

Pitfalls in AI – UNGA78 2023

Anish KURIEN

Acting Director: TUT Hub of AI-ISA

15th September 2023



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



Introduction

- An article by McKinsey & Company highlighted that as with any transformative trends our society has faced, the rise of artificial intelligence has **both major opportunities** as well as **significant challenges**.
- It is further stated that AI has the **potential to boost** overall economic productivity significantly.
- Taking into account transition costs and competition effects, the potential of AI in various sectors is projected to **add nearly \$13 trillion** to the total output by 2030.
- **Boost global GDP** by about **1.2% per year**.¹

1 - The promise and pitfalls of AI, available online: <https://www.mckinsey.com/mgi/overview/in-the-news/the-promise-and-pitfalls-of-ai>



Some Potential Risks...

- **Deepfakes, Pseudo images & Misinformation**

- Initially created for entertainment purposes, deepfake technologies have been made publicly available through software such as FakeApp, Reface, and DeepFaceLab.
- The counter is the work being done to detect deepfakes (Ex. Intel's deepfake detector, work by DARPA to identify whether video or audio has been manipulated)²

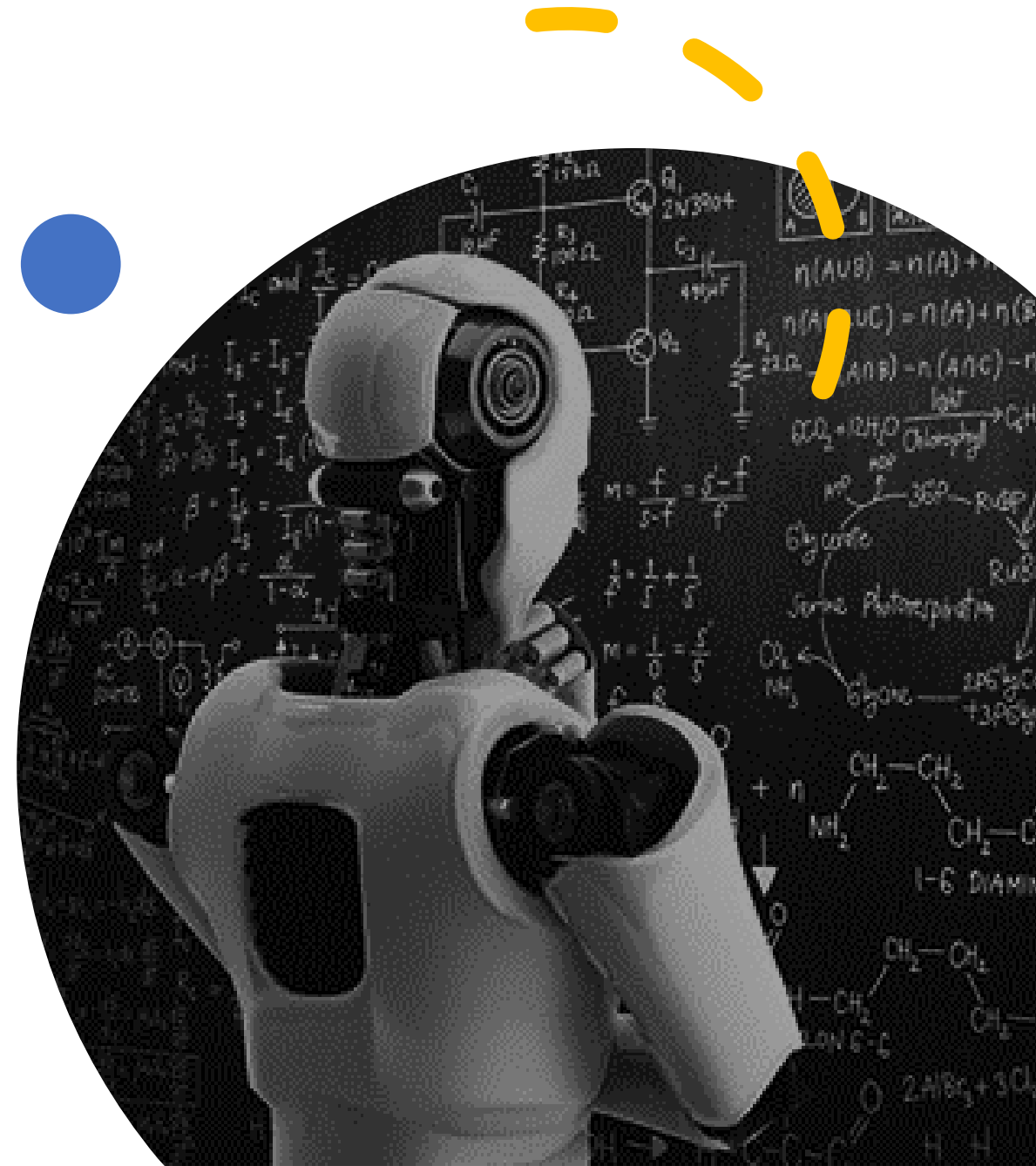


² - Bill Gates, The risks of AI are real but manageable, available online: <https://www.gatesnotes.com/The-risks-of-AI-are-real-but-manageable>

Some Potential Risks...

- **AI inherits our biases and makes things up**
 - When an AI confidently makes a claim that is not true due to the inability of a machine to understand the context of the request, creates “**Hallucinations**”²
 - AI models inherit whatever prejudices are baked into the text they’re trained on.
 - AI models can be taught to distinguish fact from fiction
 - (Ex. OpenAI work, work by the Alan Turing Institute and the National Institute of Standards and Technology working on the bias problem²)

² - Bill Gates, The risks of AI are real but manageable, available online <https://www.gatesnotes.com/The-risks-of-AI-are-real-but-manageable>



Some Potential Risks...

