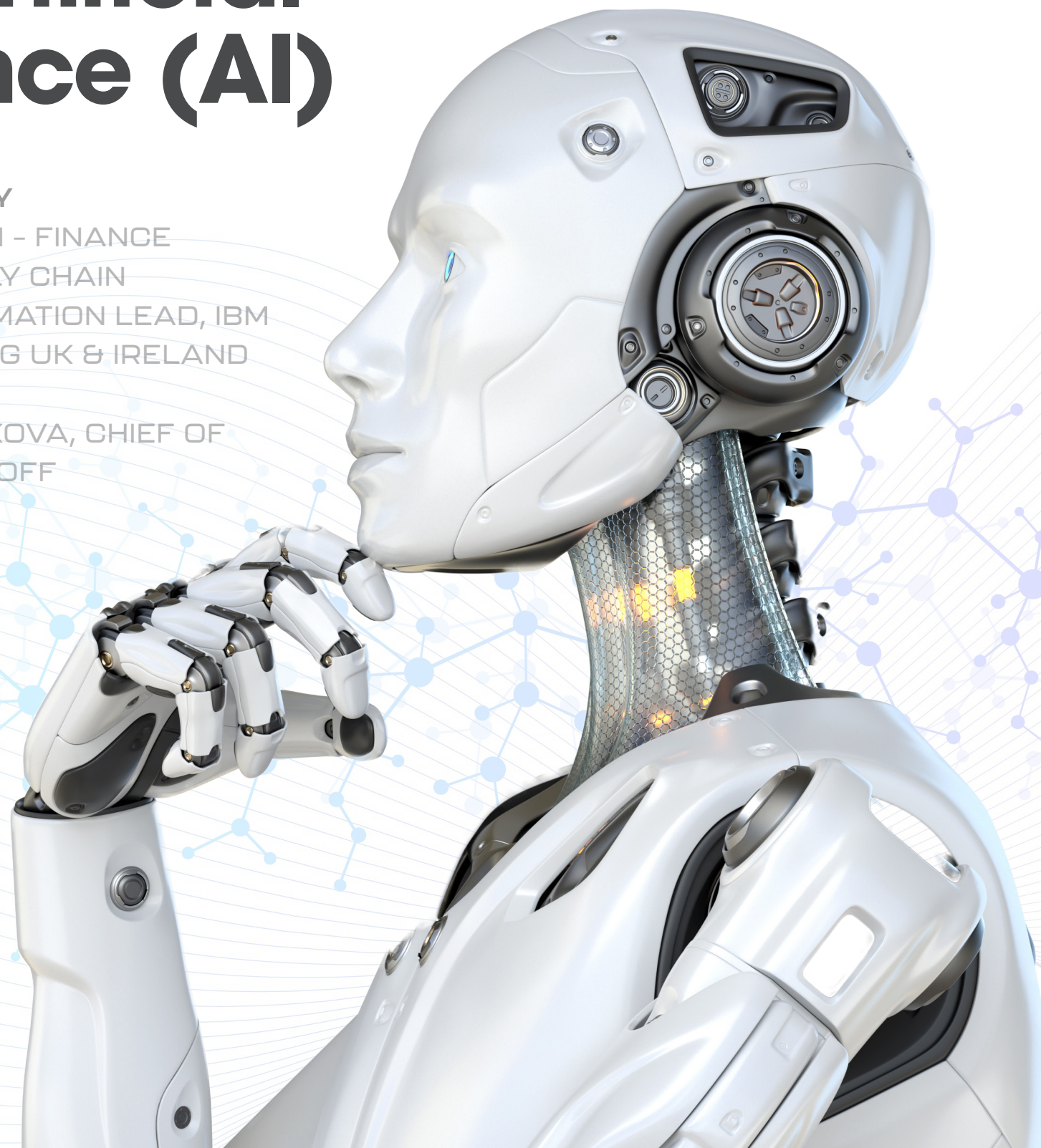# IBM ®

Ethical Artificial Intelligence (AI)

WRITTEN BY

BOB BOOTH – FINANCE AND SUPPLY CHAIN TRANSFORMATION LEAD, IBM CONSULTING UK & IRELAND
&
ALEX ISSAKOVA, CHIEF OF STAFF, LIFTOFF

Introduction to the authors

Alex Issakova is Chief of Staff at a Silicon Valley-based Adtech company Liftoff. Bob Booth is a partner in IBM Consulting's transformation business. Bob and Alex have partnered to research the topic of Ethical AI. Alex has done extensive academic literature research and she and Bob have conducted a number of interviews with people at the sharp end of AI, from which this paper has been written.

