use case. These include:

1. Time to decision (how fast does the company require information);
2. Accuracy of decision (how precise does it need to be?); and
3. Volume of decision (how many times does information need to be gathered and how many data points are required?) (see Figure 7).

[6] New York Times, 'Giving Viewers What They Want', available at: https://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html?pagewanted=all&_r=0
[7] CNN Money, "' House of Cards' made Netflix a powerhouse. What now?", available at: https://money.cnn.com/2017/11/01/investing/netflix-stock-house-of-cards-kevin-spacey/index.html

FIGURE 7: GOAL DIMENSIONS



1 Time to Decision
Speed at which a company requires information in order to make a business decision

2 Accuracy of Decision
Extent of data precision that is necessary to ensure that the business decision is made effectively and efficiently

3 Volume of Decision
Frequency of data collection per decision or volume of data points needed per decision

Source: Quinlan & Associates analysis

Let's conceptualise this in a simple case study: if a bank wishes to update a customer on their investment portfolio each quarter, the bank will need to collect portfolio movements from the last three months, parse it for the customer in question, and visualise it for them. This is a low frequency process that does not need to be collected at speed. However, it should be incredibly precise, given that the accuracy of the reported portfolio could affect a customer's future decisions.

If we were to map these requirements into the three dimensions, we would require a data system that pulls data infrequently from a few sources, is highly accurate, and scales to meet the size of the bank's customer base. The bank would, in essence, not need to invest in an overly complex system that pulls data in real-time or conducts in-depth analysis or visualisation. It would, however, need a system that scales quickly to new customers and can handle a large inflow of information in a small window of time. While oversimplified, this logical thinking can be applied to large or small business use cases; building for what is needed with a bit of headroom, saving significant overheads, and freeing capital to be allocated to other projects as the company sees fit.

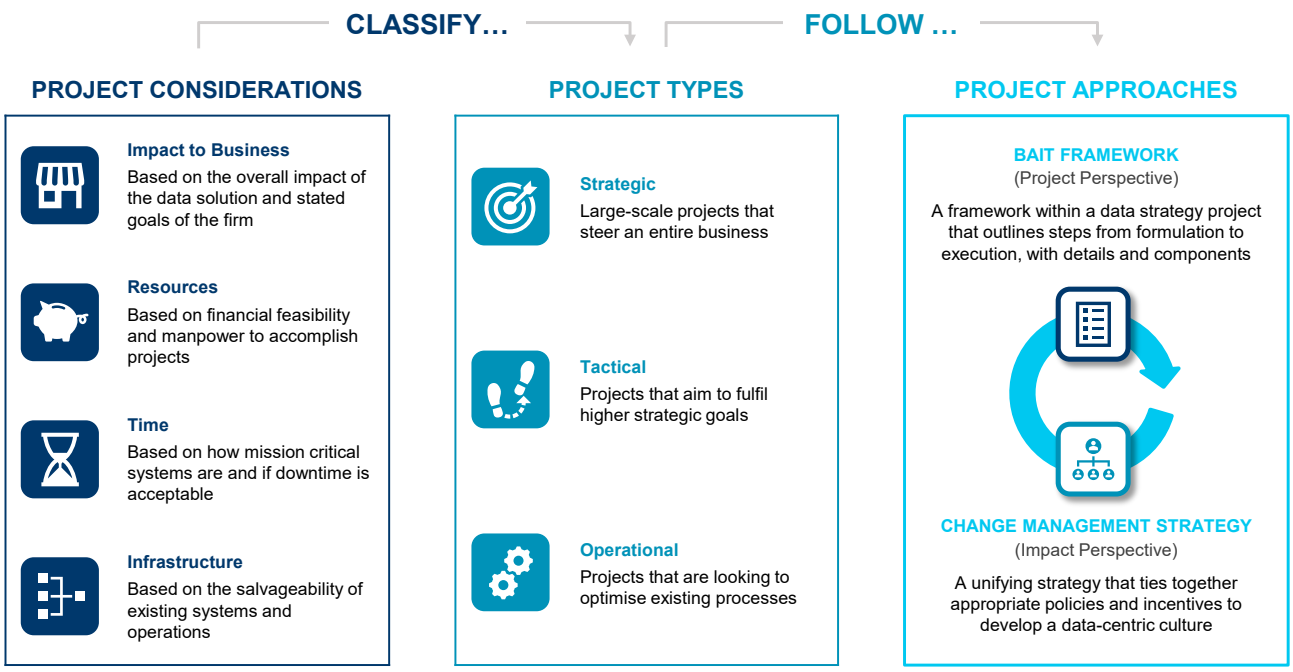
SECTION 2
DATA STRATEGY OVERVIEW

Data strategy projects begin, first and foremost, with a solid foundation of business planning and an understanding of the company's long-term goals. This means delving deep into the rationale behind a company looking to utilise data for business-critical decisions and exploring why it may be necessary.

A data strategy project that begins with the technology first is doomed to fail as companies will attempt to shoehorn it into their current business practices.

There are several tools that companies can use to determine the size and scope of a project while bringing lasting adoption of data to an organisation. These are: the project classification, the BAIT (Business, Application, Information, Technology) framework, and change management strategies (see Figure 8).

FIGURE 8: DATA STRATEGY OVERVIEW



Source: PALO IT, Quinlan & Associates

1. PROJECT CLASSIFICATION

Many executives may not have a clear picture on how much work goes into a data strategy project. To address this, we have developed a classification system that matches the size and scope of a project, along with the capital and resources needed, to determine how large of a commitment each type of project requires. The biggest first cut would be how mature or clear the current business strategy is, given that technology is only an enabler to a company's broader strategic objectives.

KEY CONSIDERATIONS

There are several key considerations that could affect the type of data project selected: impact to business, available resources, time required, and existing infrastructure.

Impact to business stipulates to what degree a given issue affects the organisational strategy as a whole. For example, a simple optimisation of a specific data system would not affect an insurance company's overall direction. However, the shift to a digital-first company would drastically change the firm's business strategy and indicate the need for a large-scale strategic project.

Resources, whether manpower or financial, are a key limitation for many management teams, as full transformations require considerably more capital than optimisation projects. Time is another factor to consider; if a rapid deployment or patch to an existing mission critical system is necessary, then optimisation of existing systems might be more desirable rather than longer transformation projects.

Finally, existing infrastructure should be taken into consideration as some data systems may be beyond salvaging, making it more costly to optimise than to transform.

1.1. STRATEGIC PROJECTS

As the name suggests, strategic projects steer the entire business. These are more commonly known as 'digital transformation' projects, given their sheer size and scope of work. To be effective, these projects must start with defining and solidifying a company's current position and strategic goals, while also considering factors that are integral to its long-term future. Subsequently, firms need to develop a list of objectives and supporting goals, which come with a budget. This informs the organisation's supporting technology and data system choices. We cannot stress enough that this should never be done the other way around.

As an example, think of a traditional insurance company. Many incumbent players still operate under paper claims or laborious manual digital claims, which typically take 2-3 weeks before customers hear back from the insurance company or receive their pay-out. This represents an unacceptable user experience in today's digital world and insurance companies know this; their manual processes and suboptimal internal systems are limited because they were not built from the outset to quickly compute and manage claims (or manage follow ups automatically). These manual processes create significant overheads and minimise the company's ability to create new, tailored insurance products, limiting their revenue growth potential.

Insurtech companies with none of the legacy systems are free to run circles around the traditional insurance industry players, offering microinsurance or tailored premiums based on granular risk calculations, allowing them to tap into a long tail of customers and process claims in a rapid fashion. As a matter of long-term survivability, it is becoming clear that traditional insurance companies are being forced to invest in their digital capabilities before digitally native InsurTechs eat their lunch for good.

A traditional insurance company looking to make a leap into the digital space quickly to defend their market share is a large strategic goal; one that requires complete organisational buy-in, significant investments, and an overhaul of existing systems. Hence, business analysts and IT specialists alike would be involved in the complete BAIT process, set tactical and operational project scopes, and couple it with in-depth change management strategies (along with robust risk management strategies) to ensure implementation success.

1.2. TACTICAL PROJECTS

Tactical projects are a step below strategic, encompassing lower-level organisational goals that bring a company closer to its strategic vision. They also include the structuring of operational projects and goals.

While it may be tempting for companies to deploy tactical projects, due to their limited scope and relatively smaller budgets, we recommend against this if the company's broader business strategy has not been established. This is one of the core reasons why dark data and wastage around data investments occur. All too often, we have seen executives charge headfirst into a tactical project to solve a smaller business goal, which ends up conflicting with other legacy systems or current practices without clear demarcations of why a system is being implemented at the highest level of the organisation.

In keeping with the insurance company example: a tactical project would be akin to the company looking to develop a comprehensive risk calculation engine based on collected historical data and other third-party sources to create tailored premiums and bespoke products for their long-tail customers. In this case, the strategy is already well thought out, with its budgetary and regulatory limitations set, so the project would only include the AIT portions of BAIT, simplifying the project and reducing the resources needed to implement it. Some activities would include build or buy considerations, along with integration with the company's data catalogue and other pre-existing systems.

1.3. OPERATIONAL PROJECTS

Operational projects are largely focused on process optimisation. They typically occur when a company's data fundamentals are already squared away and most of the relevant infrastructure is in place. Relative to strategic and tactical projects, operational projects are considered the smallest in terms of scope and are more technically focused.

Looking again at the example of an insurance company wanting to create a risk calculation engine, this would operationally require a data filtering system that sieves the relevant information needed for the risk engine to create a risk score. If the company already has a data filtering system in place but experiences latency issues when parsing information to the final risk engine, the project's scope would be limited to determining the root cause of the latency issue, optimising for costs associated with the delay and identifying the optimal technical solution(s) required to address it.

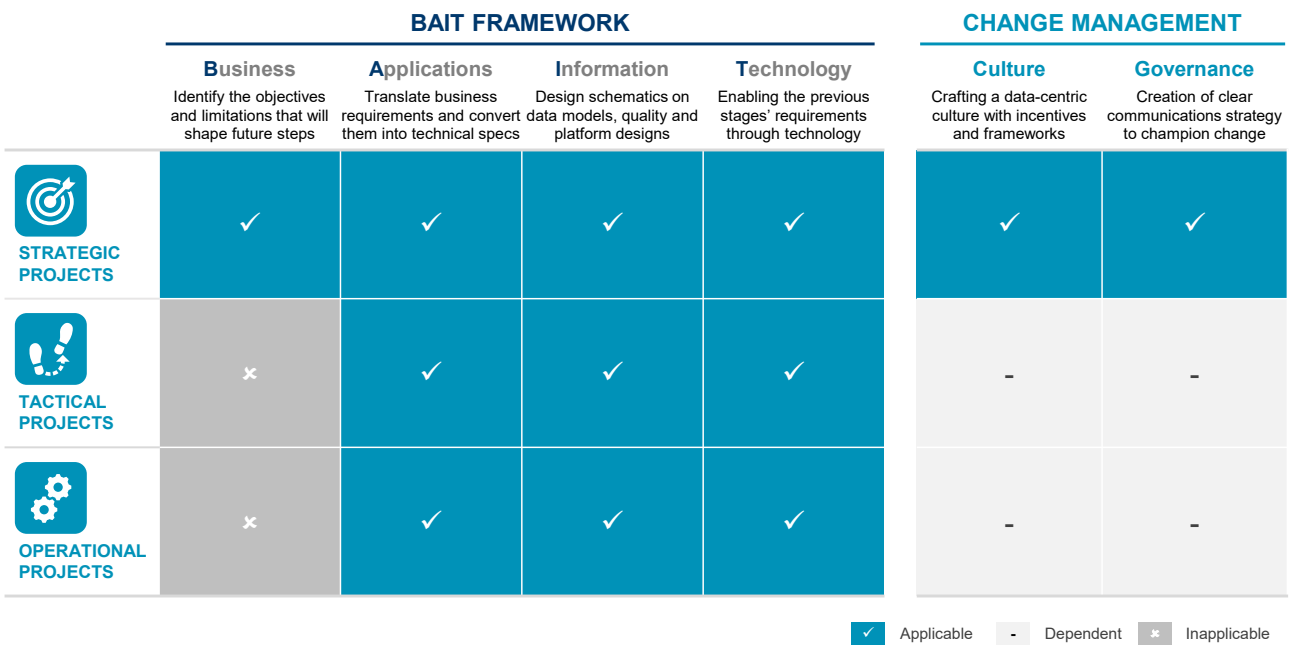
2. BAIT FRAMEWORK

BAIT is a framework that helps to showcase how a data strategy project should be approached from formulation to execution, while detailing the individual components necessary for a successful project (see Figure 9).

The BAIT framework consists of four parts, all of which represent a vital step in the overall process:

1. Business
2. Application
3. Information; and
4. Technology.

FIGURE 9: PROJECT TYPE & FRAMEWORK MATRIX



Source: Quinlan & Associates analysis

1. BUSINESS

The BAIT framework begins with business problem diagnosis. This is one of the most important steps in the process, as it requires understanding the fundamental nature of the problem the business faces. For example, is the organisation looking to optimise a current practice or refine a feedback loop? Or is it looking beyond its current capabilities to create new business lines or business models?

Organisational resources and obstacles need to be identified in this step. This requires Business Analysts / IT Specialists to work alongside senior management in formulating objectives based on the company's ambitions and identify obstacles early to prepare for the subsequent steps. Failure to meet analyses of objectives, legal issues, or costs that apply to an organisation could cause serious delays to the project and/or deliver subpar results for the given investment.

2. APPLICATION

The Application stage involves translating business requirements into technical functionalities, to be fulfilled by applications developed by the IT consulting team. The application stage serves as an interface between the business and its information, meant for viewing and performing operations on data as per business requirements. Additionally, this is the stage where proof-of-concepts are built and tested with the company, creating a feedback loop for refinement, ensuring the outcome is tailored to stakeholder and business needs.

3. INFORMATION

The Information stage takes the technical capabilities needed from the Application stage down to the data modelling level. This involves the complete discovery and mapping of a business's current data locations (including rediscovering previously mentioned dark data hidden throughout the organisation), current data quality, its owners, and the current interoperability of each business unit's data systems and operations. This is done to create the relevant data models, systems, maps, and data platform designs that are needed to inform subsequent technology selection. If executed holistically, an organisation would create data governance policies and systems that produce trusted data (i.e. information which is immediately usable for analysis by end business users) for anyone in the company.

4. TECHNOLOGY

Last, but certainly not least, the Technology stage involves matching the most suitable technologies to the applications and data platform designs, as stipulated by the Application and Information stages.

It is important to note that this would be within the Business stage's scope of timeline, budgeting, existing infrastructure, and objectives, which need to be fulfilled. In essence, technology needs to be considered last and depends on all of the previous stages. Once again, it is important to stress that technology is merely a tool to help businesses meet their objectives; executives should be wary of "quick fixes" or "promises" that frontier technologies can bring.

CHANGE MANAGEMENT STRATEGY

Tying together a data strategy requires more than just the most advanced hardware, software, or talent; it requires a unifying change management strategy.

Transformations and changes often face resistance from anyone subjected to it, and a data strategy project is no different. Without it, a data investment may not reach its full potential as employees retain old habits, failing to utilise new data tools and processes made available to them.

As such, it is crucial for an organisation to develop appropriate policies and incentives and create a data-centric culture, reinforced with clear governance models. This would encompass four key areas: (1) a communications strategy from the top; (2) actionable frameworks for employees; (3) incentive schemes to promote best-practice data procedures; and (4) agile governance to reflect a fast-changing regulatory environment.

Every change requires a champion within an organisation to see it through. Senior management should endorse and sponsor an entity to oversee the change management strategy with autonomous decision-making powers. This is typically known as an internal enablement committee. Enablement committees are needed at every stage of the project; from sponsoring changes to signing off on budgets to monitoring the overall progress of the company's data strategy. Both the committee and change management strategies are further elaborated in Section 9: "Change Management Strategy".

SUMMARY

Companies need to carefully tailor their data strategy with these three tools in mind to match their business objectives. This should be paired with near perfect execution of holistic data systems for all business decisions, as doing otherwise would result in, at best, a loss in productivity or, at worst, direct monetary losses.

While large-scale data projects may not be a frequent occurrence for many companies, it is important for firms to engage specialists, who can help steer companies away from common pitfalls when implementing data projects at scale, ensuring the best chance of successful delivery.

TYING TOGETHER A DATA STRATEGY REQUIRES MORE THAN JUST THE MOST ADVANCED HARDWARE, SOFTWARE, OR TALENT; IT REQUIRES A UNIFYING CHANGE MANAGEMENT STRATEGY

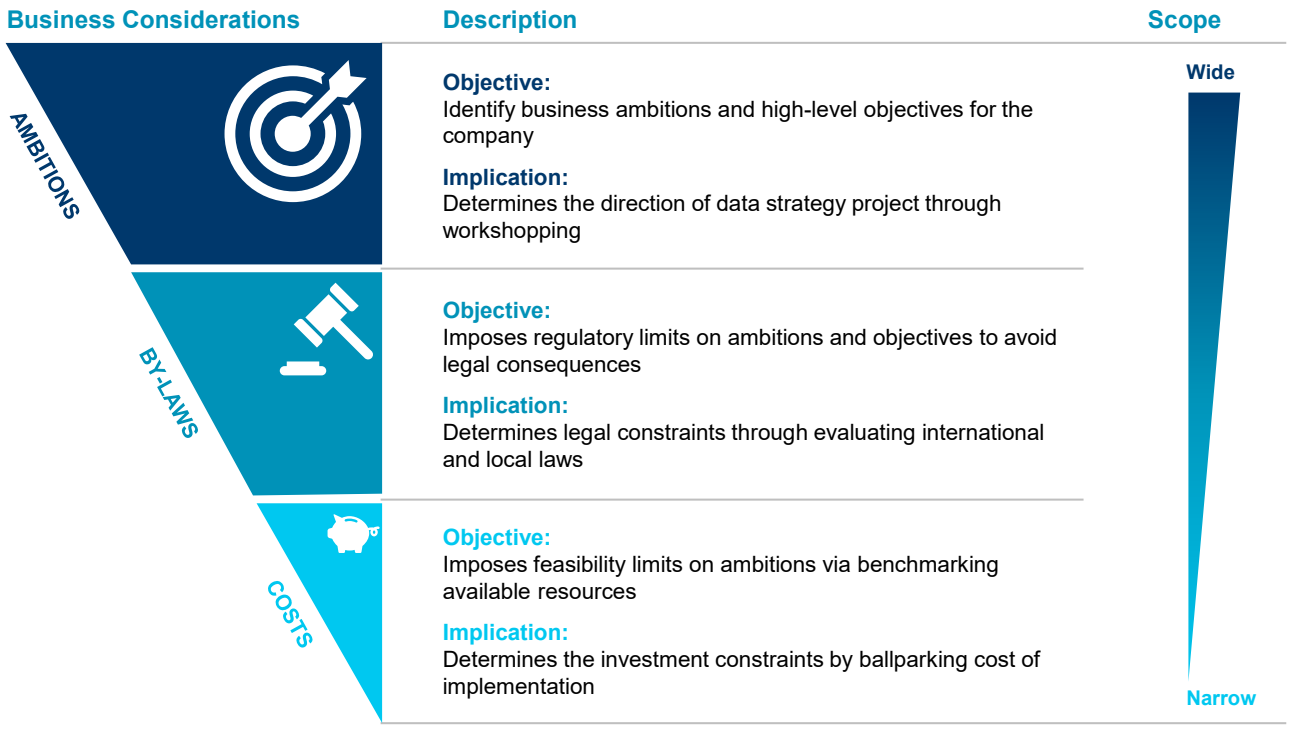
SECTION 3
DATA STRATEGY – BAIT (BUSINESS)

1. BUSINESS STRATEGY FIRST

To avoid the issue that so many data initiatives suffer from (i.e. the lack of a defined goal), developing a robust business strategy is a crucial first step in shaping the direction of a data strategy project. To achieve this, Business Analysts ("BAs") need to conduct a high-level assessment of the business segments involved through the ABC framework: Ambitions, By-laws, and Costs.

