Delivering Value From Your Data With AI

What exactly is AI and how can it benefit your organisation?

THE AI OPPORTUNITY

Businesses, research institutions, and government agencies have been exponentially increasing the amount of data that they collect. IDC1 predicts that by 2025 the world may be generating 163 zettabytes or 163 trillion gigabytes of data largely through mobile devices, distributed computing, social media, and IoT. It is therefore not surprising that organisations of all types are looking to implement artificial intelligence (AI) to deliver value from their data.

According to research from Accenture[2] , by 2035 AI may increase labour productivity by 40% and corporate profits by 38%. With so many industries experiencing slow or stalled growth, it is no surprise business leaders are looking to infuse their organisations with AI as a way of lowering costs and increasing productivity. Some are even likening AI as a new industrial revolution heralding an age of rapid innovation and growth.

Government agencies, under pressure to deliver more with less, are also looking to AI to reduce costs, prevent fraud, and improve public services. Research institutions have already begun exploring and utilising AI to develop new insights from the proliferation of scientific data.

Though the economic and utilitarian arguments in favour of AI are very compelling, decision makers and stakeholders are often unsure how to best implement this new technology. Embarking on a high-level strategy to introduce AI is insufficient as there is no one size fits all solution. Every application, market, and dataset is different. It is therefore important to understand the different options that are available in order to properly plan for a strategy that aims to deliver value through AI.

[1] IDC/Seagate, DATA AGE 2025, https://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf
[2] Accenture, HOW AI BOOSTS INDUSTRY PROFITS AND INNOVATION, https://www.accenture.com/t20171005T065812Z__w__/us-en/_acnmedia/Accenture/next-gen-5/insight-ai-industry-growth/pdf/Accenture-AI-Industry-Growth-Full-Report.pdfla=en?la=en

XENON®
High Performance Computing

NARROWING IN ON AI

Science fiction often depicts robots and computers capable of performing all sorts of human tasks. This type of AI, researched since the 1950s, is referred to as strong or general AI. As of today, no machine has yet to fully pass Alan Turing's test for general AI in which a human observer fails to distinguish between a human and a computer through a text conversation. Though it may become a reality one day, general AI is currently not a viable option for today's organisations.

Recent advances have instead centred around dedicated applications such as image classification, natural language processing, and data analysis. Instead of applying a technology across entire organisations, a more focused weak or narrow AI is addressing specific tasks within business functions such as marketing, finance, and customer support or to computational intensive research such as genetics and climate modelling. Rather than completely replacing humans, narrow AI is augmenting existing roles.
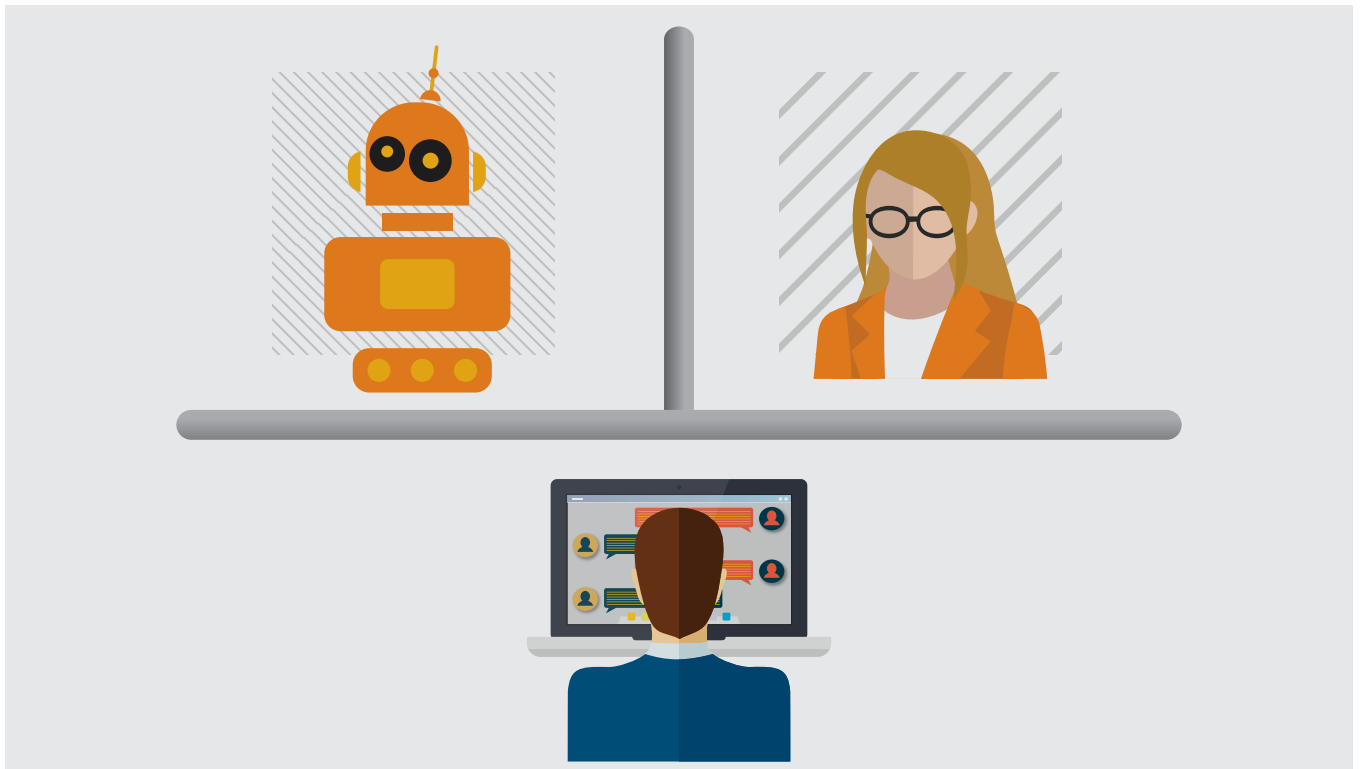
MACHINES THAT LEARN

