

Introduction

AI Age and the Data-First World

‘AI is one of the most important things humanity is working on. It is more profound than, I dunno, electricity or fire.’

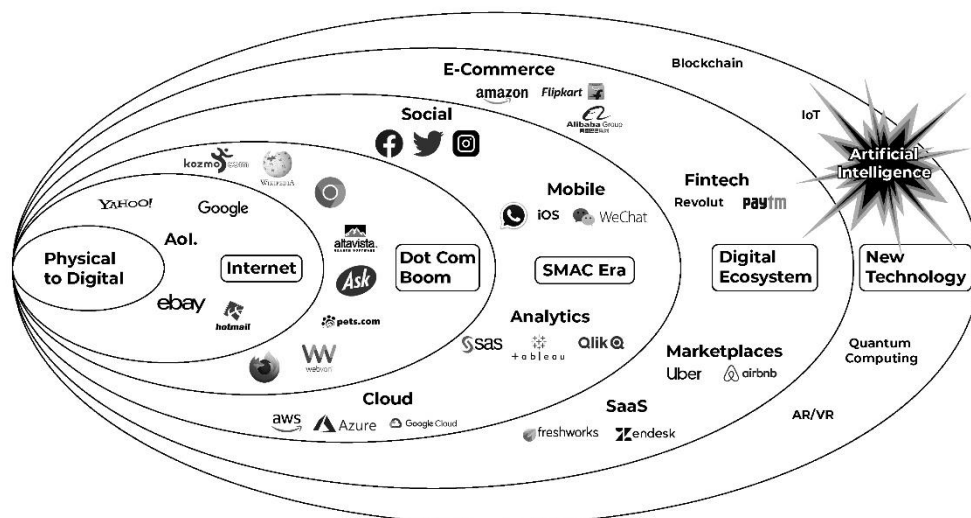
—Sundar Pichai, CEO of Alphabet and Google

The AI age is here

In my previous book, *Winning in the Digital Age*, I defined the digital age as a wide-ranging set of technology trends that have evolved over the years, impacting every aspect of life and business. For me, digital age is a broader set of events happening, a continuously evolving set of bubbles, starting with digitization, to the age of the internet, to the dot-com boom, then the SMAC (social, mobile, analytics and cloud) era, the digital ecosystem era and finally the era of new technologies like artificial intelligence (AI), blockchain, internet of things (IoT), quantum computing, etc. These and many other new technologies have been growing significantly over the past few years. But one key technology— AI—has witnessed disproportionate growth in recent times.

When you look at the infographic below, it becomes evident that the advancements across various domains within the digital age, from digitization to the new technologies era, would not have been possible without the exponential growth of data. So, a significant aspect of this digital age is data, which has brought us to the data-first world, where organizations recognize the immense value of data and aim to prioritize it as a core asset. Now, the data-first world has laid a strong foundation for the AI age by providing abundant and diverse data, further amplified by the advent of Generative AI (Gen AI) which literally unlocked the data of the world. And therefore, within the digital age, the data-first world and the AI age are significantly intersecting and reinforcing each other. Digital technologies have led to the explosion of data and in fact the foundation of the AI age. And AI in itself can significantly enhance many aspects of the data management cycle and also enhance the digital experience. This interplay between digital, data and AI is truly fascinating and leading to the rapid onset of the AI age.

AI has exploded in the last few years to become one of the key drivers of the Digital Age



Of course, artificial intelligence is not a new concept. It has been around for more than seventy years, with the concept emerging as far back as the 1950s. And over the years significant progress has been made in this field. But with all its promises and possibilities, scaling AI has been a major issue. This is because for AI to be successful, three components are crucial: data, computational power, and algorithms. While algorithms and computational power are more solvable, data has remained the most critical challenge, becoming more complex by the day. An AI model is only as good as the data that powers it. In the enterprise context, despite the availability of Big Data, organizations often encounter challenges in accessing sufficient and relevant data. Although the concept of Big Data suggests abundant data resources, accessing the right data that aligns with specific business needs can be a complex task. While digital natives like Amazon and Netflix with massive customer behaviour data have an advantage, for most companies across industries, it isn't easy to get the data required, clean it, and build AI models on it. This is why, in most cases, AI has not grown beyond proof-of-concept or the experimentation stage to scale, other than a few specific use cases like personalization.