Introduction to AI ethics

The last decade has seen an explosion of AI tools that promise to simplify our lives by making complex decisions objectively and at scale. AI is being increasingly used to recommend and make societally impactful decisions in fields like hiring, the criminal justice system, banking and others.

AI could make our lives exponentially better through personalised film recommendations, automated critical business decision making, offering driverless cars, and support for people with various disabilities. However, AI is not a magic wand that will solve all of our problems and there are various examples identified in our research and interviews of AI getting it wrong. Facebook's chat bot that became a right-wing extremist is one of the most famous examples.

We used the concept of AI Fairness (impartial and just treatment or behaviour without favouritism or discrimination) to represent the concept of Ethical AI. In this article we will consider why the CPO should care about Ethical AI and what should she or he do about it in the short and longer term.

What is AI now and in the future?

Artificial Intelligence or AI enables computers and machines to mimic the perception, learning, problem-solving, and decision-making capabilities of the human mind, passing the Turing test if you will.

That is, AI is an application (machine learning, deep learning etc) with a data model with some input and learned data, that receives inputs (sense), makes decisions (decide), takes an action to recommend or carry out a task (act). The AI then learns from the process, sometimes with human help, which is called attended, or as often the case, without human help, which is termed unattended, which is the final step of learn.

What are the components of AI and how do they combine into products?

If we reflect on some of the components of AI then it helps to understand some of the risks when these are combined into the procurement infrastructure:

- Machine learning – is a type of artificial intelligence (AI) that gathers insight from structured and unstructured data such as material line item data or contract information, without being explicitly programmed to do so, and is one of the most common forms of AI.
- Deep (structured) learning - is part of a broader family of machine learning methods based on artificial neural networks, which is a mathematical algorithm that is trained to come to the same result or prediction that a human expert would when provided with the same information, based on patterns it discerns from data or images.
- Computer Vison and Speech – field of artificial intelligence (AI) that

enables computers and systems to derive meaningful information from digital images, videos and other visual inputs – and take actions or make recommendations based on that information. In 2012 the error rate of computer was 75%, but 2021 it was claimed to be down to less than 1%.

- Natural Language Processing (NLP) – Natural language processing is the ability of a computer to understand the speech synthesised and respond with a conversation that seems as if it were human. The simplest chat bots are little more than question and answer machines, where the input has to be structured and the answers limited. Real NLP allows for real unstructured

conversations. Tools such as Siri and Alexa sit in-between these levels, and provide voice access to very powerful Apple and Google assisted search engines.

AI is now deployed as:

- Stand-alone tools built for a purpose, such as a conversational chat bot, or recognising failures on a production line, in which case the data scientists involved tend to consider ethics as a topic, at least from the larger and more recognised firms.
- Inherently included in start-up technologies and in other enterprise systems ranging from Microsoft Office to systems like SAP or Oracle.
- Some combination of the two.

What are the main risks from AI?

AI, although around since the 50s has only exploded recently and is therefore a very new industry. This newness of profession, although making it an extremely exciting and ever-changing field to work in, sometimes comes with a certain 'wild west' feel where many of the much-needed checks and balances are yet to be developed.

From our research we identified a number of risks, and here are a subset most relevant to the CPO:

• Lack of developer team diversity

Some AI applications are being written by teams of very similar developers who have a limited view of the world and they may be encoding that limited view into AI. This creates a risk of human bias, an inclination in favour of or against a person, a group, an idea, or a thing, sometimes in a way that is unfair. All humans have biases that can influence their opinions, decisions, and actions. Because AI is built and trained by humans, these human biases can unwittingly permeate AI models, causing them to perpetuate inequities, resulting in algorithmic bias.

