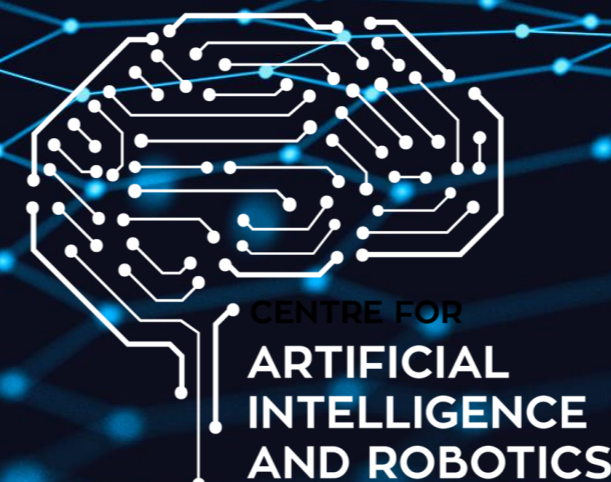


MALICIOUS USES AND ABUSES OF ARTIFICIAL INTELLIGENCE (AI)

David Sancho, Maria Eira, Aglika Klayn
Nov 2021



Background

Aim

- To provide an in-depth assessment of the present and possible future malicious uses and abuses of AI and related technologies
- To enhance awareness and preparedness to respond to such threats

Scope

- Malicious uses of AI
- Malicious abuses of AI
- Deepfakes case study

Methodology

- Research and contributions from the three entities
- Workshop held in March 2020



Malicious Uses and Abuses of Artificial Intelligence

Trend Micro Research
United Nations Interregional Crime and Justice Research Institute (UNICRI)
Europol's European Cybercrime Centre (EC3)



PART I

Present State of Malicious Uses and Abuses of AI

Present State of Malicious Uses and Abuses of AI

AI
Malware*

Abusing
AI Cloud
Services

Abusing
Smart
Assistants

AI
Password
Guessing

AI
CAPTCHA
Breaking

AI-Aided
Encryption
*

Trends Found on Underground Forums

Human
Impersonation on
Social Networking
Platforms

Social
Engineering*

Online Game
Cheats

AI Cryptocurrency
Trading

AI Hacking*

Examples from Underground Forums

NULLED
EXPECT THE UNEXPECTED

THE BEST SPOTIFY BOT (SHAMAM KODKODU) | 334K FOLLOWS | 12 FOLLOWS | NO BROWSER & DROPPER VERSION AVAILABLE

Spotify

Black Hat World
Black Hat Sea Forum

Home Forums Partnerships Members Account Upgrades Advertis Marketplaces

Follow Me As I Create An AI Trading Bot For Bitcoin / Crypto

Before crypto got big, my software engineering efforts were mainly focused on marketing. A couple years ago, I got to work for an ICO algorithmically inclined, went deep into trading maths.

I have made several successes using mostly straight-forward algebra, I've created several forecasters and advanced machine learning models I know of and applied them to the problem of trading. Crypto is highly volatile, it is for this reason that a bot should be able to trade.

I do have some proprietary efforts going on but I thought to myself "Someone has to put some sources out with a dam, why not me?"

Initial Thoughts

- + CCXT for OH-LV data & trade orders.
- + Websockets? Will look into it.
- + Most volume is probably at Binance exchange. Let's use it.
- + This will be done in Python of course.
- + Let's start with some LSTM variants. Differentiable Neural Computer maybe? hehe...
- + What about the optimizer agent? Q-Learning? Genetic? Bayesian? Let's try some.
- + Get a Github repo going. That will come first.
- + Google Colaboratory can help everyone follow along. Not everyone has a Titan GPU like me.

Let's do this thing.

GitHub Link: [https://github.com/BlackHattersWorld/BlackHattersWorld](#)

Thanks + 1

UNKNOWN CHEATS
LEADING THE BOMB HIGGING SCENE SINCE 2000

DOWNLOAD WARZONE CHEATS

I made the 1.6.8.0.0 file for the game

League AI: An AI Playing League of Legends using

And Me (DaxDax) on the Google SUIT

Romanian Security Team

WEAPONIZING AND GAMIFYING AI FOR WIFI HACKING: PRESENTING PWNAGOTCHI 1.0.0

This is the idea of a summer project that started out of boredom and that evolved into something incredible. Fun and unique. It is also the story of how that project went from being discussed on a patch by just two people to having a community made of almost 700 members (and counting) that gathered, polished it and released today's release possible.

