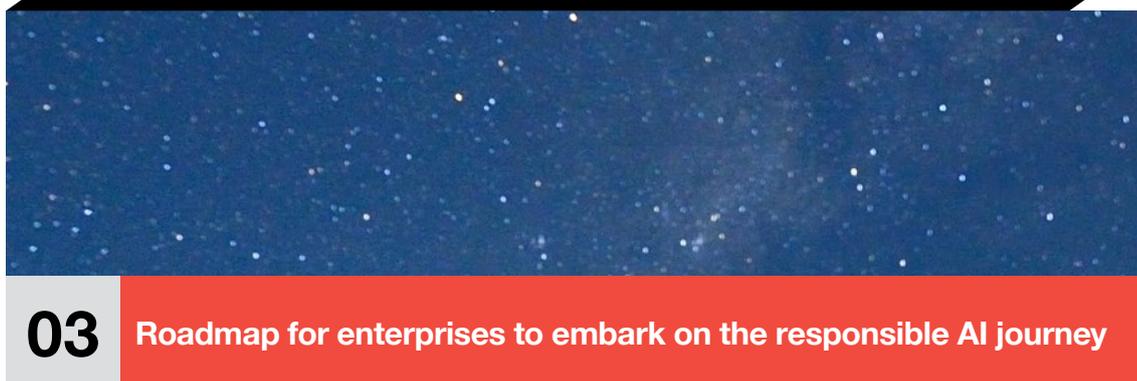


With AI's great power comes great responsibility

October 2019



Table of contents



Foreword



Sudipta Ghosh
Partner and Leader
Data and Analytics
PwC India



Anand Rao
Partner, Global AI Lead
Data and Analytics
PwC US

Artificial Intelligence can be defined as a collection of technologies which are capable of sensing, thinking and acting like rational human beings. In recent years, every individual, organisation and government has seen numerous instances of AI enter various facets of their lives – the most common applications being chatbots, facial recognition, image classification, aggregators, recommendation engines and targeted marketing. AI has the potential to solve complex problems effectively at scale. However, great possibilities carry great risks. Badly designed AI can cause more harm than good. As such, we need to be absolutely sure that AI systems behave in expected ways, can be trusted, and are easily understood by business teams.

AI then is a wellspring of both optimism and trepidation. Business pioneers often ask: What effect will AI have on my business, and is our plan of action vulnerable to disruption? Also, as these pioneers hope to profit from AI openings, they query: Where would it be a good idea for us to invest, and what sort of capacities would empower us to perform better? Cutting across these concerns is the need to build AI systems that are responsible and understandable, and thus ensure customer trust.

Different stakeholders have different queries regarding an organisation's implementation of AI – from how it's created to how it's governed. Our Responsible AI Toolkit is a suite of frameworks, processes and tools intended to equip one to tackle AI in a moral and capable way – from strategy through execution. Responsible AI can build trust at the time of starting or scaling AI and at the same time address ensuing challenges and risks.

This paper based on a PwC survey helps understand the mindset of decision makers of organisations in India towards AI from the perspectives of adoption, performance and risks. It elaborates on the need for a comprehensive responsible AI framework and toolkit for its widespread adoption. It also underlines how the potential of AI can be unlocked if a structured approach is taken towards addressing the associated risks.

Driving into a responsible AI future

01

From automation to augmentation and beyond, Artificial Intelligence (AI) is changing how business gets done. Companies are using AI to automate tasks that humans used to do, such as detecting frauds, vetting résumés and loan applications, thereby freeing people up for more complex and high-end work. Doctors are using AI to diagnose certain health conditions faster and more accurately. Also, chatbots are being used in place of customer service representatives to address simple queries from customers.¹

According to PwC's 22nd Annual Global CEO Survey, 69% of CEOs in India strongly believe that "AI will have a larger impact on the world than the Internet revolution." Further, 42% of global CEOs foresee that AI will significantly change the way they do business in the next five years, while in India, 86% of CEOs echo this view.² CEOs from Indian organisations strongly believe that AI can serve as a catalyst for business as well as larger social transformation.

Many organisations in India have started building their Analytics Centres of Excellence to lay a strong foundation for their organisations. And, around 50% of Indian CEOs plan to introduce AI initiatives over the next three years.

AI evidently has the potential to add significant value to a wide range of sectors. However, in order to scale AI initiatives, organisations will have to ensure these solutions fulfil the following conditions:

- Ethically sound and full compliance with regulations in all respects
- Underpinned by a robust foundation for end-to-end governance
- Supported by strong performance pillars addressing bias and fairness, interpretability and explainability, robustness and security³

AI not only processes data but also learns and makes decisions based on it. On careful implementation, this technology can bring about positive changes through its ability to add value.⁴ However, AI has the potential to amplify and perpetuate underlying risks (Exhibit 1), societal discrimination and inequalities. The AI-related risks may either lie in the data or may be inherent in the design of the AI model.

Exhibit 1 - Risks related to AI are spread across multiple dimensions



1 A practical guide to Responsible Artificial Intelligence (AI), PwC (2019)

2 PwC's 22nd Annual Global CEO Survey

3 A practical guide to Responsible Artificial Intelligence (AI), PwC (2019)

4 This is what the world's CEOs really think of AI, World Economic Forum (2018)



Economic potential hinges on trustworthiness

AI tends to be viewed with suspicion as its decision-making process is opaque. This is known as the 'black box problem'.⁵ Businesses that use AI to assist humans need to understand why their system is providing a particular output so that the final decision taken is a transparent and informed one.⁶ For instance, medical systems that use AI to diagnose patients would need to explain their decisions. Without transparency and trust, the economic benefits promised by AI cannot be fully realised.⁷



Securing the purpose and performance of AI

When an AI system suddenly starts behaving in a way that is out of sync with its purpose as a result of manipulation or compromise by other humans or AI systems, the risks are not only financial but also reputational, with far-reaching consequences. Hence, AI-enabled systems need to be at the forefront of security and need to incorporate guidelines that ensure protection from tampering.