But since the launch of GPT-3 in November 2022, the story has changed. Gen AI has sparked a massive unlock in AI capabilities. How? In September 2021, all the data on the internet was downloaded. And by accessing and working with this immense data set, powered by sophisticated neural network models, Gen AI has unlocked a higher level of intelligence and generative capabilities. And because we now have a way to literally leverage the data of the world, the challenge around availability of data has been significantly addressed. It has led to a significant leap in AI capabilities, moving beyond traditional interpretive capabilities of AI, like pattern recognition, content categorization, etc., to the generative capabilities of Gen AI, like creating new and original content. With that the 'AI age' is now very real, happening right now. The genie is out of the bottle!

Like Malcom Gladwell said, and I quote, 'The tipping point is that magic moment when an idea, trend, or social behaviour crosses a threshold, tips, and spreads like wildfire.' The advent of Gen AI is that tipping point. It has opened new opportunities to leverage the collective wisdom of crowds and tap into the infinite possibilities of data. Gen AI also contributes to the expansion of data sources, by adding new content to the pool of data available for analysis and innovation. Furthermore, Gen AI

offers a clear advantage by providing a starting point for problem-solving, reducing the ambiguity of where to begin.

So, the entire trajectory of the AI age hinges on data. Because data is that underlying element that would help in contextualizing and maximizing the potential of AI. Combining Gen AI models with the vast amount of data within an organization would allow for the creation of specific and actionable insights that are relevant to the context of the business. This integration would help drive both efficiency and innovation within the organization.

Combining Gen AI with individual data would enhance the user experience, making it more relevant for individuals, whether it's personalized product recommendations, customized news feeds, or personalized marketing messages. By analysing data on every individual, such as their preferences, behaviours, and past interactions, etc., AI can generate tailored content and suggestions that are specifically designed for them. So, while Gen AI models alone may not be enough, contextualizing it by integrating it with enterprise and individual data, as well as existing models and systems, can unlock synergies and allow for the creation of more comprehensive and impactful AI-driven solutions.

I foresee months and years ahead marked by tremendous experimentation and development that would result in significant progress, not just in Gen AI but in other aspects of AI as well. Industry- or function-specific Gen AI models will emerge, meshing with enterprise and individual data, enabling organizations to unleash the true potential of AI.

Hence, this book delves into the entirety of data, uncovering opportunities, challenges, and solutions across enterprise, individual, and macro levels. This exploration I believe is critical to uncover the pivotal role of data, which is the necessary foundation for organizations to thrive in the data-first world and triumph in the AI age.

My journey into the data-first world

The past twenty-five years of my professional and personal life have been a fascinating journey of experiments with data. Often caught in the thick of things, where data played a critical role in business success, I spearheaded some highly innovative and radical ways to make an impactful change with data. My experiences have led me to a firm conviction that data is one of the greatest opportunities in the AI age and data-first world that we are in, but equally it presents some very difficult challenges as well. And despite the undeniable importance of data, it is a topic that is not well understood.

This book is a synthesis of my experiments and experiences with data, where I have tried to create a practical playbook that should be helpful to both seasoned executives and young professionals and students on how to win in a data-first world. These experiences include building perhaps the most unique and innovative knowledge centre in the world, for McKinsey— a top management consulting firm, leading global offshore operations and setting up the strategy and planning and analytics functions at Fidelity International—one of the largest active asset manager in the world, leading Flipkart—an ecommerce major and the biggest start-up in India at the time, at one of its most tumultuous times as a chief operating officer (COO), and driving multiple digital initiatives for a number of Fortune 500 companies as the CEO of Incedo Inc.—a US-based consulting, data science and technology services firm, for more than a quarter of a century now. But before I delve deeper into my journey, let me first set the context for why I was convinced I must dedicate an entire book on the topic—DATA.

The current era of Digital Revolution, beginning somewhere in the late twentieth century, is marked by the widespread adoption of connected digital technologies that have significantly transformed various aspects of our lives and businesses. The digital economy now contributes to more than 15 per cent of global gross domestic product (GDP), growing 2.5x faster than overall GDP, on average, over the past decade.¹ As my fingers fly across the keyboard, four of the top five publicly traded companies in the world are technology companies, each with a market capitalization over 1 trillion (US) dollars.²

Today, we stand at a crucial juncture in the Digital Revolution happening all around us. Decades of technological progress has democratized access to digital technologies. Mobile phones and computers have ingrained digital into our daily lives, influencing every aspect of it, whether we like it or not. Historically, every period of human endeavour has had these pivotal moments, defining human progress, from the crafting of fire the very first time to the reusable rockets of SpaceX. But most of us typically overlook these pivotal moments in pursuit of immediate concerns and endeavours, only to recognize their significance through hindsight and the passage of time.

