Technology and Security

A general overview

Jerry Wang & Jonathan Swart

Agenda

- Introduction
- What is AI?
- The Threat of Algorithms and Terrorism
- Classifying AI Threat Types
- Fact or Science Fiction?
- Conclusion (Kahoot)

Introduction

Biggest concern of AI Usage by Business Leaders in AI Sector

1. AI being used for malicious reasons

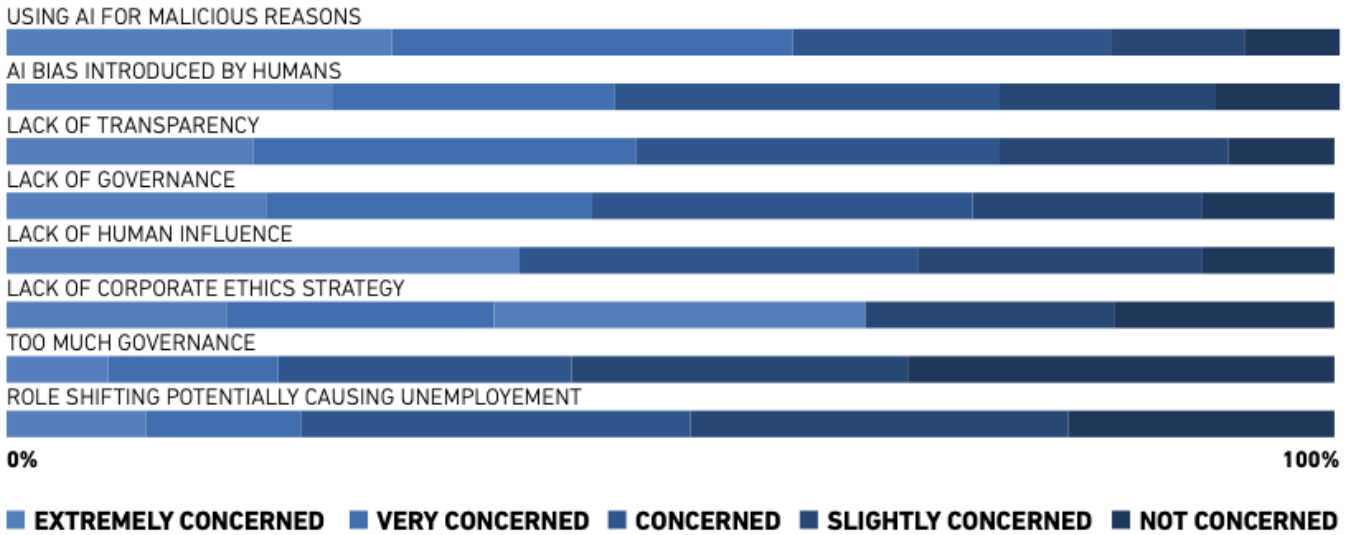
Concerns about AI falling into the wrong hands (e.g. Terrorist Groups)

2. AI bias introduced by humans

it can perpetuate societal inequalities, lead to discriminatory outcomes, reinforce harmful stereotypesand stifle innovation

3. Lack of transparency

concerns about accountability, making it difficult for users and even developers to understand how decisions are made and to identify and correct biases or errors in the algorithms



Introduction

United Nations Counter-Terrorism Centre (UNCCT)

Survey carried out to determine the concern from UN-Experts about the threat of malicious usage of AI

44%

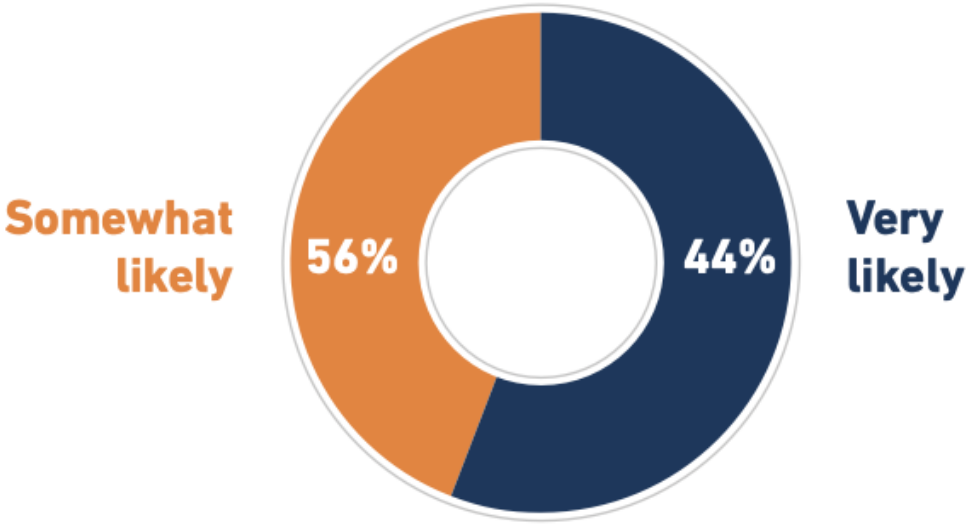
Find it very likely that AI will be used for malicious purposes

56%

Find it somewhat likely that AI will be used for malicious purposes

0%

Believe that AI will never be used for malicious purposes



Introduction

What are the main reasons for the concern?

Democratization of AI

Once an exceptionally advanced Tech sector with very few users but now accessible for all humans, no matter their degree, expertise or motive

Asymmetry in Terrorism

"we have to be lucky every time, they only have to be lucky once"