Consider the fact that multiple criminal justice systems and law enforcement agencies are using AI to solve crimes and deliver justice more effectively.^{8,9} An AI system that is continuously fed biased data on arrests will, over time, shift its recommendations and overestimate or underestimate the propensity of different groups to commit a crime. This can happen if data leakage or data shifting occurs during training. Organisations would therefore need to periodically ensure that their AI models are performing as expected.¹⁰



Societal and ethical risks

AI has the potential to remove the subjective biases that exist in society, so it makes sense to use this technology in areas where objective treatment of an issue is required. However, since AI models and datasets are developed by humans, there are significant chances of conscious and subconscious biases creeping into AI systems. A simple online search provides multiple examples of how AI can reinforce biases in recruitment. For example, in financial services, AI advisors may discriminate against certain sections of population while deriving the credit score¹¹. Similarly, AI-based facial recognition tools have been found to work with significantly reduced accuracy when trying to identify people of colour¹². To mitigate such risks, AI needs to be made fair in order to preserve the diverse realities and needs of different end users across the world.¹³



Controlling AI

In a recent case, fraudsters used AI to mimic a CEO's voice and siphoned off over EUR 200,000¹⁴. This is just one example of how organisational governance methods need to evolve in order to identify and protect businesses effectively. In another example, researchers from the University of Melbourne used AI to re-identify individuals from an anonymised health dataset.¹⁵ Organisations need to ensure that their AI-enabled systems cannot be misused by others – be they humans or other AI systems – to cause irreversible damage.¹⁶

5 The Artificial Intelligence Black Box and The Failure of Intent and Causation, Harvard Journal of Law & Technology, Harvard Journal of Law & Technology (2018)

6 Sizing the prize, PwC (2017)

7 2019 AI Predictions, PwC (2019)

8 5 Ways AI Is Used by Lawmakers for Crime Prevention in India, Analytics India Magazine (2019)

9 Should We Be Afraid of AI in the Criminal-Justice System?, The Atlantic (2019)

10 India plans to adopt China-style facial recognition in policing, despite having no data privacy laws, Bloomberg (2019)

11 Credit Scores Could Soon Get Even Creepier and More Biased

12 Why facial recognition software has trouble recognizing people of color

13 Some AI just shouldn't exist, Vox (2019)

14 Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case, Wall Street Journal (2019)

15 Health Data in An Open World, arXiv (2017)

16 State of AI: Artificial Intelligence, the Military and Increasingly Autonomous Weapons, Future of Life Institute (2019)

How responsible is Artificial Intelligence? India highlights

02

To gain an understanding of the outlook towards AI in India, PwC conducted a comprehensive global survey between May and September 2019 eliciting nearly 1,000 responses from CXOs and decision makers from India and other regions. The respondents spanned across industries such as technology, media and telecom, financial services, professional services, health, industrial products, consumer markets, government, and utilities. The survey also included respondents across various business functions such as IT, finance, operations, marketing, customer service, sales, human resources, legal, risk and compliance.

The aim of the study was to understand the perception of the following aspects of AI Implementation:

- Adoption: Are AI solutions being widely deployed and used at scale?
- Performance: Are AI solutions working reliably in a desired manner?
- Risk: Are there any ethical, legal or accountability challenges for malfunctioning AI?

Adoption



A robust explainable approach for the development of AI-based solutions is an important driver for its adoption.



Exhibit 2 - Implementation and approach to AI

	Global	India
 Have implemented AI in some form	65%	62%
 Have no formal approach to identify AI risks	36%	53%
 Have no tools to assess security flaws in their AI systems	16%	29%

It is heartening to note that India (62%) is not very far behind global (65%) in terms of implementation of AI. However, the worrying part is that Indian respondents (53%) significantly outnumber their global counterparts (36%) in admitting that they have no formal approach to identify AI risks. In addition, 29% of the respondents from India feel that they have no tools to access security flaws in their AI systems (Exhibit 2). This suggests that the enthusiasm to implement AI projects is very likely to run into headwinds unless organisations adopt a robust framework for using AI responsibly.

In order to understand the theme of adoption a little better, we analysed the primary factors that inhibit businesses from adopting and applying AI systems. While globally, respondents said budget constraints (40%) and lack of availability (39%) of appropriate skills prevented them from adopting AI (Exhibit 3), the respondents from Indian organisations cited lack of understanding of how AI arrives at a decision (58%), and insufficient data (45%) as the primary roadblocks (Exhibit 4). This shows that in addition to the traditional challenges of financial constraints (budgets), skills and data availability for adoption of AI at scale, there is a clear need for 'explainability' of AI solutions. More importantly, explainable AI is considered even more critical for Indian organisations than their global counterparts.

