The European cybersecurity landscape and threats

Weaponizing AI

Dr. George Sharkov
Assoc. Professor, Institute for ICT, Bulgarian Academy of Sciences
CEO, European Software Institute CEE
Head of Cyber Security & Resilience Lab (Sofia Tech Park)

gesha@esicenter.bg
www.iict.bas.bg
www.esicenter.bg

THE EVOLUTION OF DIGITAL DEPENDENCY:
SOFTWARE (AI/ML) NOT ONLY "EATING" BUT "PROGRAMING" THE WORLD

2011 THE WALL STREET JOURNAL.

Home World U.S. Politics Economy Business Tech Markets Opinion Arts Lif

ESSAY

Why Software Is Eating The World

By MARC ANDREESSEN
August 20, 2011

This week, Hewlett-Packard (where I am on the boa
jettisoning its struggling PC business in favor of inve
where it sees better potential for growth. Meanwhil
cellphone handset maker Motorola Mobility. Both m
both moves are also in line with a trend I've observe
about the future growth of the American and world
turmoil in the stock market.

In an interview with WSJ's Kevin Delaney,

In short,

More than
dot-com b
companie
sparking
their rapi
valuations, and even the occasional successful

2016

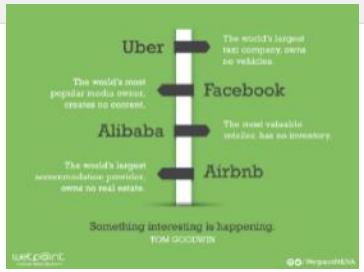
ANDREESSEN HOROWITZ
Software Is Eating the World

MACHINE & DEEP LEARNING

a16z Podcast: Software Programs the World

with Marc Andreessen, Ben Horowitz, Scott Kupor, and Sonal Chokshi

"All of a sudden you can program the world" — it's the continuation of the software eating the world thesis we put out over five years ago, and of the trajectory of past and current technology shifts. So what are those shifts? What tech trends and platforms do we find most interesting on the heels of raising our fifth fund? Are we just building on and extending existing platforms though, or will there be new platforms; and if so, what will they be? Well, distributed systems for one...



2018

WORLD ECONOMIC FORUM
COMMITTED TO IMPROVING THE STATE OF THE WORLD

Global Agenda Council on Risk & Resilience

Resilience Insights

1. Building Resilience to Water Crise
2. Building Resilience to Large-Scale Involuntary Migration
3. Building Resilience to Large-Scale Cyberattacks

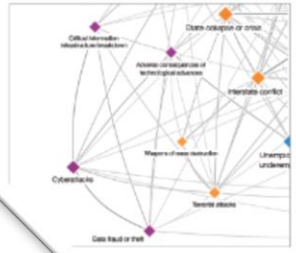
2021

"Cybersecurity failure" is one of the risks that worsened the most through COVID-19

2022

Cyber Resilience Pledge

Building Resilience to Large-Scale Cyberattacks



internet, automation of knowledge work, the Internet of Things and cloud technology will be the most disruptive".[50] While this innovation will result in new efficiencies and capabilities, it will also introduce new vulnerabilities, allowing attackers to quickly evolve their tactics and exploit unaddressed system and network weaknesses.

Further compounding the risk is today's hyperconnected global environment, where people and things, critical infrastructures and economies are increasingly digitally connected – anytime and anywhere. According to this year's Global Risks Report 2016, *"As the Internet of Things leads to more connections between people and machines, cyber dependency due to increasing digital interconnection of people, things and organizations –*

and ca
reconsi
resilien
require
advancing the understanding of and the disciplines that contribute to cyber resilience. This section posits a number of suggestions on how to improve the cyber resilience of organizations. Some require action by governments and some can be taken by all entities – public or private/big or small.