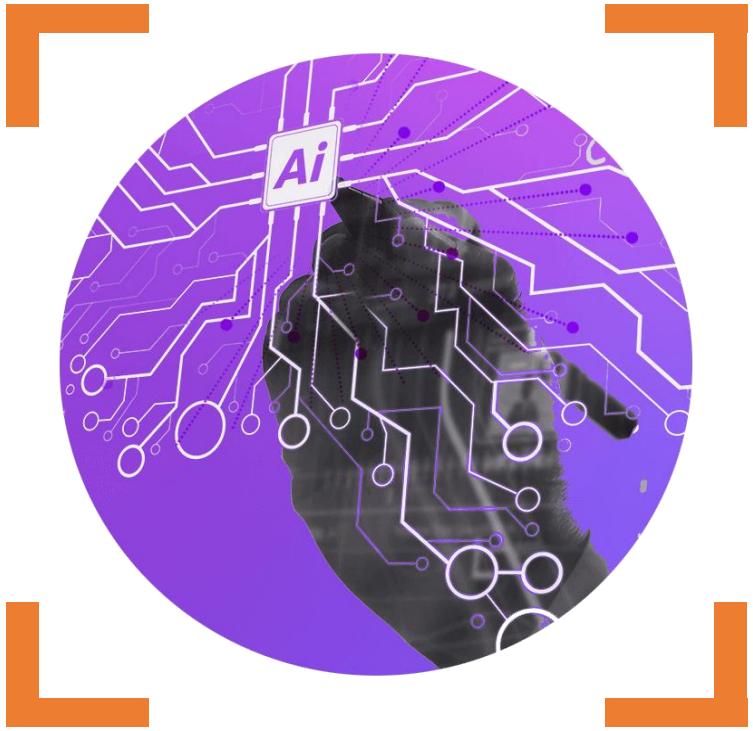
Scalability of AI

Attacks can be automated and deployed on a larger scale with minimal manual intervention

Growing society dependence

Growing society dependence on AI leads to more vulnerability like in healthcare, energy providers or biological and nuclear facilities

What is AI?



Computer systems able to perform tasks normally requiring human intelligence

- Visual perception
- Speech recognition
- Translation between languages
- Decision-making
- Problem-solving

- Software, robots, self-driving cars

What is AI?

Machine Learning

Deep Learning

- ConvNets, NLP, GAN, RNN

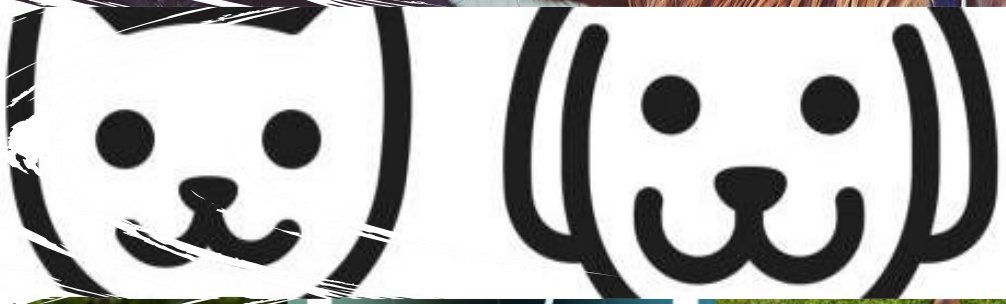
Machine Learning

Allows computers to learn from data without being explicitly programmed

1. Data Collection
2. Training
3. Model Creation
4. Prediction
5. Testing and Improvement

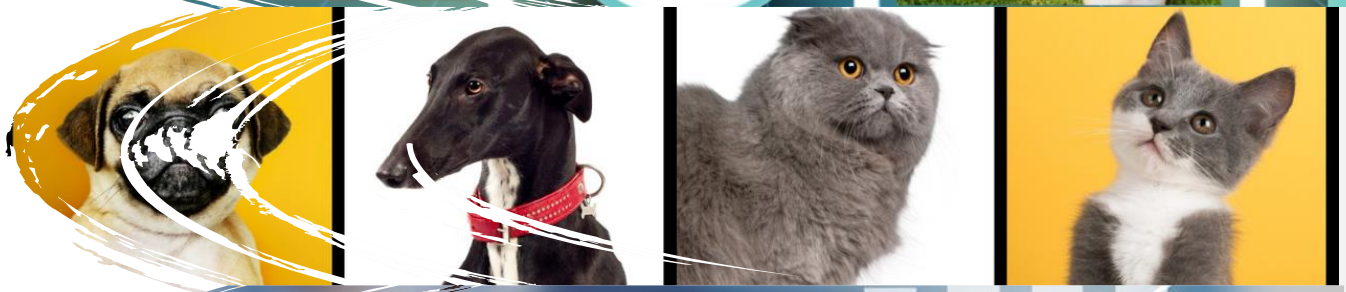
1.

2.



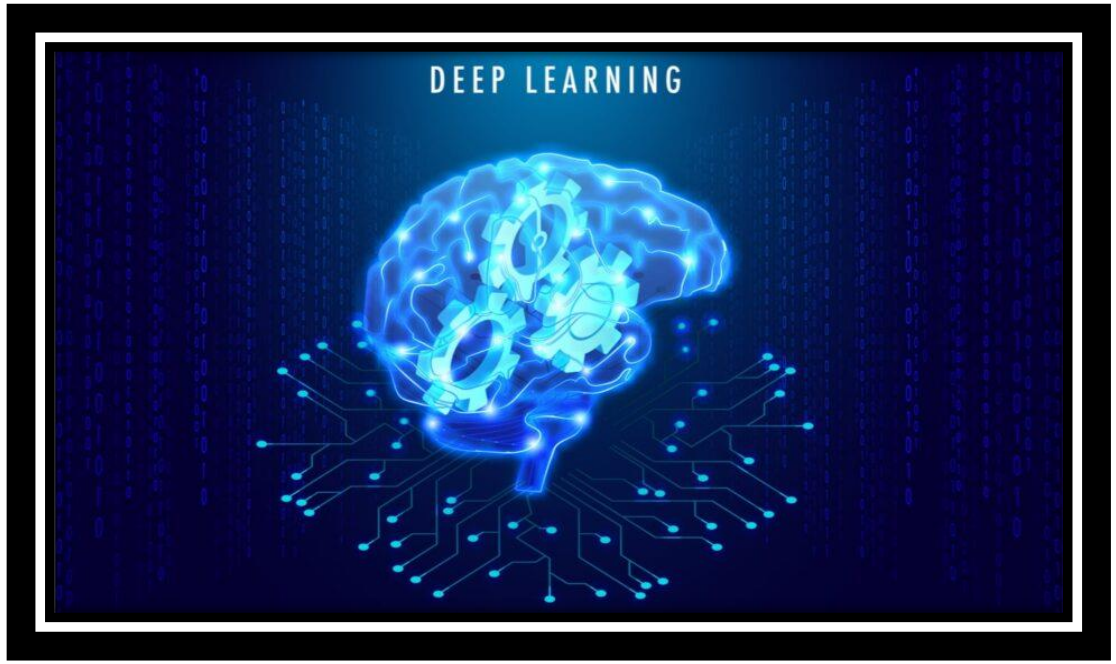
3.