Top 5 inhibitors to adoption of AI solutions

Exhibit 3 - Top 5 AI Adoption Inhibitors - Global

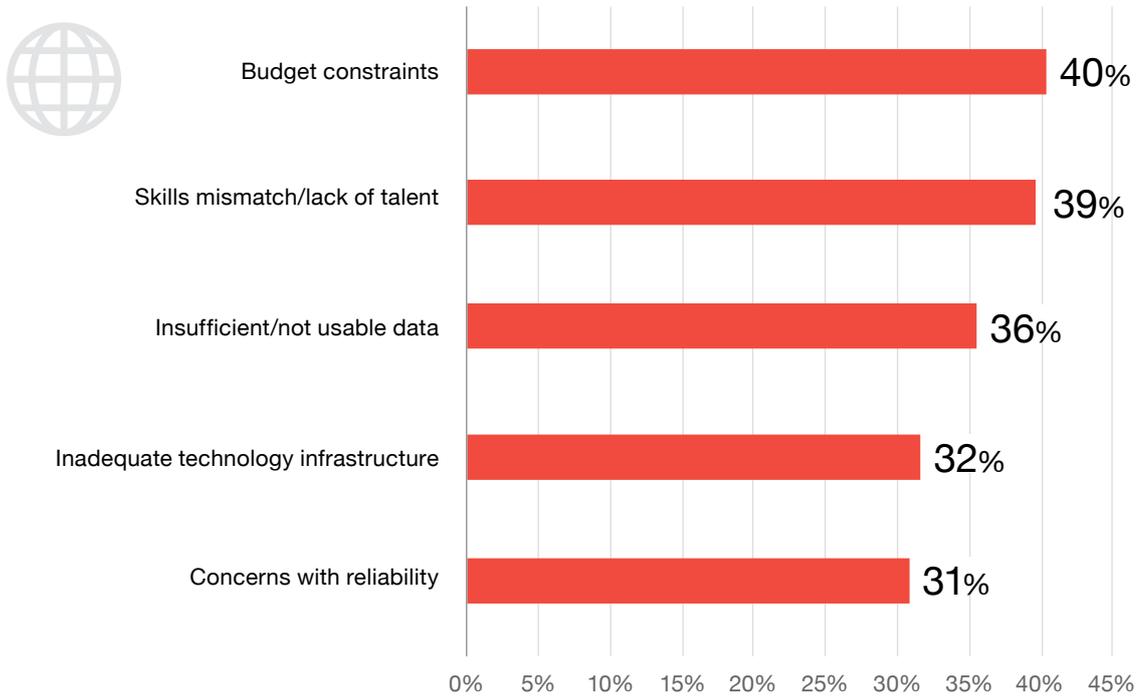
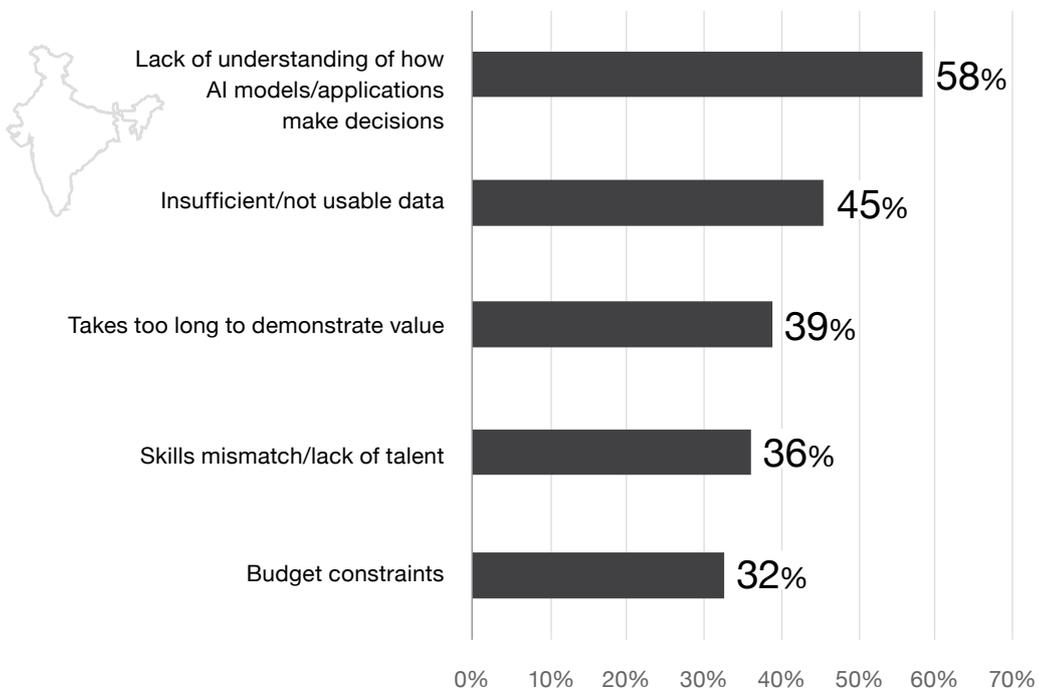


Exhibit 4 - Top 5 AI Adoption Inhibitors - India



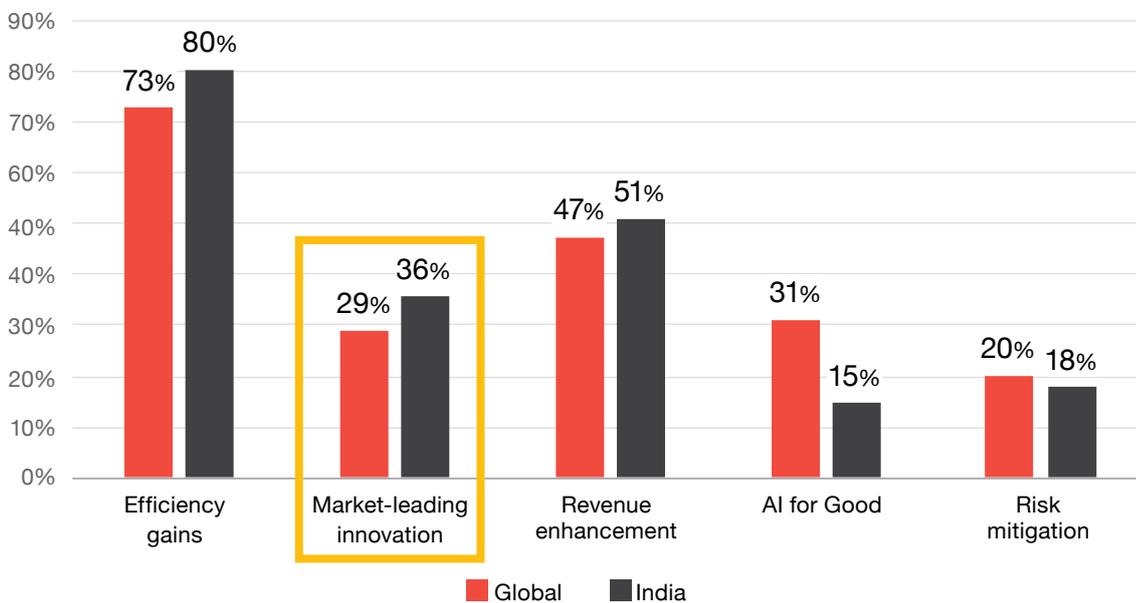


Enterprises in India strongly believe that AI solutions can help them to become market leaders in innovations and significantly improve organisational performance. This highlights the importance and potential value of AI to Indian enterprises.



Compared to global (40%), fewer Indian executives (32%) expressed concerns about budgets to implement AI, signifying a greater willingness to invest in AI. To understand this better, we analysed the primary driver for them to invest in AI. Efficiency gains emerged as the biggest driver for investing in AI, with nearly 73% of global respondents citing it as the primary factor for investing in AI. For India, nearly 80% of respondents cited efficiency gains as the biggest driver (Exhibit 5). Surprisingly, when compared to global respondents, a significantly greater number of respondents from India cited market-leading innovation as their primary driver to invest in AI. This could be a sign that organisations in India are investing in AI to create a competitive advantage rather than just using it for revenue enhancement or efficiency improvement.