FLORIN: You can download the 1.0.0.0 file from here, then just follow the instructions.

If you want the long version instead, of back view and enjoy the ride. Let me know if you're going to be using it for your usual blog posts, but it's not worth a (big) tip and for a (big) tip over here.

Hack the Planet!

PART II

Deepfakes

Deepfakes

01

Different types of deepfakes & the technology behind them

02

Current state of abuses of deepfakes & concrete examples

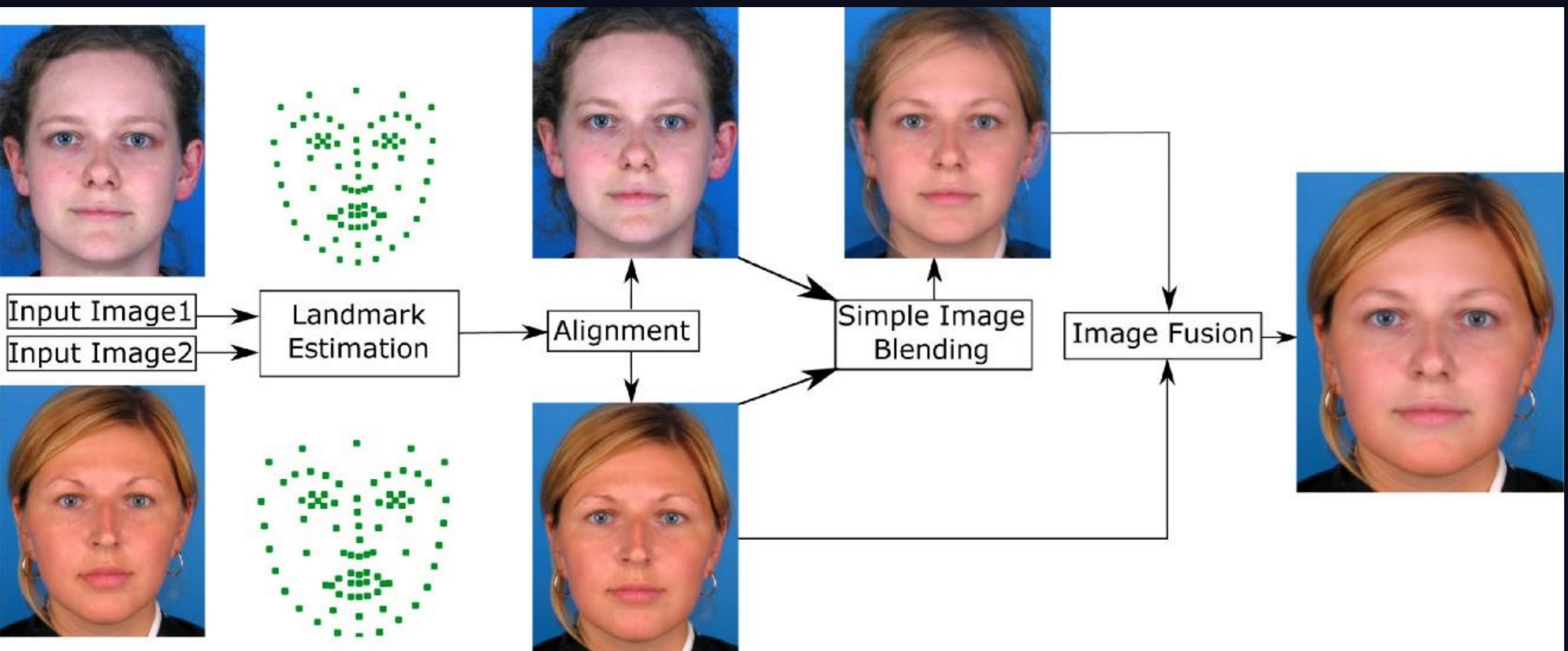
03

Possible future threats of deepfakes

04

Countering deepfakes approaches

Face Morphing and Fraudulent IDs



