

AI STRATEGIES SHOULD START WITH DATA STRATEGIES



FOLLOWING FIRTH
APRIL 2024



As the hype around Artificial Intelligence (AI) continues to gather steam, companies of all sizes and from all sectors are investing in the technology. According to a Forbes Advisor survey, more than half of businesses are using AI tools to improve and perfect business operations and to help with cybersecurity and fraud management. Fewer are harnessing AI for more specific applications, such as customer relationship management, inventory management, and content production, but these are growing in popularity. Some companies are even using AI for recruitment and talent sourcing.

Unfortunately, organisations are learning that paying attention to the adage “garbage in, garbage out” has never been more important. As implementations of AI become more prevalent – and more complex – companies are finding that the quality and quantity of data used by the AI system is directly responsible for the kind of results they are getting.

This may seem obvious, but many businesses base their data warehousing strategies on month-end or year-end. In a world where a generative AI application is expected to deliver real-time results, this immediately sets the organisation up for failure. The businesses AI data strategy has to be to deliver the data to be used for AI to any downstream repository from any operational systems in time for the AI engine to not only “analyse” the data, and to execute predictions, but implement them in time to make a difference or change outcomes to an inflight process.



Garbage in, garbage out

All aspects of AI are dependent on massive data sets, so it makes sense that the biggest risk to AI lies in bad data. Bad quality data will not only create a bad output, but it will also train the model incorrectly for all future computation and predictions. If a company includes unstructured, nonstandard and incomplete data in its AI models, the results will be completely unusable in the worst-case scenario, or incorrect in the best-case scenario.

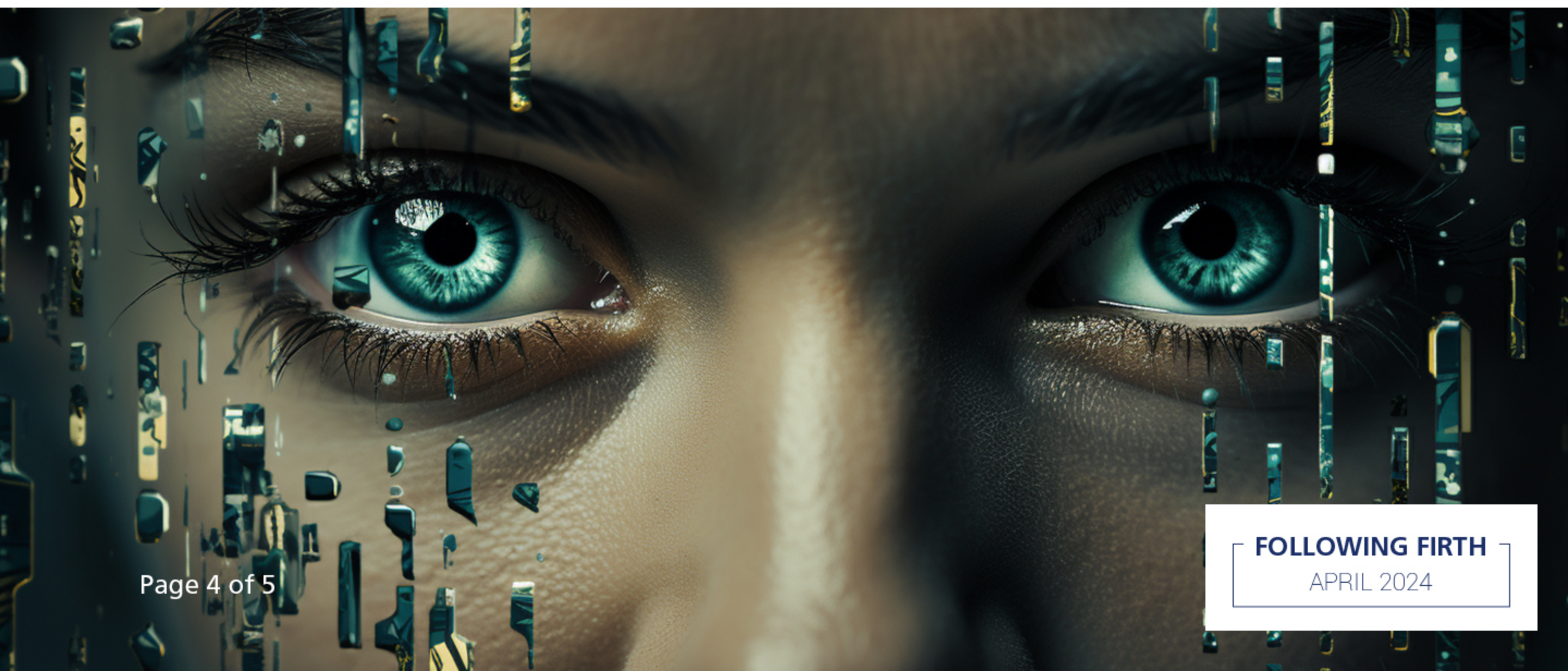
How an AI model is designed is obviously vital to the performance of the system, but the model relies on data in order to complete its tasks. The more diverse and comprehensive the data, the better the AI can perform.

With its ability to process and analyse massive datasets in real-time, AI should be allowing companies to transcend the traditional limitations of data analytics. Instead, organisations are encountering new challenges as a result of their rush to jump on the AI bandwagon. Even the concept of real-time data collection can become a challenge, as it takes time for the data to reach the server it's stored and processed on, and that minor delay is often not factored into AI's models.

Paying attention to privacy

Data privacy is another aspect many companies don't consider when starting on their AI journeys. AI systems gather and use a lot of data to learn, but while the bulk of the data might be collected from intentional sources, like when customers provide their personal information, a great deal is also being collected from unintentional sources, where the AI collects data without individuals realising it.

Since AI can unintentionally gather our personal data without us being aware of it, and since the data might end up being used in ways not always expected, regulations – and organisations' ability to comply with regulations – are struggling to keep up. For example, some call centres have started using AI-driven voice recognition to help identify a customer's mood to better tailor their interactions with the customer. If they call in a bad mood, the AI can identify the fact that they are angry or frustrated, allowing the customer service agent to start the conversation in a more pacifying manner than they would with someone who has been identified as in a good mood.



It's therefore vital that we find a balance between personalisation and privacy. AI's ability to generate DeepFakes adds an additional layer to this issue. Companies have to evaluate all of the implications for privacy in the way they use AI, looking closely at how the data is used, stored, and accessed. Privacy regulations like POPI and GDPR restrict the collection and use of customer data, requiring the careful handling of information to ensure privacy, but while AI systems may start off compliant, there are hundreds of different ways that data privacy can become an issue as the AI learns and evolves.

As we continue to advance in our AI implementations, understanding this dynamic will be key to harnessing the true transformative power of artificial intelligence. People tend to believe that AI is a magic wand that will solve all our data quality and trust issues, but the role of AI is to find insights from good quality data, not necessarily attempt to fix decades-old data management processes. From predictive analytics to anomaly detection, AI algorithms can uncover patterns that human analysts might miss, enabling companies to make proactive and data-driven decisions, but the first step is to acknowledge the fact that the organisation should start with a data strategy, not an AI strategy.



RICHARD FIRTH

CHAIRMAN & CEO, MIP
HOLDINGS