Recommendations

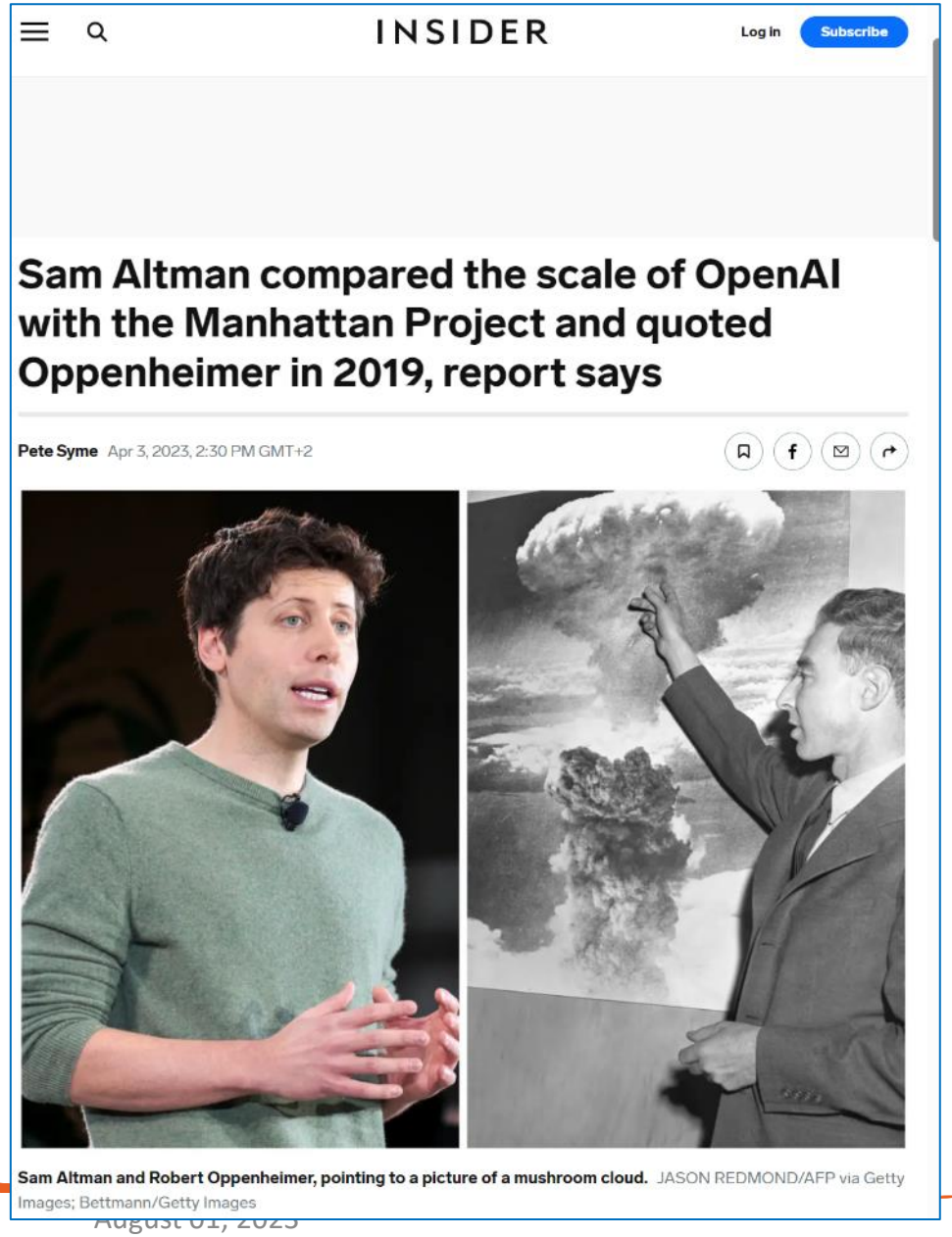
A. Increase Understanding of Risk of Large-Scale Cyberattacks and other Cyber Threats

As described above, it is clear that the dramatic pace of technological

... distributed systems — encompassing cloud and SaaS; A.I., machine learning, deep learning; and quantum computing — to the role of hardware; future interfaces; and data, big and small.