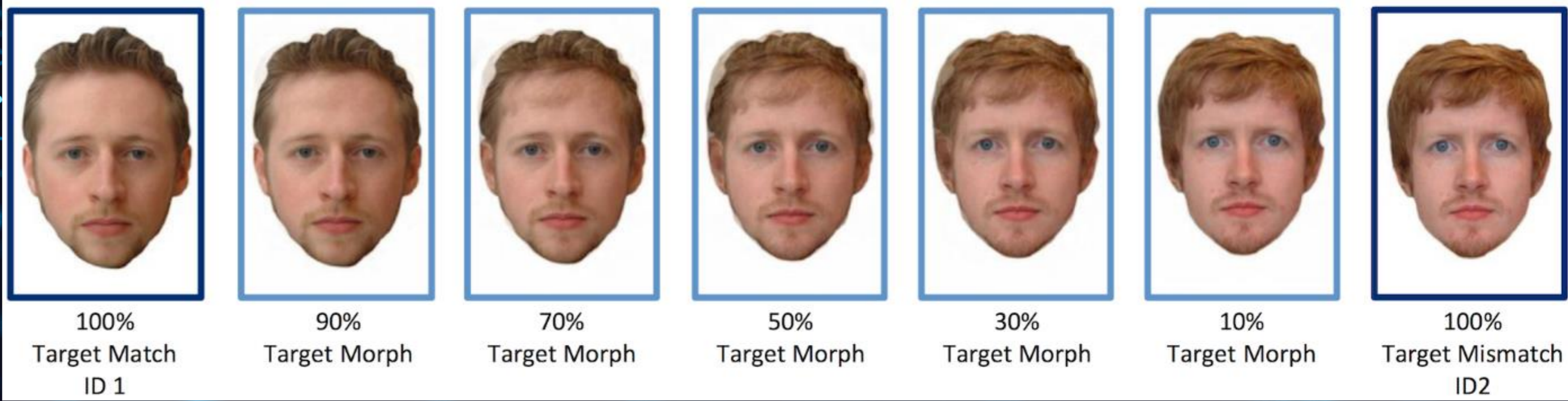
United Kingdom of Great Britain and Northern Ireland
 Passport
 Type/Type: C
 Code of Issuing State/Catégorie du titulaire: GBR
 Passport No./Passport No: P
 Surname/Nom (1):
 Given Name/Prénoms (2):
 Nationality/Nationalité (3): BRITISH CITIZEN
 Date of Birth/Date de naissance (4):
 Sex/Sexe (5): Place of Birth/Lieu de naissance (6):
 Date of Issue/Date de délivrance (7): Authority/Autorité (8): UKPA
 Date of expiry/Date d'expiration (9):
 Holder's signature/Signature du titulaire (10):

Press 1 for Match (Experiment 1 & 2)
 Press 2 for Mismatch (Experiment 1 & 2)
 Press 3 for Morph (Experiment 2)

