JOHNSTON CARMICHAEL

Walking the AI tightrope – Balancing the risks and rewards

Walking the AI tightrope – Balancing the risks and rewards

The AI market is burgeoning, and the race is on for businesses to capitalise on the transformative potential of AI and deliver a meaningful return on investment. Yet for investors and businesses it can still feel like walking a tightrope!

The market continues to grow at pace, with key players like OpenAI and Anthropic further enhancing their models, whilst incumbent technology giants such as Microsoft, Nvidia, Meta and Apple continue to make major investments. The competition has, however, taken an increasingly global dynamic as the technology's strategic geopolitical importance is recognised.

The US is currently in the strongest position and seeks to maintain that, with a planned $500bn investment in the Stargate Project to develop US AI infrastructure. However, the release of DeepSeek, China's response to ChatGPT, developed on the relative cheap, underscored the immaturity and therefore volatility of the AI market. On launch, $600bn of market value was wiped from the US stock market on 27 January and, AI chip giant, Nvidia's share price briefly tumbled 17%.

Illustrating reluctance to rely on either US or Chinese AI; Europe, Australia, South Korea and India, amongst others, have significantly invested in their own AI capabilities. At the recent AI Action Summit, the European Commission outlined an initiative to mobilise €200bn for investment in AI, including a new European fund of €20bn for AI gigafactories. This followed Emmanuel Macron unveiling plans for €109bn of private sector investments in AI, taking an active role in leading the charge as Europe seeks to take a position on the global AI stage.

Amongst the headlines and sensationalism there is a rush for organisations to get in on the 'AI act' and proofs of concepts and tests are springing up, often in a relatively unstructured and haphazard fashion.

Risk & reward

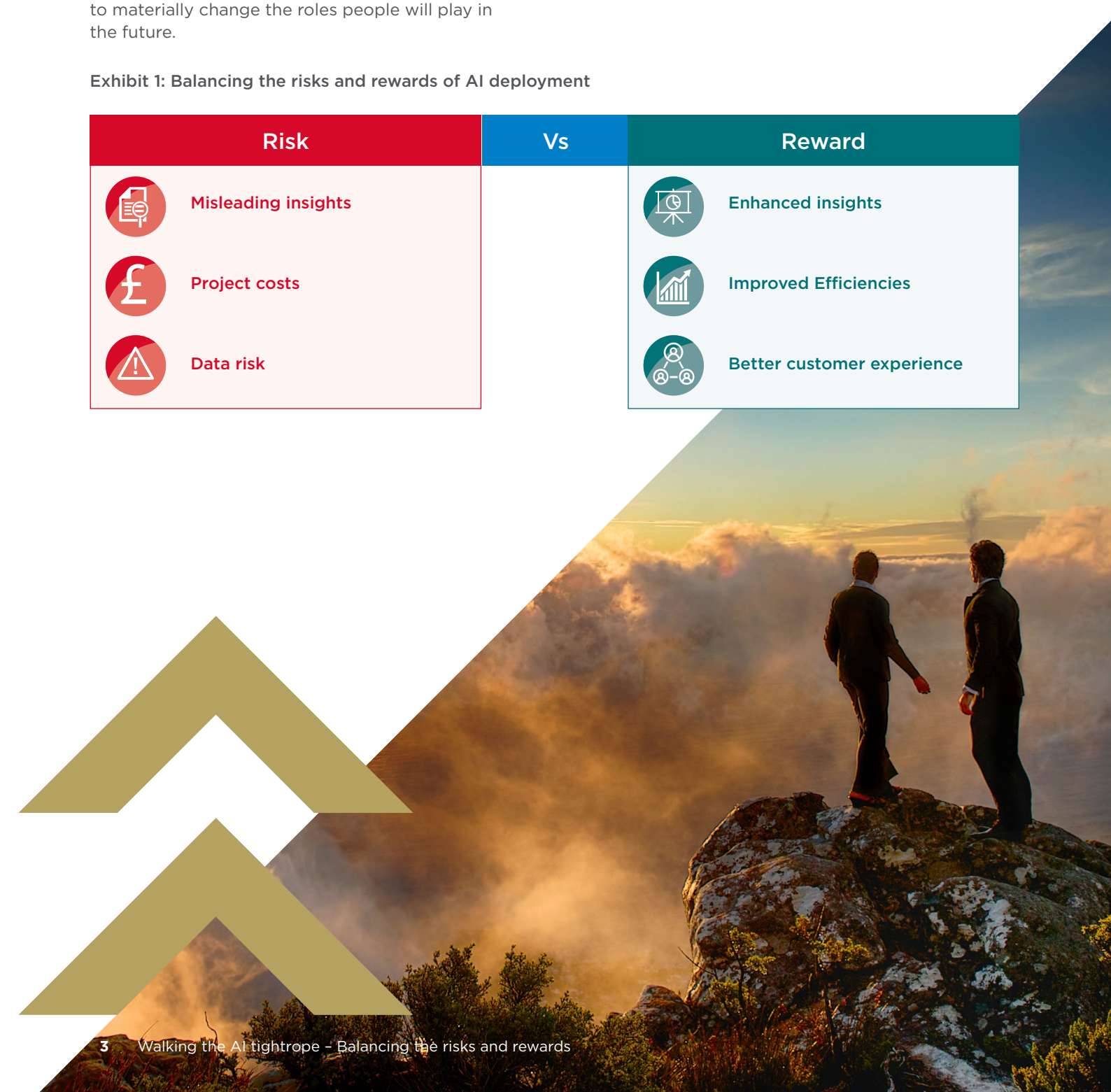
Whilst some organisations merely see AI as a tick box exercise, simply employing ready-to-use models (i.e. ChatGPT or Copilot) to perform minimal tasks such as taking meeting minutes, enlightened organisations are determining how best to deploy AI to address their specific challenges, realise value and manage the associated risks.

