Artificial Intelligence

Are We Prepared?

Mac McMillan

MA, LCHIME, LFHIMSS, CDH-E

Founder, CEO, Board Member, Advisor, Co-Author

Agenda

Introduction

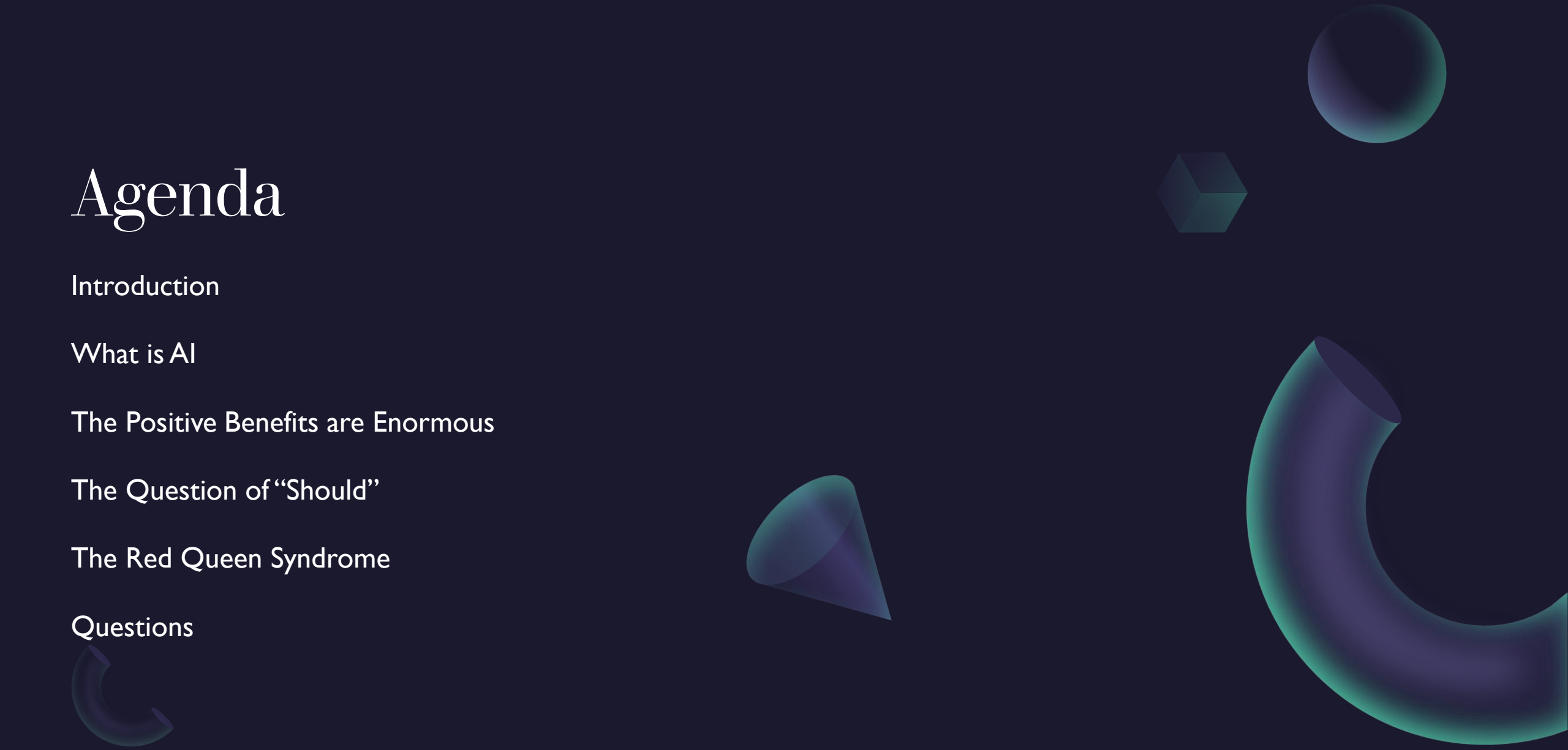
What is AI

The Positive Benefits are Enormous

The Question of “Should”

The Red Queen Syndrome

Questions



"This is the challenge that AI faces," he says. "We've got systems that work really well, but the ethical problems they create are burgeoning."