4.



5.

Deep Learning



- **Sub-field of Machine Learning**
- Inspired by the human brain – several layers of the neural network
- Seek to learn from large amounts of data by performing a task repeatedly
- Each time making minor modifications to its internal features to improve the outcome.
- Powerful and good at handling complex data

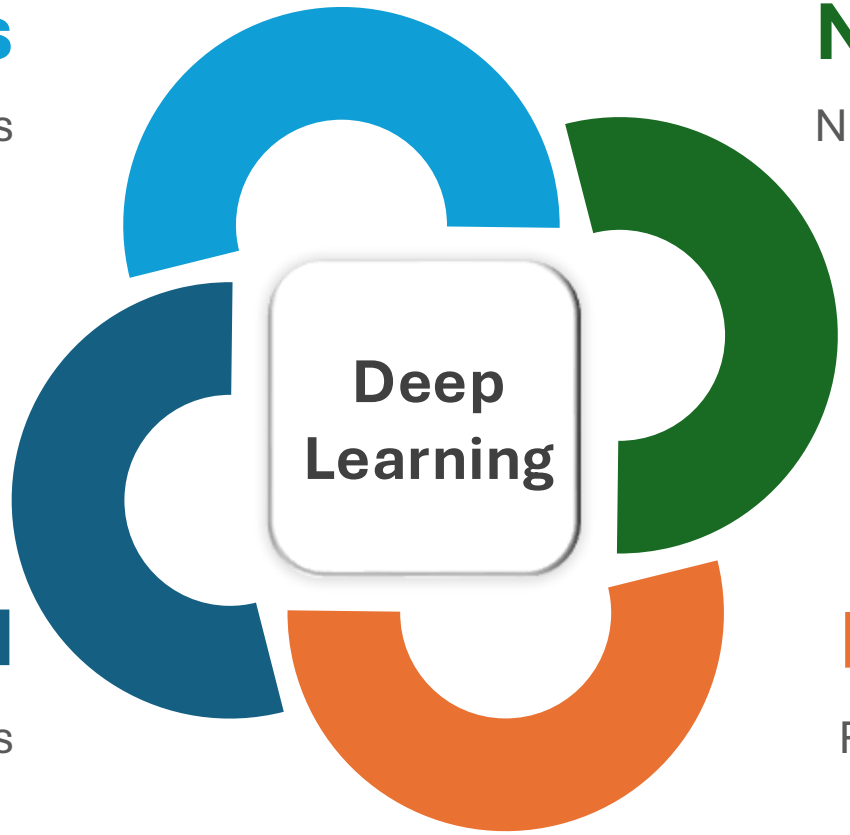
Deep Learning Architecture

CNNs/ConvNets

Convolutional Neural Networks

NLP

Natural Language Processing



GAN

Generative Adversarial Networks

RNN

Recurrent Neural Networks

ARTIFICIAL INTELLIGENCE

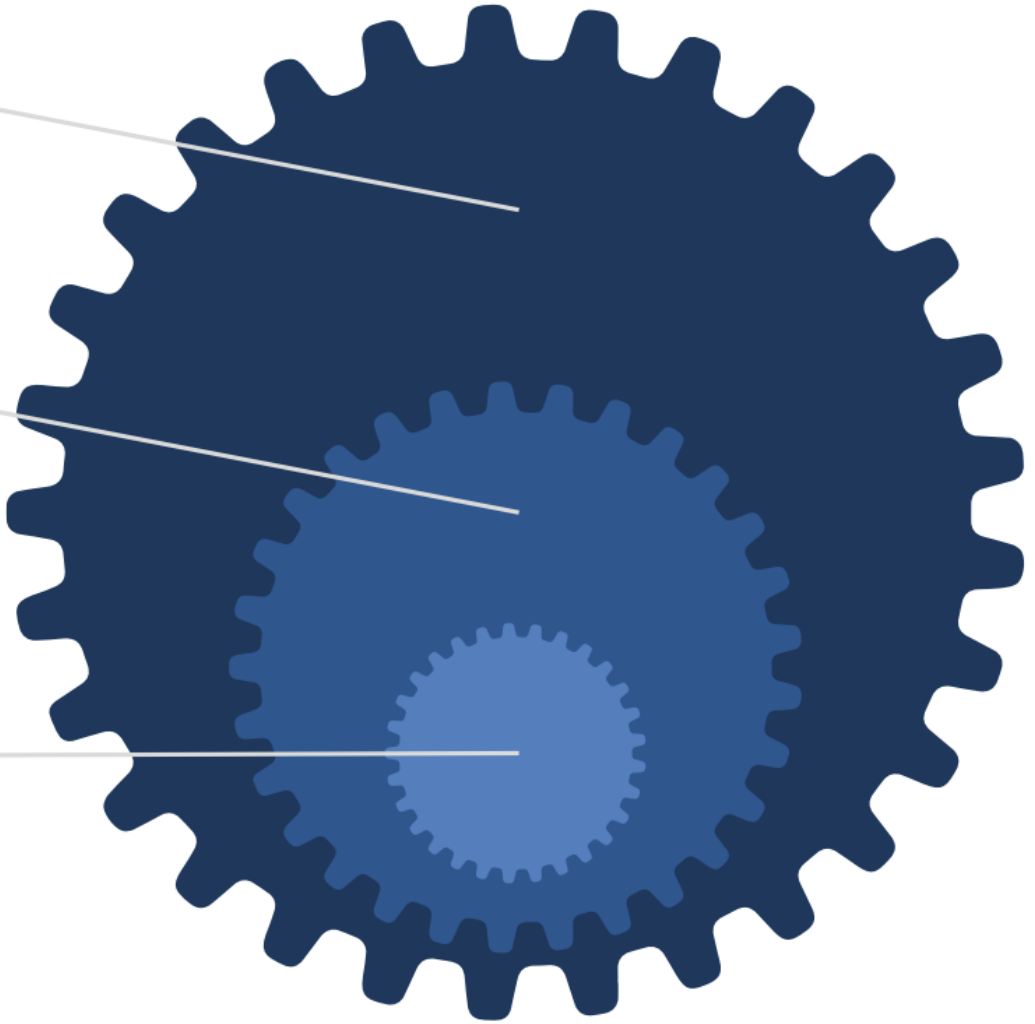
A program that can sense, reason, act and adapt