Exhibit 5 - Priorities for AI Investment

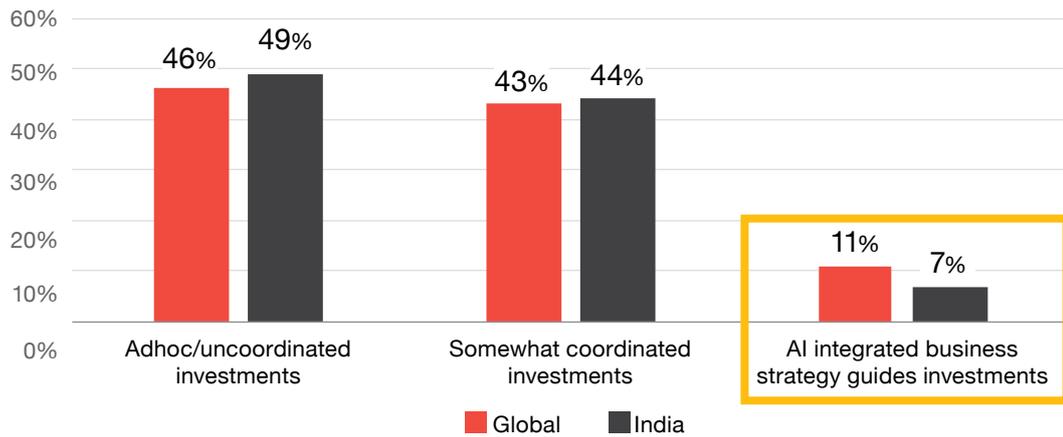


Very few organisations were found to be incorporating AI as an integral part of their business strategy, thus preventing them from realising the true potential of AI.



When we compare the AI investment strategy in terms of well-planned initiatives, 11% of the global respondents said they have a well-defined strategy for making AI-related investments, whereas in the case of Indian respondents, only 7% make AI investments as part of their business strategy (Exhibit 6).

Exhibit 6 - How are investment decisions in technology, such as AI, made?



Performance

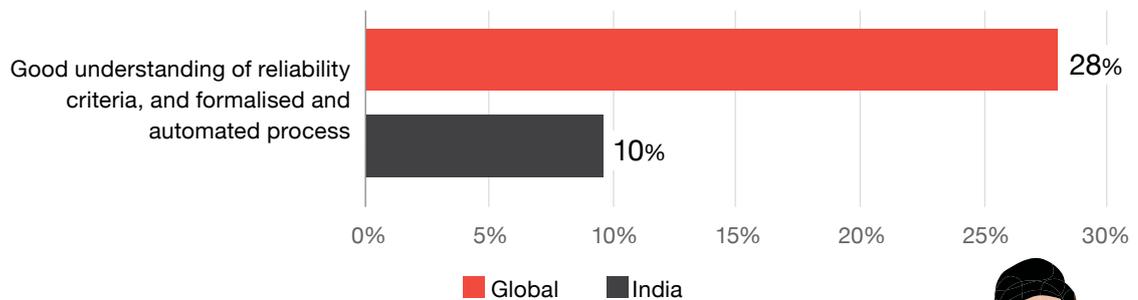


Organisations are still not confident about implementing industrial AI solutions which can be embedded in the workflow of applications and help in the day-to-day decision-making process.



An overwhelming majority of decision makers globally as well as from India confessed that they may not have robust tools or processes for ensuring reliability of their AI solutions. Interestingly, only 10% of Indian respondents were confident about the reliability of their AI applications (Exhibit 7).

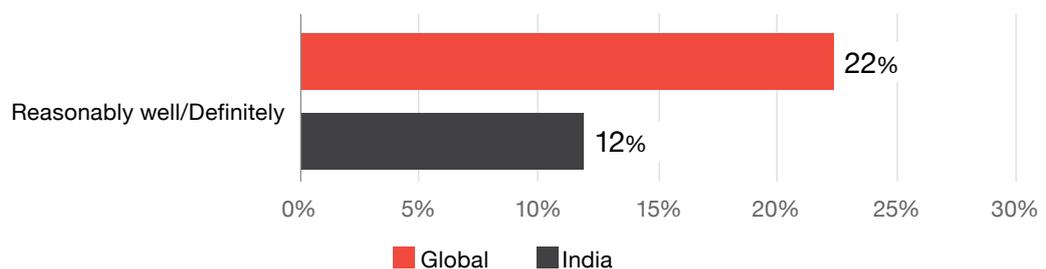
Exhibit 7 - How is your AI monitored for reliable performance?



As compared to global, Indian organisations are far behind in their understanding of the reliability criteria for deploying AI solutions at scale.



Exhibit 8 - Could you reliably deploy AI solutions at scale?



Risk



While the primary concern for both Global and India respondents is data privacy, for India respondents accountability is very critical.

Data privacy is a major ethical concern both globally as well as in India (Exhibit 9). Given the increased regulations around personally identifiable information (PII) and the increased focus on generating insights from ‘customer’ and ‘employee’ data, some of these challenges are here to stay. In order to ensure that legitimate AI initiatives are not derailed during inception because of this risk perception, organisations would need to ensure that such risks are identified at an early stage and communicated to relevant stakeholders (both internal and external) with a robust mitigation approach.