In addition to pinpointing the company vision, mission, and overarching objectives, ABC also helps to identify areas of development to ensure the project does not encounter any unexpected obstacles (see Figure 10).

FIGURE 10: ABC FRAMEWORK



Source: Quinlan & Associates proprietary framework

1.1. AMBITIONS

First and foremost, it is important to understand an organisation's overall ambitions, including the specific objectives that it hopes to reach (i.e. the "why?"). This is usually accomplished by conducting workshops in which business analysts meet with key stakeholders to review their existing business model, specific areas of development, and key business objectives for the next 1 to 5 years. This period also includes a research phase, in which BAs should aim to develop a detailed understanding of the company, its competitive landscape, and the wider industry in which it operates.

After workshopping with stakeholders, BAs should be able to identify whether the project is being created in response to an existing issue or a prediction of future outcomes. Identifying

ambitions will also determine if the data project is intended to add or improve a current practice.

As this step will define the future pathway for the project, creating a discussion on key drivers and risks is also critical. Developing a high-level direction and identifying objectives from the company's ambitions will inform the purpose of the data project, such as deciding the key metric(s) or output(s) that are needed for success.

1.2. BY-LAWS

By-law considerations will further shape the direction of the company's data strategy, requiring in-depth analysis of the regulatory landscape locally, regionally, and internationally, depending on which jurisdictions the company operates in.

Some examples would be data collection or usage laws that stipulate the granularity of data collected or anonymisation procedures needed before using it for business purposes. These by-laws will also depend on the industry in which a company operates and what needs are being addressed through the data strategy project. For example, if a company wants to collect customer data and sell it to another firm in a B2B manner, it may run into legal issues within an operating region.

Legal considerations help to determine if the company's ambitions are well founded or whether they are even possible, barring all other variables. Once a data project is deemed legal and open for consideration, a company should then move to understand the costs that come with implementing their ambitions.

1.3. COSTS

After understanding the company's ambitions and legal constraints, the BAs can then narrow the scope by considering the logical investment needed for such a data project. This is based on both a "ballpark of cost of implementation", as well as the company's existing resources, such as its available capital, technologies, and manpower. From this, the BA can evaluate the feasibility of a firm's ambitions and benchmark it against anticipated costs and resource requirements.

The scale of change that is desired with new data outputs is also dependent on the cost of change. For example, if a firm's objective is to develop a new line of products developed from the new insights, BAs must ensure the client has the necessary capabilities for such a transformation. If company resources are insufficient in meeting the cost of transformation, the BAs may need to consider steering the data project towards an alternative solution that fulfils the original business objective within the company's means. If these objectives can be fulfilled, the BA can proceed to the next stage of state analysis.

In essence, ABC creates the high-level business direction and outlines limitations that will shape the entire the data strategy project. Following the ABC framework will help prevent common misalignments between the data project and business goals / capabilities.

2. DATA DOMAINS OF BUSINESS

After outlining the business objectives and limitations, companies need to understand which data domain(s) within a business is relevant to the respective objectives, ensuring they can identify the key areas of a business that would be affected by the scope of its data strategy.

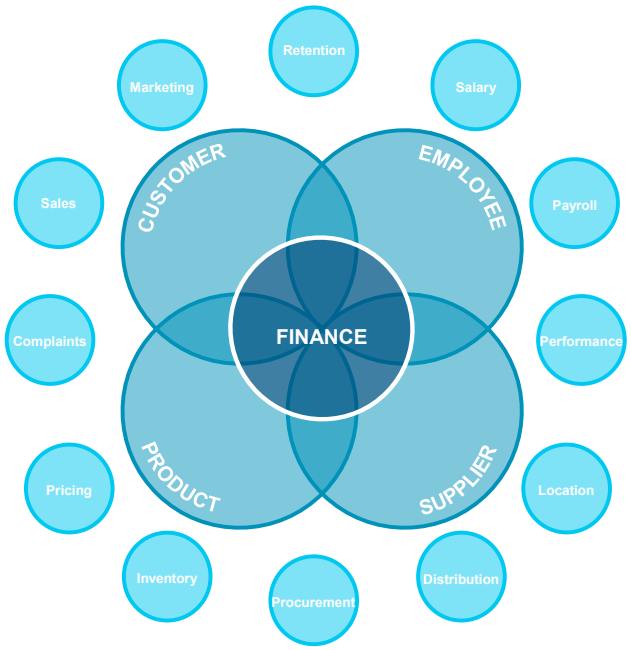
A data domain is a function within a business that can encompass multiple business units in the pursuit of specific commercial use case (see Figure 11). Take “customers”, for example: what does “customer data” entail? It could be simple identifiers like a customer ID or contact information. However, if the customer domain stopped here, it would hardly be useful to a business user.

Holistic customer profiles contain information ranging from sales data, online presence, complaints, loyalty information, marketing, financial history etc. These diverse data points are typically scattered across the business; from sales and marketing to customer service to finance. These would all connect under the purview of the “Customer Domain”, which cannot be siloed into a single business unit or department. This is why data experts use domains, i.e., to strip away the confusion around duplicating data, where it is housed, and who owns it.

You may wonder why ‘Finance’ sits at the centre as a “Master Domain”, above all others. The reason for this is simple: businesses, at the end of the day, are primarily evaluated by one key metric – money in and money out. Everything within an organisation contributes in one way or another to value, which is measured in terms of financials, and data usage is no exception.

Take customer information, where everything about a customer invariably ends up back as a form of revenue contribution (or cost) to the company. Business leaders act to either drive revenue or reduce costs to keep a company afloat. Finance informs procurement decisions, employee benefits, or product R&D, reinforcing its place at the centre of any business.

FIGURE 11: DATA DOMAINS



Domain Type & Description

Master Domain
Finance is considered the master domain as it is the key indicator of any business and all other domains reference or revert to it

Primary Domains
A primary domain encompasses a major business function or multiple assets which small secondary domains reference

Secondary Domains
Relevant data retrieved from a specific business unit or data system which collectively makes up a primary domain

As business activities ultimately drive profit & loss, finance department should ideally be the master domain

Finance
(Master Domain)

	Customer (Primary Domain)	Employee (Primary Domain)	Product (Primary Domain)	Supplier (Primary Domain)
1 Marketing	✓	✗	✗	✗
2 Complaints	✓	✗	✓	✗
3 Sales	✓	✗	✓	✗
4 Retention	✓	✓	✗	✗
5 Procurement	✗	✗	✓	✓
6 Pricing	✗	✗	✓	✗
7 Distribution	✗	✗	✓	✓
8 Performance	✗	✓	✗	✗
9 Payroll	✗	✓	✗	✓

✓ Applicable ✗ Not Applicable

Note: A secondary domain that falls between two primary domains means that the data would offer insights to both
Source: PALO IT, Quinlan & Associates

Data domains typically have a data steward; a stakeholder who is responsible for constructing an accurate domain by ensuring that the data required: (1) exists; (2) is of a high-quality standard; and (3) is connected to other domains that require the data. This is covered in depth within the data governance segment of this report. The main takeaway here is that there are data domains that cross all boundaries of an organisation.

Understanding domains and noting its level of implementation is critical for the initial stages of data mapping and will generate a preliminary understanding of the company's data gaps for subsequent remediation. If a business has not thoroughly thought through its domains and mapped them to its operating model, it will show up as deficits in data efficiency. This would inform experts on a company's overall data maturity.

IF A BUSINESS HAS NOT THOROUGHLY THOUGHT THROUGH ITS DOMAINS AND MAPPED THEM TO ITS OPERATING MODEL, IT WILL SHOW UP AS DEFICITS IN DATA EFFICIENCY

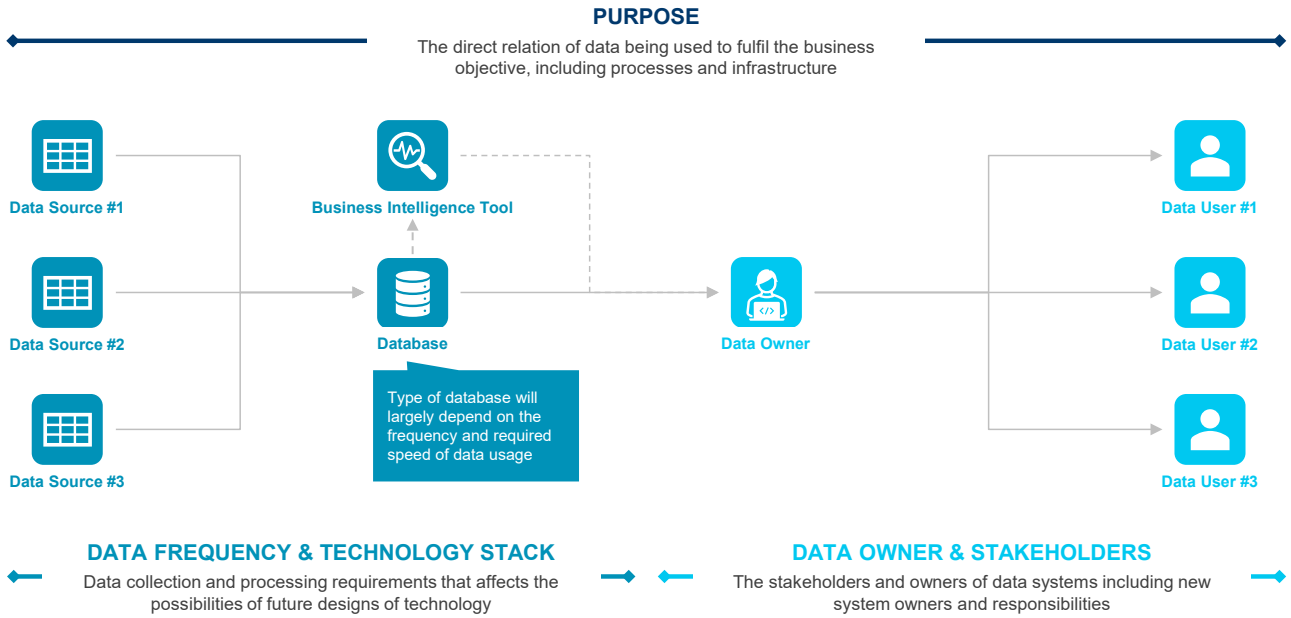
3. DATA CURRENT & FUTURE STATE ANALYSIS

After identifying a company's core business objectives, obstacles, and mapping its affected domains, a current state diagnosis on a potential data project is conducted. This encompasses basic tenets, such as the purpose of data collection and usage, the frequency of data usage, the technology stack, and current stakeholder obstacles.

There are four key components that need to be gathered to establish a current state and future state. This process allows an organisation to conduct an effective "gap analysis", which creates a potential timeline and scope of work (see Figure 12).

1. Purpose of the data strategy project;
2. Frequency of data usage;
3. Technology stack used or possible replacements; and
4. Current and future stakeholders.

FIGURE 12: STATE ANALYSIS COMPONENTS



Source: PALO IT, Quinlan & Associates

3.1. PURPOSE

It is important to note that 'Purpose' is distinct from 'Business Ambitions', as stated in the ABC framework. Purpose, in its simplest form, relates to how the data is being used to fulfil a business objective, as directed by business ambitions. This encompasses data processes such as collection, storage, usage, and disposal, as well as data systems such as infrastructure or architecture. For example, if Company A is looking to increase revenue, this is a 'business ambition' and objective of the organisation. The company also collects data on how much product is sold by its respective distributors to help direct product supply to match demand, increasing revenue; this is the data's 'purpose'.

In this stage of the process, specialists seek to understand what data systems or processes are currently in place (if any) to fulfil the overarching business objective. This is incredibly valuable, as it provides a high-level picture of how effective or efficient a company's current data operations are. Following this, understanding a future use case of data under the umbrella of meeting business objectives is critical for steering the scope of the data strategy project.

3.2. FREQUENCY

'Frequency' refers to the regularity of data collection and/or processing of current data systems.

For example, Company A collects and processes sales data from distributors on a monthly basis. Given the low-volume of data collection and processing, this would imply that the current system relies on batch-processing (i.e. pulling data after it has been consolidated over time) as opposed to real-time processing (i.e. near instantaneous pulling of data).

Understanding data collection and processing frequency helps specialists to factor in other possibilities, such as potential partnerships needed, especially if data collection involves third parties when developing potential data architecture designs (or relationship models for future state designs).

3.3. TECHNOLOGY STACK

All IT systems operate on top of a technology stack, which is defined as a set of technologies an organisation uses to build capabilities and services that support business activities.

Typical enterprise IT systems are composed of diverse topics, ranging from hardware infrastructure to programming languages, frameworks, and third-party tools used not only by engineers, but also by any other internal and external stakeholders, regardless of the business nature.

Data projects include four key areas of interest, all of which are used to store, process, and access data:

1. Infrastructure;
2. Data management;
3. Systems of record; and
4. Systems of engagement.

3.3.1. INFRASTRUCTURE

Companies rely on hardware infrastructure such cloud computing services or on-premises data centres to host and consolidate data, where necessary. Some popular examples include Amazon Web Services, Microsoft Azure, Snowflake, Microsoft Servers, and Linux Servers.

Choice of infrastructure is dependent on cost, regulations, and use cases. However, the current configuration of infrastructure greatly influences the potential design of applications hosted on top of it. For example, a legacy system may require a full migration from on-premises to the cloud to support a new business initiative, which will greatly affect the timeline and complexity of a data project.

3.3.2. DATA MANAGEMENT

Data management is the combination of technologies and protocols that maximises the business value of data throughout the entire lifecycle of data creation, access, modification, and storage.

Businesses aim to increase consistency and confidence in decision making, improve regulatory and security posture, and maximise the income generation potential of data by combining different software, frameworks, and governance practices in this layer. It is especially important to understand the specific business objectives and processes that are supported by the different ways data is collected and processed inside the data management systems.

Common topics revolving around data management include data warehouses, data lakes, business intelligence, and data governance.

3.3.3. SYSTEMS OF RECORD

Systems of record is defined as a collection of systems that serve as a shared source of truth for the business units to make accurate and consistent decisions across the organisation. Examples include customer relations management (CRM), enterprise resource planning (ERP), and human resource management (HRM). These systems are most often consumed by employees and auditors as part of the business processes throughout the entire business value chain.

3.3.4. SYSTEMS OF ENGAGEMENT

Systems of engagement represent the actual touchpoints for data users. These systems provide functionalities that equip customers, partners, stakeholders, and employees with the tools to interact with the business. Examples include email services, point of sales (POS) systems, employee collaboration platforms, and client mobile or web applications.

These four components make up a basic data technology stack and require significant resources dedicated to each layer. It is here where specialists take note of both what is currently being adopted and if transformation is required to meet data use cases for the business objective.

3.4. STAKEHOLDERS

Understanding the stakeholders and owners of each respective data system is critical in ensuring a smooth data strategy project.

Data systems that are siloed away under separate departments and owners can indicate potential administrative friction if consolidation occurs under planned future state systems. Stakeholder analysis should also understand what responsibilities are currently being held by existing data system or domain owners. This is particularly important for a company that has not adequately explored their business domains and mapped them to respective data systems. This would help inform how a data project would change or add new responsibilities to existing data or domain owners, including who would take responsibility for new systems.

Specialist training and education are critical in ensuring that targeted domain owners are ready to maximise the potential of future state systems. Additionally, specialists would take note of any mission-critical systems that may affect other users besides those stipulated in the scope of work. For example, a bank's treasury data system is the heart of any bank, and downtime for a treasury system would affect nearly every other system within a bank. Such considerations are important as they affect implementation timelines and potential stakeholder buy-in.

4. CAPABILITY GAP ANALYSIS

After gathering both current and desired future state data systems, a gap analysis should be completed with the overarching business ambition in mind, with additional consideration given to legal and/or resource constraints.

If a current state outlines where a business is today and a future state identifies where it would like to be tomorrow, a gap analysis provides the high-level roadmap for the journey ahead. In essence, the output of this exercise should be a list of business requirements that are mapped to functionalities for both current and future state designs (see Figure 13).

FIGURE 13: GAP ANALYSIS (ILLUSTRATIVE)

Gap to be explored	Technology	Talent
	• Based on future and current designs provided • Critical functions are prioritised	• Based on talent requirements needed for high-level future state blueprint
Example	• Need for real-time data pulls and processing require a future-state database application • Migrating from on premises to cloud solutions	• Need for a new domain owner which would manage future-state customer systems • Selected internally or hired externally

Source: PALO IT, Quinlan & Associates

4.1. TECHNOLOGY GAP

Building upon the existing technology and architecture designs (if any) of an organisation, specialists identify the areas of improvement and prioritise them according to project importance. It is at this point where possible solutions to the gap are drafted and explored. Areas of note include what technologies need to be deployed in order to achieve the future state data system.

A basic example could be a company moving from an on premises, quarterly batch data collection system to a high velocity, real-time data collection system. This shift would require key technology infrastructures, namely: (1) a flexible “plug and play” system for applications to connect to; and (2) a scalable data storage solution. In this case, migrating to a scalable cloud solution could alleviate this issue. This is an example of a technology gap being identified through both the current and future state analysis and existing architecture designs.

4.2. TALENT GAP

Simply building a data system for a business is not enough to ensure its long-term viability. After all, it is people who both manage data and serve as data custodians if systems go awry. Selecting qualified individuals within the organisation or external hiring talent to meet these responsibilities is key to ensuring that a new data system lasts and is adequately cared for.