... why simulations matter... and what do we make of our current reality if we are all really living in a simulation as Elon Musk believes?

AI – NEW DISRUPTIVE OR DESTRUCTIVE TECHNOLOGIES?



INSIDER

Log In Subscribe

Sam Altman compared the scale of OpenAI with the Manhattan Project and quoted Oppenheimer in 2019, report says

Pete Syme Apr 3, 2023, 2:30 PM GMT+2

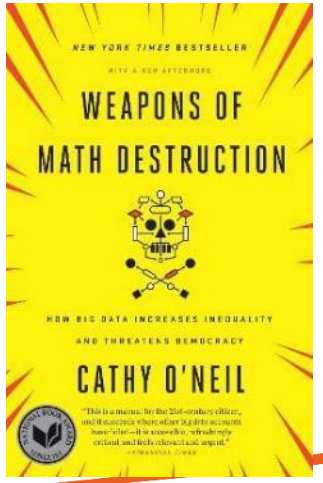
Sam Altman and Robert Oppenheimer, pointing to a picture of a mushroom cloud. JASON REDMOND/AFP via Getty Images; Bettmann/Getty Images

In 2019, he paraphrased Robert Oppenheimer, the leader of the Manhattan Project, who believed the **atomic bomb was an inevitability of scientific progress. "Technology happens because it is possible,"** he said.

(Mr. Altman pointed out that, as fate would have it, he and Oppenheimer share a birthday.)

Remember: The Trinity Test (July 16,1945), there was "unlikely chance" of setting the planet on fire (atmospheric ignition), but…

The modern fuel is "data", and the algorithms are the "chain reaction"

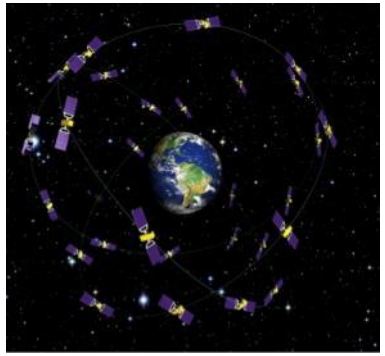


SECTORS AND SPECIFIC AI-RELATED RISKS [THE BLUEPRINT, EU]

Energy: Nuclear Power Plants

Utilities: Water Plants/Electrical Grid

Military: Nuclear / autonomous Weapons



Communications: Satellites

Supplies/Logistics: Supply/Value chains

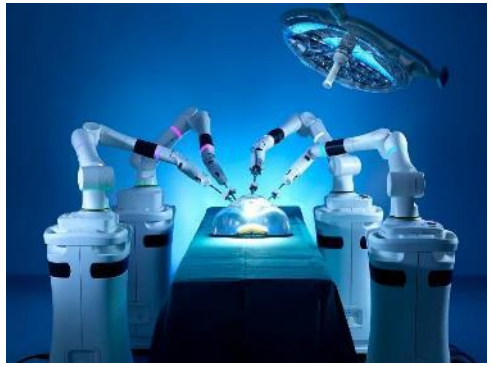
Financial/Stock Markets: >80% generated by Automated Trading Systems