MACHINE LEARNING

Algorithms whose performance improves as they are exposed to more data over time

DEEP LEARNING

Subsets of machine learning in which multilayered neural networks learn from vast amounts of data



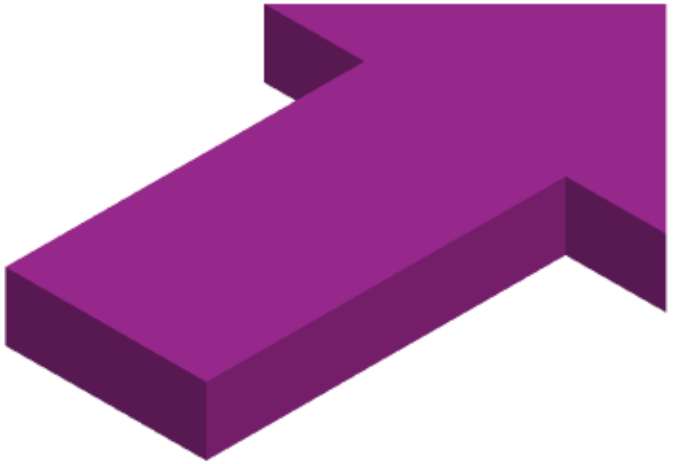
Types of AI systems

Narrow AI

Artificial General Intelligence (AGI)

Artificial Super Intelligence (ASI)

An Evolving Technology

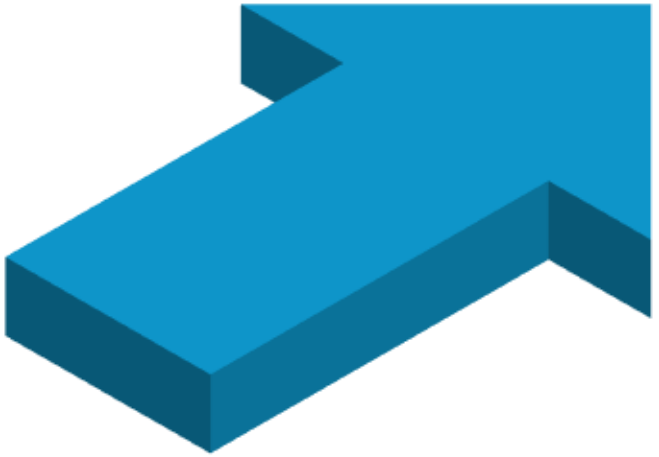


Rapid Expansion

Finance, medicine

Drones, self-driving cars
Art, Chess

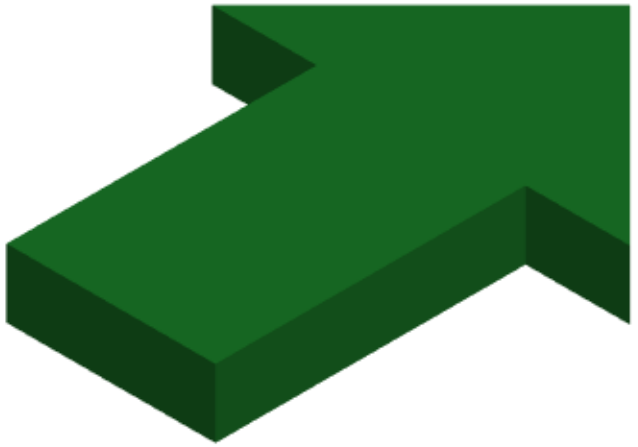
Even predicting Judicial Decisions



Machine Learning

More sophisticated algorithms

More training data



Robotics

Enhanced senses and dexterity

Take over cognitive tasks

Threat of Algorithms and Terrorism

Terrorism via Internet and Social Media

Past history with the internet and social media

- Incite violence, claim responsibility for attacks, raise funds, purchase weapons, make tutorials for members, etc.
- Europol's Referral Action Day, Facebook

Terrorists' Adaptations

- Encrypted platforms eg. Riot, Signal, Wire, etc.
- Bypassing normal detection methods

Technology as a Weapon

- Developed significantly
- Knives ➡ rifles ➡ vehicle attacks ➡ biological weapons

Technology for Logistics

- Transportation and Logistics
- Helps scale terrorists' operations

Technology for Communications

- Faster and more covert
- Spread videos and information to foster terror
- Recruitment and efficiency
- Phones, internet, social media and the dark web

Technological Advancements

- Preference for accessible, DIY, low technicality tools
- GPS, Internet, Bitcoin, the dark web
- Recent Advancements eg. Drones
- AI as an Instrument of terrorism