To ensure the right person is selected or hired for specific roles, the gap analysis should: (1) outline the role description and responsibilities; (2) identify key credentials and history of data ownership and; (3) allocate a budget allocation for said role. These individuals selected would then formulate and own new practices with regards to these data systems, ensuring it reaches its maximum potential for the organisation and for those who interact with it. Specialists should seek these advocates within the organisation or new hires during the early stages of the project. This is crucial when establishing new data operations, as resistance to change is a major hurdle companies need to overcome during any overhaul or process change.

The end result of the Business stage should be a well-defined business strategy with a complete list of business requirements that the future state data system must achieve. This list would be a part of a pair of documents, which also includes future state roles and responsibilities.

SECTION 4
DATA STRATEGY – BAIT (APPLICATION)

Application is the stage where both business analysts and IT specialists work in tandem to translate the business objectives, gaps identified, and business requirements into functional requirements, which are then fulfilled by corresponding applications.

REQUIREMENTS MAPPED TO APPLICATIONS

An application is defined as a type of software that allows a company to perform specific tasks; from simply fetching information to advanced data manipulation. It is the interface that equips business users with a means of viewing – and performing operations on – information stored in data systems.

At this stage, business requirements should be converted into functional requirements for applications to fulfil. An example of a business requirement would be “Chief Financial Officer and Finance Department want to see company P&L trends across all business domains in real time”. We will be using this example throughout the rest of this report for ease of digestion.

This business requirement can be split into multiple functional requirements, which would include:

- "Financial data is filtered by the relevant domain and domain owner”
- “Filtered financial data is pushed to the finance domain visualisation engine”
- “Visualisation engine updates visuals if new financial data is available”

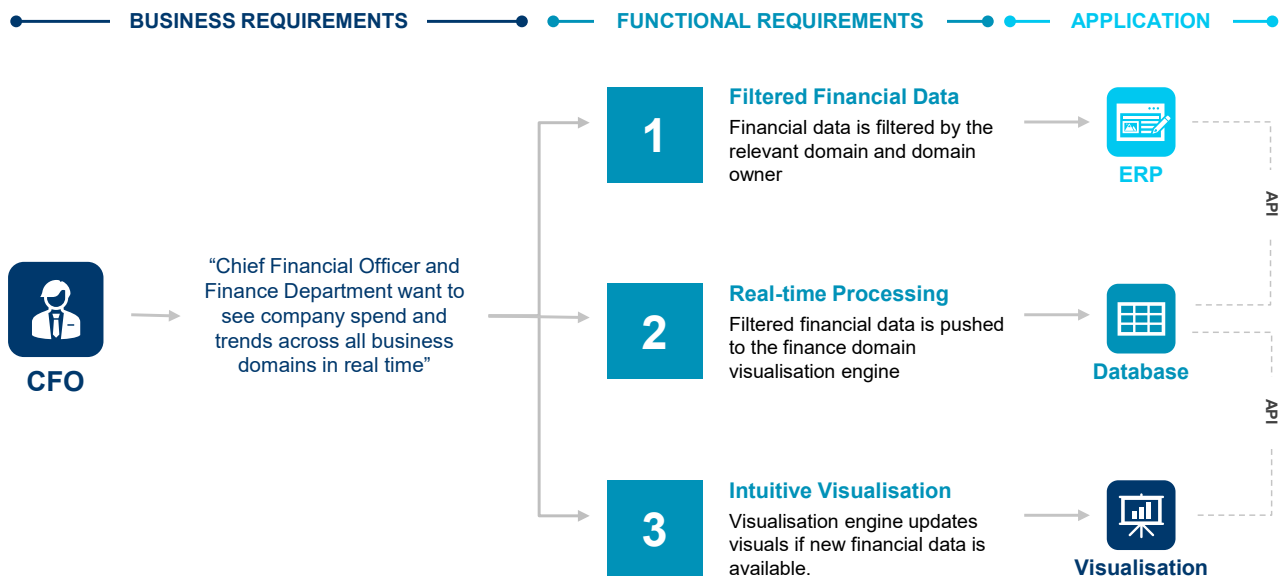
These three functional requirements would be sequentially mapped to a specific application(s) until all business requirements are met (see Figure 14).

In the above example, a visualisation engine is considered as an application that visualises processed financial data. The relevant databases that store financial data also require an application to register new data as an “event”, filter the new data, and “push” the filtered data to the visualisation engine, making this an “event-driven” process. Given that these would be considered two separate applications, an integration layer would be needed between them. An example would be an Application Programming Interface (API), which is a set of definitions and protocols for integrating application software (elaborated further in subsequent Information and Technology stages).

Business analysts and specialists would also look at follow-on technical essentials and the implications that arise out of these requirements. For example, to see everything in one place, a real-time financial monitoring dashboard for an end-user like a CFO would be required. “Pushes” from upstream data sources would constitute an “event”, which needs to be collected, collated, and processed together before visualisation. Finally, as this dashboard of information would steer company decisions at the C-suite level, the data would need to be highly reliable, validated, and ‘trusted’ before it reaches the dashboard for viewing.

In addition to understanding the nuances behind such requirements; functional requirements should include fringe business requirements such as disaster recovery, i.e., in the event of a catastrophic failure, applications may be in place to secure data and backup systems.

FIGURE 14: REQUIREMENTS TO APPLICATIONS EXAMPLE



Source: PALO IT, Quinlan & Associates

These applications and their functions would all be stipulated next to each functional requirement before being considered "complete" and appropriately mapped. This ensures business analysts and IT specialists can dutifully ensure that all business requirements are met, while setting up a firm foundation for designing data platforms and subsequent architecture blueprints, respectively.

FUNCTIONAL REQUIREMENTS SHOULD INCLUDE FRINGE BUSINESS REQUIREMENTS SUCH AS DISASTER RECOVERY

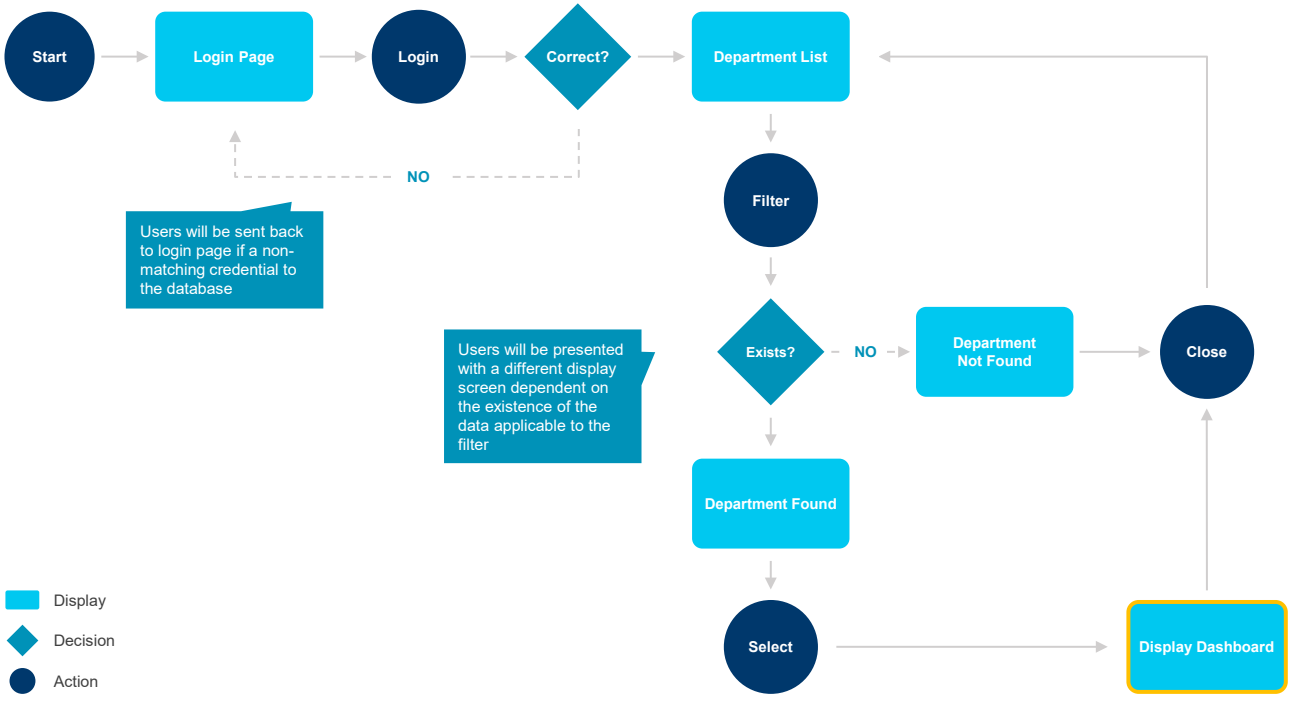
USER FLOW MAPPING

After determining the functional requirements of the project and the main applications supporting it, the next step would be mapping how an end-user interacts with the application (e.g. a financial monitoring dashboard). This user flow mapping is needed as there could be many pathways a user could take to reach a desired outcome. To ensure that the system operates in a way that is desirable, user flow should be mapped in a systematic or sequential way. This will become important when selecting or designing the final dashboard application in the later stages of the project.

User flows are visually represented by a flowchart (see Figure 15), beginning with a user's initial interaction with the application (e.g. opening the financial monitoring dashboard) and ending with a final action or outcome (e.g. viewing the dashboard populated with updated financial data).

Depicting this process allows IT specialists to evaluate the user experience (for subsequent user interface designs if being built in-house), ensuring that the relevant data is pulled, processed, and displayed in a desirable manner.

FIGURE 15: USER FLOW MAP*



Note: Based on Real-time Financial Monitoring Dashboard example
Source: Quinlan & Associates

DATA STRATEGY – APPLICATION ROADMAP

1. BUY / BUILD CONSIDERATIONS

Once a high-level list of applications is drawn up, any leftover gaps from missing functional requirements or gap analysis should be shored up.

Potential solutions to the gaps identified would need to undergo a cost-benefit analysis to determine whether to build up existing systems or buy one outright from the market. This analysis should combine both business and IT perspectives, given business analysts understand resource limitations of the business while IT specialists understand the baseline capabilities needed for a future state technical system. Key considerations will include: (1) affordability; (2) flexibility of the application in the event changes are necessary in the future; (3) personalisation of features; (4) scalability; and (5) the time needed to launch the application.

The time taken to either build/enhance an existing system or migrate to an entirely new bought system should be incorporated into timelines for implementation (see Figure 16). In the case of a financial dashboard, there could be visualisation applications which are already readily available via subscription such as Tableau; a far simpler option than building one in-house.

FIGURE 16: BUY / BUILD CONSIDERATIONS

	1 BUY	2 BUILD
	Direct Purchase / Subscription	**In-House Development**
DESCRIPTION	A purchase or subscription of a pre-existing application or service from a third party which could be modified to fit the project's needs and requirements	A complete end-to-end customised development process of an application which is completed in-house, tailored to a project's specific needs and requirements

CRITERIA	DESCRIPTION
Affordability	• Extent of one-off or ongoing financial costs required to maintain the application
Flexibility	• Level of difficulty in adjusting components of the application to suit the company's future needs
Personalisation	• Availability of customisation options to create specific features in the application according to the company's needs
Scalability	• Ability, ease, and level of cost-efficiency to expand the capacity of the application
Time-to-Market	• Estimated length of time it takes for the application to be offered readily to users

Source: Quinlan & Associates

2. TIMELINE

A timeline of the data strategy project should be created to ensure the project meets business time constraints, while also giving specialists a time-based tracking tool.

There are several dimensions that a timeline must include: (1) task and description; (2) owner of the task (detailing who is responsible, accountable, consulted, and informed on the task); (3) time required and milestones; (4) sequential tasks (if any); and (5) resources needed. This should be done at every level and type of project, be it strategic, tactical, or operational. This is particularly useful for project sponsors such as enablement committees, project teams, and employees to align on. Sponsors and enablement committees will be elaborated further in Section 9: Change Management Strategies.

The timeline should include everything from the kick-off of the data strategy project till implementation completion. Good practices include adding buffer time for certain tasks, such as creating a proof-of-concept or minimum viable product, or the iterative technology stage, where feedback is required from end-users. This ensures that a desired, finalised project is made in accordance with an organisation's preferences.

MINIMUM VIABLE PRODUCT / APPLICATION SELECTION

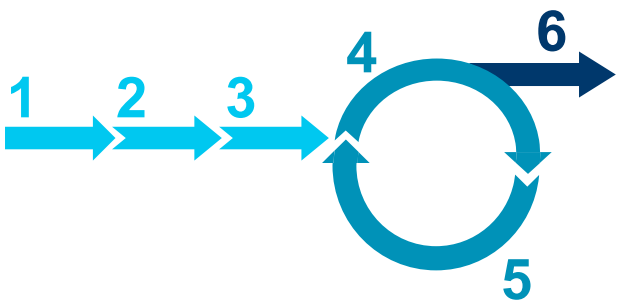
A minimum viable product ("MVP") is a basic, launchable version of an application or product that supports minimal yet "must-have" features.

An MVP is created with the idea that rapid iterations will be made based on feedback from end users. Iterations include bug fixes and introduce features that are discovered as critical along the way to a full launch. This feedback loop means MVPs are extremely useful in managing senior management expectations, given they can see an application being built piece by piece.

MVPs typically consists of a six-step cycle. The good news is, should specialists complete the Business stage correctly while completely mapping out functionalities, steps 1 to 3 of the MVP cycle can be easily bypassed, saving time and manpower. For example, using the real-time data visualisation case above; if an MVP of this singular business requirement were to be produced, each functional requirement would be tackled and created sequentially. The MVP cycle continues until a fully functional product is formed and launched with client approval.

An MVP cycle is applicable both when applications are built in-house or when a third-party provider demonstrates an application for the end user to evaluate and select; in the case of the financial monitoring dashboard, the finance department and the CFO (see Figure 17). The only difference between them is the final step of launch or approval by the end-user.

FIGURE 17: MINIMUM VIABLE PRODUCT



Initiation
Map out business requirements that will determine core functional requirements of the MVP

Development
Decide on building in-house vs. seeking a 3rd party provider to build an application and test it out to end users for evaluation

Implementation
Integrate the product into either planned or existing data architecture upon approval of final application

1 Determine Value Proposition
Determining the exact value that would be derived from a new application

2 Map User Flow
Designing high-level user interaction flow from front to back processes

3 Prioritise MVP Features
Prioritizing functionalities of the MVP with high-dependency systems first

4 Build / Explore
Building out core functionalities / Exploring 3rd party application providers

5 Review & Repeat
Consolidating user feedback and incorporating it, repeat steps 4 and 5

6 Launch & Integrate
Set aside the approved final application of product for subsequent integration

Source: PALO IT, Quinlan & Associates

1. DETERMINE VALUE PROPOSITION

Specialists and the end-user should align on the value that would be derived from a new data application; in the case of a dashboard, to display information in a clearly digestible, visualised manner. If an application does not help meet business objectives in a meaningful way, additional research should be conducted to find alternative solutions that meet business needs.

2. MAP USER FLOW

Once the value proposition is agreed upon, designing an application and how users interact with it – from front-to-back – should be mapped at a high level, based on the functionalities required. The steps required to reach the main objective of an application should be laid out in basic tasks. In the case of the financial monitoring dashboard, the process of logging in and filtering data are key steps a user will take before the information is visualised.

3. PRIORITISE MVP FEATURES

By this stage, a list of features and functional requirements will have been generated. IT specialists should prioritise core applications before building occurs. All features and functional requirements should be appropriately labelled based on order of priority: high, medium, and low, with high dependency systems such as production databases going first.

In the case of a real-time dashboard, priority should be given to the visualisation touchpoints that are visible to the user, the gathering of financial data, and the processing of disparate financial data sources to create a visual. Once priorities are set and labelled, IT specialists can identify the components needed to start

creating an MVP or create the relevant evaluation criteria needed to seek external providers, if preferred.

4. BUILD / EXPLORE

This is the main stage of the MVP process and depends on all prior designs and requirements being appropriately addressed. It is here where IT specialists create a fully functional MVP that covers only the prioritised core features. Functionality would be prioritised over form in all cases; the MVP should be barebones, and the user interface fleshed out after an end-user agrees to the baseline functionalities.

In an application selection process, analysts will explore various third-party applications by reaching out to external providers with the list of requirements and schedule demos for the end-user, as per the agreed timeline.

5. REVIEW & REPEAT

Upon fulfilment of the core priorities, the MVP is delivered to end-users for feedback, validation, and approval. If approval is not secured, developers will improve functionalities based on feedback and continually create new iterations for review. Only upon end-user approval will the team shift to lower priority features. This cycle repeats until a fully-fledged MVP of the final application is created.

In the case of selecting a third-party application, this stage would simply include the back-and-forth feature enhancements (if required) by both the end-users and external providers. If an application does not meet the required standards, analysts return to the exploration phase, seeking out other third-party providers. Only when all requirements are fulfilled (or the gap is adequately filled) can negotiations occur. Negotiations around duration of use and future support are included alongside subscription or purchase fees. This process helps set management expectations, with rapid feedback sessions ensuring the final product is launched as desired.

6. LAUNCH & INTEGRATE

Whether built in-house or procured externally, the final application needs to be launched upon approval and integrated into existing – or planned – data architecture.

It is important to note that this feature or functionality prioritisation takes place at every level of a project; be it strategic, tactical, or operational. Take the financial monitoring dashboard example; on the face of it, this type of project would be considered tactical in nature. However, if the company's overarching business strategy necessitated a rapid reduction in costs and the reallocation of the firm's assets, the dashboard would be considered a critical feature, such that it could be re-prioritised as a major strategic project.

This has several implications. Recalling our goal dimensions:

1. **Time to Decision:** a financial dashboard would greatly reduce decision-making time due to its aggregation of high-quality, organisation-wide data;

2. **Accuracy of Decision:** financial data has to be incredibly precise and reliable; and

3. **Volume of Decision:** the dashboard must collate a significant amount of information across the organisation, which can come in various different formats that need to be standardised before visualisation at a moment's notice.

While this may sound obvious, it is vital to understand these dimensions, given their major implications when moving into the subsequent Information or Technology layer (whose only purpose is to satisfy these dimensions and meet these functions).

At this point, the business and application layers would have yielded the following:

- Clear business strategy with a list of key objectives and results expected
- Business requirements built around meeting business objectives
- List of current and proposed future applications for subsequent integration (in this case; a real-time monitoring dashboard application, built either in-house or procured externally)
- Roadmap of applications and their integration timeline
- For each application listed, a brief list of technical capabilities required, e.g., a real-time monitoring dashboard with:
  - Ability to ad-hoc query upon a data lake;
  - Event processing on financial transactions from multiple data sources; and
  - High-reliability data validation

MVPS ARE EXTREMELY USEFUL IN MANAGING SENIOR MANAGEMENT EXPECTATIONS, GIVEN THEY CAN SEE AN APPLICATION BEING BUILT PIECE BY PIECE

SECTION 5
DATA STRATEGY – BAIT (INFORMATION)

The information layer involves business analysts, domain owners, and data specialists collaborating to discover the location, quality, stakeholders, and interoperability of data created and managed in each department, project, and operation; this is called data discovery.