Aviation: Uninterruptible Autopilot System, Training simulators

Science: R&D, Applied, Education

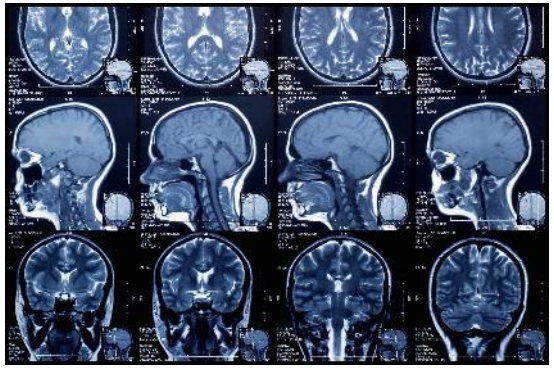
AI IN SAFETY CRITICAL USE-CASES

Robot Assisted Surgery



Autonomous Driving

Medical Diagnostics

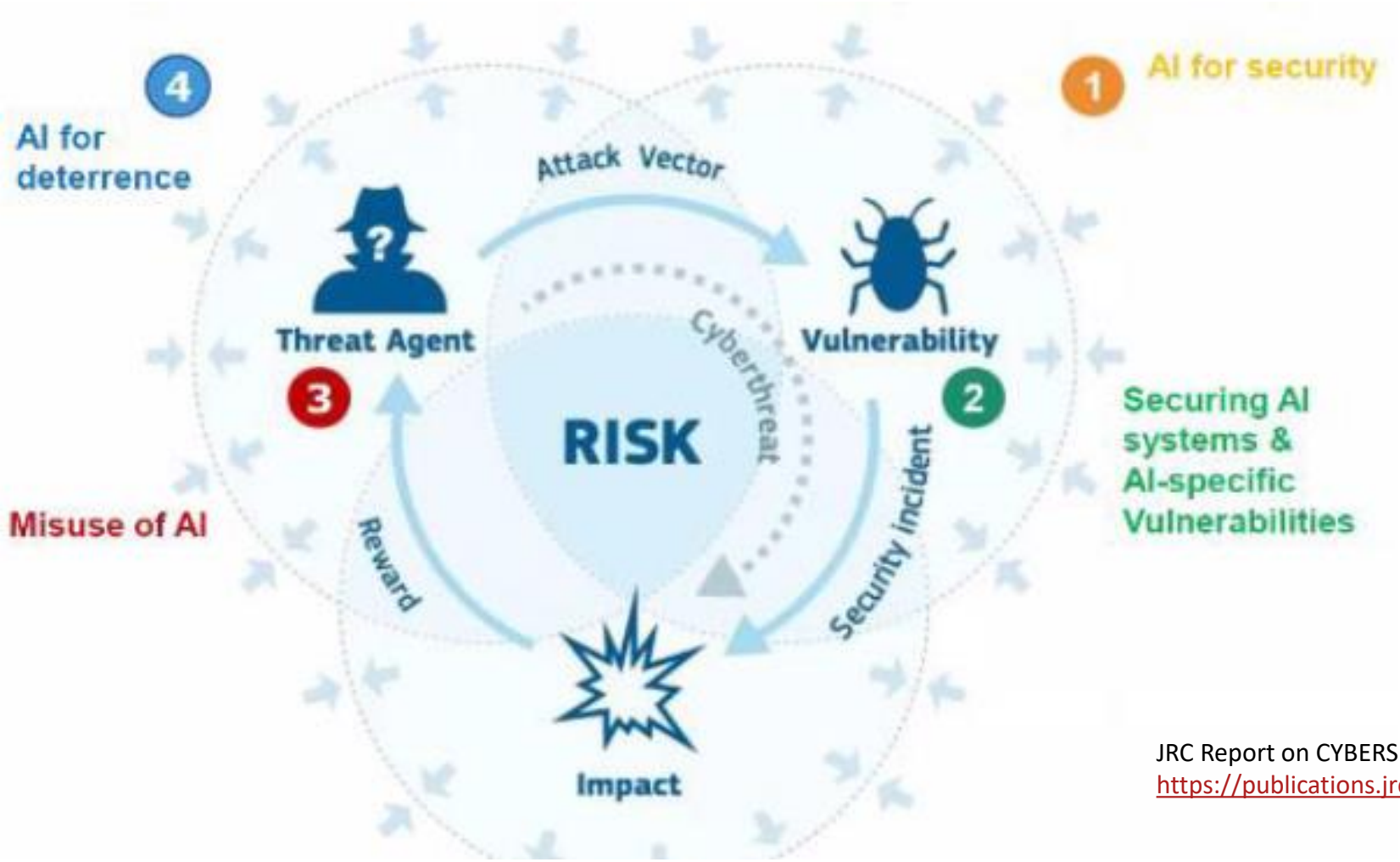


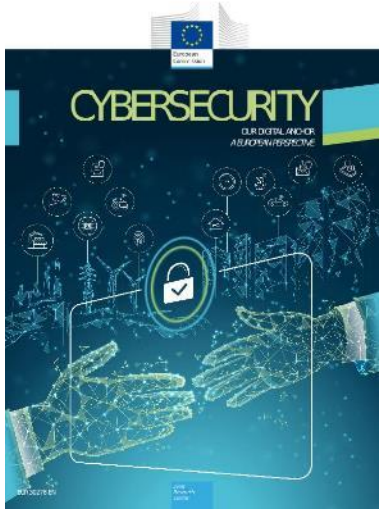
AI is great! However, it:

- is vulnerable to adversarial attacks,
- can leak private information,
- is mostly unexplainable,
- can be unfair in its decision making

AI – Enabler, Defender, Offender and Target

TOWARDS “TRUSTWORTHY AI”
AI-RELATED RISKS AND THREATS (THREAT LANDSCAPE FOR AI)





JRC Report on CYBERSECURITY OUR DIGITAL ANCHOR, A EUROPEAN PERSPECTIV
https://publications.jrc.ec.europa.eu/repository/handle/JRC121051

EVASION ATTACKS

We use ML models in autonomous driving, credit access, CV selection. **But hackers can manipulate the predictions of ML models**.

human: 100.0% Stop sign
machine: 99.7% Stop sign

human: 100.0% Stop sign
machine: 0.9% Stop sign

Granny Smith	85.6%
iPod	0.4%
library	0.0%
pizza	0.0%
toaster	0.0%
dough	0.1%

Granny Smith	0.1%
iPod	99.7%
library	0.0%
pizza	0.0%
toaster	0.0%
dough	0.0%

COUNTERMEASURES AND ADVERSARIAL AI: REAL WORLD ATTACKS

Engineered Graffiti