In my previous book *Winning in the Digital Age*, I shared my learnings of the past twenty-five years working both as a consultant serving other organizations and senior executive leading business and digital transformation at various organizations. It was an attempt to bring clarity to the complex and vast topic of digital, by sharing practical insights and best practices on the various aspects of digital transformation.

During this journey, the most prominent aspect of the Digital Revolution that emerged for me was DATA. I have been a wide-eyed witness to the phenomenal growth of data and the unprecedented opportunities that it brought about with it. I have seen and in fact been one of the early voices in propagating the tremendous opportunities that data and analytics could bring to an organization. I have seen how data and analytics have transformed the world of consulting and beyond and have been blown away by the sheer force that data has become today.

But just like every story has two-sides, I have also been a teary-eyed witness to the havoc that inability to manage data well can cause in an organization. As data grew in size, shape and form, it became a relentless force in itself. A force that is becoming harder to contain each day. It was like trying to stop a Tsunami with your bare hands. To the extent that it had the potential to overwhelm and even wreck the entire organization.

This is when I realized that data can easily create a Catch-22 situation—a difficult paradox with no apparent way out. Things can easily get out of hand if one does not know how to handle data the right way, and one can find themselves in situations that are often impossible to salvage. Too dramatic? Well, let me take you through my journey over the years and I will let you decide for yourself.

[Recognizing the power of data and analytics at McKinsey](#)

My data journey started as a consultant at McKinsey in 1996. Those were the days when email accounts were not commonplace or easily accessible and mobile phones were a luxury. Yes, I know it is hard for the current generation, who probably grew up streaming cartoons on YouTube, etc., on their parent's mobile phones, to believe life existed without phones. My son who is a product of the current generation found it astounding and asked me once, 'How did you contact someone in case of an emergency?' Well, we did, and we did ok. But I am sure most millennials and all baby boomers can relate. Anyway, I digress.

So, as a consultant in McKinsey, my main job really was to crunch data; 1996 was a data-scarce world, and McKinsey was a leader in strategy consulting owing to its capability to dig out data from nowhere, make sense of it and package it for clients. And that's what people like me, from the best business schools of the country, were hired to do.

For example, if we had to do a tractor market sizing in Brazil, believe it or not, it used to be a three-month-long project, because such kind of data wasn't readily available anywhere in the pre-internet era. And this was literally a brain-bending exercise. We would spend days collecting information from various global locations, often travelling across multiple countries, then collate it and size the market based on multiple approaches. To build confidence in the analysis, we used methods like triangulation—using three different approaches to get to some estimate and then narrow it down to a number that appeared to be the most logical. The catch here was that no matter how logical it was, it was still nothing more than an educated guess. And some of the best minds were employed to do just that. Such was the reality of the data-scarce world.

At McKinsey, my specialization was operations diagnostics and benchmarking, travelling all over the world, comparing a company's internal data with external best practices, extracting insights to highlight improvement opportunities for the company. So, I was dealing with data and analysis day in and day out. After four years of doing that, I left the company to build my own start-up for the next two years. But as fate would have it, things didn't work out as planned and I got an offer to join back at McKinsey to head the McKinsey knowledge centre (McKC) that was set up in 1998. But at that point it was quite fledgling, with just a few people trying to make something of it. My job was to lead this effort and to transform this back-office kind of set-up that provided overnight research support to a few teams in North America into a true knowledge centre. But the firm didn't have any concrete plan to achieve that.

When I came onboard, I realized that just in the span of two years that I was away from McKinsey, from 2000 to 2002, an information revolution had happened. I was literally blown away by how much had changed in the way data was now being generated and consumed. And the key force behind it was the explosive growth of data because of 'The Internet'. And within a short span of 2–3 years when I wasn't looking, the world of information had completely transformed. The information that we toiled for 3–4 months to gather, something that required highly skilled, highly paid professionals to be employed, was now available at the click of a button and could be done by anyone who had Internet access and basic Google search skills.

This was a hard-hitting moment for me. A moment of revelation that McKC, which so far was considered as the poor cousin of consulting, suddenly had the keys to the treasury. It was now the fuel that fired the consulting engine and had the ability to either empower or dis-intermediate the consulting teams. We had very quickly moved from an information-scarce world to an information-abundant world. I saw the opportunity that came with data abundance. Data became a source of knowledge that could be translated into insights very quickly and effectively. Voila, I had found the secret ingredient to pivot McKC into a one of a kind, best-in-class knowledge centre. How?