The final output of the information layer should yield a system that contains the following:

1. Data models and metadata rules;
2. Data catalogue;
3. Data platform design; and
4. Organisation-wide data training for both man and machine.

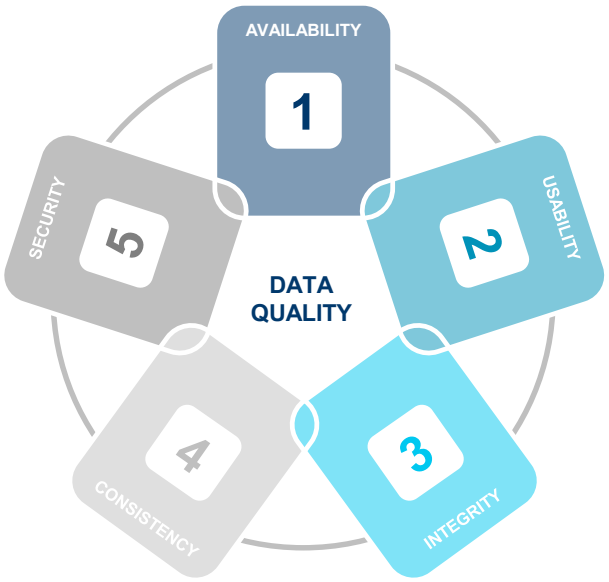
All four components help to ensure that the company generates high quality, trusted data; in other words, data that is validated and can be readily utilised by any end-user or machine to generate business-ready insights. These are encapsulated in an organisation's data maturity, which will be explored later in this section.

INDICATORS OF DATA QUALITY

An overarching metric that serves as a benchmark for how advanced a company is in its data journey is data quality.

To measure the quality of data a company generates, we can use the five pillars of data quality (see Figure 18). This would be monitored continuously throughout the project, with data stewards accountable for data quality within their respective domains.

FIGURE 18: FIVE PILLARS OF DATA QUALITY



1 **AVAILABILITY**
Extent of data accessibility in terms of timeliness and reliability

2 **USABILITY**
Data readiness in terms of presentation and comprehensiveness in offering insights

3 **INTEGRITY**
Accuracy of data with no incorrect figures to provide precise representations

4 **CONSISTENCY**
Content of underlying datasets are the same across multiple connected systems

5 **SECURITY**
Data is protected from unauthorised access and corruption

Source: PALO IT, Quinlan & Associates

DATA AVAILABILITY

Availability refers to the timeliness and reliability of the access to and use of data. For example, the availability of an order management system and its data will be mission-critical for any business. The time it takes to fetch and process data from this system would be a measurement for data availability.

DATA USABILITY

Data usability is the ability of data to enable users to derive useful information from it. For example, low-resolution photographs of invoices will be less usable for accounting clerks than invoice details collected into CSV files. Requests for additional data or further manipulation of data points can serve as a measurement of data usability within an organisation.

DATA INTEGRITY

Data integrity depends on the uniformity of shape and accuracy of data when stored under a specific domain. For example, is a dataset's schema or distribution precise and in its respective container, with no incorrect figures or misrepresentations? The number of changes to a dataset's container or underlying data points without input from a specialist can be a measurement of integrity.

DATA CONSISTENCY

Data consistency looks at whether connected data matches across different systems. For example, data generated and managed within banking transaction processes must represent identical information in internal audit systems and on automatic teller machine displays. The number of times data needs to be duplicated into a connected system is a measurement of consistency.

DATA SECURITY

Data security is the ability to protect data assets from unauthorised access and preventing data corruption throughout the entire data lifecycle. The number of security breaches and failed attacks are measurements of data security.

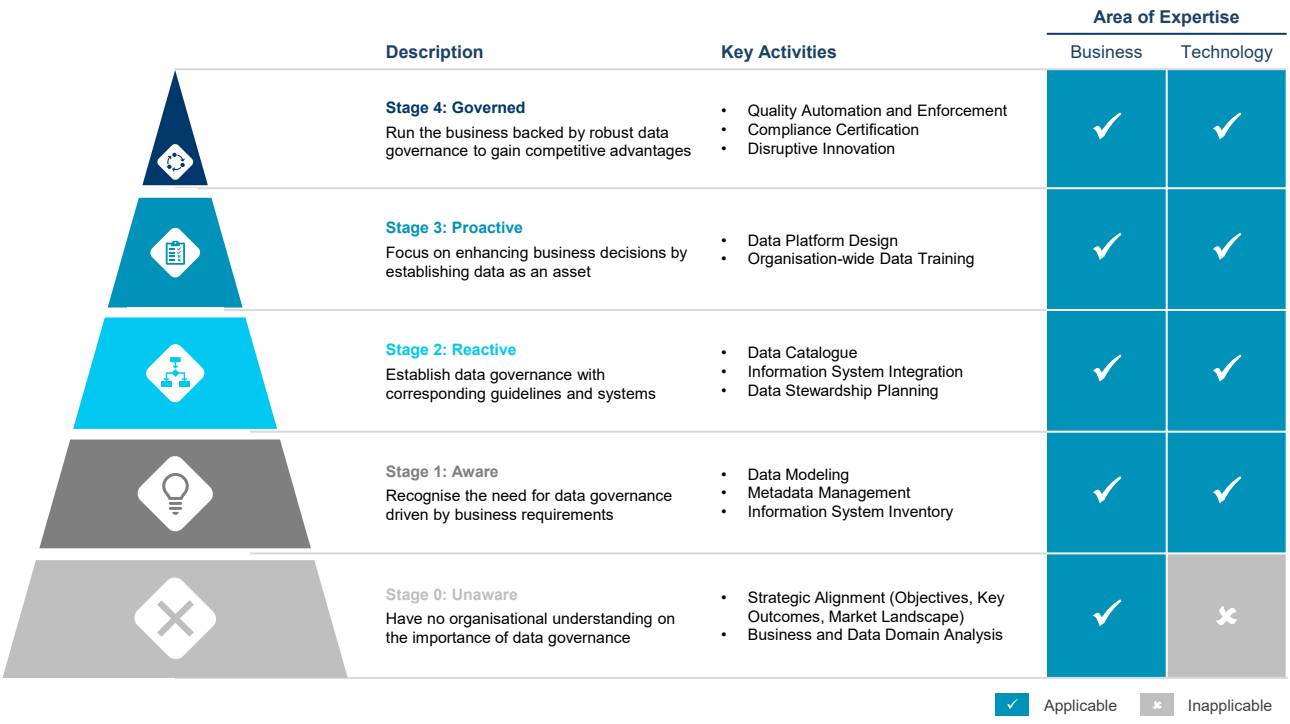
These five pillars serve as measuring sticks for the quality of data generated across an organisation. This is important for business users to note as it helps showcase progress to management in an otherwise abstract journey towards data maturity.

DATA ORGANISATION MATURITY

Discovering and mapping every system and linkage within an organisation is vital for the technical integration of applications and ensuring that data is well placed, tagged, and tracked for usage. The process of discovery increases in difficulty if the company has not considered data models or systems for communication, maps, or governance policies from the start of operations.

To assess the ability of an organisation to fully discover – or rediscover – its data, we have developed an organisational data maturity model, identifying the key activities and areas of expertise required for each stage (see Figure 19)

FIGURE 19: ORGANISATIONAL DATA MATURITY STAGES



Source: Gartner, PALO IT, Quinlan & Associates

We believe it is in the best interests of all organisations, especially mid-to-large-sized corporates, to be 'Governed' – in short, being capable of discovering all the data situated within their systems, generating and processing it under unified guidelines, and consuming it with trust (and ease) to inform business decisions.

A well-governed data organisation is built upon models, systems, and policies, and tempered with robust integration to ensure the business produces trusted data with minimal maintenance costs. If done correctly, every layer will contribute towards pulling dark data from the shadows, ensuring that the company maps every inch of their data architecture to find any and all data that can be leveraged for its operations, cost optimisation, and future revenue opportunities.

STAGE 0: UNAWARE

There are many examples of enterprises in stage 0 of their data maturity, often relying upon legacy mainframe systems and operations that have existed for multiple decades.

Companies in this stage find it difficult to utilise their data assets for two main reasons: (1) there is no reliable and unified record of available data inside their systems; and (2) it takes an unreasonable amount of time to clean and prepare the data due to inconsistency in data structures. This pushes analysts to brute-force the discovery and curation process.

Consequently, these businesses must rely on legacy reporting systems that are too inflexible (and unreliable) for modern business needs. While this sounds like it is limited to large incumbents, the truth is, most SMEs still rely on simple applications like Microsoft Excel to keep track of major business functions. These can be considered 'legacy' companies as well.

The first step towards any company becoming data governed is understanding the risks of remaining as a legacy organisation in their industry. As outlined in the earlier Business layer, if a firm's ambitions involve matching or outpacing its competitors, data needs to be at the centre of its business strategy. In the case of the financial monitoring dashboard, companies that can track and detect trends in their financial health across the entire organisation are best equipped to develop granular tactics in their day-to-day operations and allocate resources more efficiently. It would only be logical for that company's CFO to understand their firm's holistic financial performance at a glance.

To move out from stage 0, businesses must first be willing to commit to transformation / optimisation. Once again, we emphasise that this process is mainly about defining the strategic value of using data, mapping high-level objectives, key outcomes, and operations with data domains, to ensure success in subsequent stages of the maturity model.

STAGE 1: AWARE

To become data-aware, companies need to focus on making an inventory of information systems, build linkages between existing and future data structures, and create a unified language across the organisation on how to interpret data domains, with data models as a support tool that evolve with the business over time.

DATA MODELING FROM BUSINESS USE CASES

Data models are visual representations of an organisation's information systems, showcasing connections between data points and structures. These visual aids help identify all the possible entities (i.e. concepts, events, or sources) within a company that could potentially produce or hold information. Additionally, it helps to identify the key attributes of each respective entity that is used in system communication and a common set of 'formats and structures' for data communication. This is the 'unified language', so to speak.

Once a data model is created, it can be directly applied to the creation of computational algorithms or mathematical equations; a language to bridge the gap between business operations and information systems.

Data models are built around business use cases. Rules and requirements of a business use case are defined upfront in the business and application layers. These requirements are baked into the design of a new data model or adopted into an existing one. Without understanding the scope of a business need, a company cannot create a data model. If a data model represented a book written in a specific language, it would be akin to writing a book with no narrative or plot. While this is possible, it is pointless as it does not convey information in a clear way.

Applying our financial monitoring dashboard's data model to the book analogy; a book on 'procurement' would include all the financial data properties of a procurement system, in no particular order (e.g. Procurement Invoice Number, Currency, Amount of Currency, Date of Procurement, etc.). This book would contain all of the attributes – and information necessary – to tell a story around procurement trends at a specific point in time.

METADATA MANAGEMENT

Metadata is information that describes or summarises the context around an underlying dataset. If data models are dynamic books, metadata are the indicators on a book like 'Author', 'Publishing Date' or 'Blurb', that helps one to find the specific book they're looking for.

Metadata management is a crucial set of storage rules that tags data in universal ways, laying the groundwork to build a system that equips users with the ability to efficiently index and search for data assets within an organisation. Without serial numbers or categories for books, it is difficult – or even impossible – to cluster similar datasets together or locate them effectively. For the financial monitoring dashboard case, crucial metadata relates to financial transactions, which would include information such as: (1) data storage location; (2) type of transaction; (3) type of accounting; and (4) transaction owner.

Without oversimplifying the process too much, business operations garnered from previous stages of BAIT would be communicated to data architects. They would collaborate with business analysts to build metamodels that draw out high-level concepts such as domains,

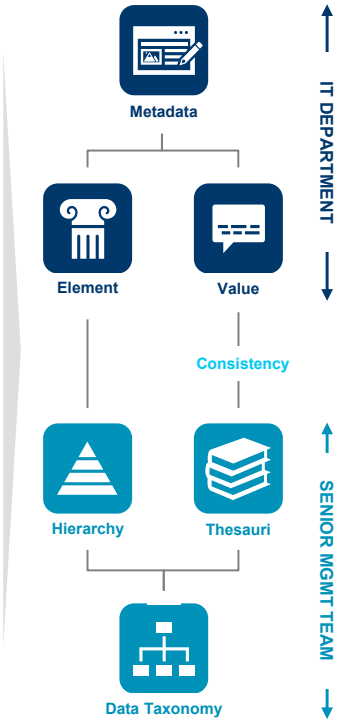
entities, and key indexes to show the relationships / interactions between business operations. For example, a simple process of product purchasing involves entities such as 'product' and 'supplier' and generates indexes such as 'unit price' and 'transactions'. This is summarised under the 'procurement domain', involving both the 'product & supplier domains', indicating that this piece of data has been generated from two primary domains of interest.

This process makes it easier for data specialists to discover, clean, and take inventory of an entire organisation's data operations from a single map. Again, this is a critical process, as not building metamodels can generate pockets of dark data which, as highlighted earlier, are highly undesirable to a business.

So, what exactly does metadata look like? Metadata should consist of elements (i.e. the list of criteria tracked and recorded by the company) and values (i.e. the text-values assigned to each element). One of the most popular schemas for metadata is the Dublin Core Metadata Element Set (see Figure 20), provided and managed by an open community called the Dublin Core Metadata Initiative ("DCMI"). This schema has been widely adopted by corporates, as standardisation enables information interoperability across organisations.

FIGURE 20: DUBLIN CORE METADATA ELEMENTS & DATA TAXONOMY

	Elements	Description	Value
DESCRIPTIVE	Title	A name given to the resource.	Data Report
	Subject	The topic of the resource.	Corporate Data Strategy
	Description	An account of the resource.	Very long report
	Language	A language of the resource.	English
	Relation	A related resource.	Strategy Consulting
	Identifier	An unambiguous reference to the resource within a given context.	Quinlan
ADMINISTRATIVE	Creator	An entity primarily responsible for making the resource.	Quinlan & Associates
	Contributor	An entity responsible for making contributions to the resource.	Quinlan & Associates
	Publisher	An entity responsible for making the resource available.	Quinlan & Associates
	Rights	Information about rights held in and over the resource.	Quinlan & Associates
	Date	A point of time associated with an event in the lifecycle of the resource.	31 Dec 2020
	Coverage	The spatial or temporal topic / the spatial applicability of the resources	End-to-End Lifecycle
	Source	A related resource from which the described resource is derived.	Diverse sources
	Format	The file format, physical medium, or dimensions of the resource.	PDF
	Type	The nature or genre of the resource.	Thought Leadership



Source: DCMI, Quinlan & Associates analysis

In the Dublin Core Metadata Element Set, the elements are categorised into descriptive and administrative elements, with descriptive elements describing the content of the file and administrative elements tracking data exhaust of the content.

SUMMARY

An organisation equipped with a firm grasp of data models and metadata management would be able to create a complete inventory of their company's information systems. While simple in theory, reality shows that it takes a significant amount of time to take stock of a company's existing data operations, especially as they begin to model complex business problems and processes. It is here that firms can better identify high-level bottlenecks, as well as any data that the business is highly dependent on to improve its ongoing operations and identify future opportunities.

While lengthy, this process is well worth the time, given it allows companies to quantitatively label the costs, revenues, and risks associated with each and every data point, providing them with the ability to better react to business events that data pushes into the fore.

STAGE 2: REACTIVE

After a business takes stock of its existing systems, data models, and metadata to create a basic set of rules to identify information, they will need to consolidate it into a larger, usable system for everyone.

In the reactive stage, data specialists collaborate with business domains to combine all the relevant information systems together under each respective domain, train and allocate data stewards, and create the beginnings of a cross-organisational governance system.

If done correctly, the reactive stage would enable business domains and departments to generate enriched reports that collate data from this unified system, allowing executives to analyse the performance of their operational decisions. A simple dashboard that displays product revenue and past R&D costs per product is an example of cross-domain reporting, allowing a company to gauge which R&D investments have borne the most fruit for the company.

To create the abovementioned reporting process, two main components need to be addressed: (1) building a data catalogue; and (2) creating a comprehensive data stewardship plan.

BUILDING A DATA CATALOGUE

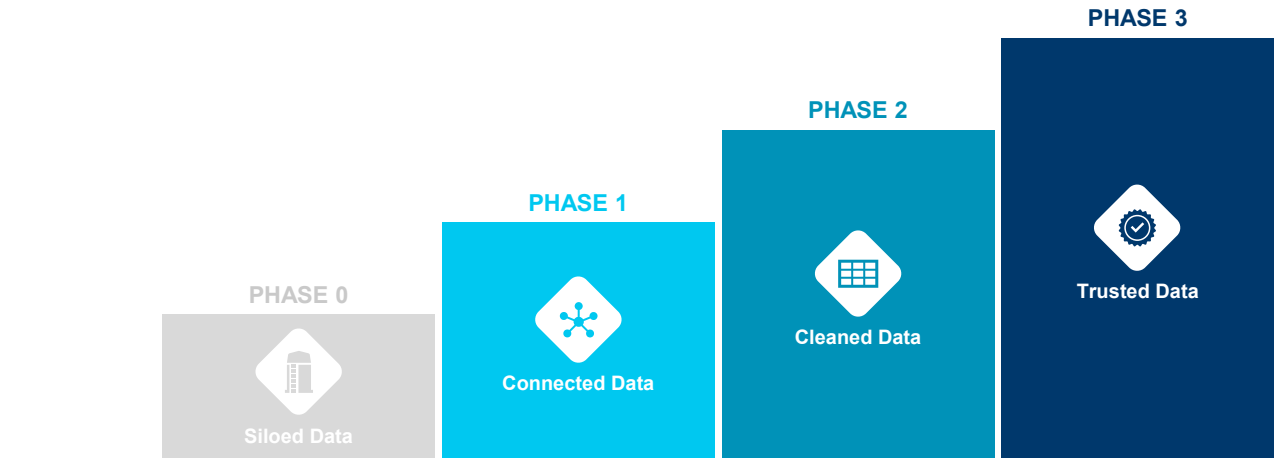
A data catalogue is a system that uses metadata to create an informative and searchable inventory of all data assets within an organisation. Similar to bookshelf labels in a library, it allows users throughout the business to filter and search for enterprise data assets without IT specialist intervention.

Building this catalogue requires both business and technology domain knowledge. This is because it is difficult for data specialists to fully grasp business terminology; on the other hand, a financial analyst would have a sound understanding of financial metrics since they live and breathe the topic on a daily basis.