Exhibit 9 - Key ethical considerations

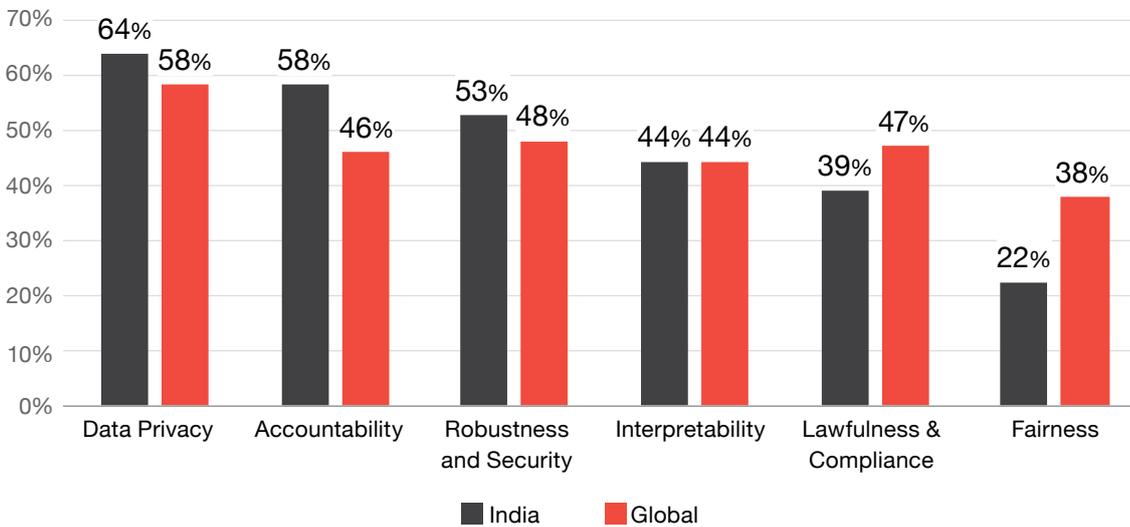
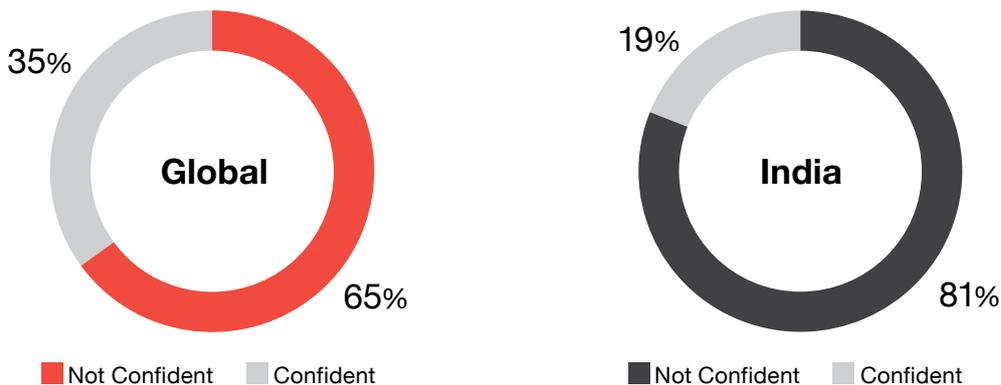


Exhibit 10 - How confident are you to detect and shut down a malfunctioning AI system?



While 65% of global organisations are not confident of detecting a malfunctioning AI system, for India the figure goes up to 81%.



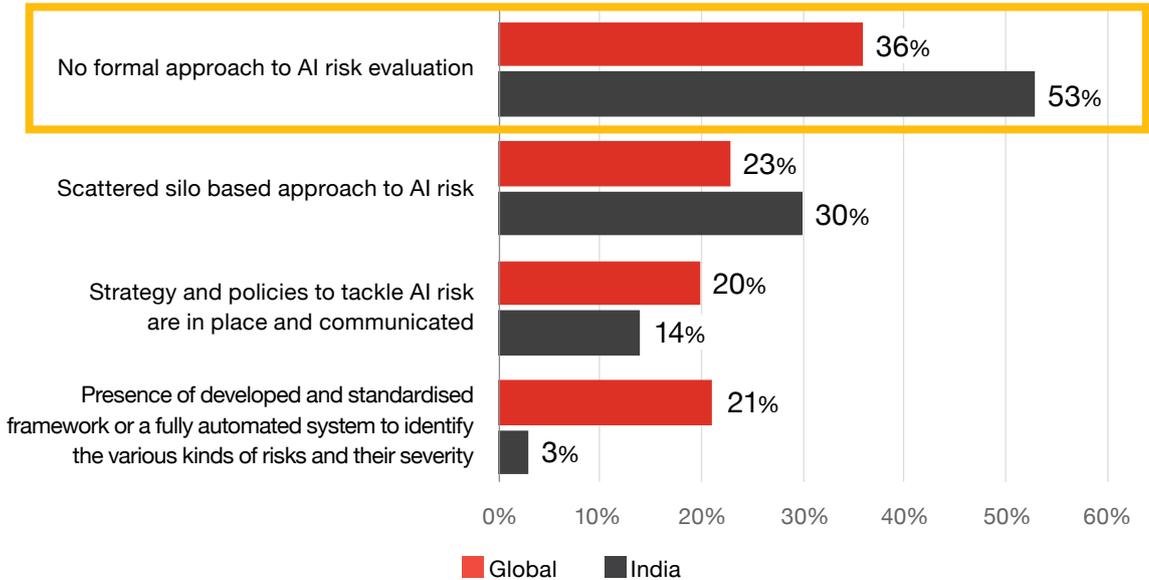
With 81% of the respondents lacking confidence with respect to detecting a malfunctioning AI system in India (Exhibit 10), the risks and threats posed by AI require a robust framework for identification.



An overwhelming majority of respondents from Indian organisations (53%) do not have a formal approach to AI risk evaluation. This does not bode well for addressing the unintended consequences of AI solutions.



Exhibit 11 - How are risks identified?



Globally, 38% of the respondents have a good understanding of bias criteria and have tools for accessing ongoing biases, whereas only 29% Indian respondents have this capability (Exhibit 12). Further, 55% of the global respondents are confident that they have the right tools and methodologies to meet the compliance requirements, while in India only 33% believe they have the right tools to ensure such requirements (Exhibit 13).

Exhibit 12 - How is your AI assessed for bias?