Classifying AI-Threats

Novel Challenges can arise in all AI lifecycles

01

Design of AI

Ethical considerations in the developement of AI systems

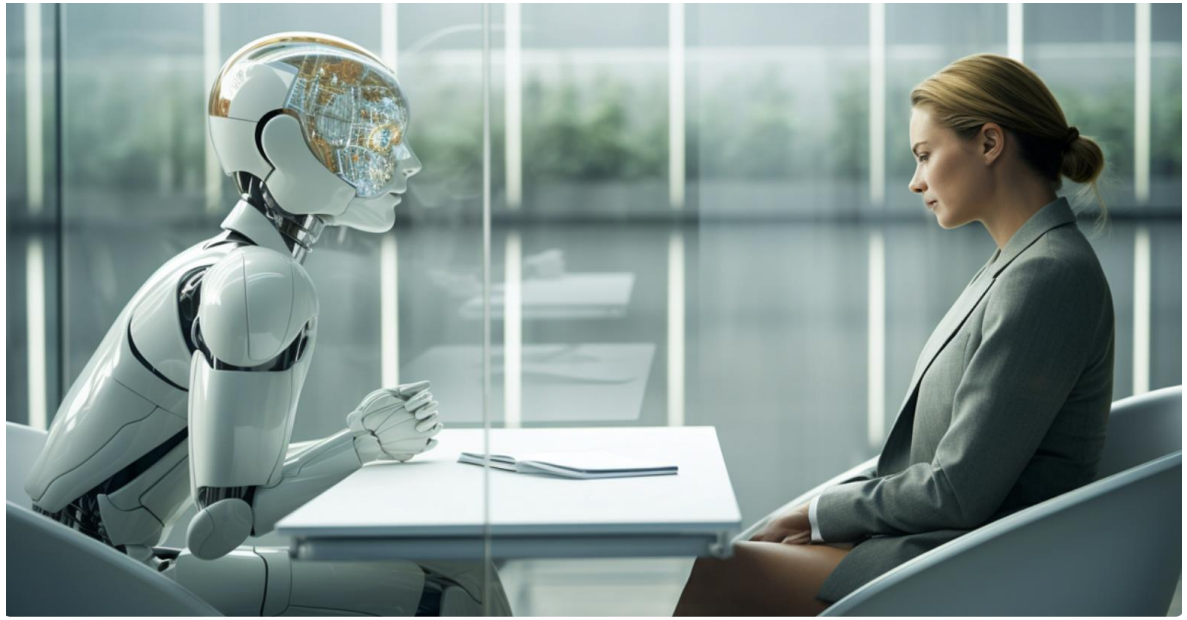
02

Deployment of AI

Challenges faced in the application of AI

Challanges faced by legitimate actors

- For legitimate actors using AI, the main concern is the potential for human rights infringements, including violations of the right to privacy, equality, gender equality, and non-discrimination,
- Can happen intentionally and unintentionally → such as through the use of unconsciously biased data to train machine learning algorithms, resulting in unfair decisions and discrimination.



Classifying AI-Threats

- Cyber Threats
- Physical Threats
- Political Threats

Classifying AI Threats

1. Cyber Threats @

The growing area of concern in cybersecurity stems from the inherent vulnerabilities in cyberspace and the asymmetrical nature of threats posed by cyber attacks, where smaller, less-resourced entities can effectively challenge more powerful adversaries using cyber means, mirroring the dynamics of asymmetrical warfare traditionally characterized by significant disparities in military capabilities and strategies.

Examples

Phishing

involves tricking individuals into providing sensitive information, such as passwords or credit card numbers, typically through deceptive emails or websites that appear legitimate

Man-in-the-middle

attacker secretly intercepts and relays communication between two parties, often to steal sensitive information or manipulate the data being transmitted without either party realizing it

Ransomware

malicious software that encrypts a victim's files or locks them out of their system, demanding a ransom payment to restore access to the data, often causing significant disruption and financial loss

DDoS attacks

malicious attempts to overwhelm a target's online services by flooding it with traffic from multiple compromised systems, rendering the service unavailable to legitimate users

Classifying AI Threats

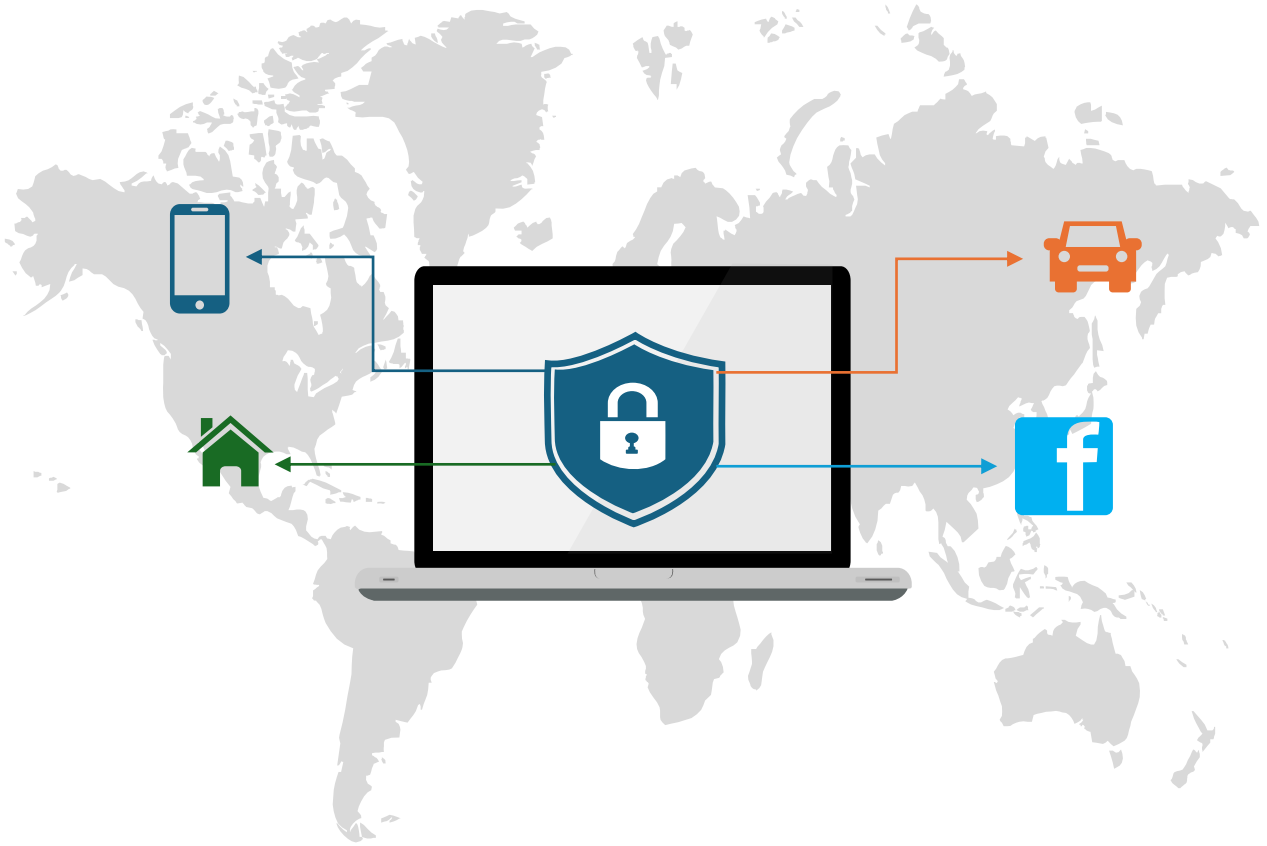
1. Physical Threats

The Internet of Things (IoT)

Increased interconnection within in the human infrastructure
(e.g. self-driving cars, drones or Smart home devices)