The product would be an optimised catalogue and interface that is tailored to accounting / financial models and terminologies, such that finance users can generate key performance indices without having to depend on data specialists to query the relevant data. Additionally, it allows business end users to generate, test, and implement new metrics that are relevant to them as time passes, which continually adds value to the data in the process.

Building a fully functional data catalogue is accomplished in multiple phases, showcasing how data is connected, cleaned, and validated to develop ‘trusted’ information for business use (see Figure 21).

FIGURE 21: DATA CATALOGUE PHASES



	PHASE 0	PHASE 1	PHASE 2	PHASE 3
Relevant Audience	1. Limited Individuals within the Process and Teams	1. IT Specialists 2. Data Architects	1. Data Analysts 2. Data Scientists / Engineers	1. Business Users 2. External Data Customers
Description	Raw data of each department and team are locked into its own silo, which leads to a lack of integration and delays in data discovery and access	Raw data from various systems are connected into the first version of a data catalogue, which allows developers to develop data sharing structures	Raw and incomplete data is cleaned, with major outliers removed but output has not been well-categorised to accommodate precise searches	Validated and labeled clean data to enable search and utilisation of relevant data which is understood and interpreted similarly across all domains of the business

Source: Microsoft, PALO IT, Quinlan & Associates

1. CONNECTING DATA

After taking an inventory of existing systems, IT and data specialists group information systems by departments according to use cases, following which they connect relevant systems according to the data model designs.

This process normally includes: (1) building or expanding APIs to expose windows of communication to external systems; (2) connecting them through communication brokers, such as message and event buses; and (3) building an API catalogue that lists out different instructions, alongside expected responses from queries, to allow developers to build business applications that interact directly with the underlying information systems without having to connect each one individually. This integration saves a tremendous amount of time and effort that would otherwise need to be expended each time a new data system or application needs to be connected.
Furthermore, data scientists and engineers can start connecting their respective data workbench software to the information systems to fetch, clean, experiment, and publish various insights once the integration is complete.

This concludes the shift in moving an organisation from phase 0 to phase 1, streamlining the flow of information instead of having to manually collect data from disparate sources.

2. CLEANING DATA

Clean data generally refers to datasets that are free from incomplete, incorrect, or inaccurate data. Major outliers are removed at this stage to lay the foundations for subsequent usage by business users. This is done via data workbenches by data specialists, who fetch

sample datasets from each source to identify unnecessary or inaccurate data points through data cleansing algorithms.
Data engineers transform the algorithms into automated scripts that streamline the cleansing process of each data source. Cleansed data is then registered into the data catalogue for general consumption by data users across the company. This does not mean, however, that it is ready for everyone to use; clean data must still be validated and modified to fit specific business users in the organisation before it can be considered 'trusted'. This concludes the movement from phase 1 to 2.

3. VALIDATING DATA AND BUILDING TRUST

Trusted data differs from clean data depending on its intended users. While the audience for cleaned data is strictly data specialists, trusted data is generated for direct consumption by business users and stakeholders, which requires the data to be validated thoroughly and transparently by the users themselves. For example, a clean dataset of profit and loss figures from a specific product department could be 'trusted' by the product team, but product-specific jargon / formats could prevent finance or customer teams from using it freely, creating the need to validate the data directly with product owners before use.

Data validation benchmarks five main metrics to ensure trust in data from both internal (main domain) and external (anyone outside the main domain) users:

1. **Thorough**: is the data clean, complete, and consistent across the entire system?
2. **Transparent:** is the data accessible and understandable?
3. **Timely:** is the data up to date and readily available to people who need it?
4. **Traceable:** does the data able reveal its origin and how it has been used?
5. **Tested:** has the data been rated and certified by other users?

These metrics can be built into the data catalogue and enriched by expanding the metadata over time. Trust is built as more data use cases are validated by business users.

So, what would this look like? In reality, it could be as simple as a feedback rating popup, where a user can rate the data on its validity or trustworthiness. In the case of our financial monitoring dashboard, a customer domain user who is fetching complaints from the product domain could rate the validity based on their use. As the volume of validation metrics increases, the product dataset in question eventually reaches a point where it becomes statistically reliable. From there, it can be applied to more categories of data generated and consumed by the organisation via methods such as advanced analytics or other processing engines.

More importantly, this validation process eventually creates a data catalogue that becomes the single source of truth in the organisation. This simplifies privacy and compliance requirements, removing risks involved in cross-department data sharing while promoting data transparency and a strong culture around data quality within the organisation.

A data catalogue's aim is to set the foundations for free and trusted data sharing across all domains. This free flow of information ultimately leads to the creation of new insights in novel places, spurring innovation from stakeholders and data consumers. This concludes the final phase of a data catalogue's maturity.

DATA STEWARDSHIP PLANNING

The second step in creating a 'reactive' data organisation is to build the beginnings of a comprehensive data governance model.

We've touched on the topic of data stewards in data domains in earlier sections, but at its core is the creation of accountable, data-savvy individuals who can influence their respective domains from the ground up. To support these newly fledged data stewards is a committee that sponsors the governance process and organises an operating model to propagate policies, monitor progress, respond to issues, and enforce governance standards. This committee forms part of the company's change management strategy and is responsible for setting up systems and processes to consciously record data, information interpretations, decisions, and outcomes generated throughout the organisational data value chain. This would support the transformation process necessary to build and maintain new guidelines and culture.

Appointed domain owners will identify key risk indicators within and between their domains, ensuring data quality and interoperability of domain information. Data stewards bring operational policies, technologies, and models together to maximise data quality within the organisation. Their main responsibilities include: (1) cross-domain data sharing; (2) monitoring compliance and security awareness; (3) communication and data training of their respective domain members; (4) managing the data value chain; and (5) data end-user support. All of these are crucial for increasing the consistency, trust, and quality of data consumed, which in turn affects the outputs generated throughout a data organisation's journey towards maturity.

This initial committee and batch of data stewards will continuously develop measurements to monitor various risks and performance indicators emerging from the tighter integration of information systems. They would also note learnings from this initial foray into data integration for subsequent autonomous information quality improvements in later stages.

By the end of stage 2, companies would be equipped with basic cross-domain information sharing, using the data stewards as nodes of information collection and dissemination. With a consistent and well-maintained data catalogue in place, even employees with a basic level of data querying can help themselves to information via the instructions laid out in the catalogue. This marks the beginnings of a mature data organisation, with stewards being able to react quickly to domain-specific issues with the right analytics and tools to convince senior management or make informed business decisions on their own.

STAGE 3: PROACTIVE

The third stage of the data maturity model is when an organisation is able to proactively make business decisions using data at scale.

Stages 0 to 2 of the data organisation maturity model are focused on getting a company to a basic level of functionality in terms of data discovery, searchability, and integration. If the previous stages brought library-like functionality to a company, where certain individuals are domain experts (akin to librarians), stage 3 would modernise the library and make it scalable, with everyone acting as a personal librarian or curator of data within the organisation. This scale is accomplished in two manners: (1) data platform design; and (2) organisation-wide data training.

DATA PLATFORM DESIGN

A data platform is a solution designed to ingest, process, analyse, and present data generated by systems, processes, and infrastructures of a digital organisation. Simply put, it provides end-to-end data management, instead of point-to-point solutions, which are impossible to scale after a certain number of connections.

A platform is theoretically future proof, as it allows new applications to 'plug and play' directly into it without having to change many other critical systems. Additionally, it pools all data services into a single hub that can be monitored and governed as one service entity instead of having multiple departmental systems to communicate with. This reduces integration complexity and streamlines the flow of data in an organisation.

For an end-user like a CFO, a platform design enables real-time monitoring abilities that are required for a financial dashboard. The output of Data Platform Designs are functional and non-functional requirements of a platform, which are used for the actual architectural designs in the subsequent Technology layer.

Given the upside, you may rightfully be wondering: "why isn't every organisation built as a platform from the start?" The answer is simple: platforms require major financial investments that don't reap significant benefits unless there are a considerable amount of data users or applications that need connections. In a company's early stages, it is far more resource-efficient from both a cost and time perspective to subscribe to a single cloud provider or server rack and develop from there. However, at some point, every medium or large company should consider the leap to a platform if they are to make data-driven decisions effectively in the future.

ORGANISATION-WIDE DATA TRAINING

With every application and information system available for use, the power now rests with employees to use them effectively.

Training in the prior stages has been isolated to a few individual stewards. The next step is training every employee to manage and maintain good data habits for their decisions. This is a herculean task that involves long-term operational and cultural commitment. The good news is, after reaching a critical mass of employees who understand how to use data to drive their decisions, habits tend to propagate to new employees as part of company culture. We will explore how to take the first step in this long-term change in depth in Section 9, Change Management Strategies.

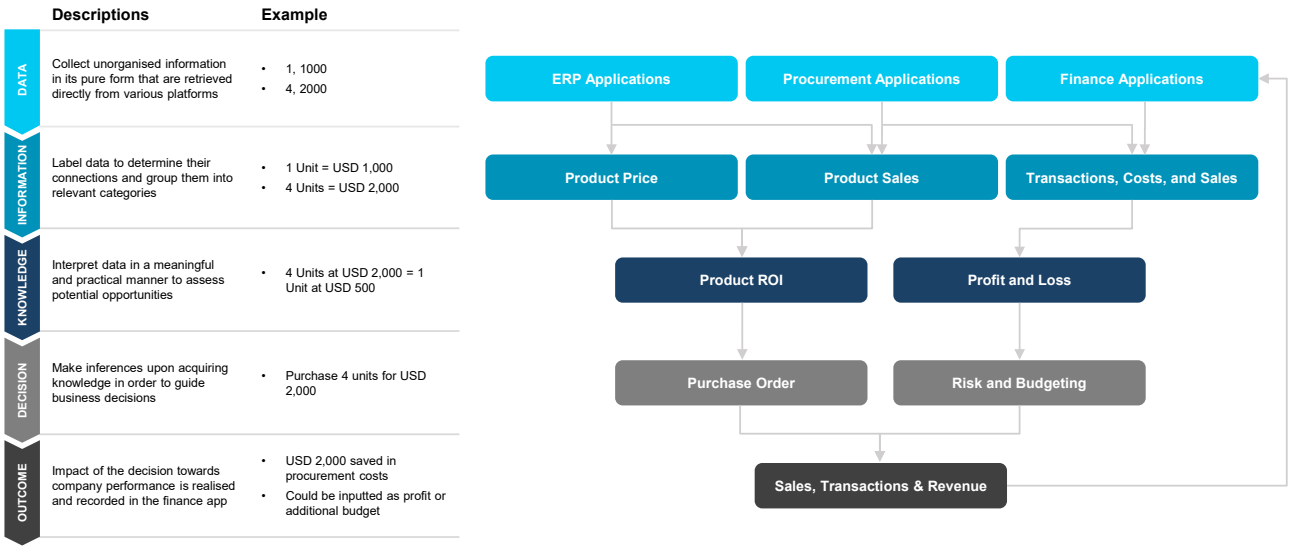
Training does not simply entail 'how' to use a particular application to generate insight. The

desired outcome is to have employees who are well versed in querying data and deriving value from it autonomously. An analogy would be teaching critical thinking skills to a student population; it would be the equivalent of showing employees how to find these data 'books' in the 'library', and how to cite and use the information within them to generate new insights for their own business units.

Ultimately, an employee should be able to analyse their outcomes from a decision they've made based on data they personally collected and processed to establish data as an asset. After all, data does not inherently generate revenue itself unless the organisation and its members collaborate to find ways to leverage customer, product, and financial data to actualise revenue (see Figure 22).

While it is no doubt simpler to centralise insights to a few analysts, this training process teaches all employees to trust the data, instead of just a handful of analysts. This seemingly minor detail is the difference between trust based on humans and trust based on data; the latter would empower every employee to contribute directly to the direction of a company in their everyday work and yield far more productivity and creative solutions in the long run.

FIGURE 22: DATA TO OUTCOME FLOW



Source: PALO IT, Quinlan & Associates

Stage 3 would create a data library at scale, streamlining data operations into a single hub, with the ability to interchange applications and systems at a much lower cost (after the initial investment). It would also give rise to a new generation of tech-savvy workers, who can autonomously pull data from systems to drive their decisions and manage their outcomes on a daily basis.

STAGE 4: GOVERNED

A governed data organisation is the final frontier for any aspiring data-driven company. At this stage of evolution, data quality is enforced automatically and evolves with business use cases over time. This is also the only stage where advanced analytics and artificial intelligence truly reaps the most benefits from a company's data investments.

As highlighted at the start of the report, we made a case around how much wastage there is in the data project space; this is primarily due to the businesses being too immature in their data journey to fully leverage their AI or analytical investments. Foregoing the earlier, less attractive, and time-intensive stages of the data organisation maturity model only limits machine learning applications from reaching their full potential. If a business cannot showcase a holistic data catalogue with a competently trained employee base, the chances of it housing high quality data is low. This is precisely what advanced analytics requires to produce high quality insights that inform business decisions. Otherwise, it's simply garbage in, garbage out.

QUALITY AUTOMATION & ENFORCEMENT

The final stage of a data organisation's maturity is when a business is able to automate its data stewardship and curation processes with minimal human input.

Prior stages have looked towards employees, rules, and cultural standards to maintain data quality across the business. After a period of operating at a 'proactive' capacity, any business would have gained a history of data catalogue changes, which include maintaining existing metadata definitions or evolving them over time. This creates patterns around which definitions are constant, informing software on overall operating parameters and allowing it to take over the responsibility of data stewardship. This would be akin to automating a librarian or curator of data.

An automated librarian could be trained to monitor data inputs, categorise them into their respective domains, and validate them with minimal manpower. The only manpower requirements would be parsing ambiguous data points through a business domain user, who would rate the usefulness or accuracy of the data points, which would then be considered by the machine. If a specific data point is receiving consistently low scores on data quality by employees, it could be discarded or disregarded from further analysis. This allows the machine to learn and evolve alongside the business over time; a form of data governance-level machine learning.

Why is this important? As a business grows, the manual process of updating a data catalogue or individual metamodels becomes cumbersome. Additionally, as many more employees are hired, the level of effort required to train and reduce bad inputs increases exponentially. To alleviate this, machines must become a part of the quality enforcement process. This ensures that data remains trusted and usable across the organisation, regardless of the size of the workforce or the company's operations.

If data quality is well enforced and automated, our CFO can rest easy knowing that the information displayed on the financial monitoring dashboard has already been validated, stringently filtered for anomalies, and will show up in the same desired format every time. This fulfils the cycle of 'trust' in data and can inform business decisions directly without the need for further checks.

COMPLIANCE CERTIFICATION

If executed well, this final stage will create the flexibility a company needs to alter the parameters of their data operations to fit regional or international data privacy standards. The most widely known would be Europe's GDPR. An example of such flexibility would be an internal machine (or automated librarian, in this case) receiving regulatory requirements around the 'right to be forgotten', after which it could automatically dispose of data points relating to a specific customer who demanded to be 'forgotten' from the company, without further input from employees.

To further the example, say a customer has asked for the 'right to restrict processing'. The automated librarian would exclude data points from being processed or shown on the financial monitoring dashboard, keeping the company GDPR compliant with minimal action from staff. An automated librarian would make up a majority of the 'governed' part of an organisation, freeing up precious manpower and resources for other value-adding tasks.

DISRUPTIVE INNOVATION

Disruptive innovation is a buzzword often thrown around in earnings calls or as clickbait articles on social media. The truth of the matter is that innovation is uncommon. What is even less common is disruptive innovation by way of data. This is shown by just the handful of companies who have mastered their data value chains to create competitive edges over a sea of their peers in the world.

Take a Google or a Netflix. The amount of effort these firms have placed into harnessing data, which could be used to inform their business strategies, are awe inspiring. While it's easy to point at their search engine algorithm or recommendation engines in an attempt to replicate it, both were only possible through consistent and careful calibration of their data strategy, business objectives, applications, and data curation. The fact is, most companies in the world are nowhere close to this level of care around their data operations, leading to considerable waste.

The good news is, at this stage, data is generally trusted by both human and machine. This allows AI and machine learning to be applied in earnest, creating new insights and business opportunities by illuminating previous blind spots that employees may have missed. Examples of this would include being able to read market trends or anomalies in an organisation from big data processing, giving a company the opportunity to respond appropriately.

Tying this back to our financial monitoring dashboard, the CFO would benefit greatly from an analytics engine in the background that can highlight spending trends or inefficient resource allocation across the company on their dashboard. This concludes the final stage of a data organisation's journey towards maturity.

SUMMARY

The information layer of a data strategy project would yield the following:

- Data models and metadata rules, which identify a common set of features and attributes for data communication
- Data catalogue, which is a searchable inventory system of all data assets within an organisation
- Data platform design, which maps all applications and systems to one another in a key capabilities list
- Organisation-wide data training for both man and machine to ensure that all systems and models work as intended to produce trusted data

These outputs would provide enough information to complete the final layer of a data strategy project, selecting the appropriate technology and investments needed to drive a company's business strategies.

THE TRUTH OF THE MATTER IS INNOVATION IS UNCOMMON. WHAT IS EVEN LESS COMMON IS DISRUPTIVE INNOVATION BY WAY OF DATA

SECTION 6
DATA STRATEGY – BAIT (TECHNOLOGY)

The technology layer involves IT architects, engineers, and analysts collaborating together to design and build an optimised tech infrastructure to support confirmed application and information requirements defined in the previous layers.

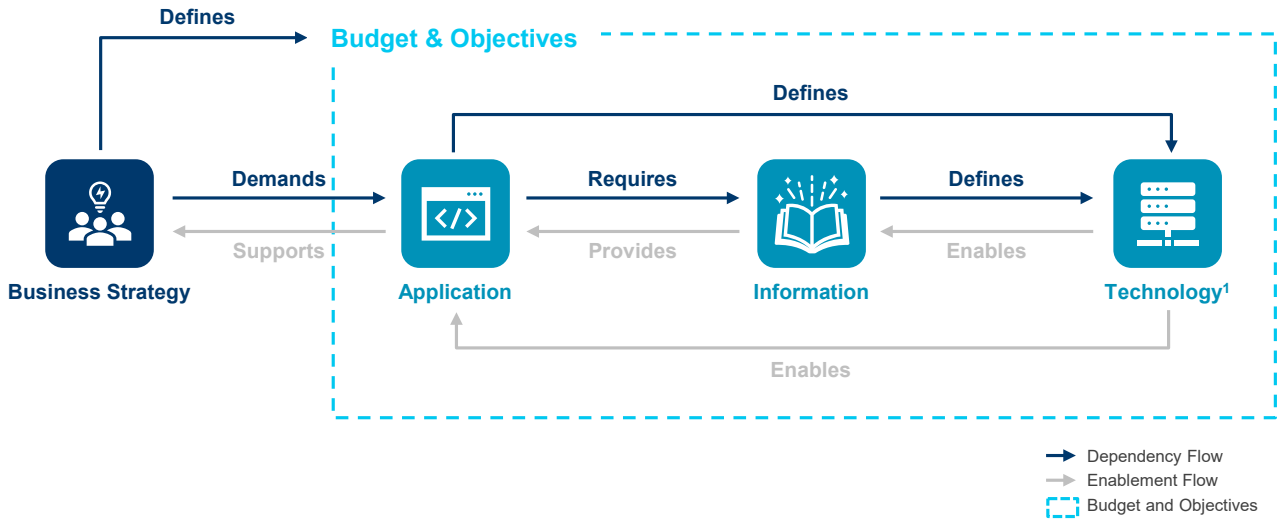
Here, engineers would dissect business objectives and capabilities required under the Business layer to define the size, complexity, and selection of technologies. This would be in line with budgetary and operational limitations set by the preceding layers. As we have stressed earlier, technology is merely an enabler that delivers functionalities with integrity, availability, reliability, and security in mind. It should not be the opposite, where a hyped technology is shoehorned into a business use case. This is a recipe for a failed data project with minimal upside.