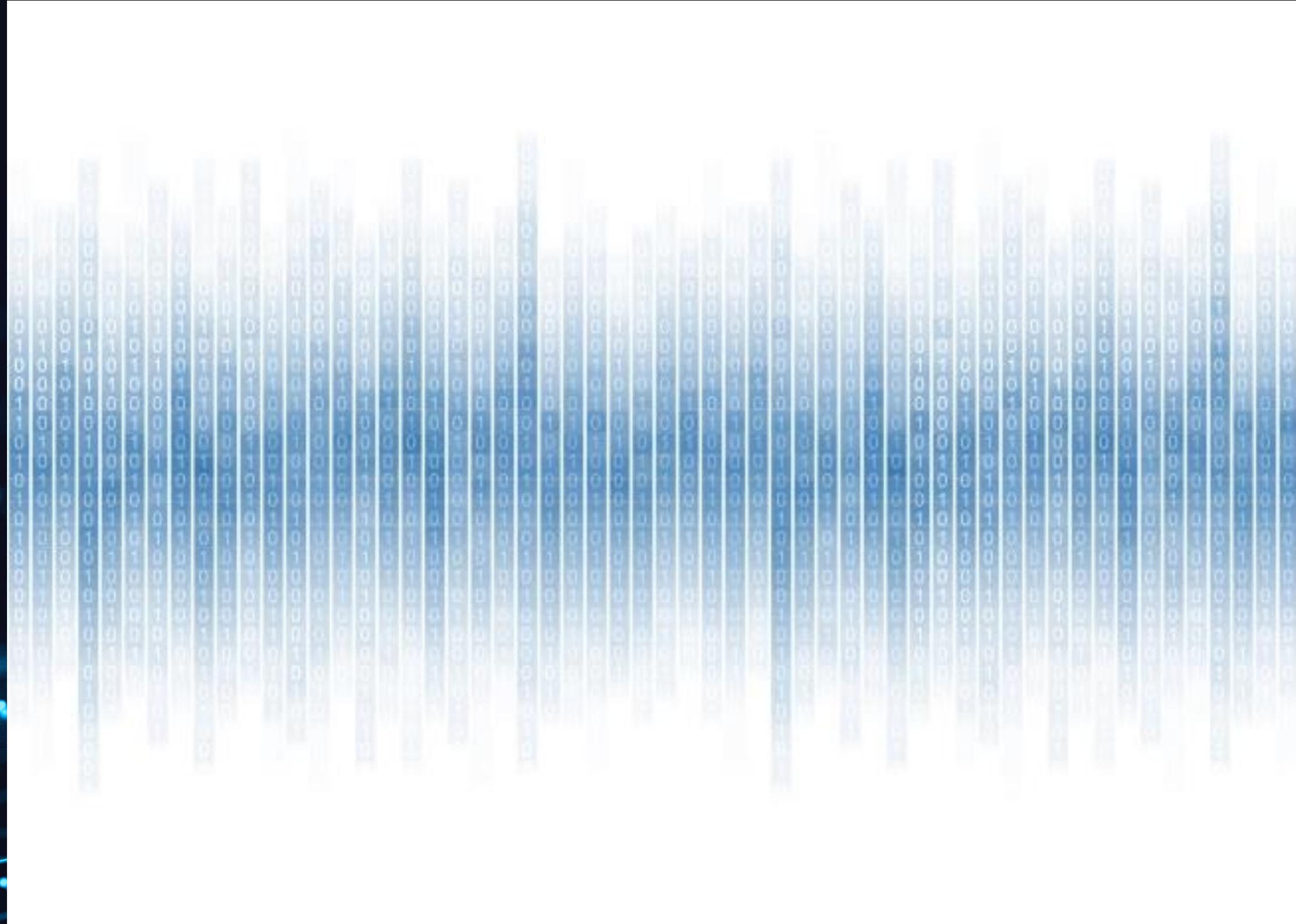
Passport Inspection Desk



Example Passport Photos



Deepfake Voice Fraud Case Examples



AI-based voice fraud leads to \$243,000 losses for energy company



AI voice cloning used in a \$35 million bank heist

Possible Future Deepfake Threats

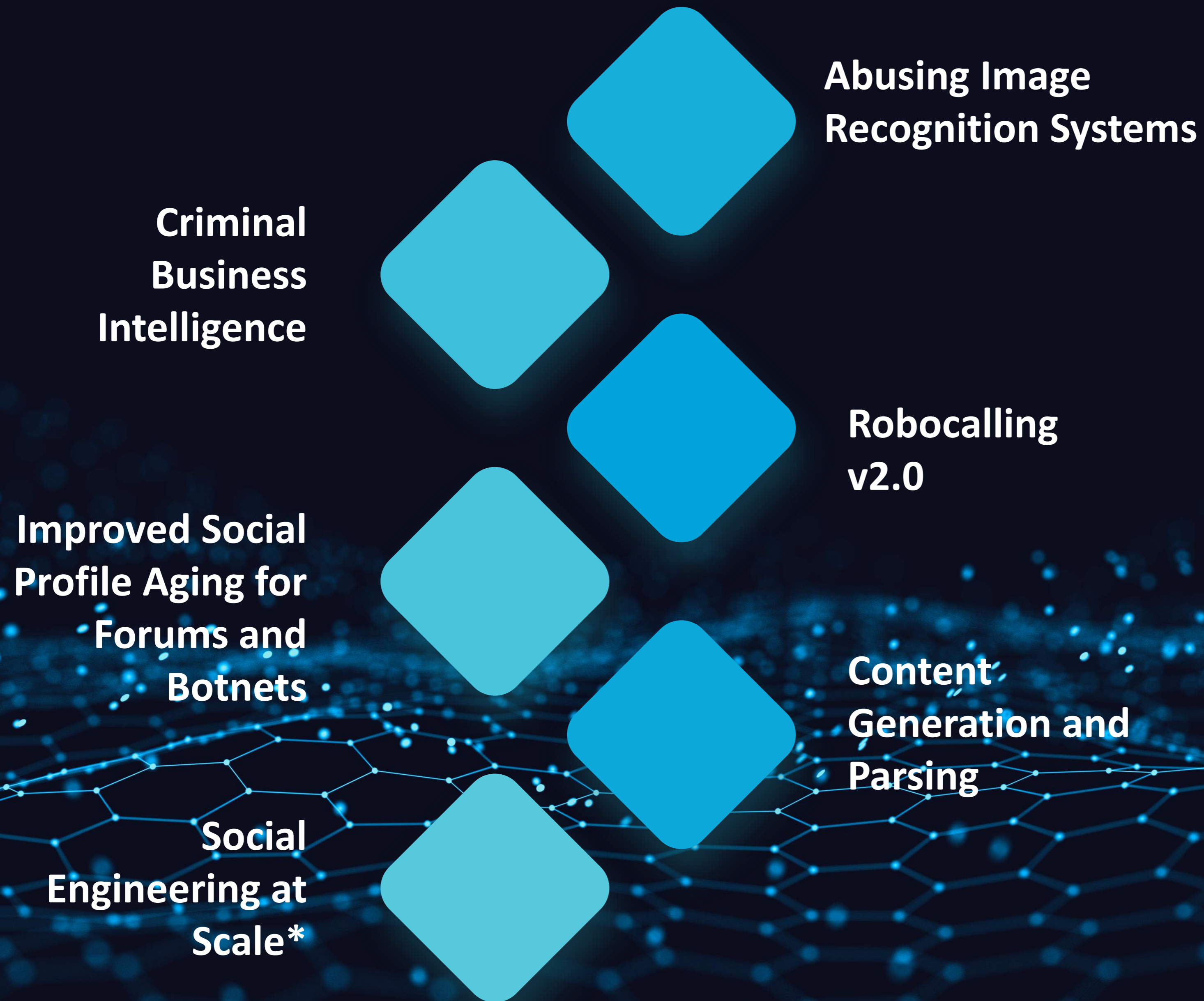


- Disinformation campaigns
- Securities fraud
- Extortion
- Online crimes against children
- Obstruction of justice
- Cryptojacking
- Illicit markets

PART III

Future Scenarios & Recommendations

Future Scenarios of Malicious Uses and Abuses of AI



**Escaping an
Image
Recognition
System**

**Remote
Machine
Learning Sets
Pollution**

**AI-Enabled
Stock Market
Manipulation**

**Business Process
Compromise**



**Insider Attacks
(Banking &
Trading Floor AI)**

**Local Library
Poisoning by
Resident
Malware**

**AI-Supported
Ransomware***