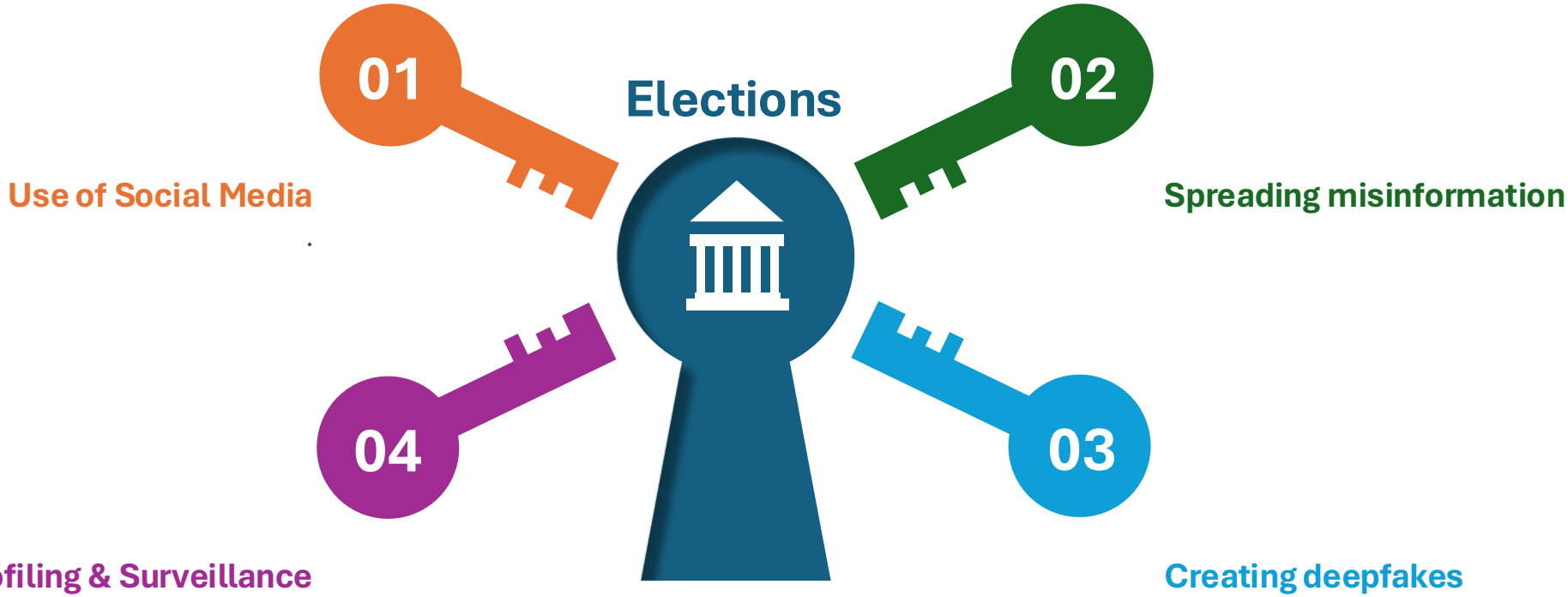
Risks

Highly increased potential for attacks in within sensitive areas



Classifying AI Threats

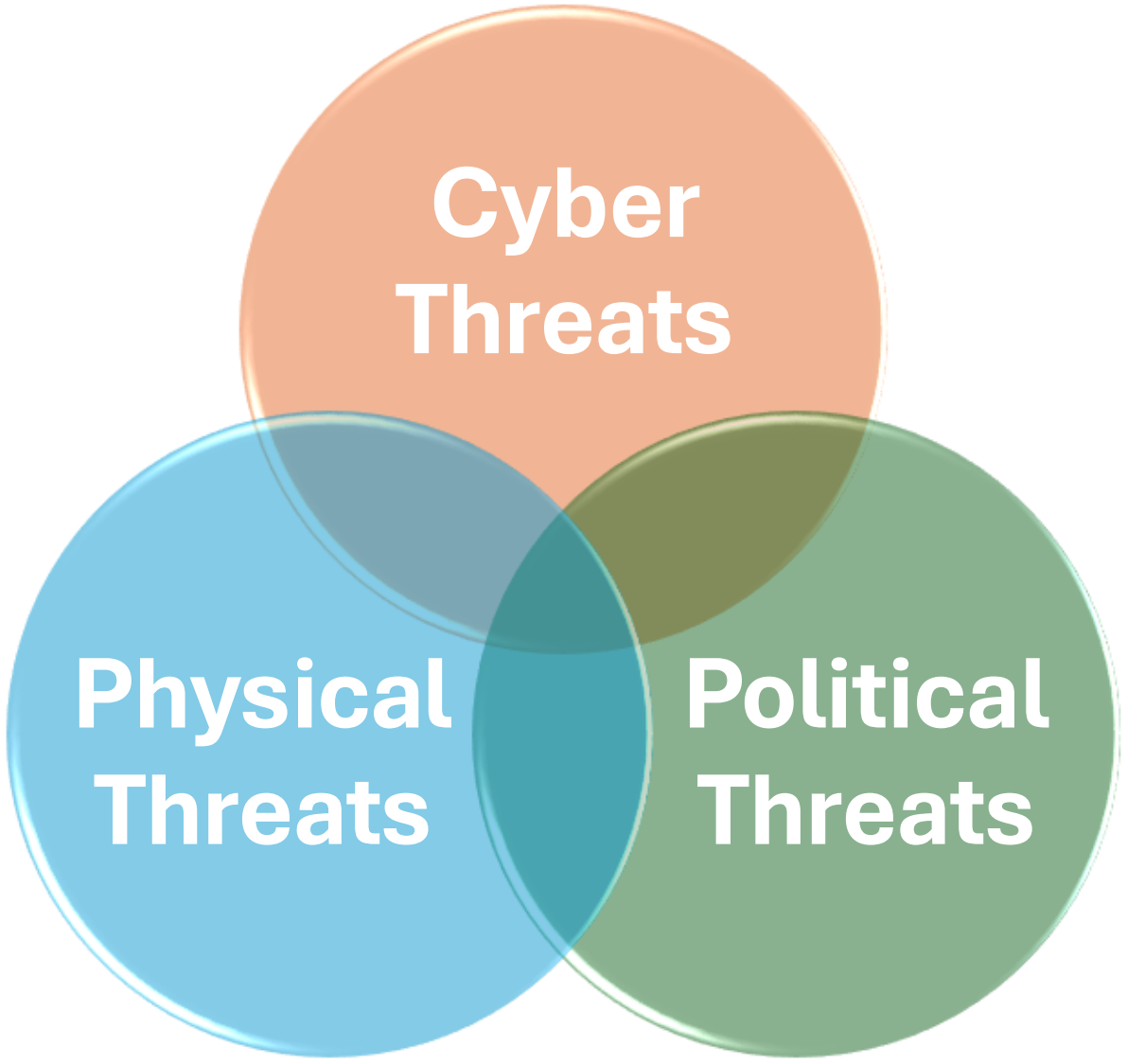
3. Political Threats



Classifying AI Threats

Relation between the Threats

These categories are interconnected, as AI-enabled hacking can target cyber-physical systems, leading to potential physical harm, while cyber and physical attacks may also be executed for political objectives.