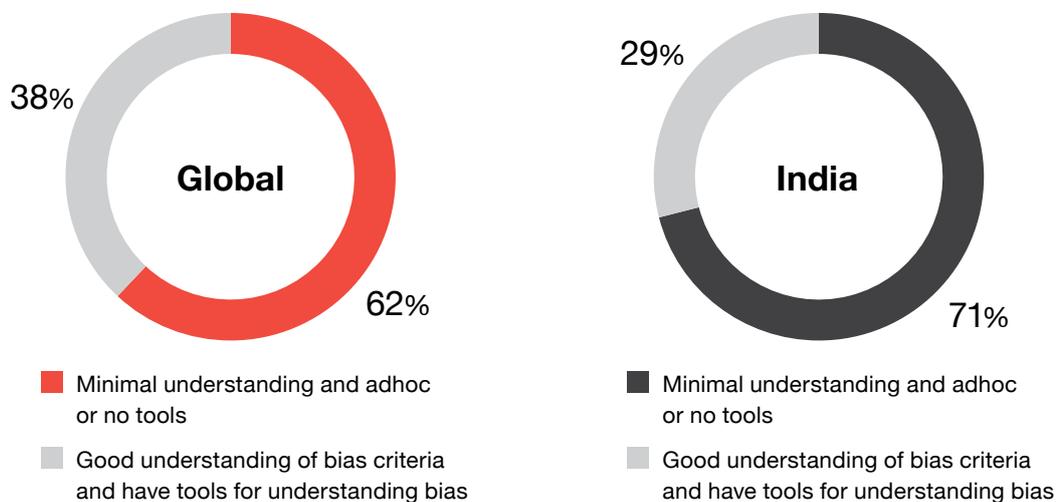
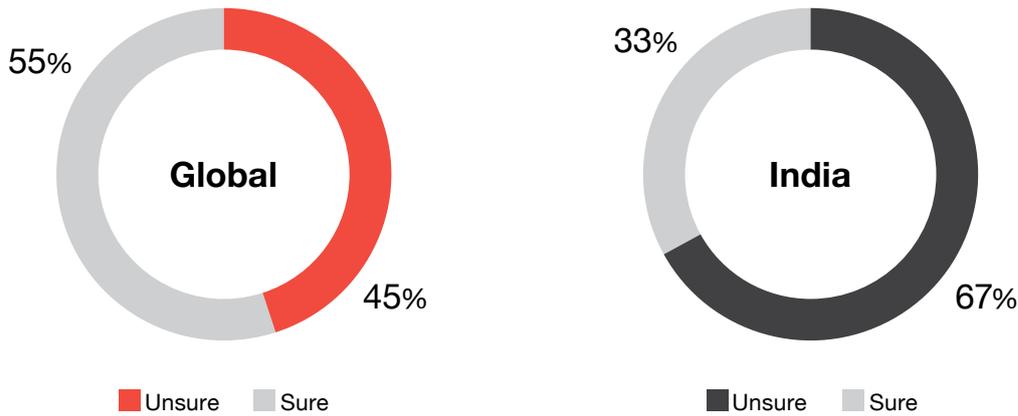


Exhibit 13 - Currently, how would you best describe your organisation's ability to demonstrate regulatory compliance with respect to AI?

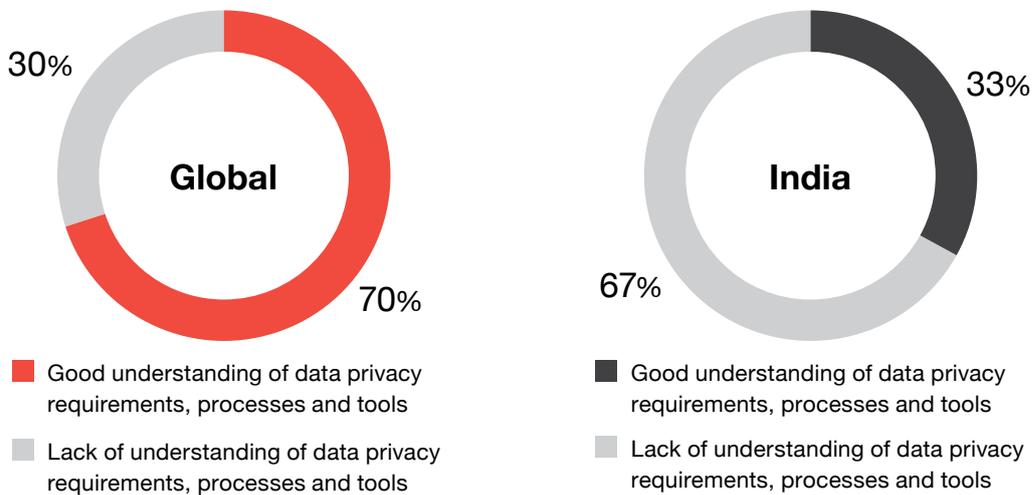


67% of the organisations we surveyed in India are unsure of the regulatory compliance with respect to AI and have minimal understanding of tools needed for maintenance of data integrity.



More than 70% of the global respondents have an understanding of data privacy requirements, processes and tools, as compared to only 33% of India respondents (Exhibit 14).

Exhibit 14 - How well do you understand data privacy requirements, processes and tools?



The concerns around the assessment of various risks related to usage of AI in India have a significant impact on the scalable deployment of AI solutions. A robust responsible AI (RAI) framework can help organisations assess potential threats and safeguard risks.

Overall, the survey clearly highlights that RAI is an important driver for building trust in AI solutions, leading to higher adoption across all sectors and business functions.

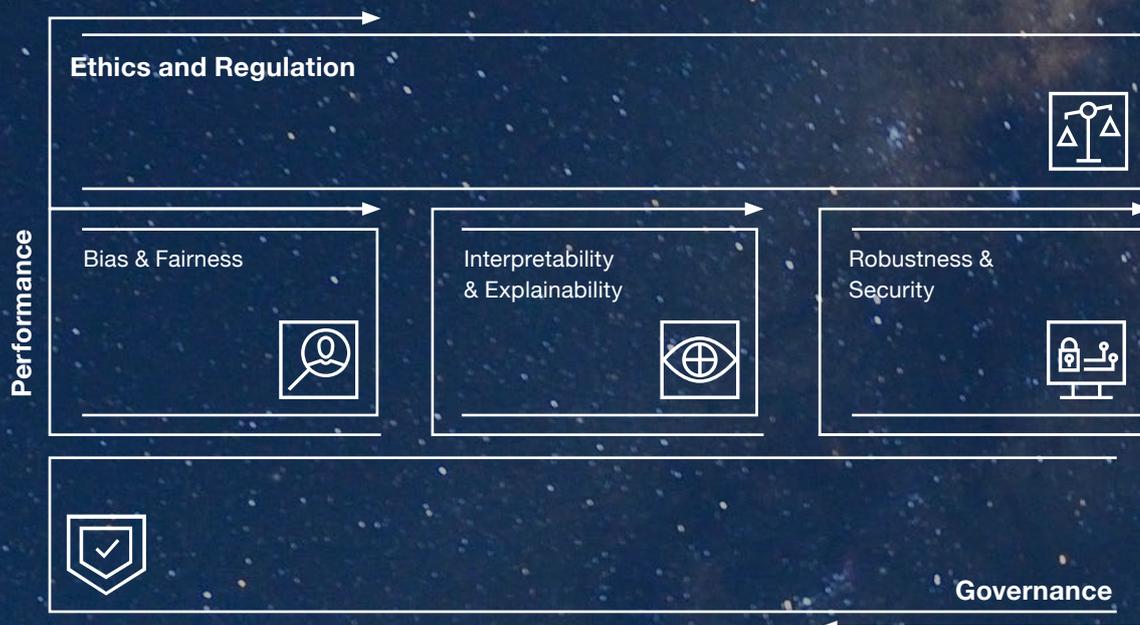
Indian enterprises need a robust RAI framework to develop and implement AI strategies that can help them create an ecosystem that is effective, makes the business stronger and adds value to society. Every effort towards AI needs to be responsible, or we risk perpetuating the very problems we are trying to solve.