In the 1950s, computing pioneer Arthur Samuel applied an alpha-beta pruning algorithm to an IBM vacuum tube machine so that it could teach itself to play checkers (draughts) by "thinking" ahead to possible moves and eliminating those that would favour its human opponent. This was the beginning of a field of AI referred to as machine learning.

An early and now common machine learning application is spam filtering which relies on a technique known as a naive Bayes classifier to identify keywords that have been flagged as junk. Other widely used machine learning techniques include support vector machines that

analyse data along hyperplanes, matrix factorisation for recommendation engines and random forest algorithms for predictive analytics.

Off-the-shelf software can include machine learning algorithms for analysing data sets and predicting potential outcomes. Predictive analytics is commonly used for credit rating, reducing customer churn and identifying potential health risks. Off the-shelf-software can be a good option for organisations looking to adopt AI for data science, without the need to develop their own algorithms.

DEEP NEURAL NETWORKS

Most of the recent AI advances have taken place within a branch of machine learning known as deep learning. Instead of linear algorithms, deep learning relies on artificial neural networks consisting of multiple interlinked nodes, much like a biological neural network found in the brain.

Deep neural networks start with an input layer with a set of nodes for taking in specific data points such as fields within a table or pixels from an image. These inputs are then processed through a series of hidden layers and nodes before reaching the output layer. The hidden layers and nodes are reshaped over time to improve the accuracy between inputs and outputs through supervised, unsupervised, and semi-supervised training techniques.
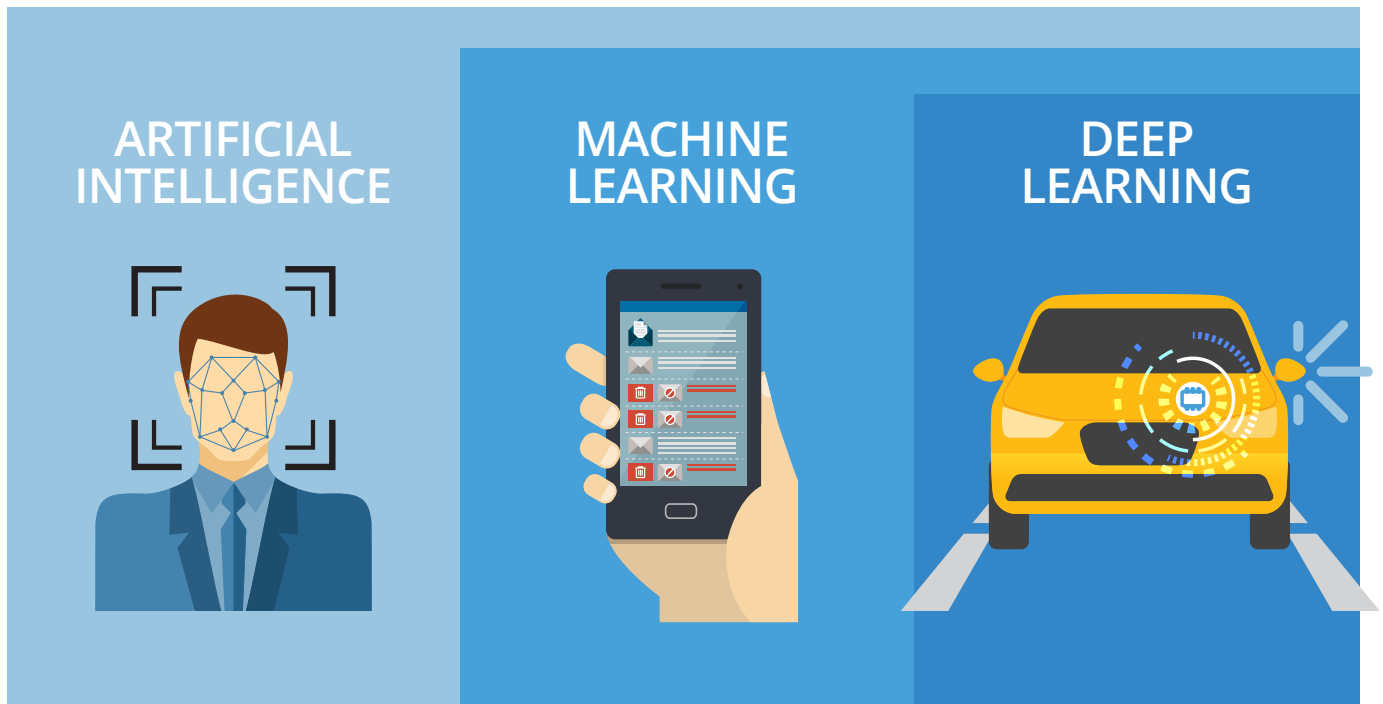
MORE THAN ONE WAY TO TRAIN A MACHINE

Unlike human brains which are able to perform a multitude of tasks, artificial neural networks are trained to perform a specific task faster and more accurately than humans. Training a neural network for deep learning requires processing large data sets which may consist of text, numbers, photographs, or sound. Neural network training can either be supervised with the help of humans, unsupervised, or a mixture of both.