Let's consider the tractor market sizing scenario again. Now, data was available on not just the market size, but also the number of players, market share of every player, etc. All the information was just a mouse-click away. In addition to that, since this data kept growing, we now had the opportunity to capture it both longitudinally—over several years, and cross-sectionally—across multiple players. So, the role of analysts evolved from searching data to bringing it all together and analysing it to generate deeper insights. The Knowledge Centre had become more than just an information dissemination centre. It was now a Research and Information Centre. We started

realizing the value from data because now we had enough information that we could do research on top of it. There was no stopping us now. As the research capabilities grew bigger and more sophisticated, the Research and Information Centre was renamed as the 'Knowledge Network'. And as data continued to grow, we started to bring it all together from various industry verticals and functional areas in a systematic manner, which became the foundation for what we call 'Proprietary Knowledge' and the creation of proprietary databases at McKinsey. As a result, the expertise that was earlier in the minds of people could now be institutionalized. For example, now we were able to develop diagnostics that would systematically bring the benchmarking data for each industry vertical together, which is the starting point for many projects that McKinsey takes up today. We were able to build deep expertise and research capabilities that truly transformed the way one of the world's best consulting firms engages with its clients today. This is how we graduated from data to analysis to insights and finally to truly generating value from data.

And that brings me to another interesting story about setting up analytics capabilities at McKC. In the data-scarce world, problem-solving was predominantly done using deductive logic—the hypothesis route. But now suddenly we had access to bytes and bytes of customer data that could be used to do precise segmentation. And this data was growing by the minute. Suddenly we were dealing with not just 10 but 10,000 data points. Analysing data at such a large scale required technologies more advanced than excel. And lo and behold! We had technologies like statistical analysis system (SAS) and statistical package for social science (SPSS) then, that could deal with up to 2 billion rows of data at a time.³ And this was another eureka moment for me as I recognized a whole new world of opportunity that these technologies could open for us. We no longer had to rely on intelligent guesstimates. We could analyse large sets of data faster, with less manpower, with greater accuracy and could perform varied types of analysis which we had never imagined was possible. I also realized that as data kept growing, it would become impossible to generate comprehensive insights without upscaling to newer and more sophisticated analytical tools, and we could face the risk of being left behind as a consulting firm.

Often true with any megatrend, even the best and the most experienced minds may not be able to foresee the potential of an idea before it becomes a phenomenon. Also moving away from the tried and tested formula that has worked for years, to try something new and radical isn't that easy for all, which was also the case in a great firm like McKinsey at that time. But I was convinced that setting up a specialized analytics capability at scale was in fact an important pivot point for McKinsey. So, I pitched for it multiple times, before multiple stakeholders at McKinsey, including going to the Shareholders Council, the highest decision-making forum within McKinsey, with detailed study and numerous proof points. Yet, it was difficult for the top brass at McKinsey to fathom that this new breed of analysts sitting in India can achieve, or maybe even be better at, something that the best minds from the likes of Harvard could not do. Then again, I wasn't ready to give up yet. Even at the McKinsey Practice Olympics, the firm's initiative in 2005–06, where colleagues would submit proposals for new knowledge and innovation initiatives, I presented my perspective on how analytics was going to change the face of consulting in the future. Alas, there too I was met with a lukewarm response. Undeterred in my conviction, I continued to pursue the idea. In the absence of topdown support, we took the route of bottom-up experimentation, and this opportunity certainly existed in a decentralized global firm like McKinsey. It took a while, but we did end up building a series of specialized analytics capabilities, most of them completely new for McKinsey, and helping it create new areas of engagement with its clients. Analytics and other data-driven capabilities that I mentioned earlier helped not just create the most innovative knowledge centre of its time, but fundamentally reshaped how a great consulting firm like McKinsey serves its clients. Today, digital

and analytics, the combination of proprietary knowledge and the platforms, make up for a pretty big chunk of McKinsey's global revenues.

That is the story of how data and analytics, in just the short span of 20–25 years, changed the consulting landscape, impacting it in such profound ways. A firm like McKinsey never really planned for it initially, and many organizations like them, that owing to the sheer force of the changes that data was bringing about, had to completely transform their business models to adapt, survive or stay ahead of the curve.

Dealing the dark side of data at Flipkart

Now that I have talked about the good side of data, let me also unveil the dark side of data. This story began when I joined Flipkart in early 2016, which is one of India's dominant e-commerce giants. But back then too, it was considered as a poster child of the Indian start-up explosion. And the culture at Flipkart was markedly different from what I was exposed to. While McKinsey had more of a hypothesis driven approach, Flipkart was absolutely obsessed with data. Everything was data-driven, with every employee exposed to and dealing with thousands of data points every day. Over 100 people joined our daily stand-up meetings, and everyone talked in numbers. And me, an IITian, an IIM grad, an ex-McKinsey consultant, a CXO, who considered myself pretty good at data analytics, was struggling with the endless sea of numbers.