Jack Clark, Co-Director AI Index Steering Committee,
Stanford Institute for Human-Centered Artificial Intelligence

The Trust Deficit

In December 2021 using the TruthfulQA benchmark it was determined that at that time most models trained on the Internet were truthful only 43% of the time.

Stanford Institute for Human-Centered Artificial Intelligence

More than 60% of patients surveyed expressed a lack of trust in AI in their healthcare. This skepticism is rooted in concerns over data privacy, potential biases, and the lack of transparency in AI models.

Journal of Medical Internet Research 2024

The Ethical Dilemma (Predicament)

UAVs save lives/money/improve performance

U.S. Commitment to Protecting Lives

- Law of Land Warfare states we must seek to preserve civilian lives, only engaging combatants.
- In asymmetric warfare it is unclear "who" the combatant is.
- Autonomous drones cannot always discriminate foe from friendly.
- D0D Directive states that autonomous drones cannot deploy weapon systems unilaterally, a human must be in the decision loop.

What is AI?

A Historical Perspective

Isaac Asimov (Az-im-ov), A Science Fiction Writer

Asimov, who coined the term "robotics", delved into the moral imperative of creating computers that were or could be smarter than humans, *I Robot, 1950*.

- 1st Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- 2nd Law: A robot must obey the orders given it by human beings except where such orders would conflict with the first law.
- 3rd Law: A robot must protect its own existence as long as such protection does not conflict with the first or second laws.

The history of AI reaches all the way back to the 1930s:

- Alan Turing, mathematician/scientist who broke the enigma code in WWII, coined the term "Machine Learning"
- Christopher Strachey wrote first successful AI program in 1951
- John McCarthy, another British scientist coined the term "artificial intelligence" in 1956.
- 2009 Ecole Polytechnique Federale of Lausanne, Switzerland experiment with robots.

Understanding Artificial Intelligence

- **Artificial Narrow Intelligence** – ANI are systems engineered for a particular task or domain. Examples are personal assistant software like SIRI, Alexa or Chat GPT.
- **Artificial General Intelligence** – AGI exhibit human-like intelligence and versatility, capable of learning, adapting and comprehension across multiple domains, but lack many aspects of human cognition.
- **Artificial Super Intelligence** – ASI systems include cognitive capabilities that exceed human comprehension and capabilities; exponential learning, unbiased decision making, global insights and enhancements of its own intelligence.

AGI = Singularity

Singularity is the state achieved through Artificial General Intelligence (AGI). AGI refers to AI systems with the capacity to understand and learn new concepts, adapt to different situations, and solve complex problems as well as, or even better than, their human counterparts. AGI is considered **"Transformative."**

AI Is Far From Perfect Yet

- Brittleness – confounding examples
- Embedded Bias – automated discrimination en masse
- Catastrophic forgetting – loss of memory with new information
- Explainability – the black box syndrome
- Quantifying Uncertainty – bad conclusion multiplied
- Common Sense – when an adjective becomes bad
- Math – getting better (train, train, train)

Everyone is excited about the possibilities

Radiology: MRI and CT Scans

AI Enabled Camera Technology

Can detect anatomical landmarks in a patient to enable fast, accurate and consistent positioning during imaging, thereby reducing radiation dosage and improving CT quality. As many as 80% of images taken are delayed in read out. AI can assist in erasing this delay by providing better images and preliminary read outs to aid medical staff and help patients.

Phillips White Paper

Reducing Post Operative Adverse Events

UC Davis uses AI enabled technology to monitor all their patients, 8 vital signs, every minute, every day, 14,400 data points, with real time alerts.