Roadmap for enterprises to embark on the responsible AI journey

03

PwC has developed a comprehensive framework to address various issues concerning the implementation of AI solutions (Exhibit 15).

Exhibit 15 - RAI framework



Ethics and regulation

India ranked third in the Asia Pacific region on AI tech readiness.¹⁷ The country also has the highest data consumption rate per smartphone and this figure is expected to double by 2024.¹⁸ As we move from the information era to the AI era, the ethical and regulatory aspects of developing solutions using AI will assume greater importance. Translating ethical principles into concrete actions is difficult for businesses. But, by factoring in the ethical context and relevant principles when designing and operationalising AI models, organisations may be able to deal with the ethical implications of these models. Organisations also need to monitor the regulatory environment in which they operate and understand how emerging regulations will shape future business practices for robust mitigation of ethical risks.¹⁹

¹⁷ India third in APAC in AI tech readiness, Economic Times (2019)

¹⁸ India's data usage per smartphone highest in world at 9.8GB/month: Report, Livemint (2019)

¹⁹ A practical guide to Responsible Artificial Intelligence (AI), PwC (2019)



Bias and fairness

One of the major risks associated with the AI system is bias. Some recent cases where machines have provided prejudiced outcomes are: (1) racial bias in criminal justice systems and (2) gender discrimination in hiring talent. A bias arises when there is conscious or unconscious discrimination against a defined group, and it can creep into an AI system as a result of skewed data or an algorithm that does not account for skewed data. For example, an AI system that reviews job applications by learning from an organisation's historical data could end up promoting a particular gender or race if that group was well presented in the organisation's hiring in the past.

Humans interpret bias through the subjective frame of fairness which is a social belief with powerful local nuances, and many different and conflicting definitions. There are at least 20 mathematical definitions of fairness, and when one is selected, some aspects of the others are violated. In any case, it is not possible to be fair to all stakeholders with every decision.²⁰ However, it is possible to tune AI systems to alleviate bias and facilitate decisions that are as fair as possible, adhere to a firm's corporate code of ethics, and also follow anti-discriminatory regulations.

Exhibit 16



In order to be fair, businesses need to determine their degree of comfort in the selections they make, and they need to balance these against the associated trade-offs which might be negative for some stakeholders.



Interpretability and explainability

Lack of interpretability and explainability in AI-powered system/application exposes an organisation to financial, operational and reputational risks. Interpretability helps different stakeholders including regulators, auditors, end users and data scientists to discern the cause and effect of a decision. Explainability helps to understand why the AI system does what it does.

If we consider a car as an analogy, knowing what functions each component performs is interpretability and going into the mechanics is explainability. To build RAI systems, people must know the unknowns of the models as well as the data which is used for training, the reasons behind each decision, and provide rational explanations to the stakeholders.

For example, if one were to invest in an AI application that helps to evaluate loan applications and better gauge the credit worthiness of applicants, certain parameters need to be considered. If the bank can ensure appropriate levels of explainability and transparency, it can help build confidence in the AI system among loan officers and applicants alike.

20 Fairness Definitions Explained, 2018 ACM/IEEE International Workshop on Software Fairness



Robustness and security

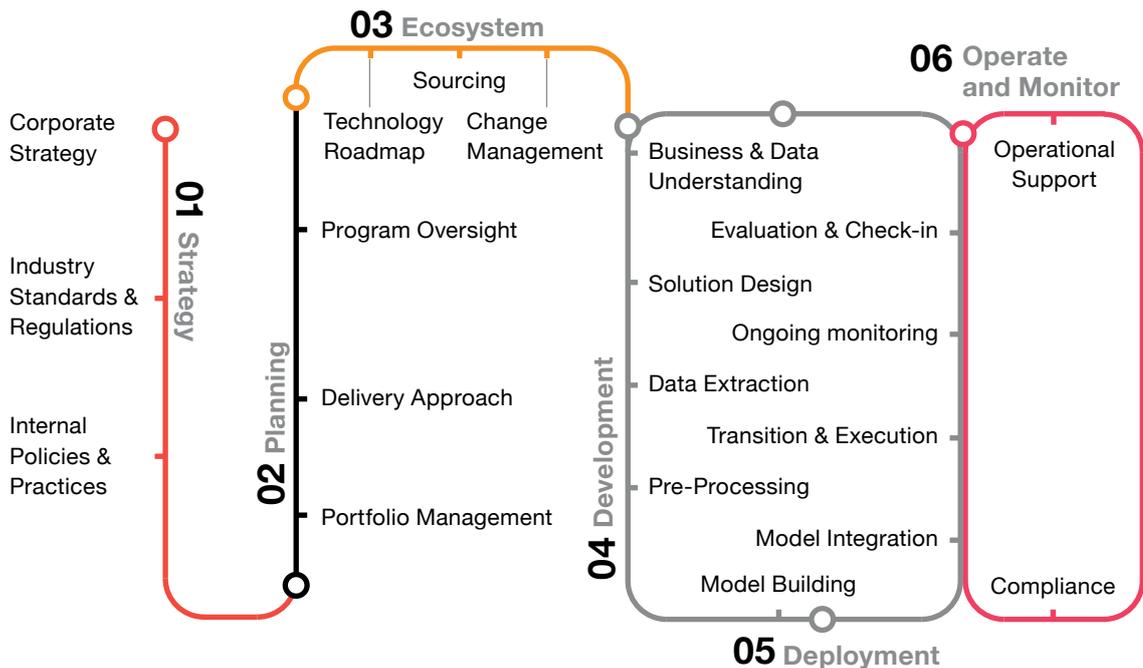
Like everything else, AI comes with a set of concerns and limitations. While it brings in a great deal of predictability, there are issues related to stability, robustness and security. A good and robust RAI toolkit must produce consistent outputs for the same type of inputs and under dynamic conditions. At the same time, the outputs need to be reasonable and justifiable. For example, an algorithm which assigns different credit scores to two different persons who apply for a loan at different times although they have the same level of income is not robust. Concerns related to data privacy and confidentiality can be resolved with the use of synthetic data derived from client data.