Our biggest project at that point was the Flipkart Data Platform (FDP), to bring all the Flipkart data onto one platform, creating a single source of truth and building a single view of the customers. Flipkart was obsessed with becoming truly data-driven by bringing all the information at one place to be viewed equally by everyone working there. We believed that this was the answer to generating unique and innovative insights. Now, between 2014 and 2016, Flipkart was on an acquisition spree. It acquired Myntra and Jabong, two e-commerce companies, to add to its fashion and lifestyle segment and acquired a stake in MapmyIndia to enhance its delivery operations. Year 2016 was also momentous for Flipkart as it crossed a significant milestone of acquiring 100 million customers that year.⁴ With such a significant customer base and expansion in operations, imagine the amount of data that was being generated at that time. As a result, the situation quickly spun out of control. By the end of 2016, Flipkart was going through the biggest crisis since its inception. As a result, a major internal shuffle ensued at the top management level, Amazon was slowly invading the e-commerce space, the company was losing more than a billion dollars every year and was on the brink of becoming cash strapped. In essence, the 7-billion-dollar company was literally coping with an existential crisis at this point. But the biggest problem the management team was often debating was not that we were losing money, or that Amazon had beat us, or that we could not find investors. The biggest problem for us was that the data platform had become a complete bottleneck, which significantly impaired decision-making and execution in the company.

A single source of truth and a single view of their customer was expected to act like a turbocharger, boosting us to dash ahead in the e-commerce race. But contrary to that belief, it ended up stalling the company's decision-making and operations, bringing everything to an abrupt halt. The single source of truth now turned into the biggest roadblock for the company.

Here is the harsh reality. The one key error in judgement that most organizations and their data and technology experts are making is underestimating the pace at which data is growing and will continue to grow in future. As I explain later in Chapter 2, in detail, data is growing in geometric progression whereas technologies that are being built to deal with this data are growing in arithmetic progression. So, the gap will continue to widen, with technology

always failing to catch up.

Therefore, learning it the hard way, I have now ceased to believe in the myth of a single source of truth or a single view of the customer. It is a very difficult feat to achieve.

In Flipkart, we realized it the hard way too. The best engineers of the world—Flipkart in those days hired the *crème de la crème*—got down to solving it. But alas, ‘all the king’s horses and all the king’s men couldn’t put Humpty together again’. We realized that we must let go of this impossible task and tackle the data problem in a completely different way. Luckily, we had some practical minds who could step away from the endless technology puzzle that we were trapped in. And that’s when we realized that we cannot solve the data problem only as an engineering problem. It required a radically different approach. Since the data was too large and varied, quality was always an issue. And the Flipkart business was dependent on taking real-time actions based on real-time data. The challenge was how do we get the right data and accurate data each time. And this was the moment when another myth of Big Data that the law of large numbers will take care of quality was also dispelled for me. It did not work in real life. It was an impossible task to manage the quality of the truck loads of data that was pouring in every day.

The answer was to start small. We narrowed down the problem to make it more manageable. We decided to first tackle just four critical decisions—we called them use cases that we wanted to drive. To drive these decisions, we identified the number of data pipelines required for just those, which were somewhere around 100 data pipelines, versus 7,000 data pipelines we were dealing with earlier. These pipelines were then cleaned and curated for those use cases, and since this was a manageable task, we could make it work. And then we continued to replicate the same mode across a greater number of use cases as well, and slowly and gradually we dug ourselves out of a very big ‘Data’ hole.

[And the data journey continues](#)

My data journey continued as I joined as the CEO of Incedo Inc. in late 2017. And believe it or not, I was flabbergasted to see that one of our biggest clients, a global financial institution, was also falling into the same trap of trying to create a single source of truth, by consolidating all their data from a risk and compliance perspective. Hundreds of millions of dollars had already been spent on this project over the years. But eventually the project was scrapped since they were running after the unachievable. I stepped on the scene a bit too late to stop the disaster from happening. But here was another proof that such big bang approaches rarely work and that is the reason why Big Data initiatives that cost a lot of money still often fail to deliver the expected impact.

In contrast, another client of Incedo, a telecom player, was trying to bring all their customer data together to create ‘a data lake of all data lakes’. And I could clearly see them rapidly cruising towards a dangerous whirlpool of never-ending spend on data infrastructure build-out. But here we could help salvage the situation. We did a lot of groundwork and built a case to prove that it would be better to take a more focused, business use-case centric approach. Basis this the mega infrastructure build-out that was being envisaged was moderated.