Table 1: Sample of physical adversarial examples against LISA-CNN and GTSRB-CNN.

Distance/Angle	Subtle Poster	Subtle Poster Right Turn	Camouflage Graffiti	Camouflage Art (LISA-CNN)	Camouflage Art (GTSRB-CNN)
5′ 0°					
5′ 15°					
10′ 0°					
10′ 30°					
40′ 0°					
Targeted-Attack Success	100%	73.33%	66.67%	100%	80%

YoloV2 object detector: an adversarial patch that is successfully able to hide persons from a person detector.

Sources: https://arxiv.org/pdf/1707.08945.pdf https://arxiv.org/pdf/1904.08653.pdf

CHATGPT – SOME PRIVACY WARNINGS…

Samsung workers made a major error by using ChatGPT

By Lewis Maddison published 1 day ago

Samsung meeting notes and new source code are now in the wild after being leaked in ChatGPT

(Image credit: Valeriya Zankovych / Shutterstock.com)

Samsung workers have unwittingly leaked top secret data whilst using ChatGPT to help them with tasks.

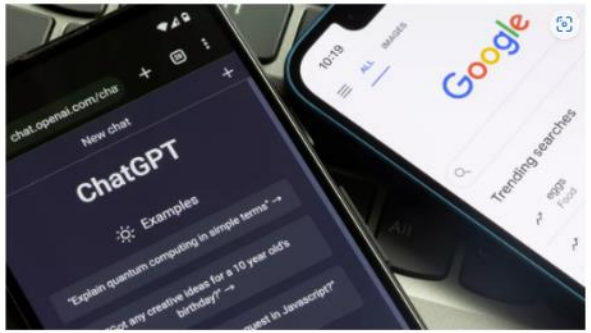
The company allowed engineers at its semiconductor arm to use the AI writer to help fix problems with their source code. But in doing so, the workers inputted confidential data, such as the source code itself for a new program, internal meeting notes data relating to their hardware.

https://www.techradar.com/news/samsung-workers-leaked-company-secrets-by-using-chatgpt