- **Students won't learn to write because AI will do the work for them.**
 - Many teachers are worried about the ways in which AI will undermine their work with students.²
 - Work by Cherie Shields in Education week – **“Teachers will have to embrace AI technology as another tool students have access to”**;
 - Acknowledging AI's existence and helping students work with it could revolutionize how we teach”.²
 - Current work ongoing to develop tools to determine whether something was written by a person or by a computer – determining “Original Work”

² - Bill Gates, The risks of AI are real but manageable, available online: <https://www.gatesnotes.com/The-risks-of-AI-are-real-but-manageable>



Some Potential Risks...

- **The AI revolution benefits are not likely to be shared equitably..**
 - This will lead to the reinforcement of the already widening digital divides that fuel economic inequalities globally.
 - Firstly, Innovative and leading-edge **companies** that are able to fully adopt AI technologies have the capacity to **double their cash flow** between now and 2030.
 - What about **those that cannot?**¹



1 - The promise and pitfalls of AI, available online: <https://www.mckinsey.com/mgi/overview/in-the-news/the-promise-and-pitfalls-of-ai>

Some Potential Risks...

- The AI revolution benefits are not likely to be shared equitably..
 - Thirdly, the **growing divide** among countries becoming more apparent (developed vs emerging countries)
 - The Developed world are able to establish themselves as AI leaders and better capture economic benefits compared with emerging economies.¹



1 - The promise and pitfalls of AI, available online: <https://www.mckinsey.com/mgi/overview/in-the-news/the-promise-and-pitfalls-of-ai>

The TUT Hub of the AI-ISA

- The Presidential Commission on the 4th Industrial Revolution (PC4IR) was established to provide a guide on how South Africa as a country can *take advantage of the emerging digital technologies at a global scale*.
- One of the key recommendations of the PC4IR report the *establishment of an Artificial Intelligence (AI) Institute*;
- The mandate of the institute must 1) be embedded *within the state*; 2) must *generate technology applications* in key government sectors; and 3) should *include training* to the whole of society.



Focus at the TUT Hub of AI-ISA

- The main focus of the Tshwane University of Technology Hub is centered on *seven areas* that it has expertise on.
- The projects are being driven to be matured in order for them *to have national and global impact.*
- The proposed domains are considered as the *starting point* to define the areas of focus for the TUT Hub.
- Each domain is being further developed to align them *to catalytic projects defined by Government in the national strategy.*



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



Focus of the TUT Hub of AI-ISA

Item No	Focus Area
1	AI for Automotive Sector
2	AI in Transport (inclusive of the usage of Drones)
3	4IR in Manufacturing (value-chain)
4	AI Tourism (inclusive of Virtual Reality)
5	AI in Farming and Food processing (Agriculture 5.0)
6	AI for Telecommunications
7	AI in Healthcare

10



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



Linkages to Government Catalytic Projects

- The linkages of the above focus areas of the TUT Hub of the AI-ISA to **Government Identified Catalytic Projects** is provided below

Government Approved Catalytic Projects	TUT AI Hub Focus Areas
Project 2: AI Capacity Building for Public Servants	AI Capacity building for public servants in all seven (7) Hub 2 focus areas
Project 3: AI Motor Industry Infrastructure Enhancement Programme	<ul style="list-style-type: none"> - AI in the Automotive sector. - Transport (including the use of drones) - 4IR in Manufacturing (value-chain).
Project 6: Modernising Public Services	<ul style="list-style-type: none"> - AI in Transport (including the use of drones) - Tourism (including Virtual Reality) - AI for Telecommunication operators and carriers - AI in Healthcare ¹¹
Project 8: AI In Farming and Food Production	- AI in Farming/Agriculture and food processing (Agriculture 5.0)



The TUT Hub of the AI-ISA - Farming

- The **Intelligent chemical spraying/liquid fertilizer** project focuses precision operations and farm management that uses intelligence-based recommender systems using drones and artificial intelligence approaches.
- The **Cell phone-based intelligent plant diagnosis system** uses cell phone images to identify plant diseases using image processing and classification with robust artificial intelligence algorithms.



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



The TUT Hub of the AI-ISA – Healthcare (1)

- The **Emotion Computing Project** focuses on the detection of human emotions using multimodal-fine-grained emotion recognition based on deep learning, which considers different modalities and the geometry of data to develop more efficient learning algorithms for pattern recognition of human emotions, activities, and behaviour.



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



The TUT Hub of the AI-ISA – 4IR in Manufacturing (1)

- The **Gamified learning project** prepares manufacturing enterprises for the evolving industrial revolutions by bringing interactive ways of learning to key performance indicators and resources using Neural Networks.
- X Reality Lab's **Assistive AI learning project** uses AI fortified Augmented Reality, Virtual Reality, Mixed Reality and Other Realities to enhance worker learning, relearning, and continuous-learning.



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



Conclusion

- As the scope of its impact on society continues to unfold, business and government organizations are *still racing to react to its introduction*
- It is important that we assess the ways in which *AI can be adopted in organizations* and to *prepare them for a future* that is already here.
- There is a *need to re-think the use of AI* in organizations and to consider that they become a core within organizations³.

3 – *What is generative AI, what are foundation models, and why do they matter?* available online: <https://www.ibm.com/blog/what-is-generative-ai-what-are-foundation-models-and-why-do-they-matter/>



Thank you!
kurienam@tut.ac.za



**communications
& digital technologies**

Department:
Communications & Digital Technologies
REPUBLIC OF SOUTH AFRICA



**Tshwane University
of Technology**

We empower people



**UNIVERSITY
OF
JOHANNESBURG**