1 in 5 Patients

In medical-surgical areas of hospitals nationwide will experience a post-operative adverse event. Manual monitoring of patient vital signs is time consuming and prone to error. Automated continuous real time alerts can reduce this risk.

Interview with David Lombarsky, MD

Medical Device Serviceability

- **AI Enabled Monitoring**

- Of IOMT can avoid downtime by identifying and allowing resolution of 30% more medical device service cases, preventing avoidable interruptions to clinical practices and unnecessary delays for patients.

Understanding the Brain

Using LLMs for Neuroscientific Research

Can give neuroscientists the ability to bridge the siloed areas of neuroscience to uncover new information that humans are incapable of finding alone. LLMs in genetics and neuroscience can be combined to develop new drugs focused on promising candidate molecules to stop neurodegeneration.

The Journal Neuron

Improving Care In Time

- **Productivity and Efficiency**

- In 2050 1 in 4 people will be over the age of 65. In order for systems and care to keep pace they will have to be more efficient and more productive. AI has the potential to accomplish this and allow care givers to do what they do best, focus on patients, and improve morale and retention.

The World Health Organization

But the question is “should” we...

An Inevitable Collision: AI With Ethics

The **moral principles** that organizations use to guide responsible and fair development and use of AI. This means taking a safe, secure, humane, and environmentally friendly approach to how we employ AI.

The challenge is counterintuitive. As AI capabilities improve the harm they can create where fairness and bias are concerned get worse.

Job Displacement

AI Will Change Our Workplace

Automation is expected to lead to the displacement of 300 to 800 Million workers worldwide by 2030. Another 375 Million more will have to switch to a new occupation. Most affected: professional, scientific and technical service workers. 3 out of 4 workers have this concern. Organizations, education, government and individuals affected need to develop mechanisms to **provide transitional pathways.**

Privacy

Training, Surveillance

Both training and operating AI models require vast amounts of data to be collected and analyzed, creating storage and access concerns. New privacy preserving technologies can mitigate this risk. Law enforcement and other agencies us AI to monitor and track movements of persons of interest, these can infringe on people's privacy rights. Rules regarding the training and deployment of these models need to be developed.

CompTIA Community

Bias

Gender, Ethnic, etc.

Bias is a huge challenge with generative AI models because they are only as good as the data sets they are trained on and the algorithms they follow. For example; there are ethical biases that undermine certain individual's ability to give birth following a cesarean section, there is bias with children in ERs, bias in legal systems for determining recidivism, etc. Bias can happen indirectly or directly. AI models need to be trained on large diverse populations to improve accuracy and reliability, and the data they use must be curated and protected.

Security

AI Is A Threat

From AI-powered ransomware to weaponized phishing to enhanced malware that can learn from its environment, evade detection, mutate, obfuscate and mimic. One attack, Deepfakes, are particularly capable at circumventing security controls and duping users. Deepfakes are easily created with AI, and not so easily detected. This level of deception with this capability for harm can have devastating results. And, AI itself is at risk, corrupt the algorithm, corrupt the data set, you end up with bad conclusions.

Explainability, Transparency and Accountability

When You Just Don't Know

Already in generative AI use cases doctors are learning that they cannot explain how the tool reached a particular conclusion which could lead to uncomfortable conversations with patients. In some fields, like healthcare and law enforcement, understanding the how, the influencing factors, is as important as the conclusion. Developers need to be able to explain in plain language how their models work and how the data will be used. The challenge is that Deep Learning is for now in many cases still very much a Black Box.

The Neuro Surgeon's Conundrum

Misinformation

When Subterfuge Looks Real

AI is particularly adept at creating and spreading misinformation while making it look legitimate. Misinformation messages utilizing multiple well-respected personalities as Deep Fakes, the Internet and Social Media and you can quickly flood the airwaves with content on a particular topic that could undermine people's confidence, plant false conclusions, affect public opinion and create real reputational harm.