Again, with machine learning (ML) and AI activities becoming increasingly common, there is a higher risk of manipulation. Algorithms can be written that train the AI model into providing wrong or unethical results. For example, opportunists may exploit vulnerabilities in the system for unfair benefits. These days, ML models are also considered as an intellectual property due to the efforts required to train them. So, if it is possible to replicate such a model using an open Application Programming Interface (API), the owner may face a great deal of embarrassment and difficulties. The ease or difficulty of developing an ML model using open API can be determined using a security comparison software. While tools to tackle these issues are available, ultimately, the implementation must be a combination of business-directed decisions and technical know-how.²¹



AI governance

Exhibit 17 - PwC's AI governance framework



Benefits of AI may be realised when an appropriate governance framework and dimensions are in place, and humans and machines can collaborate effectively. This will allow humans to turn their attention to more strategic pursuits and use their knowledge and experience to govern the workings of AI systems. End-to-end enterprise governance is the foundation for RAI in PwC. PwC's governance framework helps cover all dimensions of RAI and build an ecosystem that creates trust in the minds of stakeholders and aligns external or internal policies and processes.

AI governance covers the entire AI lifecycle from strategy and planning to monitoring of deployed solutions across the organisation. It also takes into account existing capabilities and the vendor ecosystem, along with the unique model development process, and model monitoring and compliance.²²

²¹ Robust and secure artificial intelligence: A primer for the C-suite, PwC

²² A practical guide to Responsible Artificial Intelligence (AI), PwC (2019)

PwC's Responsible AI Toolkit

We, at PwC, recognise the importance of responsibly unleashing the potential of AI. AI-based innovation should add value to business and the community at large, with informed governance to mitigate and manage limitations, vulnerabilities and the potential pitfalls of AI. The responsible AI framework we have developed provides a practical solution to ensure effective stewardship of the outcomes.

This technology-enabled toolkit consists of a plethora of flexible and scalable capabilities curated to enable and support the development and assessment of high-quality, explainable, transparent and ethical AI applications. The toolkit's diagnostics are designed to help different stakeholders generate trust and inspire confidence. It provides a set of assets curated to accelerate the evaluation of data, models and their trade-offs, taking into account the associated risks and relevance. The frameworks, leading practices, and practice aids can help organisations define their individual requirements for Responsible AI governance.

We believe that integration of the RAI toolkit with AI-related initiatives will enable businesses to accelerate innovation and realise their vision.





Building a firm foundation for the future

The speed at which AI will develop is expected to be faster than Moore's law.²³ In such a scenario, we need to ensure that AI acts in the interests of society at each stage of development. RAI helps mitigate many of the risks – foreseen or unforeseen – associated with AI:

- RAI ensures that AI-enabled systems being used by an organisation conform to the **governance** processes of that organisation and other **regulations** as mandated by law.
- It reduces the risk of **deterioration or shift** in the AI application's **performance** over time.
- The RAI approach ensures that the **best practices** for development and operation are used and that AI applications are effective.
- RAI ensures that AI applications are **ethical, not biased, and resonate with the company's vision and values**.
- RAI improves the accountability and quality of AI models.
- RAI adds to the **explainability** of modelling processes, thus improving the understanding of business decisions by executives.

Once a firm foundation has been set, organisations can take advantage of the potential of AI. India has a natural advantage to adopt quickly, as it carries no legacy of the past. RAI helps formulate effective operating models for AI that minimise risks and maximise rewards. It accelerates innovation and the potential to create value – which could be at stake if AI is implemented in the wrong way.²⁴ What is therefore important is a structured approach to developing AI-driven solutions such that they are not only effective but also follow a predictive trajectory without unintended consequences. Only then can AI's great power be unleashed responsibly at scale.

²³ Why AI progress is faster than Moore's Law - the age of the algorithm, Hackernoon (2018)

²⁴ Sizing the prize, PwC

About PwC

At PwC, our purpose is to build trust in society and solve important problems. We're a network of firms in 155 countries with over 284,000 people who are committed to delivering quality in assurance, advisory and tax services. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

Find out more about PwC India and tell us what matters to you by visiting us at www.pwc.in

To learn more, take PwC's free Responsible AI Diagnostic.



Or check our Responsible AI website.

www.pwc.com/rai

Contact us

Arnab Basu

Joint Technology Consulting Leader, PwC India
arnab.basu@pwc.com

Pawan Kumar S

Joint Technology Consulting Leader, PwC India
pawan.k.s@pwc.com

Sudipta Ghosh

Data and Analytics Leader, PwC India
sudipta.ghosh@pwc.com

Raman Bhushan

Advanced Analytics Leader, PwC India
raman.bhushan@pwc.com

Authors

Sudipta Ghosh, Anand Rao, Raman Bhushan, Saurabh Bansal, Rudrani Das, Abhishek Ranjan, Zeeshan Ashraf, Mohammad Huzaifa Sabir and Shaoni Gupta.

Contributors

Ilana Golbin, Jasha Bowe, Jason Warke and Alex Zong

Editorial support

Vishnupriya Sengupta and Dion Dsouza

Design

Faaiz Gul and Harshpal Singh

pwc.in

Data Classification: DC0 (Public)

In this document, PwC refers to PricewaterhouseCoopers Private Limited (a limited liability company in India having Corporate Identity Number or CIN : U74140WB1983PTC036093), which is a member firm of PricewaterhouseCoopers International Limited (PwCIL), each member firm of which is a separate legal entity.

This document does not constitute professional advice. The information in this document has been obtained or derived from sources believed by PricewaterhouseCoopers Private Limited (PwCPL) to be reliable but PwCPL does not represent that this information is accurate or complete. Any opinions or estimates contained in this document represent the judgment of PwCPL at this time and are subject to change without notice. Readers of this publication are advised to seek their own professional advice before taking any course of action or decision, for which they are entirely responsible, based on the contents of this publication. PwCPL neither accepts or assumes any responsibility or liability to any reader of this publication in respect of the information contained within it or for any decisions readers may take or decide not to or fail to take.

© 2021 PricewaterhouseCoopers Private Limited. All rights reserved.

HS/September 2019 - M&C739