With supervised training, data is labelled to indicate the desired answer. For example, if the goal is to recognise photos of cats and dogs then the neural network is fed with thousands of cat and dog images, each labelled accordingly. Every time the data set has been processed, known as an epoch, the neural network gets reshaped. By applying a technique known as backwards propagation, the weight of individual nodes and layers are changed depending on the role they played in reaching the correct answer. After millions of epochs, the neural network is eventually reshaped to the point that the response becomes highly accurate even when presented with new data.

In some cases, it is possible to transfer a trained neural network from one application to another, especially when dealing with similar outputs such as dogs or cats and racoons or possums, or for adding additional outputs such as dogs, cats, or racoons. By transferring a pre-trained neural network, the subsequent training process can be accelerated and deliver more accurate results.

Once the level of accuracy is sufficient a neural network can then be ported to a less powerful and more efficient machine such as an enterprise server, a mobile phone app, or an embedded computing device. This process, referred to as inference, may also involve further optimisation by pruning nodes or fusing layers. In some cases the training and inference platform can be the same, in other cases they may be completely different devices, depending on the application.

With unsupervised learning, a computer analyses an unlabelled set of data or images and tries to identify relationships, discover anomalies, or extract features. Though there is no guarantee that results are pertinent, unsupervised learning can highlight data and correlations which would have otherwise remained hidden to humans. Though currently limited to tasks such as data mining and clustering, unsupervised learning may lead to advanced research applications and perhaps even eventually towards general AI.

For semi-supervised learning a portion of the data is labelled for training purposes while the rest is left unstructured. This approach can speed up the learning process and improve results, especially when dealing with complex data. Yet another approach involves adversarial networks in which one computer competes with another to improve overall results.