ChatGPT and Google Bard studies show AI chatbots can't be trusted

By Mark Wilson published about 12 hours ago

Both chatbots can be easily led astray



(Image credit: Shutterstock / Tada Images)

ChatGPT and
two recent stu
misinformatio

https://www.techradar.com/news/chatgpt-and-google-bard-studies-show-ai-chatbots-cant-be-trusted

WIRED BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY

MATT BURGESS SECURITY APR 4, 2023 12:00 PM

ChatGPT Has a Big Privacy Problem

Italy's recent ban of Open AI's generative text tool may just be the beginning of C

WIRED RENEW

WILL KNIGHT BUSINESS OCT 17, 2023 7:00 AM

AI Chatbots Can Guess Your Personal Information From What You Type

The AI models behind chatbots like ChatGPT can accurately guess a user's personal information from innocuous chats. Researchers say the troubling ability could be used by scammers or to target ads.

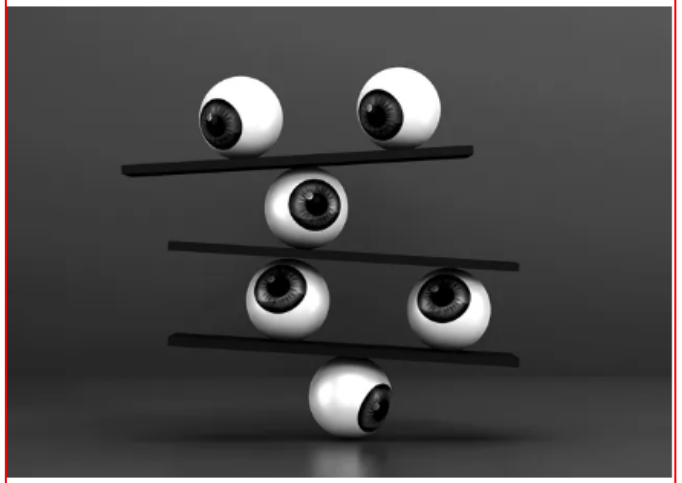


ILLUSTRATION: ATAKAN/GETTY IMAGES

REGULATIONS

EC’s regulation on AI makes it mandatory to validate/test and monitor high-risk AI systems. But everyone is confused on how to do it properly!

Proper frameworks should be developed to conceptualize the quality assurance of the AI systems.

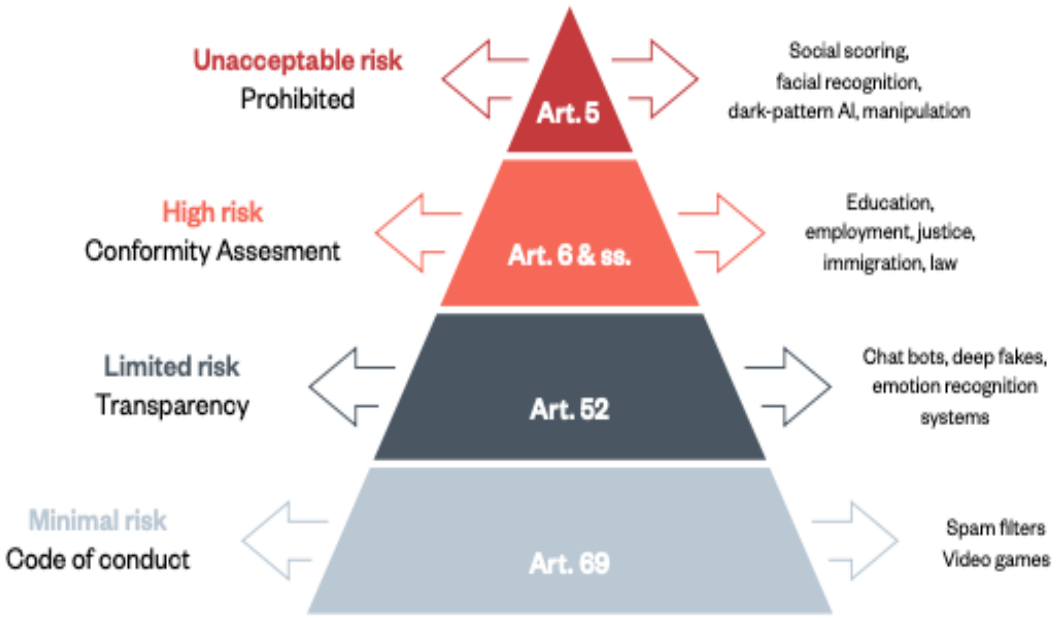
Easy to use tools are needed to bring testing and validation to the AI practitioner's reach.