• Lack of maturity in data standards

The challenges with groups of developers that are not diverse also extends to data sets. When you consider what use case is needed, for example in the case of talent such as a hiring use case, you need to make sure that the data set you are training the system with represents the diverse set of talent that you want to hire and not just the set of talent that you used to hire or even worse a set of talent that developers have identified themselves without business involvement.

• Complexity of building and running AI

Building AI systems can be a complex process particularly when AI systems are combined. Some applications are built as an amalgamation of different AI tools and teams and one person may not be responsible for checking that the system is working in a bias free manner. Also AI applications are not finished after the first build, they are only really just starting. AI systems continue to learn after they go live and need constant business monitoring. This means that the classic approach of application build using a large knowledgeable team, typically a mix of onshore and offshore, followed by a skinny offshore team for application development will not cut it. You need a support function that can assure the ongoing fairness of the data, the applications and the decisions made.

• Lack of transparency, explainability and right of appeal

Some AI applications make recommendations to humans, and some make automated decisions. With current AI development, not all systems are:
• Transparent. In AI, transparency refers to sharing appropriate information with stakeholders about how an AI system has been designed and developed.
• Explainable. Explainability refers to an AI system's ability to provide insights that humans need to explain how decisions were made.
• Right of appeal. There may be no right of appeal for someone to understand how the decision was made even if the owners of the AI process are capable of doing so.
• Lack of business readiness. Managers who are users of AI-based tools are often unskilled in noticing issues with AI tools

and they may not fully understand how AI works, what the risks are and how to make the right decisions given the tools available.

Why should the CPO care?

In our research we adopted the phrase societally/socially impactful use cases. In societally/socially impactful use cases a decision, if wrongly or unfairly made, could materially affect someone, their wealth, access to their family, and or their physical or mental well-being. Examples outside procurement include AI enabled decisions that could affect whether someone is employed, receives government benefits, is accused or convicted of crimes or has access to their children. Clearly the CPO is not faced with these decisions day to day, but when we look across the Yearbook, we see some technologies that drive category strategy, support negotiation, report carbon and identify risk. These decisions will start to impact the suppliers we select and that may become societally/socially impacting. Even where it is not socially impactful, the CPO and their teams will be called upon by stakeholders and suppliers to explain decisions.

As a result of this the CPO and their teams should start to think about Ethics and AI. At this stage we suggest 3 simple steps:

1. Appoint an owner. Appoint someone in your team who is a system and data native to take responsibility for the ethical AI to become your SME. Have that person work with colleagues in IT, legal etc and gain skills on these topics. IBM is a leader in ethical AI and this link may be a good starting point for ethical AI.

2. Add Ethics to your technology selection process. When you are selecting tools for procurement, and indeed when supporting the business with a wider IT selection, make a point of asking suppliers what types AI is included, where the data sets are derived from and whether the system is explainable. Ensure that where you have AI suggest that the design team and data teams are thinking about ensuring that the systems are built and stay fair.

3. Experiment with systems that include AI. As part of your experimentation use cases for 2023 pick a package that includes AI, monitor the results in design, test and live and assess whether you are experiencing the risks above. You may or may not prove the concept of fairness, but by considering this you will learn a tremendous amount about ethical AI

In conclusion, we would encourage you to positive and curious about the potential. Each of the risks expressed above has also have real positive potential. The flip side of risk, the possibility of creating inherently fair systems that are transparent and explainable that allow humans to make decisions free from bias is hugely exciting. We believe that fair and explainable decision making has the potential to be a formal ESG measure within the next decade. Taking some interest in this now will give you a competitive advantage in the future.

We are in the infancy of AI, and within 20 or 30 years we are likely to have technology that surpasses the human mind, and so we need to think now about how to build a profession and a society that can safeguard against any possible risks. ■