Organisations adopting AI will most likely rely on a form of supervised or semi-supervised learning. Others may choose off the shelf software or cloud based services that do not require developing algorithms or training neural networks. Larger organisations with multiple AI objectives may mix off the shelf software with their own AI development, depending on the tasks. No matter which type of learning is involved, organisations will need a powerful computing platform, on premises or in the cloud, that is capable of supporting AI now and into the future.
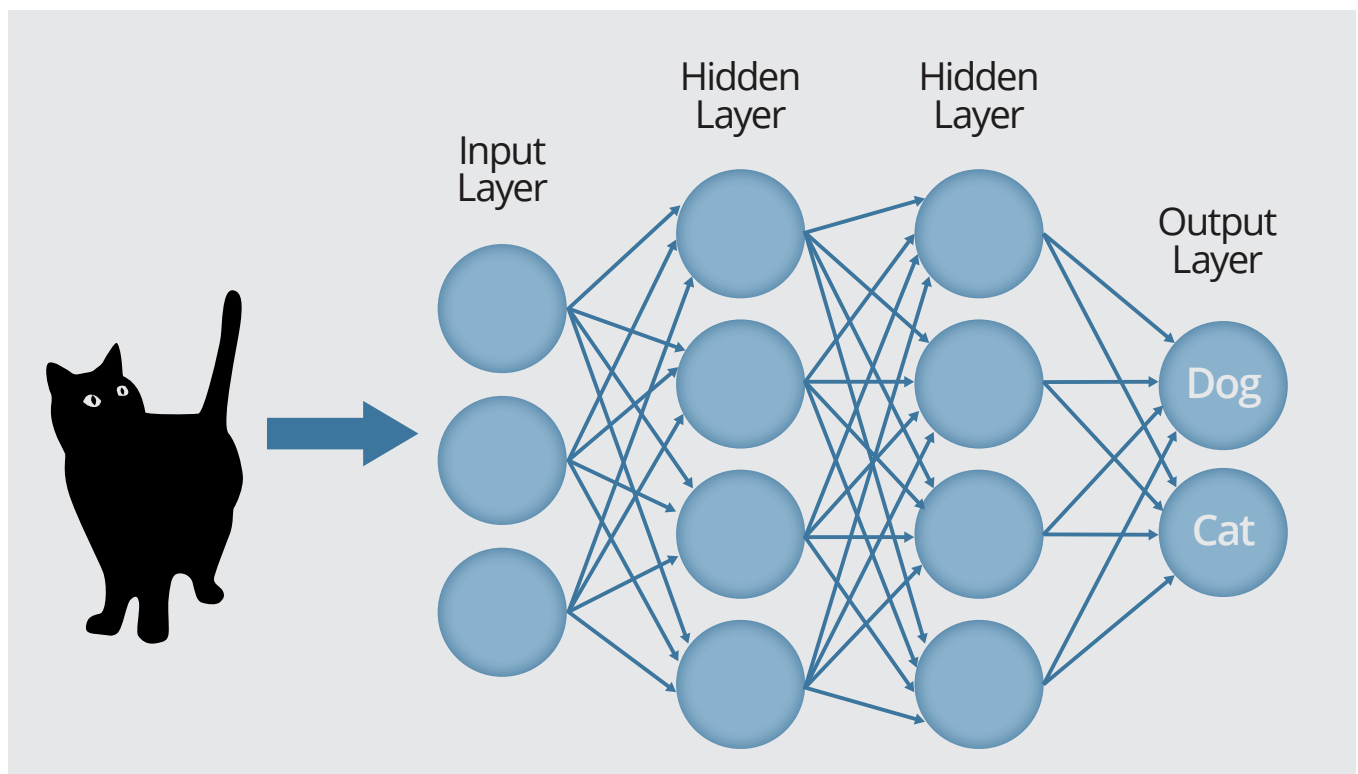
Building an AI technology platform

HARDWARE INFRASTRUCTURE

Recent advances in computing hardware have pushed AI and deep learning to the forefront. Network speeds and bandwidth have greatly increased with the general availability of 10Gb Ethernet, fiber optics, and 4G cellular networks making it easier to collect data from the field. High speed solid state drives (SSD) and other high capacity storage technologies allow organisations to store huge amounts of data while larger RAM capacities enable fast processing of large datasets.

For processing data, the latest generation of CPUs feature multiple cores which can be used in parallel for training while low power CPUs and custom ASICs can be embedded within AI inference devices. Much of the excitement, however, has been around GPU technology. Originally designed for speeding up video displays for gaming and other visual applications, GPUs are increasingly used for parallel processing of deep learning data. Since a single GPU can include hundreds of parallel processing cores, the technology is well suited for AI, especially deep learning training.