The process of technology selection begins by firstly collating Application and Information layer requirements, which will have stipulated basic applications, data models, and platform designs. This indicates to an IT architect that matching technologies and abilities, such as a data lake, event processing, and data validation modeling, are necessary to meet such requirements.

Secondly, architects will take supporting information assets such as the data catalogue and governance processes and integrate it into the creation of a technical architecture.

Finally, architects and engineers will forecast maintenance costs based on the volume of data and frequency of access to determine if the technology solution is viable for the business's budget, coming full circle back to the Business layer (see Figure 23).

FIGURE 23: BAIT HIERARCHY



Note: Includes hardware and software
Source: PALO IT, Quinlan & Associates

DATA PLATFORM ARCHITECTURE

As highlighted in the information layer, a data platform is a conceptual framework used to describe a set of tightly integrated information technologies that collectively delivers an organisation's end-to-end data needs.

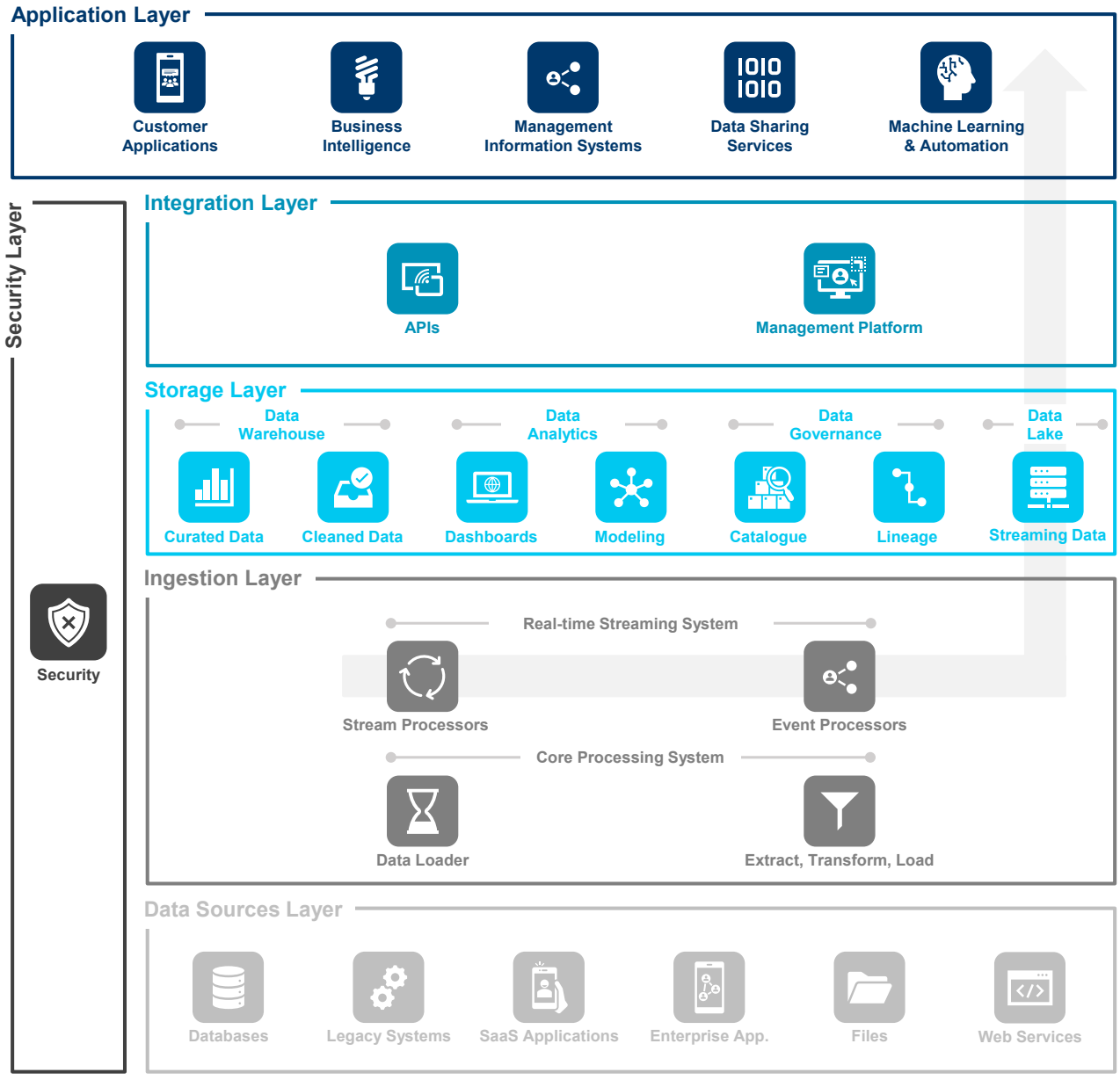
The need for such a centrally integrated platform for data becomes crucial when data consumption starts growing at an exponential speed. An architecture diagram would be the final blueprint that reflects the entire structure of an organisation's data requirements – hence the name 'architecture diagram', akin to an engineering blueprint for a 20-storey building.

As a simple example we can revert to our financial monitoring dashboard; if the organisation is spread around the globe, an architecture diagram would include the central integration of every different system in every location for the dashboard use case. Central integration would alleviate the technical complexity and cost associated with future changes or new applications needed for additional features. Future development and use cases would only require an interface between the data platform through various APIs, without any additional orchestration of point-to-point information system integration. An example of a holistic architecture diagram for the financial monitoring dashboard is shown below (see Figure 24).

To supply the best quality information to the dashboard application at the right time, the data platform includes data lifecycle functions such as acquisition, storage, and processing through various technologies. This is to curate and package the correct data to be served to the application layer at the final user interface.

A DATA PLATFORM IS A CONCEPTUAL FRAMEWORK USED TO DESCRIBE A SET OF TIGHTLY INTEGRATED INFORMATION TECHNOLOGIES THAT COLLECTIVELY DELIVERS AN ORGANISATION'S END-TO-END DATA NEEDS

FIGURE 24: DATA PLATFORM ARCHITECTURE SAMPLE



Source: PALO IT, Quinlan & Associates

A platform's design elements, such as the placement of systems, size, and cost, vary greatly depending on the business, application, and information requirements. For example, if the objective is to optimise decision-making in a product procurement process, a simple batch processing system in a data warehouse of historical transactions, inventories, and market price information would fulfil the necessary needs. On the other hand, a real-time

dashboard requires flexible and event-driven architecture, which should be fulfilled by a data lake. Simply put, procurement of the final components of an architecture design rests solely on the requirements stipulated in the earlier layers of BAIT.

PROCUREMENT DECISIONS

The selection of technology is based entirely on the budget that has been set by the business, regardless of whether a decision has been made to build or buy a technology.

Bluntly speaking, technology choices are constrained by previous business strategy decisions and by the final end-user (i.e. the CFO in our dashboard case). This makes the procurement process incredibly simple: (1) fulfil the technical functions required; (2) under the company's given budget. For most technical procurement decisions, the budget, timeline, and business objectives are already fixed. Burndown and fulfilment of the budget and timelines are closely monitored, evaluated, and re-negotiated through project owners and committees for change. Technical requirements are discovered and incorporated into the design process in parallel under strict acceptance criteria to be set and monitored by IT leads. This dance between meeting both requirements and budget are carefully choreographed between project owners and technical staff.

Before stepping into procurement decisions, IT leads should have a firm grasp of both functional and non-functional requirements needed in the end-state system.

FUNCTIONAL REQUIREMENTS

Functional requirements refer to the set of service descriptions that must be included in the procured software. Functional requirements specify what a given system must do and, just as importantly, not do. This is defined in terms of inputs into the system and expected outputs to be produced.

In our dashboard example, if a new invoice has been received and the transaction is validated, the data system must process the transaction and update the dashboard accordingly within a certain timeframe. On the other hand, if an invoice is invalid and the transaction is cancelled, the system should not update the dashboard as the data input is invalid, acting as a data invalidation process.

While most functional requirements are product features, there can be certain functional requirements that are not directly derived from user requirements. In the case of a data platform, there are functional requirements from the information layer, such as the data invalidation process in the abovementioned example, which ensures data quality. This is an important non-user functional requirement that engineers should be aware of. This reinforces the importance of understanding and accessing the needs of previous BAIT layers to design the ideal system that fulfils the exact needs of a user.

NON-FUNCTIONAL REQUIREMENTS

While functional requirements define what the system must do and must not do, non-functional requirements define how the system should execute it. This is important, given that a functional requirement can be fulfilled using a different combination of technologies, while only a few would meet the necessary non-functional requirements.

The following are common non-functional requirement criteria:

- **Resilience**: system resilience refers to the system's ability to isolate and recover from unexpected scenarios to ensure service reliability. For example, the system should be able to reject an invalid invoice early on to prevent data corruption and ultimately deliver accurate information to an end-user dashboard.

- **Security**: system security refers to the system's ability to detect, defend, and respond to adversary activities performed by both internal and external actors.

- **Scalability**: system scalability refers to the system's ability to adapt to different levels of loads arising from a variety of factors such as usage spikes, mandatory system updates, and cyberattacks, to ensure service stability.

- **Cost**: cost as a non-functional requirement indicates the ongoing cost of operations, including regular maintenance, backups, end-user support, and disaster recovery, to ensure acceptable levels of business continuity.

Non-functional requirements are crucial to the reliability and stability of services that the overall system provides and must be considered in-depth.

Functional and non-functional requirements are equally important in the procurement / design process. Once both have been adequately mapped out by an architect, the sketching of the architecture diagram can begin.

THE SELECTION OF TECHNOLOGY IS BASED ENTIRELY ON THE BUDGET THAT HAS BEEN SET BY THE BUSINESS, REGARDLESS OF WHETHER A DECISION HAS BEEN MADE TO BUILD OR BUY A TECHNOLOGY

1. DATA SOURCES LAYER

Data always has to be sourced from somewhere. In this case, data sources refer to the various sources of data existing inside an organisation, such as databases, APIs of SaaS applications, or excel spreadsheets stored in servers.

When designing the data platform, identifying each data source and determining the means of connection (and fetching mechanism) is crucial if a company wishes to utilise the data stored inside of them. For example, financial indicators can come from different data sources such as accounting software, inventory, and enterprise resource planning software, as well as excel spreadsheets used by financial analysts. Architects must effectively aggregate the desired raw data through technical solutions and map them under this layer.

2. INGESTION LAYER

The ingestion layer includes technologies used to process data fetched from the data sources layer to be consumed by analytics or internal business analysis applications. This creates readily available information in the background that can be stored for business users. This reduces the need to wait for long computations, particularly when dealing with large volumes of data, or where processing logic is complex. Ingestion also includes data cleaning and metadata injection to comply to the data model defined in the information layer.

There are two different types of processing commonly used in data platform systems:

2.1. BATCH PROCESSING

Batch processing is often called an extract, transform, and load (ETL) process, and refers to the processing and analysis of a dataset which has been stored for a period of time. Common examples of batch processing include monthly payroll systems.

Batch processing requires a large amount of computing resources as it processes data in bulk. However, the required computational performance can be low as it does not execute frequently, while the information produced from batch processing is not frequently accessed.

Batch processing is extremely useful when consolidating large transactional data and performing statistical analysis and inferences. Take a product development team as an example. The team is looking to plan purchase decisions based on the return on investment of all previous purchases. Here, engineers would develop a system that processes product purchase records and matches them to sales records in an ERP database, allowing the team to calculate its return on investment for each product on a monthly or bi-weekly basis. Following this, the curated data could be stored in a data warehouse for the product team to access without having to trigger the long and expensive computations used for earlier batch processing.

2.2. STREAM PROCESSING

As the name suggests, stream processing refers to the processing of data that flows through a system, resulting in analysis and reporting of events as it happens. Infrastructure built to digest steaming data is called event-driven architecture, because the data flows into a system when a certain event causes a change in either the internal or external environment.

Stream processing is commonly used for systems like ATM backends, fraud detection, or intrusion detection. Because this type of processing doesn't require the data to be

aggregated before calculation and computation occurs on a relatively small packet of data, it has far fewer computing resources to maintain. However, this comes at a trade-off; given the speed at which data needs to be processed, computational logic needs to be optimised to maintain latency requirements to within a few seconds to milliseconds.

The architecture also requires an intermediate system to digest and distribute incoming events to the relevant processing systems, which are called brokers. A broker ensures the incoming events are addressed by order of time and allocates events to idle resources. These systems also detect bottlenecks or surges of requests to trigger a given data system and scales resources accordingly. A broker is crucial in maintaining the reliability of the stream processing system.

In the case of our financial monitoring dashboard (whose main requirement is to operate in a real-time manner), financial transactions must be supported by a stream processing system, with the ability to push updated reports to the dashboard based on any new transactions being executed.

3. STORAGE LAYER

The storage layer refers to a set of technologies used to store multiple sources of processed data in a centralised location. Identical to the supply chain model used in logistics, the storage layer aggregates data from disparate sources, already cleaned and curated through the ingestion layer, and organises it to be stored for later use.

This process is crucial as a growing number of modern businesses are spreading out globally, where data fetching becomes incrementally slower due to the distance between storage points. With a competent storage layer in place, data is locally available for use, with end-users remaining immune from network latency issues.

This layer contains several key modules, such as the governance module, used to maintain a high-quality inventory of data. Embedded analytics tools allow data scientists to extract insights and directly build new processing models to supply end-users with at the application layer.

The two main storage facilities include: (1) data warehouses; and (2) data lakes. While we provided a conceptual outline of what a both these storage facilities are in our previous report, we will be exploring both contextually below.

3.1. DATA WAREHOUSE

A data warehouse is a storage hub for structured and relational data only. Warehouses store cleaned, raw data that is shipped to different data marts (i.e. smaller, purpose-built databases for a specific business application or use case) (see Figure 25).

FIGURE 25: DATA WAREHOUSE

Note: Extract, Transform, Load
Source: PALO IT, Quinlan & Associates

It is important to note that while duplication of data is generally undesirable, some level of duplication occurs in both the data storage and data sources layer. This is because fetching data directly from the sources layer increases accessibility time. Storage of processed data closer to the application layer is similar to stocking a local supermarket for users instead of expecting them to go directly to wholesaler, who is much farther away. Just like local supermarket, data marts can adapt to localised needs, which improves the usability and utilisation of data and its underlying infrastructure.

Data warehouses work best for a batch processing system that requires regular periods of processing within a fixed time interval. Examples of systems that utilise batch processing include payroll systems, aggregated sales reports, and inventory management systems. These use cases have already been confirmed in the earlier stages of the BAIT framework and outputted as a data model. Changing this at the technology stage is not recommended as it requires additional technology spend to resolve, potentially overstretching a firm's budget.

In the case of the real-time dashboard, a warehouse storage design is unsuitable for the company's needs. The CFO and finance department would be unable to fetch data and visualise it at speed, leaving only one other option for our real-time requirement: data lakes.

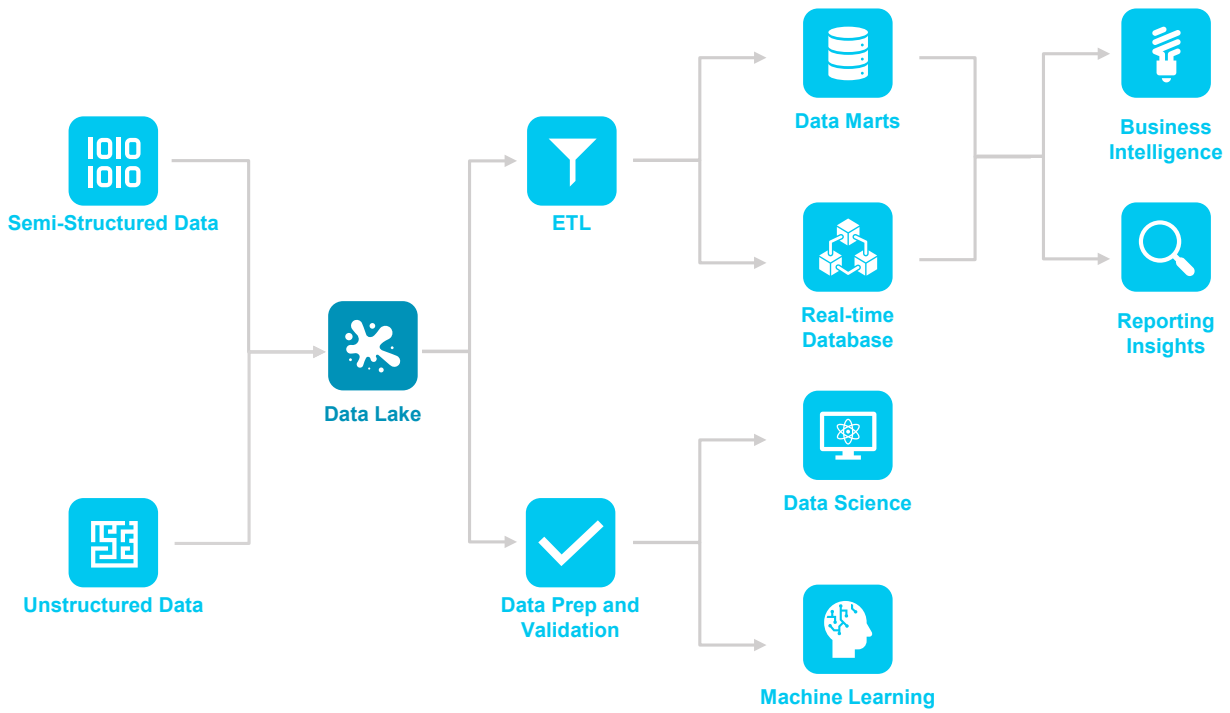
3.2. DATA LAKE

Data lakes are different from warehouses in that they are able to handle semi to unstructured data, such as documents (e.g. JSON, XML files, etc.), timestamped logs, and graph data from social networks.