Fact or Science Fiction?

Intentional Usage of AI

2020: Report by ISIL/AL-Qaida experts says that there no actual evidence that AI has been actually used by terrorists

2024: still no evidence of intentional use of AI for terrorist purposes

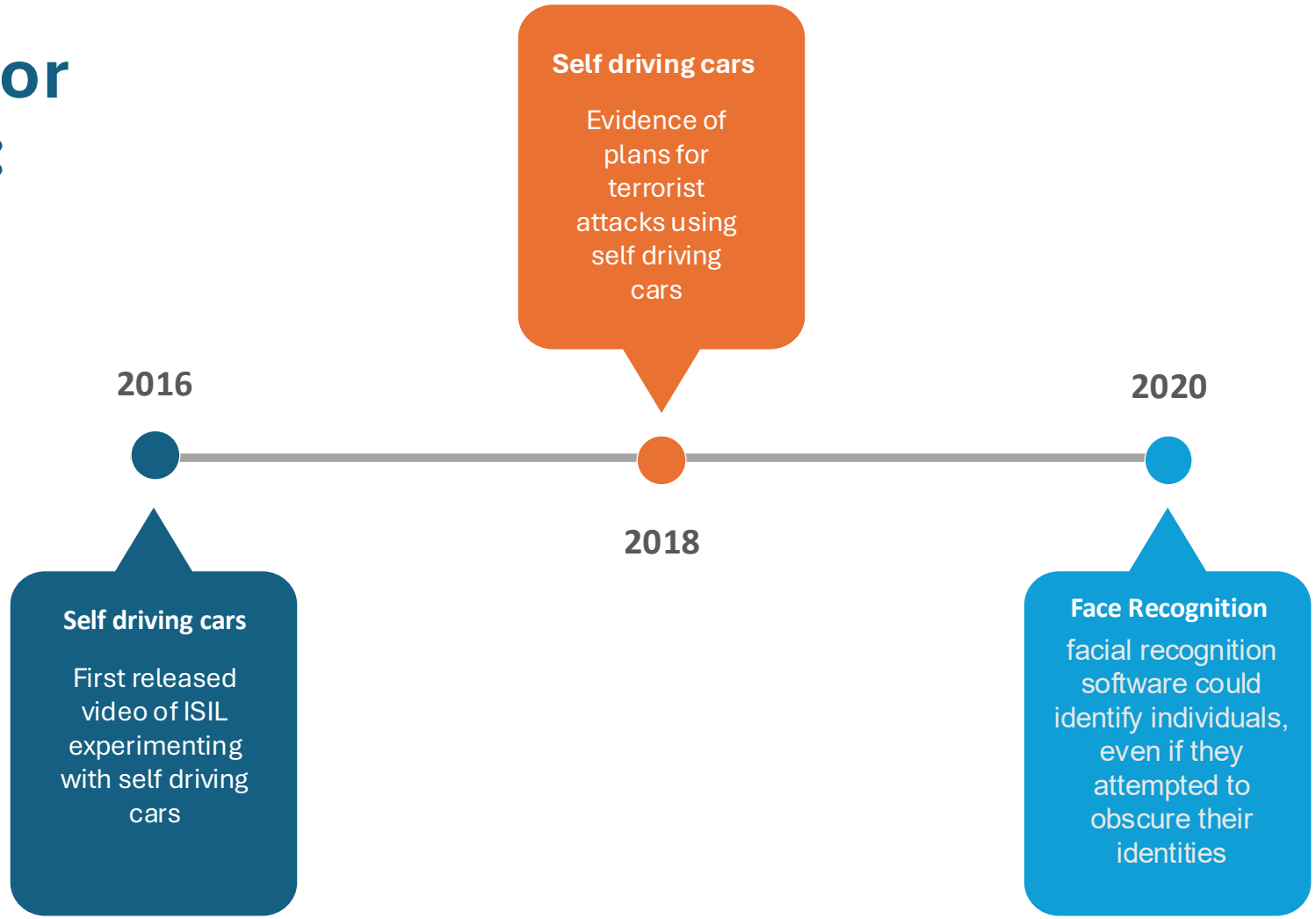
Indirect Usage of AI

Terrorist groups are leveraging AI systems created for illegitimate purposes without fully grasping their implications and may exploit existing AI technologies, such as surveillance systems, to facilitate the planning or execution of attacks, even if they do not directly use AI in their operations

There is however evidence, that these Groups have significant interest in using AI for malicious reasons

Fact or Science Fiction?

Timeline of AI use for malicious reasons:



Fact or Science Fiction?

Unmanned aerial systems (Drones)

- Unmanned aerial systems count as AI beacuse of their cetrain degree of autonomy (e.g. Global Navigation Satellite System (GNSS) which allows them to detect and avoid features)
- The Usage of Drones for malicous purposes dates back to 1995

UN: Current areas of use of Drones

01 Actual and attempted attacks

02 Disruptions

03 Surveillance and Propaganda

Fact or Science Fiction?

Examples for the Usage of Drones

1. Terrorist attacks

2016: drone equipped with explosives kills two kurdish peshmerga fighters and wounded two french soldiers

2017: drones killed/wounded 39 soldiers in one week

2017: Drones spread propaganda papers calling for terrorist attacks

2. Other non-state actors

2018: Assassination Attempt on Nicolás Maduro

2019: Attacks on Saudi-Aramco

Fact or Science Fiction?

Why are drones of such high interest?