**Escaping AI
Detection Systems
(Fraud & Voice
Recognition in
Banks)**

Recommendations

AI for Good

- AI technology as a crime-fighting tool
- Responsible AI innovation
- Exchange of best practices

Further Research*

- Forward-looking threat assessments
- Continuous mapping of AI threat landscape
- Risk management driven threat classification & response

Secure AI Design Frameworks

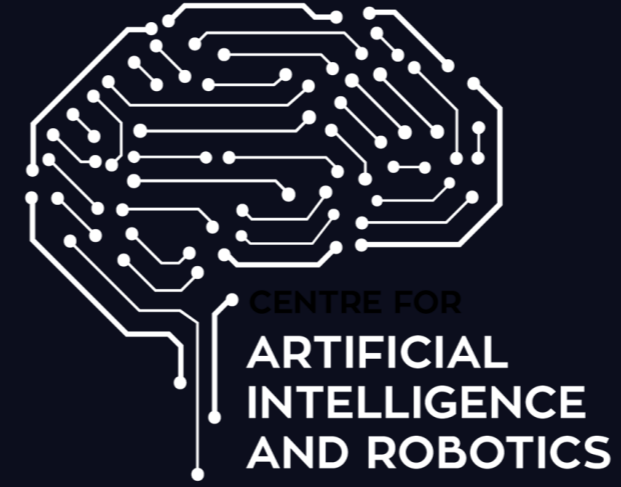
- Security-by-design & privacy-by-design
- Specific data protection frameworks
- Human-centric approach to AI
- Technical standards

Policy

- Well-informed public debate
- International technology-agnostic policies
- Long-term sustainability driven by effective oversight

Outreach

- PPPs & multidisciplinary expert groups
- Enhanced AI literacy & cyber hygiene
- Knowledge sharing on tools, tactics, & techniques



Thank you for the attention
Any questions?