[Key themes covered in this book](#)

This book on data is my effort to bring forth the learnings from such experimentations and suggest actionable strategies to harness the full potential of data especially as we march into the AI age. I do not claim to have found the holy grail, because data is such a vast and complex topic that is becoming even more complicated and therefore has not been fully understood. But over the years, I have worked tirelessly to understand how data works and believe I have succeeded in uncovering a

few key mantras to unleash the power of data. Through this journey of exploration and revelation, I intend to provide some critical insights and a structured approach to organizations, entrepreneurs and young professionals alike. And this book is not just for those who deal with data day in and day out. There is something in it for everyone. From the individual who is interacting with data every day through the internet to multimillion- or multi-billion-dollar companies that are in the midst of their digital transformation agenda to political leaders looking to solve complex national and global problems. Data is a valuable asset for everyone, and so dealing with it is a critical skill that each one of us should master to win in the data-first world.

The book is divided into three sections. **Section I** is dedicated to **Understanding the Data-First World**. The fuel powering the digital age and the upcoming AI age is data—like electricity powered the second Industrial Revolution,⁵ in this fourth Industrial Revolution (the digital age), data is a source of significant opportunity and advantage if utilized the right way, as proven by the digital natives like Google and Amazon. This section lays out the opportunities that the explosive growth of data has brought about for both organizations and individuals and delves into the challenges that emerge as organizations are now forced to deal with exponential amounts of data, while the technologies are playing catch up.

This section has five chapters. I begin by highlighting the **Data Explosion** in the digital age (Chapter 1). With the advent of the internet, data witnessed exponential growth, leading to the phenomenon of Big Data. This growth evolved across three core dimensions: volume, variety, and velocity, as data expanded in terms of size, type, and speed of generation. Then I put a spotlight on **Data, the Fuel for the Digital Age**, that powers every aspect of life and business, transforming experiences and pushing industry boundaries (Chapter 2). And as organizations gain access to vast customer data, they are able to unlock unprecedented opportunities for value creation and reimagine business models. Next, I highlight the possibility of **Value Reimagined** through the Data-Insights- Actions-Impact (DIAI) framework, enabling organizations to realize transformational value from data (Chapter 3). By generating insights and taking effective actions, organizations can deliver tangible positive outcomes for customers and themselves.

This raises an important question—despite the ubiquitous nature of data, and the unprecedented opportunities made available, why are most organizations struggling to maximize value from data? My study reveals the contradictory nature of challenges that most organizations face today —**The Data Paradox** (Chapter 4). On one hand, they are unable to deal with the overwhelming 3V (volume, variety, velocity) explosion of data, while on the other, they struggle to get relevant insights to enable data-driven decision-making. And that despite continuous attempts to deploy innovative technologies, the problems persist. On closer examination, I discovered **The Root Cause** to be a disproportionate focus on solving the data problem with technology and infrastructure, which in my experience, is not enough (Chapter 5). Most times, the root cause of the data paradox is more logical (narrowing the problem to be solved) than physical (technology infrastructure), and can have a wider impact on the organization, if it remains unresolved.

Thus, to navigate through the choppy waters of the data revolution age, organizations must understand the root cause of the problem to narrow down the issues and attempt to solve it through a structured and comprehensive approach.

In Section II, **Maximizing Value in the Data-First World**, I propose an innovative and practical way to make sense of all the chaos brought about by the explosive growth of data to maximize value realization from data. I propose a thirteen-component solution framework which I call the **Unified Solution Framework (USF)** that can very well be the thirteen mantras for organizations to succeed in

the data-first world and the AI age.

These thirteen components are divided into five layers. Layer one is the **Business Objectives**, which is the starting point for any organization. Here, I emphasize that it is critical to clearly **Define the Business Problems** (Chapter 7). Because in my experience, breaking the business problems down into logical components, prioritizing them, and mapping data requirements accordingly, in effect narrowing the data problem to be solved, is the only way to make the data process more focused and manageable.

The second layer is recognizing and understanding the rich, diverse **Data Ecosystem** made available to organizations today because of the 3V explosion of data. Here I have highlighted three key types of data—multi-source data, real-time data and proprietary data (Chapter 8–10)—which I believe drive disproportionate value generation for organizations. **Multi-source Data** is critical to generate comprehensive and more meaningful insights. **Real-time Data** is essential to drive action with speed and **Proprietary Data** enables organizations to build sustainable competitive advantage.