Unstructured data is commonly produced in vast quantities by a variety of modern applications, especially from social media platforms, which are extremely useful in supporting organisations with their marketing and forecasting capabilities. Data lakes allow streaming data to cleaned and consumed immediately by business users without having to wait until a new batch of data is processed.

Data stored in a data lake can be delivered in various ways: through a data mart, just like a warehouse; through a separate real-time database; or queried directly from the lake by data scientists and machine learning developers. Data experts, in particular, benefit the most from this storage solution as they can leverage more up-to-date data for training and validate their projects on the fly. This greatly reduces the possibility of skewing or overfitting in advanced modelling projects, which are commonly caused by bias or outdated datasets. These features make a data lake a highly versatile storage hub for all types of data sources (see Figure 26).

FIGURE 26: DATA LAKE



Source: PALO IT, Quinlan & Associates

Aside from storage versatility, data lakes are also important in information security. Organisations typically have few data lakes, most of them acting as a central 'lake' for an entire business unit. This streamlines the security process for security specialists, who can monitor access and modifications to data that occur at this layer of an overall platform, detecting abnormal activities from a bird's eye view.

The one key downside to a data lake is its complexity, which make it unsuitable for business users to tap into without prior processing.

4. INTEGRATION LAYER

An integration layer of a data platform is the akin to a storekeeper taking requests from customers to pull their desired products from the back storage room of a shop. This layer is a key component of any platform design, allowing fresh applications to plug directly into an organisation's underlying data storage, ingestion, and sources layers without having to manipulate them for compatibility. Typically, this consists of different API gateways, which provide uniform, structured, and quality-controlled data services to business users.

An API gateway is a single-entry point for defined backend APIs and microservices. A gateway should contain service and API catalogues that explain expected inputs and outputs for each resource available for a user to access. The main audience of this catalogue are business and data analysts of each domain or department looking to create a new analytics model, or engineers who are developing new business applications.

Gateways can be created centrally or separately based on application protocols. Normally, representational state transfer (REST) APIs are adequate to fulfil a majority of the functional requirements required by any business application. However, this can vary based on the business use case. In the case of the financial monitoring dashboard, a real-time dashboard needs to be able to pull data and receive live updates from the data platform. As such, bi-directional communication protocols such as WebSocket or other streaming protocols are needed on top of the basic REST APIs.

Other possible gateways include remote procedure call (RPC) gateways, which help in use cases where asynchronous communication is needed. Asynchronous communication occurs when modern systems need to communicate with older legacy systems, a common occurrence at many large organisations.

Finally, there could be a single management platform that manages these gateways together. These factors should all be considered when designing the integration layer. This is especially the case for our financial monitoring dashboard, which would need to collect data across an entire organisation, touching various business domains (and, by extension, their own data sources, ingestion, storage, and application layers).

5. APPLICATION LAYER

At the top of the data platform architecture is the application layer. This layer is composed of all business applications that consume backend data services provided by the platform and interface directly with end users. By including applications in the architectural design process, technical requirements are easily visualised and mapped to the corresponding data components within the platform itself.

Common business applications can include everything from direct customer-facing applications, business intelligence, management information systems, third-party data sharing applications, and internal machine learning & automation interfaces. Our financial monitoring dashboard would consist of most of the abovementioned common business applications.

6. SECURITY LAYER

The security layer is a vertical layer composed of technologies that support non-functional requirements relating to the protection of digital assets and information from expected or unexpected malicious events.

Keen eyes would notice that the security layer does not encompass the application layer. This is because attempting to control and monitor the security postures of every application connected to the platform is extremely resource intensive from both a cost and time perspective. Instead, the integration layer acts as a "buffer zone" between external applications and backend platform services. If API and security policies are implemented correctly, security breaches via an application vector are unlikely. A platform's security is built upon three measures, typically described in a 'CIA' triad (see Figure 27):

1. **Confidentiality:** the ability of a system to protect sensitive information from exposure to unauthorised users;

2. **Integrity:** the ability to validate and protect data from undesirable or unexpected modifications either at rest or in transit, ensuring that the authenticity and reliability of the information is guaranteed upon arrival at its destination; and

3. **Availability:** the ability for the system to protect itself so that authorised users have timely and reliable access to the information they need.

Popular cyber threats that affect availability include distributed denial of service (DDoS) attacks, which flood a system with a large volume of empty packets to block authorised users from accessing the system.

FIGURE 27: SECURITY – CIA TRIAD

CONFIDENTIALITY

The ability of a system to protect sensitive information from exposure to unauthorised users

INTEGRITY

The ability to validate and protect data from undesirable or unexpected modifications either at rest or in transit

AVAILABILITY

The system's ability to protect itself so that authorised users have timely and reliable access to data even under pressure

Source: PALO IT, Quinlan & Associates

With these three measures in mind, there are a variety of core security services that a data platform should have to create a good security posture. Basic examples of core security services include:

1. Identity and Access Management (IAM);
2. Encryption and certificates management;
3. Load balancing and failover system; and
4. Security Incident and Event Management (SIEM).

1. IDENTITY AND ACCESS MANAGEMENT

IAMs are centralised systems for user creation, authentication, authorisation, and permission control. Most companies already own similar systems for IAM, such as Microsoft's Active Directory (AD), Okta, or Google's Workspace, which manages user identities and access rights. Access rights are a basic tool used to ensure that confidentiality of a system is maintained.

2. ENCRYPTION AND CERTIFICATES MANAGEMENT

Encryption and certificates management is another safeguard for the confidentiality and integrity of information provided by a platform. As explained in our previous report, encryption tools secure data and are validated with a cryptographic key to prevent tampering. Certificates are proof of this security, which is crucial when sharing data over the public web.

There are many services to manage encryption and certificate management on a platform, most of which fulfil the baseline requirements needed to maintain a secure platform, depending on the business use case.

3. LOAD BALANCING AND FAILOVER SYSTEM

Load balancing and failover systems are a technical design pattern used to assure the availability of information when a platform is under pressure. Load balancing should be in place to absorb unexpected spikes in data requests and prioritise authorised users, even when the platform is partially down.
A failover system is a complete replica of a platform that only activates under disaster scenarios, either manmade or natural. It would, at the very least, allow critical functionalities to still be operable, keeping a company online. Failover systems can also take on requests from a load balancer when spikes in requests occur.

4. SECURITY INCIDENT AND EVENT MANAGEMENT

SIEM systems are used to predict and respond to cyber threats and security incidents. SIEMs log and trace access, modifications, and creation of information assets within the platform to detect and analyse abnormal behaviours. SIEMs also raise alerts to human operators during major incidents to kickstart the appropriate actions or disaster recovery processes.

Security is a “must have” non-functional requirement, yet it is often overlooked in favour of easier to notice functional requirements. In our financial monitoring dashboard example, it simply does not affect a user’s experience. However, the dashboard handles an organisation’s most critical piece of data – its finances. It is imperative that executives take security seriously and avoid costly breaches before they occur.

TECHNOLOGY SUMMARY

The technology layer's primary purpose is to fulfil technical requirements of a business use case and do it within the budget set by the business.

We've discussed various layers and key technologies to look out for, but beyond these fundamentals, it is entirely up to a business to determine the right tools to ensure their platform meets the expectations of prior BAIT stages. Hence the need for talented tech individuals either within or outside the business to carefully curate a list of technologies to meet the organisation's business strategies.

BAIT FRAMEWORK SUMMARY

If done correctly, companies will have successfully created a governed organisation that feeds trusted financial information to a real-time dashboard, which in turn creates valuable insights for an entire organisation.

As mentioned before, this is incredibly valuable in the long run, in which dashboard visualisations for a CFO (and his/her finance department) can help steer business strategy decisions, from resource allocation to product development.

However, this isn't the end of a data strategy project. To ensure that a company's data strategy project is well received and implemented, organisations need to give careful consideration to change management strategies and planning.

SECURITY IS A "MUST HAVE" NON-FUNCTIONAL REQUIREMENT, YET IT IS OFTEN OVERLOOKED IN FAVOUR OF EASIER TO NOTICE FUNCTIONAL REQUIREMENTS

SECTION 7
CHANGE MANAGEMENT STRATEGY

Large-scale transformation exercises often draw the ire of anyone subjected to it. As such, it is critical for any company undergoing transformation to pair it with the right change management strategy to maximise the value of their data investments.

Change management strategies include developing an appropriate foundational data-centric culture and backing it up with clear organisational governance models. These strategies should be supported by an enablement committee, which is empowered to sponsor the governance process, organise the dissemination of policies, and communicate effectively to the organisation's employees. An enablement committee is composed of business domain leaders, C-suite management, and the board of directors, who can command change in a business and lead by example.

Business and data specialists can help an enablement committee develop a suitable data culture and the creation of good habits via an organisational cultural governance framework.

CULTURAL GOVERNANCE FRAMEWORK

Data strategy projects reach their full potential when an organisation adopts a data-centric culture.

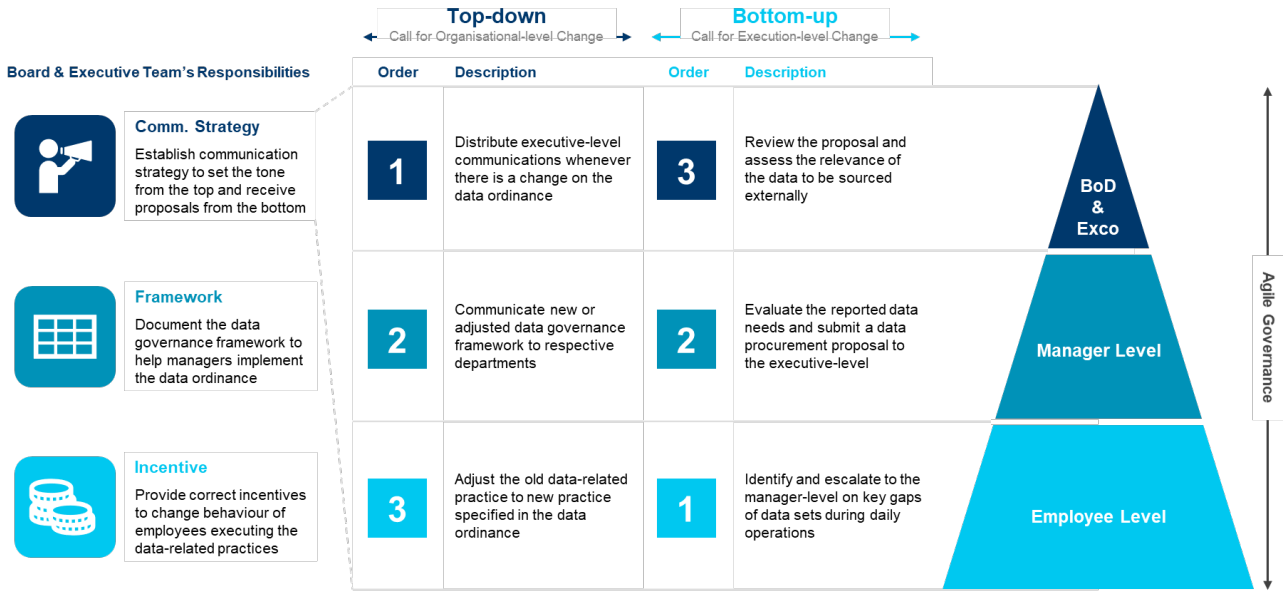
A company can have the best hardware and software available to manage data, but without an appropriate culture that nurtures employee mindsets to embrace data-driven decisions, technological investments are worthless.

The company culture must have a set of practices and controls that can guide employees while remaining agile to regulatory changes, which happens in this ever-changing field of data usage. This begins by companies delegating data responsibilities across the organisation (as opposed to keeping it siloed away under IT teams or niche departments).

The creation of an appropriate culture must start from the top management and, by extension, the enablement committee. This begins by setting the right tone that cascades down throughout the entire company. We identify four key areas that need to be addressed to foster the right culture, including:

1. Communications strategy;
2. Actionable framework,;
3. Incentive schemes; and
4. Agile governance (see Figure 28).

FIGURE 28: CULTURAL GOVERNANCE FRAMEWORK



Source: Quinlan & Associates proprietary framework

1. COMMUNICATIONS STRATEGY

The board of directors and committee members must establish a pertinent communications strategy, along with appropriate channels across corporate levels to communicate data-related requirements and targets. Communications should be enabled in a two-way manner: (1) top-down, to call for organisational-level change; and (2) bottom-up, to drive business execution-level change.

Organisational-level change includes key changes in data ordinance, front-end business intelligence tools, and back-end system infrastructure, which will all result in changes in existing business practices. As employees may be resistant to change (and as adaptations to habits may require time and resources), top-down communication is required to encourage behavioural change from a more authoritative position.

At the other end of the spectrum, execution-level employees manage data-related tasks on a daily basis. As such, they have an in-depth understanding of their data needs. These employees must be empowered and encouraged to communicate their data requirements and expectations to the management team, such that a bottom-up communications channel is also required. An open communication channel should be established for employees to escalate their requests in a timely manner.

2. ACTIONABLE FRAMEWORK

After management-level communication, mid-level managers are responsible for setting up the right framework to set expectations with the upper echelons of the organisation while appropriately steering execution-level employees.

Whenever there is any major organisational change, status updates must be communicated

clearly to the management team, which helps them to make appropriate business decisions. The most efficient and effective communication channel is to create a relevant framework with key performance indicators ("KPIs") that are agreed with the senior team, so that the only necessary information is communicated in the most concise manner.

Furthermore, since the organisation-level communication is inevitably kept generic without operational directions, the bridging role of mid-managers becomes critical to bring about meaningful change. With the help of the management team, managers should come up with department-level communication protocols that are most suitable for the department.

Managers must also develop a – or repurpose the existing – standard operating procedures to appropriately reflect the changes imposed by the management team. Without clear procedures in place, it is nearly impossible to communicate consistently to execution-level employees, which is critical in driving organisation-wide change.

3. INCENTIVE SCHEMES

A key tool to drive change, especially among execution-level employees whose operations are directly impacted by changes in data culture, are a company's incentive mechanisms.

Data-related expectations and performance metrics should be reflected in an employee's performance review and should be directly related to discretionary compensation (and even promotions). Suitable incentives, along with a relationship to data metrics, need to be established, tailored to the different levels and functions of employees across the organisation.

4. AGILE GOVERNANCE

Issues around data privacy, localisation of data, and ownership have increasingly become a concern for corporates, especially as some have tried to take advantage of regulatory grey areas to commercialise sensitive customer data. These practices have driven international and local watchdogs to impose more stringent regulatory standards in an effort to ensure robust and ethical corporate data management. Many regulatory bodies have already developed and implemented data laws, based on country-specific circumstances. With ongoing tightening of data regulations, mismanagement of data may result in severe regulatory consequences.

Data protection laws cover the entire data value chain, from sourcing to disposal, and companies must develop an in-depth understanding of the implications – and must monitor changes – to adapt their data strategy and operations accordingly.

For companies with global operations, different requirements imposed by various jurisdictions may pose significant challenges. As such, businesses must remain agile to these changes and update their governance policies to match the most stringent regulations they are subject to (i.e. they must play to the highest common denominator) if they are to avoid undesirable legal outcomes.

SECTION 8
REAL-WORLD CASE STUDY

GLOBAL LUXURY RETAIL BRAND GOES DIGITAL

We have encountered several companies on their respective data strategy journeys. One case study that stands out is a global luxury retail brand (henceforth known as “GLR”, for simplicity).

GLR provides top-tier luxury products globally through physical retail stores at prime locations across the world. Under tough competitive pressure and due to the COVID-19 pandemic, GLR was forced to accelerate plans in opening digital distribution channels in a very short period of time.

GLR’s business strategy evolved from one of physical to online distribution in a matter of months. This tremendous transformation integrated loyalty programs, communication channels, community applications, and user experiences under a single application. In the process, their data strategy reaped significant value for the company in terms of time and cost optimisation while generating new revenue streams. Let’s explore their journey.

CONTEXT

GLR has a global customer base that comprises of low volume but high value targets. The company has a strong branding strategy that has driven significant brand loyalty over several decades of the company’s existence. Despite their longstanding loyalty and brand staying power, GLR noticed changes in consumption patterns worldwide, shifting to a more digitally focused approach with a focus on travel retail.

Company management realised the need to shorten the time-to-market latency of goods on digital platforms. As an extension of this, they needed to integrate third-party data from physical malls and travel retail chains to get a 360-degree view of their customers, allowing them to create personalised recommendations whenever their customers were online. Additionally, a rapidly evolving political and economic landscape, including disruptions in global supply chains and higher manufacturing costs, meant that the company was seeing a compression in its profit margins. It needed a new strategy to optimise both cost and time.

Management understood there would be challenges around gaining more granular data insights from the organisation, given its historical reliance on physical locations and regional data hubs, which were siloed away from one another. This was all before the COVID-19 pandemic, which pushed customers further away from physical locations and drove unprecedented demand for digital experiences. As a result, GLR was forced to accelerate its data strategy planning and execution.

FIGURE 29: CASE STUDY – OVERVIEW

Source: PALO IT, Quinlan & Associates

DATA STRATEGY PROJECT

1. PROJECT CLASSIFICATION

The project was first scoped by investigating four main considerations: (1) impact to business, (2) resources, (3) time and (4) infrastructure (see Figure 29). Given that GLR was attempting to shift most of its business online and grow its data capabilities to better understand its customers, it was considered priority in terms of its business impact.

GLR management also committed to this change by dedicating a significant amount of financial and manpower resources to the project, making it high priority in terms of resource consumption. Moreover, mission critical systems like global sales data and customer management systems would be affected, meaning downtime would need to be minimised.

A brief analysis of existing infrastructure showed that key operations were working through antiquated legacy systems, calling for the need to upgrade. This meant that significant changes to core infrastructure would be needed. All of this signalled to both GLR and external consultants that this would be a large-scale strategic project.

2. BAIT FRAMEWORK

2.1. BUSINESS STAGE

On the outset, GLR worked to define its business ambitions, which was to provide a bespoke digital experience for its loyal customers. This would be achieved by collecsourting data that was being collected and stored across different products, regions, and distribution channels, in order to build a holistic picture of each of its customers. This would generate the ability to create tailored recommendations to the right customer at the right time, regardless of where the transaction occurred.

By-laws and cost limitations were considered and had minimal affect on GLR's ambitions. With these out of the way, the project shifted towards understanding the current state of GLR's data knowledge and systems. Both were considered acceptable, with the company having a general grasp of its business data domains in specific regions, though not at a unified global level. GLR also understood the purpose of collating its customer data, with a significant understanding of their technology and talent limitations in supporting a future state system. From there, a gap analysis was conducted and passed smoothly into the Application stage of BAIT.