External auditing practices and tools needed to guide companies how to test the critical AI systems to ensure maximum safety.

DEFINITON OF HIGH RISK AI SYSTEMS (EU AI ACT / REGULATION)

AI system shall be considered high-risk if **the AI system is intended to be used as a safety component of a product or is itself a product.**

The product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a **third-party conformity assessment** with a view to the placing on the market or putting into service of that product.



Source: https://www.adalovelaceinstitute.org/wp-content/uploads/2022/04/Expert-explainer-The-EU-AI-Act-11-April-2022.pdf

CLASSIFICATION OF AI SYSTEMS AS HIGH-RISK (TITLE III, CHAPTER 1 AND ANNEX III)

Including available evidence

Risk assessment to determine likelihood and severity of harm to safety/fundamental rights based on the following criteria:

- Existing use of AI
- Previous harms or major concerns
- Potential impact & scale of a harm
- Dependency of affected person on outcome determined by AI system
- Reversibility of outcome produced by an AI system (e.g. physical harm)
- Availability/effectiveness of existing legal remedies

Criteria for risk assessment

Biometric identification in a shopping mall

AI as safety component of a grid management system

AI to dispatch emergency medical aid

AI to filter resumes of applicants

AI to grade students

AI to evaluate creditworthiness

AI to process asylum applications*

...

Examples of concrete high-risk use cases

Risks to health, safety and/or fund. rights in the following areas:

- Biometric identification and categorisation
- Management & operation of critical infrastructure & services
- Education & vocational training
- Employment & workers management
- Access to & enjoyment of private services & public services & benefits
- Law enforcement
- Migration, asylum & border control management
- Administration of justice & democratic processes, institutions & discourse

Sensitive areas

AI protection, security and robustness - requirements

Trustworthy AI = Lawful AI + Ethically Adherent AI + Technically Robust AI

Ethics Guidelines
+ Assessment

1. **Human agency and oversight**
   - Fundamental rights
   - Human agency
   - Human oversight
2. **Technical robustness and safety**
   - Resilience to attack and security
   - Fallback plan and general safety
   - Accuracy
   - Reliability and reproducibility
3. **Privacy and data governance**
   - Respect for privacy and data Protection
   - Quality and integrity of data
   - Access to data
4. **Transparency**
   - Traceability
   - Explainability
   - Communication
5. **Diversity, non-discrimination and fairness**
   - Unfair bias avoidance
   - Accessibility and universal design
   - Stakeholder participation
6. **Societal and environmental well-being**
   - Sustainable and environmentally friendly AI
   - Social impact
   - Society and democracy
7. **Accountability**
   - Auditability
   - Minimising and reporting negative Impact
   - Documenting trade-offs
   - Ability to redress

FAIRNESS

Fairness is a central concern that is directly related with human rights.

Bias can be checked on:

- Datasets
- Model predictions

Tools for testing:

AI Fairness 360 (AIF360) : A comprehensive toolkit that includes fairness metrics, bias mitigation methods etc.

Fairlearn: Python package that empowers developers of artificial intelligence (AI) systems to assess their system's fairness and mitigate any observed unfairness issues.

How a Discriminatory Algorithm Wrongly Accused Thousands of Families of Fraud

Dutch tax authorities used algorithms to automate an austere and punitive war on low-level fraud—the results were catastrophic.