As I mentioned before, despite the abundance of data available today, organizations struggle to leverage it effectively due to increased complexity and scale. Keeping up with the volume, variety, and velocity (3V) of data and translating it into meaningful insights or actions requires a well thought out data architecture. So, the third layer, **Technology Infrastructure**, emphasizes the right way to bring the 3V of data together, in a form where it can be used in combination to solve multiple business problems. Here I emphasize the criticality of building a customized data stack that enables scalability and facilitates end-to-end integration of the various layers of the data management value chain, which isn't possible without building a **Modern Data Stack** (Chapter 11).

Once the data ecosystem and the technology infrastructure are established, organizations must build **Core Processes**, which is layer four, to achieve the desired returns from data initiatives. Setting up and institutionalizing these core processes is important for the successful implementation of any data initiative. Here I highlight the importance of **Agility** to successfully adapt to the evolving nature of problems in a VUCA (volatility, uncertainty, complexity, and ambiguity) world (Chapter 14). I recommend a 'two-speed approach' to start by delivering on short-term business value through speed One initiatives, while building long-term capabilities, speed Two, by bringing together speed One initiatives. Furthermore, I discuss the importance of **Data Democratization** (Chapter 15). Making data available to all employees appropriately with seamless, anytime access while empowering them with the right level of knowledge and know-how to use it effectively, is critical to enabling better, more informed decision-making and driving actions faster across an organization.

And while it is paramount to make data accessible and available to all, all the efforts towards building a data-driven organization could easily be put in jeopardy if the right **Data Security** measures are not taken (Chapter 16). And with the evolving complexity of the data threats, and the growing vulnerabilities brought about by ever expanding data and technology ecosystems, I recommend implementing a zero-trust architecture, keeping the context in mind.

And finally, **Organization and Culture**, which make up the fifth layer of the Unified Solution Framework, address the mindset and behaviour of people in a manner that fundamentally transforms the organization's DNA, thus impacting its overall ability to drive transformational value from data. Here I talk about the need to rethink organizational structures, decision-making processes, and most effective organization design to execute data initiatives, for building effective **Organizational Alignment** (Chapter 17). I also emphasize the importance of adopting a three-pronged approach to move away from gut-based decision-making, where decisions are taken based

on the highest-paid person's opinion (HiPPO) and build a robust **Data Culture** (Chapter 18). And finally, as the Big Data world becomes more complex, technologies evolve rapidly and AI goes mainstream, I see the role of specialized **Data Talent** evolving significantly as well (Chapter 19). Data talent needs to move beyond the core data skills to include aspects like deep problem-solving, critical thinking, creativity, storytelling, and deeper domain knowledge, as focus shifts to achieving business outcomes and driving impact.

The five layers that must come together to create a data-driven organization require something to hold them all together. The glue that unifies the efforts at each layer of the Unified Solution Framework are two critical elements which I call the **Integrators**. The first one, **Data Quality**, to me is the biggest concern of the Big Data world, as data continues to become more complex with huge and diverse data sets being continually generated and new use case needs emerging (Chapter 12). I propose a context-first approach to tackle quality issues. The second integrator to me has the potential to become the most effective solution to maximizing value from data. These digital business solutions, called **Data Products**, built as a vertical slice integrating various elements across the data stack to deliver a business outcome, can help organizations deliver impact at speed and in a repeatable manner (Chapter 13).

In **Section III, Data for Individuals and Beyond**, leading in with the quote 'As is the microcosm so is the macrocosm, as is the macrocosm so is the microcosm', I change gears, and both zoom in to the individual level and zoom out to the macro level of the society and nations. I explore the universal nature of the data paradox, highlighted in Chapter 4, which plagues organizations today. I emphasize on how the data paradox that is true at the enterprise level, also exists both at the individual level and the society/global level (Chapter 20). I start this section by elaborating on how Big Data has brought about **The World Hyper-personalization**, creating a 'segment of one', where every individual is unique and therefore the products and services are increasingly being customized at the individual level and not at mass level as was happening traditionally (Chapter 21). But for this to happen effectively, individuals need to engage more with the digital world, allowing organizations to better understand their customers and generate greater value through personalization.

Then I talk about how despite so much data being available to individuals, most of us are still relying on the gut-based approach to make decisions in real life, whether big or small. I find the DIAI framework, that I talk about in Chapter 3, useful at the individual level as well to effectively **use Data for Better Decision-making** (Chapter 22). Furthermore, while access to data is valuable, it can lead to information-overload for an individual and make it challenging to extract meaningful insights for decision-making. The key is to balance **Information vs Wisdom**, by building wisdom, which is proprietary to an individual, to cut through the noise by identifying patterns that matter most (Chapter 23).