2.2. APPLICATION STAGE

At this stage, data domains were re-identified, with dedicated data specialists engaged to digest the new business objectives. These objectives translated to business and functional requirements, which were subsequently mapped to supporting applications. Key observations were identified below:

1. To aggregate insights, a centralised business intelligence application was needed due to disjointed reporting systems spread across different geographies, causing confusion in reporting at the regional organisation level.

2. To fulfil the goal of managing supply chains to support a global customer base, real-time event processing around manufacturing, inventory, and delivery data was needed to react to unexpected events.

3. The need to integrate customer engagement applications, ranging from loyalty programs, point-of-sales systems, and eCommerce systems, was needed to create a streamlined holistic view of customers from a single location. Customer facing applications were tested here, and end-user approval obtained at this juncture before moving onto the Information stage.

2.3. INFORMATION STAGE

Since GLR already had a basic understanding of data domains in its operating regions, there were some data stewards already in place, which sped up the data discovery process. These stewards were matched with IT specialists into a dedicated team, which designed and re-identified key performance indexes for global domains. The team went on to update data models based on global data domain designs in tandem with architects, who had reviewed the company's existing inventory of useful information systems.

Domain analysis revealed the need to create a data platform that supported both batch and stream data processing to accelerate digitalisation and fulfil the main business objective of understanding customers from a 360-degree view. Other capabilities the data platform needed to support were the ability to integrate third-party data from real estate firms (i.e. shopping malls which GLR operated in) and duty-free retail chains to accurately interconnect consumer data from a variety of locations.

This stage was concluded with a list of necessary data platform capabilities and a set of global data culture pointers to be incorporated into GLR's change management strategy.

2.4. TECHNOLOGY STAGE

Recall that GLR's main business objective was to aggregate customer data under a single hub for ease of access and analysis; this meant that integration into a data platform was the key technology design needed to achieve this.

Several tasks had to occur simultaneously for seamless integration:

1. Cloud solutions were selected for the data platform to be built on; this was to support high speed streaming and batch processing systems with both data warehouses and lakes being used.

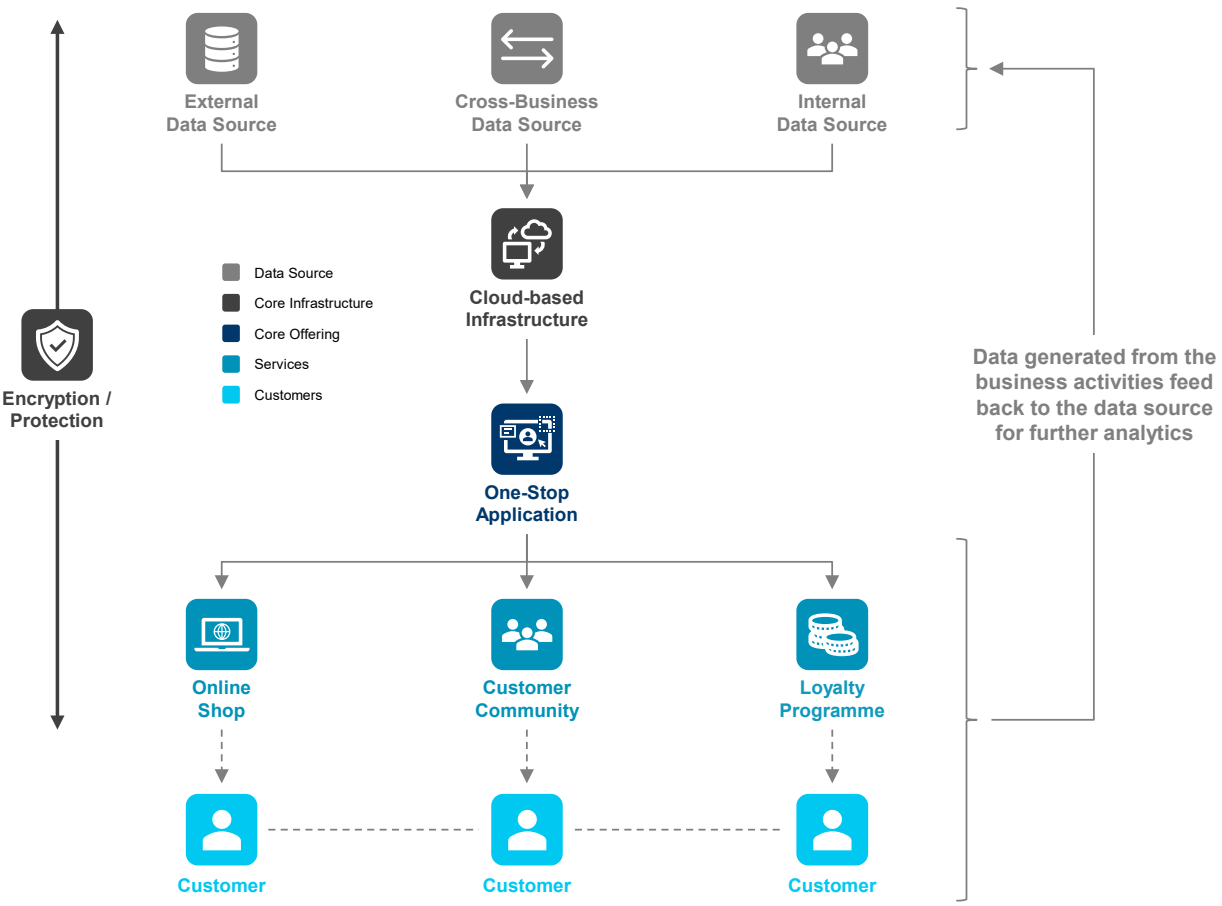
2. Legacy systems were migrated to cloud solutions for scalability, ease of integration, and security purposes, given the amount of data being housed and fetched on a regular basis

3. Analytics and business intelligence applications were streamlined into a single Microsoft PowerBI application for standardised interpretation of data and consumption across worldwide locations

4. A global information security team was formally established, with SaaS-based security incident and event management pipelines to safeguard the entire platform and its enterprise applications

All tasks were merely in support of data platform capabilities, requirements, and business strategies stipulated in prior stages of the project. Technology was selected carefully by the change committee and fell within GLR's budget. The completed streamlined data ecosystem was completed and rolled out for use earlier in the year (see Figure 30).

FIGURE 30: CASE STUDY – END RESULT ECOSYSTEM



Source: PALO IT, Quinlan & Associates

2. CHANGE MANAGEMENT STRATEGY

GLR's change management strategy comprised of a global enablement committee, consisting of board members and senior global and regional executives (in tandem with consultants), which signed off on resource allocation, communications, and project progress.

Once management understood the scope of the data strategy project, regional offices disseminated the upcoming list of the backend changes and streamlining of frontend customer service applications. This was to ensure that middle manager and ground-level employees were prepared for the changes and had time to prepare operational policy changes in advance. Data training pointers from the information stage of the project was incorporated into communications and regional employee training, coupled with locally scaled data incentive schemes. This was to ensure that data governance and new data operations would be adhered to as they were rolled out.

OUTCOMES & NEXT STEPS

At the time of writing, GLR is nearing the end of its data strategy project. Even before the end, meaningful infrastructural and operational developments have brought a majority of core enterprise applications onto a single data platform, with legacy systems replaced to enable smooth data integration.

By doing this, GLR has met several common business goals most companies seek to achieve in their data strategy journey:

1. **Time optimisation:** with all business domains and their respective KPIs integrated completely onto a single application (PowerBI), every authorised executive can monitor quarterly performance and plan local strategies around the trusted data provided (reaching stage 4 of a data organisation's maturity). This reduced latency in data-driven decision making for the company by several months.

2. **Cost optimisation:** GLR experienced increased supply chain visibility and protected the business from major supply shortages throughout the COVID-19 crisis and Sino-American trade war. The platform's newfound ability to forecast logistical and production issues based on supply chain data insulated it from the issues GLR's competitors suffered from. Additionally, the streamlining of applications, such as its loyalty program and eCommerce services, greatly reduced the time-to-market for new programs and products while reducing resources needed to launch.

3. **Inorganic revenue generation:** the centralisation of internal customer data allowed executives and product developers to identify key cross-sell opportunities while leveraging additional data to price themselves more favourably in markets of choice.

4. **Organic revenue generation:** the final integration of third-party data sources from real estate management firms and market research firms greatly improved the company's visibility on macro and micro market and consumer trends. Eventually, with enough data, insights generated from the platform could begin to inform business strategy around market entry or bespoke product development. On top of this, data linked to specific products and geographies could eventually create indices of trends that GLR could sell to research firms, supplementing their services and creating an entirely new revenue stream.

All in all, this data strategy project made it easier for GLR to monetise and track its revenues and costs at a granular level across the entire organisation. A simple integration project yielded short-term revenue and cost opportunities, while creating a foundation for long-term fundamental business model changes.

For a company like GLR, which is only just beginning to scratch the surface of what data can do, its business remains well-positioned for future market growth. And like many companies going digital, GLR clearly recognised that data is the new language of business.

SECTION 9
HOW CAN WE HELP?

BUSINESS CONSULTING

Quinlan & Associates has extensive experience working with a wide range of multinationals, SMEs, and startups on their end-to-end corporate strategy development, business model design, and implementation planning, with significant experience in the data strategy space.

We have worked with clients on a wide variety of data strategy projects, including aspects such as:

1. DATA-LED CORPORATE STRATEGY DEVELOPMENT

We support companies looking to either optimise or transform their business through leveraging data in the pursuit of their business goals. Our services include:

- Review existing or future business strategies and map business objectives and processes for adequate data utilisation
- Perform in-depth market research and competitive landscape analysis into possible edges a company could gain as part of a data strategy project
- Conduct detailed feasibility analysis around business ambitions and objectives
- Develop go-to-market strategies around new data products or services, including product design, pricing, and customer segmentation analysis

2. DATA & BUSINESS STRATEGY PREPARATION

Launching a data strategy project requires a holistic understanding of one's business capabilities. We support and prepare businesses for the arduous journey ahead through the following services:

- Define the current and future state of an organisation's data systems based on its business objectives
- Perform a capability or gap analysis based on the current and future state of the business
- Review operating models and internal operations that would inform business domains and gaps in data usage

3. BUSINESS DATA SOLUTIONS DESIGN

Some data solutions would be best outsourced or built in-house depending on your needs. We can help support the design process through the following services:

- Map business requirements to functional requirements and applications for data integration
- Conduct a buy / build analysis for applications selected, in line with your budgetary needs
- Conduct detailed vendor evaluation / benchmarking

4. GOVERNANCE, CULTURE & REGULATORY ANALYSIS

Best in industry governance practices informs company culture and has an outsized impact on whether data investments bear fruit. We support our client's shift towards a data-centric culture through the following services:

- Evaluate the data maturity of your business, create data metrics, and develop pathways to full data maturity
- Analyse the data culture of the organisation, including the development of comprehensive data governance and cultural policies, including employee training requirements
- Review regulatory data positioning and advise on adaptations to organisational compliance frameworks, based on the company's operating jurisdiction(s)

5. CORPORATE TRAINING

Conduct in-person or online corporate training and coaching programmes to equip our clients' employees with the necessary knowledge and capabilities to support cultural and mindset changes for a robust data value chain:

- Provide world-class employee training workshops (on areas including specific compliance topics and broader cultural change programmes), focusing on turning concepts into action, and committing actions to practice
- Engage managers and executives in dedicated coaching programmes, creating actionable plans for them to inspirate and champion good data business conduct within their teams, divisions and across the entire organisation
- Assess business performance improvements attributable to mindset and behaviour changes from training and coaching efforts, and further fine-tune the programmes.

IT CONSULTING

PALO IT's diverse capabilities and experience allow end-to-end data offerings throughout different corporate data lifecycle stages. With a strong emphasis on technical excellence and diversified methodologies on organizational transformation, PALO IT delivers:

- Information system audits to streamline an organization's strategic vision with underlying technical architecture
- Information system and data integration to interconnect disjointed systems and insight within an organization
- Data modelling and business intelligence development to support and enhance business-wide, data-driven, decision-making processes
- Data governance to establish a reliable and robust organizational culture and system to assure data quality
- Cloud computing, DevOps, and automation strategy to improve business-wide, time-to-market capabilities with regards to IT projects

We have overcome significant challenges and have delivered extraordinary success across numerous projects together with our clients. We know digital transformation is quite a journey, and we know there is no one-size-fits-all solution. Our team supports clients with customised approaches to adapt to their specific context and environment. Some of our services include:

1. KICKSTARTING YOUR DATA JOURNEY

If your organisation wants to become data-driven but hasn't yet taken concrete steps to move in that direction. We can help kickstart your transformation through:

- Technical assessment of existing data governance and architecture in place (if any)
- Technical visioning workshops to create technical transformation backlogs and possible timelines for execution

2. DATA ASSESSMENT

If your organisation has embarked on a transformation, but you are not entirely clear on goals or the implementation roadmap, we offer:

- Technical Architecture assessment
- Data Governance assessment (security, metadata management, data quality and ownership)

3. DATA PLATFORM CREATION

If your organisation fetches / stores data across multiple sources and is experiencing scalability issues, we can build a data platform, tailored to your business needs. We help by offering:

- Creation of a scalable, secure, and resilient platform which can handle high volumes, velocity, and variety of data
- An automated infrastructure setup (on-premises or public cloud)
- Integration of corporate responsibilities such as ESG Data Management and Reporting at scale

4. DATA INGESTION & PROCESSING FOR DATA ANALYTICS

If your organisation already has a platform but requires help building or maintaining data pipelines, we can offer the following:

- Creation of a holistic data pipeline from multiple sources
- Illuminate data recommendations through business intelligence, machine learning and artificial intelligence deployment

ABOUT QUINLAN & ASSOCIATES

Quinlan & Associates is a leading independent strategy consulting firm specialising in the financial services industry.

We are the first firm to offer end-to-end strategy consulting services. From strategy formulation to execution, to ongoing reporting, communications, and employee training, we translate cutting-edge advice into commercially executable solutions.

With our team of top-tier financial services and strategy consulting professionals and our global network of alliance partners, we give you the most up-to-date industry insights from around the world, putting you an essential step ahead of your competitors.

Quinlan & Associates. Strategy with a Difference.

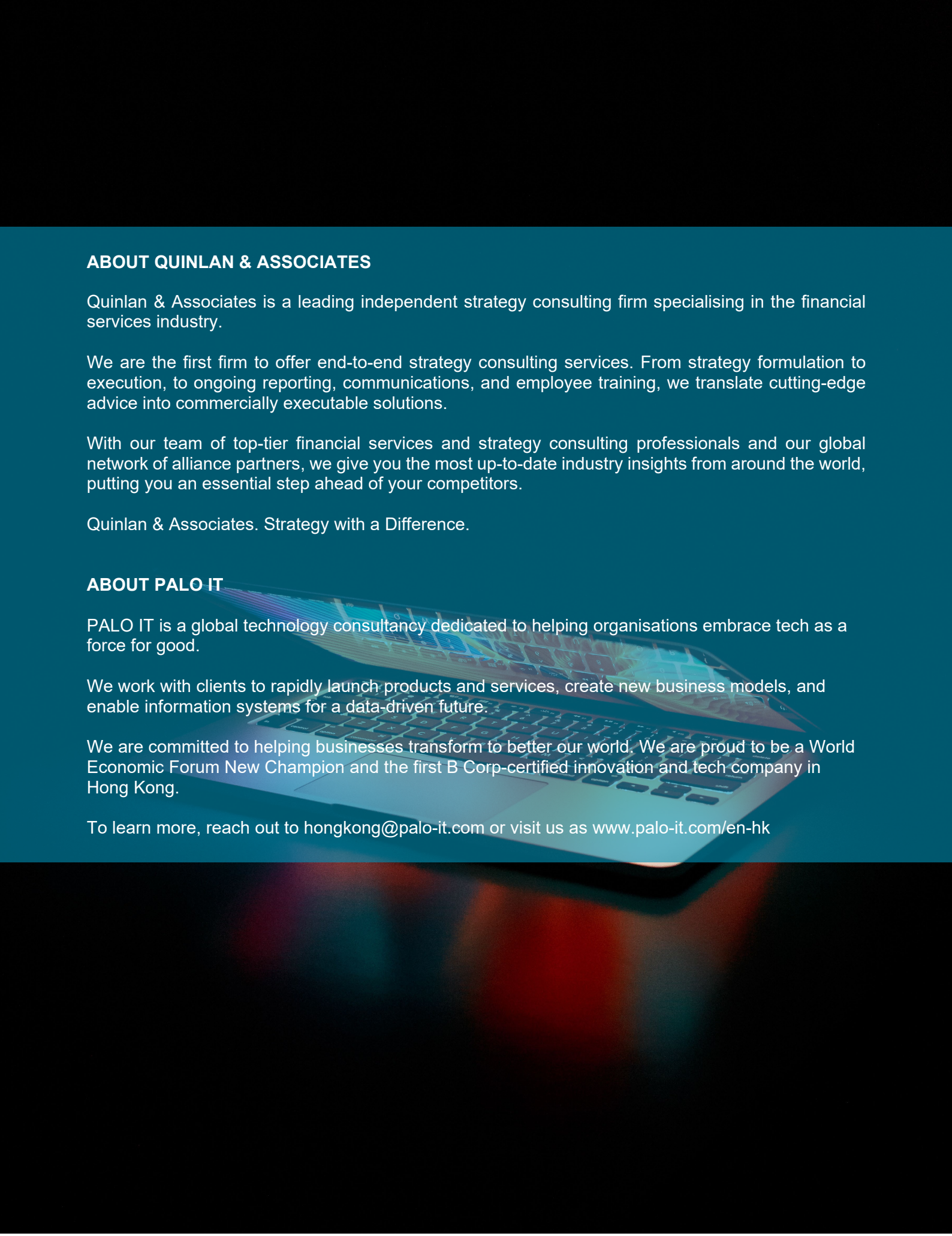
ABOUT PALO IT

PALO IT is a global technology consultancy dedicated to helping organisations embrace tech as a force for good.

We work with clients to rapidly launch products and services, create new business models, and enable information systems for a data-driven future.

We are committed to helping businesses transform to better our world. We are proud to be a World Economic Forum New Champion and the first B Corp-certified innovation and tech company in Hong Kong.

To learn more, reach out to hongkong@palo-it.com or visit us as www.palo-it.com/en-hk



@ enquiries@quinlanandassociates.com | www.quinlanandassociates.com | +852 2618 5000
Level 19, Two International Finance Centre, 8 Finance Street, Central, Hong Kong SAR