The rewards for using AI to good effect are well documented and already we are seeing that the range of potential use cases are vast. AI has been directed at supporting policy setting, process efficiency & effectiveness, regulatory compliance, system programming & deployment, and promises to materially change the roles people will play in the future.

However, delivering a return on investment requires a balance to be struck between understanding and mitigating the risk considerations, and securing the compelling rewards of AI deployment (Exhibit 1).

Exhibit 1: Balancing the risks and rewards of AI deployment

Risk	Vs	Reward
Misleading insights		Enhanced insights
Project costs		Improved Efficiencies
Data risk		Better customer experience



Misleading insights vs. Enhanced insights

AI models have the processing power and ability to analyse data and provide unrivalled insights at pace, allowing incisive decisions to be taken with confidence.

However, organisations must remain vigilant, as models built on the wrong objectives, 'bad' data or poor training will produce misleading outputs and inaccuracies ranging from biases to hallucinations to underfitting/overfitting (Exhibit 2). Ensuring models are trained on 'good' data, i.e. a sufficient volume of well structured, clean, diverse and representative data, reduces the likelihood of inaccuracies, therefore producing insights the organisation can trust.

Exhibit 2: AI insights are subject to inaccuracies which can greatly impact performance and reliability

Inaccuracies	Description	Real world examples
Biases	The model has not been trained on a diverse and fair dataset and, therefore, draws conclusions that are not relevant for a wider data population.	AI used in US hospitals favoured healthier white patients over sicker black patients because it was trained on cost data rather than care needs.
Hallucinations	The model recognises coincidental/ unintentional patterns in its training datasets and makes incorrect predictions or outputs as a result.	A lawyer who relied on ChatGPT for legal precedents became the subject of ridicule when the cases sighted did, in fact, not exist and had been hallucinated by the AI.
Underfitting/ overfitting	Underfitting refers to the model being trained on data that is too simple resulting in poor performance, whilst overfitting refers to the model experiencing high variance and consequently performing well on training data but not test data.	AI developed to assist doctors diagnose symptoms struggled with real-world application as the training data was too simplistic.

Project costs vs. delivered efficiencies

A key driver of AI implementation in businesses is to unlock significant increases in efficiency by augmenting or replacing human-intensive processes. With AI in place of humans, automated processes can be completed faster and offer scalability coupled with greater consistency. They also cost less to run than a conventional workforce.

However, many businesses have found developing and deploying AI to be a long and costly process that brings no automatic guarantee of success, with many projects never making it past the development stage. Before embarking on an AI journey, organisations should first consider whether they have the internal capabilities to develop a solution or whether they will require 3rd party support. Both options contain inherent risks:

1. In-house solutions can be low-cost alternatives to 3rd party development. However, in absence of sufficient internal capability, projects risk both failure and erosion of internal investment budget.

Case study:

Zillow

The technology-driven US real estate platform infamously developed a real estate valuation model in-house that significantly overvalued properties. The model failed to properly understand market trends leading the company to overpay upwards of $500m on properties.

2. Seeking 3rd party experience will generally increase a project's chances of success but this expertise can be costly, particularly if their solution is unproven.

Case study:

McDonald's

The fast-food giant were estimated to have spent tens of millions of dollars on a failed partnership with IBM that aimed to introduce AI to their drive-throughs. The AI failed to understand customer orders, with cases reported of the AI adding bacon to ice creams and incorrectly adding 260 chicken nuggets to orders!

Whether solutions are developed in-house, or external support is sought, the risk of failure is extremely costly, and increasingly organisations choose to run a proof of concept before 'going big'. Managing the balance between delivering long-term savings and short-term development costs is essential to ensure the project is profitable and value is realised over the course of its lifetime.

Data risk vs. Better customer experience

Leveraging AI to analyse vast amounts of data, including specific customer data, can enable organisations to provide tailored products, solutions and service to their customers. This allows them to build stronger relationships and increase customer satisfaction and loyalty.

Of course, data privacy and security is a key concern for many when the subject of AI is broached. Organisations must contend with the inherent risks of handling large volumes of sensitive data. Businesses with inadequate data security can quickly find themselves on the wrong side of regulators, opening themselves up to fines and reputational damage.

Case study:

Clearview.ai

Clearview AI, the facial recognition platform, were fined €30.5 million by the Dutch data protection watchdog for uing an 'illegal database' in September 2024. The database was comprised of billions of faces taken from social media and the internet, without the knowledge or consent of those individuals.