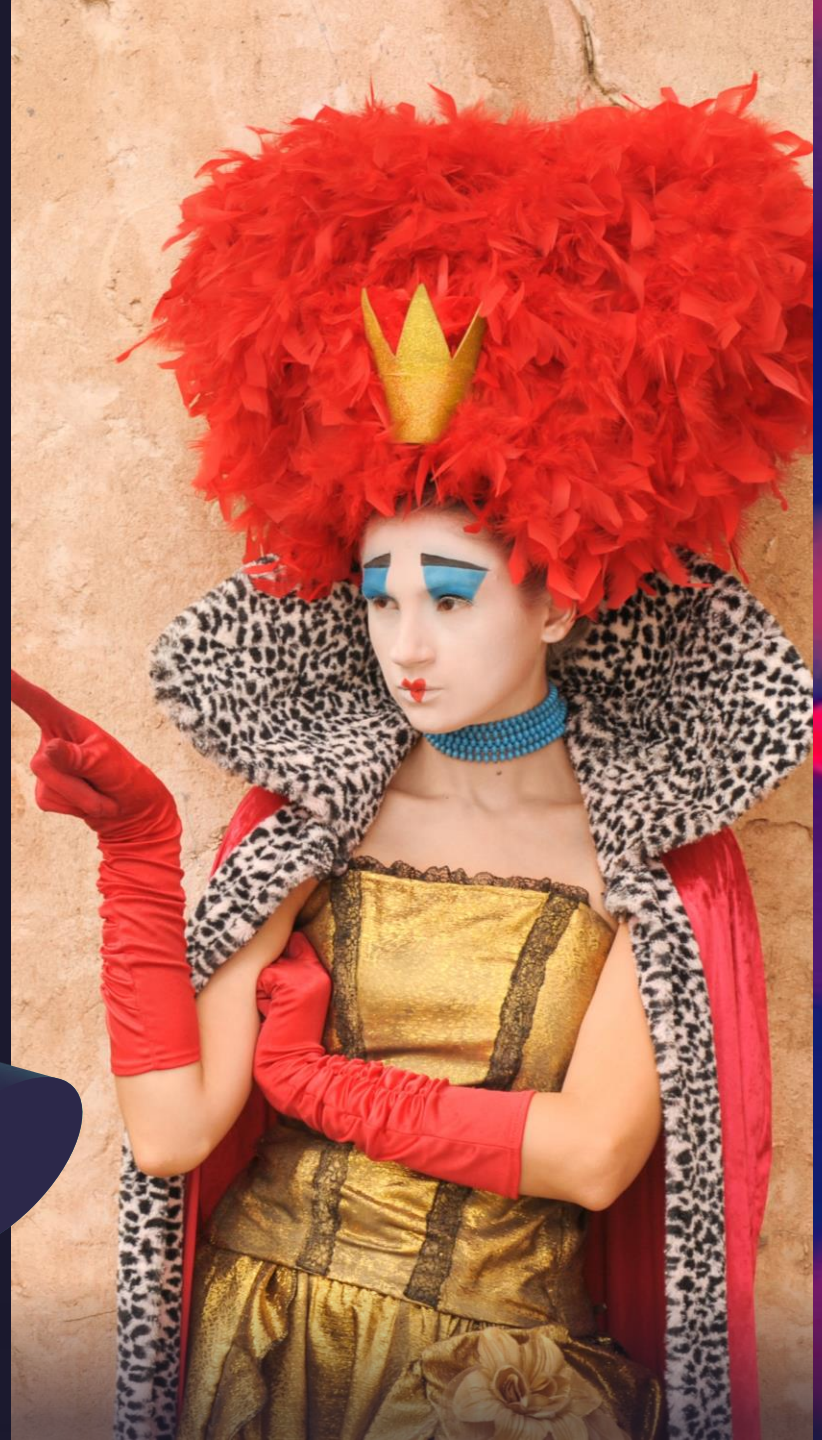
Environmental Impacts of AI

ML Is Very Hungry & Dirty

From the millions of gallons of water to the special materials needed for batteries and the process for making them ML places a heavy burden on the environment, with substantial carbon emissions contributing to potential environmental impacts. AI models use considerably more power, consume more and pollute more than standard computing.