Moving on to the paradox of **Digital Engagement vs Mental Health** (Chapter 24). Information-overload leads to noise that surpasses an individual's capacity, creating mental havoc. The solution lies in finding ways to create a stronger connection with oneself at a deeper level, to get to a right balance, which to me is spirituality. I further move on to address another paradox of the data world, **Data Sharing vs Data Privacy**, which is becoming a greater concern as we continuously engage with the digital world, generating and sharing data about ourselves every minute of every day (Chapter 25). While it is difficult to keep track of how and what data is being accessed for what purpose, we must be vigilant in protecting the few critical data points most vulnerable to misuse.

Now moving beyond individuals, I believe in the importance of **Data Collaboration for a Better World** (Chapter 26). Global data collaboration can address complex problems like climate change and

healthcare that no one country can solve alone and help create a more equitable and sustainable world. And finally, I have no doubt that a new world order will emerge as the AI age unfolds and nations that use **Data as the Source of National Competitive Advantage** will emerge as the new superpowers (Chapter 27). So, data is a topic not just for enterprises and individuals, but for national leaders and policy makers as well.

The pages that follow is the result of my many deep encounters with data over time. The more I have experienced and researched the way data works, the more convinced I got of the boundless potential of data to generate extraordinary value and that now is just the beginning. Through this book I lead you into this remarkable world of data to open your eyes to its full potential, truly making it the pivot which can help organizations win in the AI age. But as data continues to explode in all directions, the data paradox is overwhelming most organizations leading to frustrations and challenges in value creation. So, through the Unified Solution Framework that I present, my aim is to provide every individual, professional and organization a playbook to succeed in this data-first world.

I have endeavoured to delve into every component of the solution framework, attempting to demystify each concept in detail and highlighting its importance in the data-First world and its growing criticality as the AI age unfolds. As I have mentioned earlier, data and AI have a recursive relationship, so AI would play a critical role in solving for each of these components, at the same time solving for these components would be critical to unlocking the full potential of AI. In simple terms, the better organizations get at dealing with data, the more value they will be able to generate from their AI efforts and AI would play a significant role in solving some of the key issues organizations face with data. I have attempted to provide a structured and practical approach on how to assimilate each of these components into the DNA of an organization to maximize value from data in the data-first world, an approach that is industry or function-agnostic. I have also called your attention to the fact that this data paradox exists at both the individual and macro levels, and that various components of the solution framework can act as a guide to resolve the paradox at those levels as well.

It is truly exhilarating to find ourselves at the forefront of not just one, but two remarkable phenomena that are unfolding almost simultaneously. The first is the emergence of a data-first world, where data has become a central driving force, shaping industries and fuelling innovation. The second is the dawn of the AI age, propelled by the advent of Generative AI, that has created the possibility to leverage the data of the world for the first time. The convergence of these two holds immense promise and the opportunities are boundless.

And with this book I intend to equip you with both the principles and practical frameworks to tackle the challenges of the data-first world and get ready for the AI age.

With that thought, let's embark on this fascinating data journey together!

Through this book I will attempt to answer the following key questions:

1. Why and how can AI powered by data create transformational value for enterprises and individuals?
2. The world is full of paradoxes and so is the world of data. How can individuals and enterprises effectively deal with the paradoxes, to unlock the transformational value of data and AI?
3. Are there some principles that we can learn from life and apply to data and vice versa, as we navigate the new world of data enriched lives?

References

Introduction: AI Age and the Data-First World

1. Zia Hayat, 'Digital trust: How to unleash the trillion-dollar opportunity for our global economy', WEF (17 August 2022), accessed on 14 July 2023, viewable at: <https://www.weforum.org/agenda/2022/08/digitaltrust-how-to-unleash-the-trillion-dollar-opportunity-for-our-globaleconomy/>
2. 'Largest American companies by market capitalization', Companies Market Cap, accessed on 14 July 2023, viewable at: <https://companiesmarketcap.com/usa/largest-companies-in-the-usa-by-market-cap/>
3. 'What is the maximum number of columns / variables in statistics?', IBM Support (16 April 2020), accessed on 14 July 2023, viewable at: <https://www.ibm.com/support/pages/what-maximum-number-columnsvariables-statistics/>
4. Sajjad Hossan, 'The history and rise of Flipkart: Largest ecommerce company in India', Business Inspection (27 May 2023), accessed on 14 July 2023, viewable at: <https://businessinspection.com.bd/history-andrise-of-flipkart/>
5. Klaus Schwab, 'The fourth Industrial Revolution: what it means, how to respond', WEF (14 January 2016), accessed on 14 July 2023, viewable at: <https://www.weforum.org/agenda/2016/01/the-fourth-industrialrevolution-what-it-means-and-how-to-respond/>