Setting up for success

Whilst the potential for AI to deliver meaningful business benefit is huge, a robust AI development methodology will enable organisations to effectively manage the associated risks.

The most successful AI programmes have strong executive sponsorship. Establishing a business case that defines the problem statement, articulates the value to the business, contributes to the strategic vision of the organisation and engages with key stakeholders (typically across multiple functions) will better ensure buy-in.

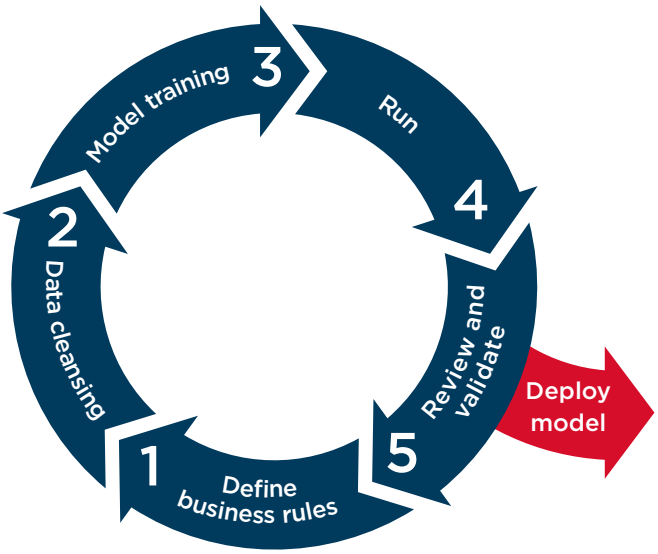
AI capabilities will continue to advance at pace and businesses should recognise that AI is not a silver bullet which will instantly transform entire processes and workflows. Instead, having a cyclical approach to learn and refine initiatives to deploy AI will enable faster time to value. When done in this way, over time, these models can be scaled to a broader range of tasks, maximising value across the organisation.

When looking at successful deployments, several key themes consistently emerge:

A cyclical approach to model development

Developing AI models is an iterative process that requires regular monitoring, retraining and optimisation. Continuous feedback loops, including post-deployment, can refine models, adapt to changing conditions, and enhance long-term performance (Exhibit 3).

Exhibit 3: A cyclical approach to AI development



Synergy between the business and AI engineers

Since many data scientists are not well-versed in the business outcomes they are building models for, enlisting subject matter experts to translate the problem statement into business requirements and to provide subsequent input during the refinement of the model will help ensure AI achieves business objectives.

Comprehensive business rules

The business rules, built around the problem statement and strategic vision of the organisation, ensure that the solution operates within defined parameters and delivers consistent, ethical and reliable outcomes.

Shifting from data-driven to data-centric

Clean, structured, and well-formatted data improves AI accuracy, reduces biases, enhances model performance, and gives added confidence in outputs.

The importance of human review

Instances of bias and hallucinations necessitate human review to validate outputs, ensure accuracy, and train the model to deliver reliable results. Reviewing outputs enables the model to learn and adapt, ultimately improving accuracy and reliability.

AI is empowering businesses to drive efficiency, innovation and growth but, as its adoption continues to accelerate, design and development effort must be matched with the ongoing learning and training of the models and a robust controls framework (Exhibit 4).

Exhibit 4: An AI Control Stack can manage risks and instil confidence across the business.

1. Governance & ethical controls: Appropriate governance structures are embedded to guide and manage AI within the organisation, and ensure robust policies, processes and procedures are in place to manage risks and address ethical concerns effectively.

2. Data controls: Data is representative, unbiased and high-quality to ensure reliable and accurate outputs.

3. Model controls: Robust and continual model oversight is in place to ensure the model is transparent, explainable and accurate. Conducting repeated testing and training cycles to validate and refine model outputs, with human input at key stages.

4. Security & compliance controls: Robust IT and cybersecurity protection measures are in place to prevent breaches or non-compliant uses of data.

Walking the tightrope

We are at the tipping point of a reality where AI will play an integral part in shaping all aspects of business and life, and this will have a profound impact on the way our societies and businesses operate. Considerate, informed and steady stewardship will be critical to manage the considerable risks involved.

Progress needs to be made by consciously walking the tightrope that realises specific business case benefits, whilst remaining alert to the risks associated with the still emergent technology including bias, discrimination, breaches of confidentiality and violation of intellectual property rights. Utilising a proven AI adoption methodology, alongside a robust control stack, will help mitigate the risks whilst allowing businesses to fully harness the transformative power of AI.

Johnston Carmichael leverages their extensive financial services experience to help clients articulate their problem statements and develop their AI use cases. We work closely with key stakeholders and technical teams, both internal and external, to develop business requirements to ensure effective deployment of AI initiatives and reduce time to value.

If you'd like to explore how AI could benefit your organisation, feel free to reach out—I'd be happy to discuss further.

Zoë Hodge
Partner and
Head of Consulting
E: zoe.hodge@jcca.co.uk

Where sharp minds meet

JOHNSTON CARMICHAEL | MOORE

jcca.co.uk