Depending on the business model, location, and sensitivity of data, organisations may choose to build an on-premises AI computing platform with a mix of CPU and GPU processing. Some cloud service providers also offer GPU powered instances which can be useful for ad-hoc training or inference of cloud based services. Organisations looking to deploy multiple AI applications may consider a hybrid platform that can draw from both private and public cloud resources using virtualisation and portable application containers.

SOFTWARE STACKS

In addition to a computing platform, organisations also need to consider which type of software they will need for their AI applications. In some cases, such as for data analytics, using an off-the-shelf software package running on a standard operating system or even as a cloud service may make the most sense as it does not require any specific machine learning development. If developing their own algorithms and neural networks, then they will need to choose from a large array of AI software components.

A typical AI software stack will include a framework as a template for developing applications. There are many open source options available and supported by leading technology vendors and research institutions. AI frameworks can invoke libraries that include neural network building blocks tuned to specific CPUs and GPUs which will themselves rely on proprietary or open source drivers. In some cases a graphic based development tool may also sit on top of this stack to enable data scientists to develop algorithms without the need to write code.

Whether using off-the shelf software, a development tool, or custom AI code, forward planning is required in building an AI computing platform that is flexible and scalable in order to meet the future needs of an organisation. An initial AI implementation is only the first step towards AI – it's a continuous process requiring ongoing monitoring and optimisation to ensure that value is being delivered across a wide range of business and service functions.
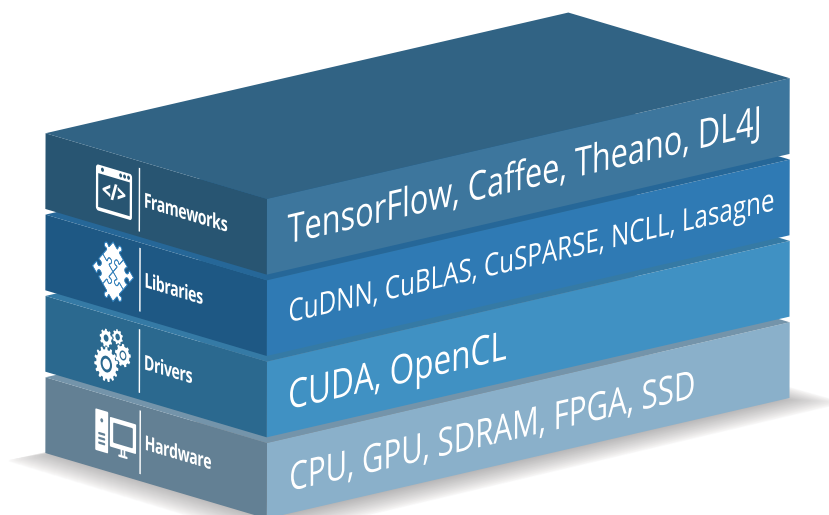
Steps to implementing an effective AI strategy

COLLECT AND BACKUP CRITICAL DATA

Most organisations have been collecting data for years if not decades through field sensors, loyalty cards, social networks, and a myriad of other data collection methods. Curating unique data can provide a competitive edge to businesses, while for governments it can add significant value to the services that they provide.

For deep learning to be useful it is imperative that data is efficiently captured, formatted and stored. It is also critical to ensure that proper backup and disaster recovery policies are in place. Investments in scalable storage, network bandwidth, and backup systems are a pre-requirement to implementing AI.

## IDENTIFY TASKS AND FUNCTIONS WHERE AI CAN DELIVER VALUE

In developing an AI strategy, it is important to recognise that machine learning and deep learning works best when focused on specific tasks. Rather than aiming for overarching goals such as increasing productivity across the board, AI implementation strategies should include specific objectives within different business functions. For example, inside sales could apply AI to lead qualification, accounting could look to speed up auditing, while customer support could be improved by resolving issues faster. The more precise or narrow the individual tasks, the more likely the overall AI strategy will be successful.

GATHER RESOURCES AND EXPERTISE

Implementing AI requires the expertise of data scientists, IT professionals, as well as functional stakeholders to make sure that tasks are correctly defined and budgets are aligned with strategic priorities. In some cases an organisation may already have data science and technical resources internally. By assessing the internal capability gap between in-house and required expertise, organisations can better plan for acquiring additional resources, including external consultants and technology experts.