- Commercial availability
- Affordability
- Convenient operation of Drones

Danger of these facts?

Apart from the dangers already mentioned, drones offer the possibility of so-called swarm attacks

Fact or Science Fiction?

Measures to reduce Drone flow

1. UN Security Council Resolution 2370 (2017)

urging the Member States to take greater collective effort to prevent terrorists from acquiring weapons

2. Global Counterterrorism Forum (GCTF)

Berlin Memorandum on Good Practices for Countering Terrorists Use of Unmanned Aerial Systems in 2019 providing guidance for the: identification, develop- ment and refinement of policies, practices, guidelines, regulations, programmes and approaches for countering the terrorist use of drones

Still, terrorist organizations rarely use drones, which remain largely unsophisticated and dependent on human control, with little evidence of AI-enabled drones being utilized or sought after

Fact or Science Fiction?

The Usage of AI Drones in Ukraine

1. Manual use of Drones

30-50%

Target strike rates for first-person view (FPV) units have fallen to between 30% to 50%. New pilots experience even lower hit rates, around 10%.

2. Predicted improvements with AI-Drones

80%

AI-operated FPV drones are predicted to achieve hit rates of around 80%, significantly improving effectiveness in targeting.

How does the System work?

The software allows the pilot to select a target via the drone's camera, and the drone autonomously completes the flight to the target

Reports indicate that the AI-guided systems have successfully destroyed three tanks and targeted logistics and field headquarters.

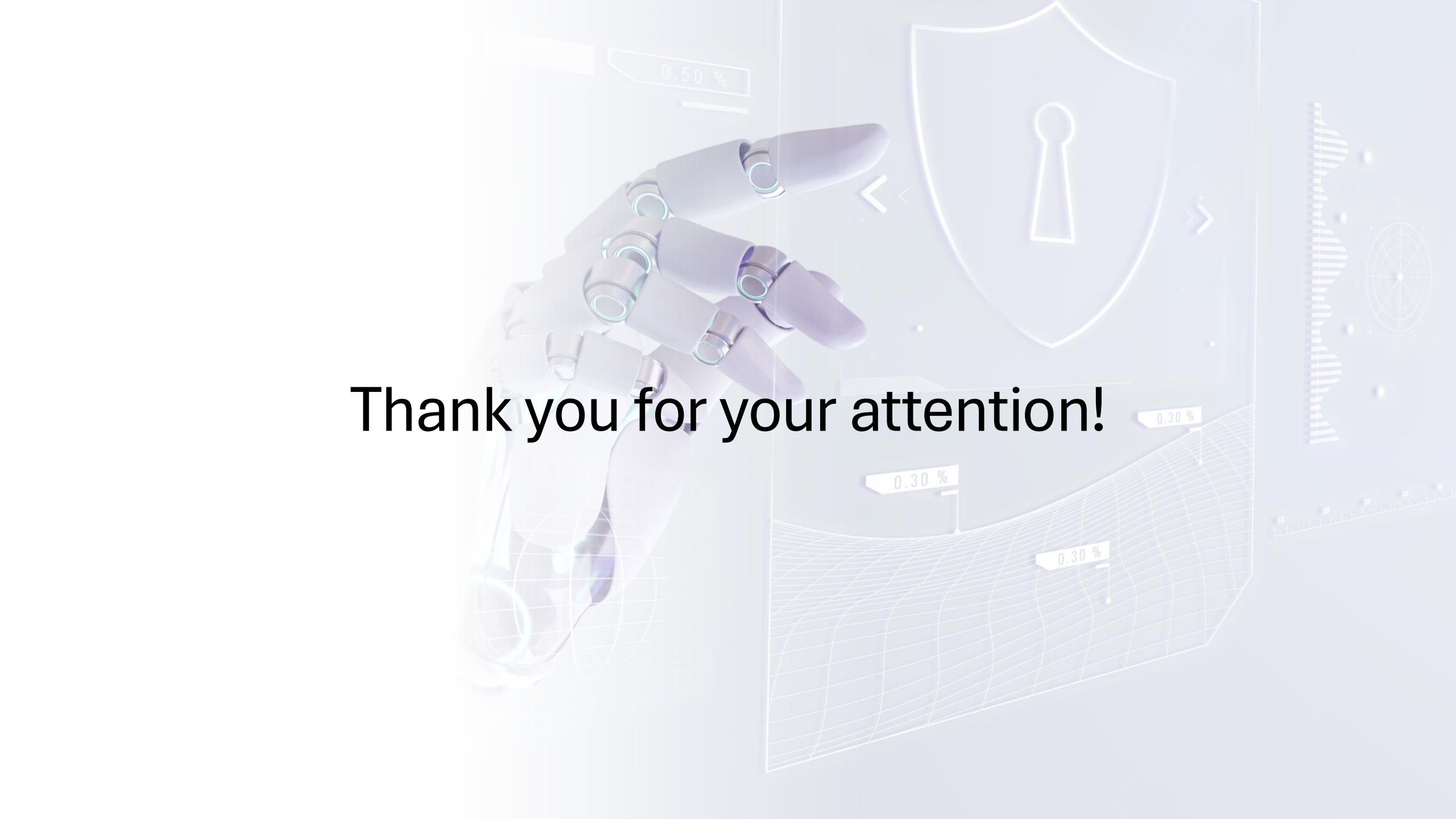
Consequences for Terrorist Groups?

The timeline for terrorist groups to adopt new technologies varies widely based on factors like the technology's nature and the group's capabilities, but historical evidence shows they often adapt quickly, especially when these technologies provide tactical advantages.

https://www.reuters.com/world/europe/ukraine-rolls-out-dozens-ai-systems-help-its-drones-hit-targets-2024-10-31/

Conclusion

- AI's two main subfields—machine learning and deep learning—are rapidly developing
- Terrorists have adapted very well to the rise of internet and new technologies
- The question is not if terrorists will start using AI, but when
- Growing Concern Over AI Misuse due to democratization and accessibility of AI
- Increasing dependence on interconnected systems expands the potential attack surface, making critical infrastructure more vulnerable to exploitation
- Especially drones, will even more likely become the future of AI warfare with alarming consequences
- To combat the evolving landscape of AI-related threats, it is essential to adopt proactive strategies, collaborate on best practices, and invest in advanced security technologies



Thank you for your attention!