Nature: Machine Learning

The Red Queen Syndrome



The Queen Cried; Faster! Faster!

In Lewis and Carroll's classic *Through the Looking Glass* the Red Queen explains to a frustrated Alice that if she wants to get anywhere she'll have to start running and continue to run because the reality in Wonderland is things move twice as fast as where she is from. The analogy is fitting for AI as this technology will move exponentially faster than anything we have ever seen. We failed to keep up with the Internet, we failed with Social Media, and we are already falling behind with AI. The question is can we run fast enough to catch up, and then maintain it indefinitely?

The Brookings Institute

Alarms Sound & Calls For Regulation

Quote from initial 1,100 signatories: "should we let machines flood our information channels with propaganda and untruth?"

There are 32,000 signatories to a letter expressing the position that we all need to take a step back from AI and figure out how to control it before proceeding. US Senator Durbin (R-IL) observed, according to the Washington Post, that it was historic that large corporations would come before Congress and plead for the Government to regulate them. Even the people who developed AI are nervous about what it can do.

Alarm Bells Continue to Sound

On Monday, April 7, 2024 Japan's largest tech company, NTT, and its largest newspaper published a manifesto calling for speedy legislation, warning that "Unless AI is restrained, in the worst case scenario, democracy and social order could collapse, resulting in wars." But, did anyone notice?

The Wall Street Journal

How Do We Answer "Should"

"We want to eventually transition parts of our tech to be stand-alone – to become fully automated and remove the doctor or the nurse from the loop", E. Ben Joseph, MD of Penn Medicine Oncology

Politico 2023

UC Davis Medical CEO, David Lombarsky, MD: stated during a discussion of how they are using AI that they have the following policy directive for their staff. "Doctors are to NEVER, ever rely on the Internet or ChatGPT (AI) for a decision – because they are just not reliable. His term for AI – Augmented Intelligence. It's a tool like any other to aid the Doctor's decision process.

Interview 2024

We Need AI Governance Now

AI governance refers to the guardrails that ensure an AI Model or system is developed safely, implemented with proper oversight and monitoring, and is ethical in its outcomes. If we are to overcome the "trust deficit" we talked about earlier we will need effective governance. Models will have to demonstrate transparency, explainability and accountability.

AI Principles

When evaluating AI it should exhibit:

- Empathy: Analysis of the model to understand social impacts
- Bias Control: Examine data sets and algorithms to ensure bias is not introduced deliberately or inadvertently.
- Transparency: How AI Models work, their logic, reasoning and outcomes are transparent/explainable.
- Accountability: AI models should provide a clear understanding of what and who went wrong.

Final Tips & Takeaways

- AI is like any other IT tool, just significantly more powerful
  - Strengthen your knowledge now
- Develop your AI governance board and policy now
  - Bring projects in line with your principles and goals
- Identify priorities for AI
  - Give generative AI time to mature
- Develop strategy to meet ethical challenges of AI
  - Enlist members from across organization to address

The Cyber Arms Race

Creating and implementing advanced technology is instrumental to a state's ability to gain advantage over adversaries, expand its reach, bolster its capabilities, and otherwise. And in the competition for digital supremacy, when it comes to this technology, AI is the icing on the cake. If a state has an offensive capability, it will leverage it to support its own objectives.

Emilio Iasiello, OODALOOP, 2024

Artificial Intelligence

Artificial Intelligence will transform our lives, and generative AI will drive that transformation. Time has always been a principle determining nemesis for cybersecurity, with AI it could be our greatest enemy.

Better start running Alice…

Thank you

Mac McMillan

979-933-5311

mac@mac-mcmillan.com