XENON. High Performance Computing

6

www.xenon.com.au
info@xenon.com.au
1300 888 030
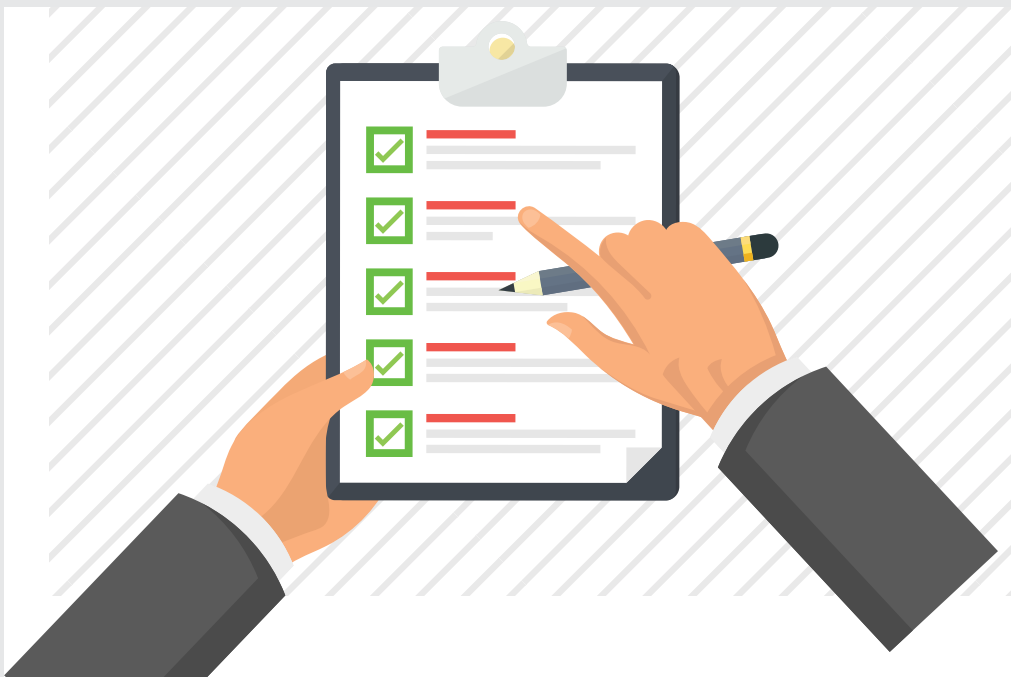
INVOLVE STAKEHOLDERS

Any strategy that introduces AI will inevitably incur changes that require the full buy-in from senior management. For this reason it is important that concerned stakeholders are kept informed and "sold" the benefits of implementing AI and the role it will play in the overall strategy. Executives should take ownership of strategic AI implementation and pro-actively communicate as to how it will contribute to business objectives or government mandates. Staff will need to be kept informed as to how it will impact their roles and be given the chance to share ideas on how to best integrate AI. If IT managers are not communicating with staff working within functions in which AI is to be deployed, then the value it will bring may be limited. It is imperative that AI is implemented with a focus on the organisational objectives to create value, rather than just on a successful deployment of the technology itself.

MEASURE RESULTS

There is no final step to implementing AI. Data, algorithms, and technology platforms will continue to evolve with AI itself driving much of the change. It is therefore critical to continually measure results of AI deployments. Key performance indicators (KPIs) need to be established from the onset. It can sometimes make sense to run pilot AI programs prior to launch. A team combining both technical experts and business leaders should be established to review AI KPIs on a regular basis. By measuring results and considering ways to optimise, scale up, and expand AI across an organisation, this will enable value to be delivered over the long term.

CONCLUSION

AI, especially deep learning, has the potential to create significant value by reducing costs, increasing productivity, and improving services. Successful implementation of AI requires advanced planning and preparation involving teams comprised of technical, data, and functional experts from both inside and outside of the organisation.

By focusing on very specific objectives that aim to augment tasks, services, and functions, an organisation can organically implement AI at a pace that best suits their overall strategy. Deploying highly scalable and flexible technology along with AI performance indicators will enable ongoing optimisation and organic infusion of AI across all parts of an organisation.

There is no easy simple one size fits all approach to AI. Though much can be learned and borrowed from implementations within other organisations, AI is also an opportunity to differentiate and lead. However, the innovation and business transformation potential of AI is so great that it can't be disentangled from an organisation's overall strategy. Delivering value with AI therefore requires bold new thinking from senior managers willing to consider all the technical considerations and from IT experts and consultants who understand that AI is more than just a new technology.