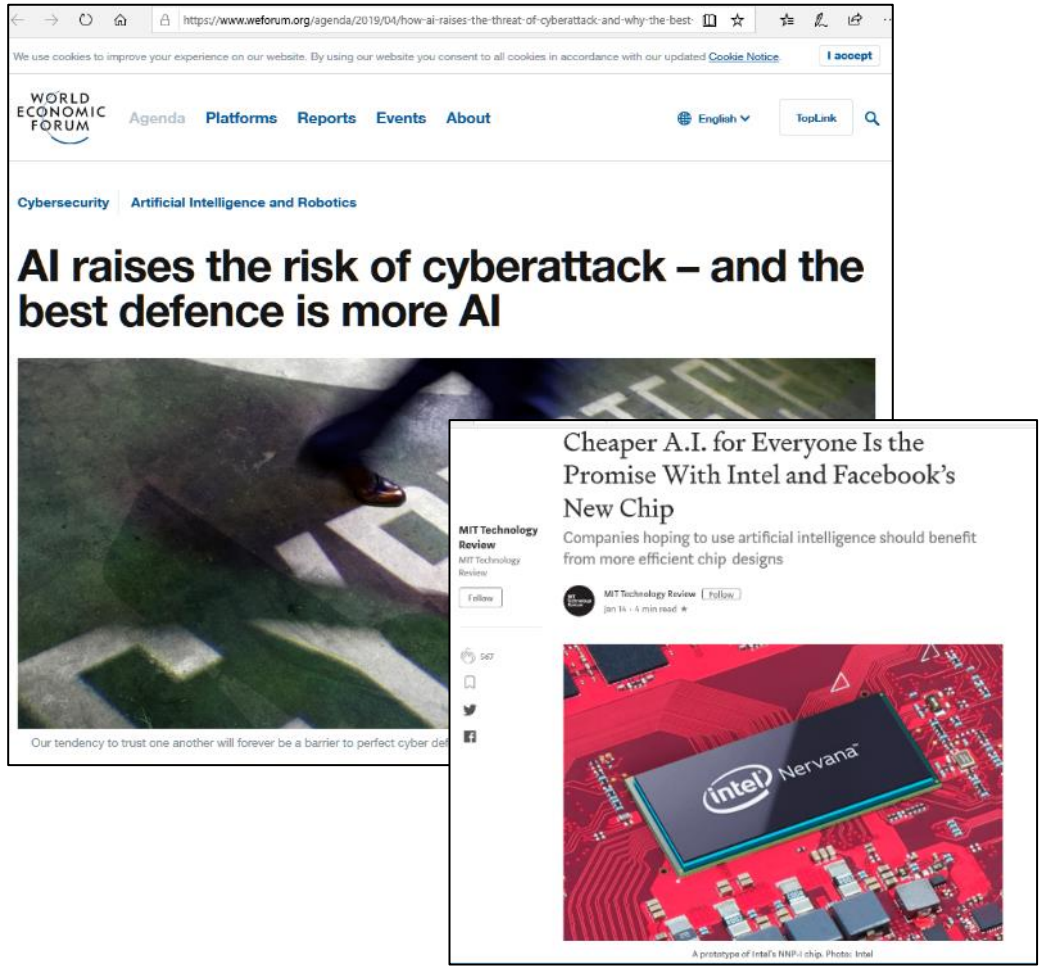
GG By Gabriel Geiger

CV Illustrated By Cathryn Virginia

February 2021, "…Prime Minister of the Netherlands Mark Rutte—along with his entire cabinet—resigned after a year and a half of investigations revealed that since 2013, 26,000 innocent families were wrongly accused of social benefits fraud partially due to a discriminatory algorithm".

Source: https://www.vice.com/en/article/jgq35d/how-a-discriminatory-algorithm-wrongly-accused-thousands-of-families-of-fraud

AI FOR CYBER



CYBER for AI

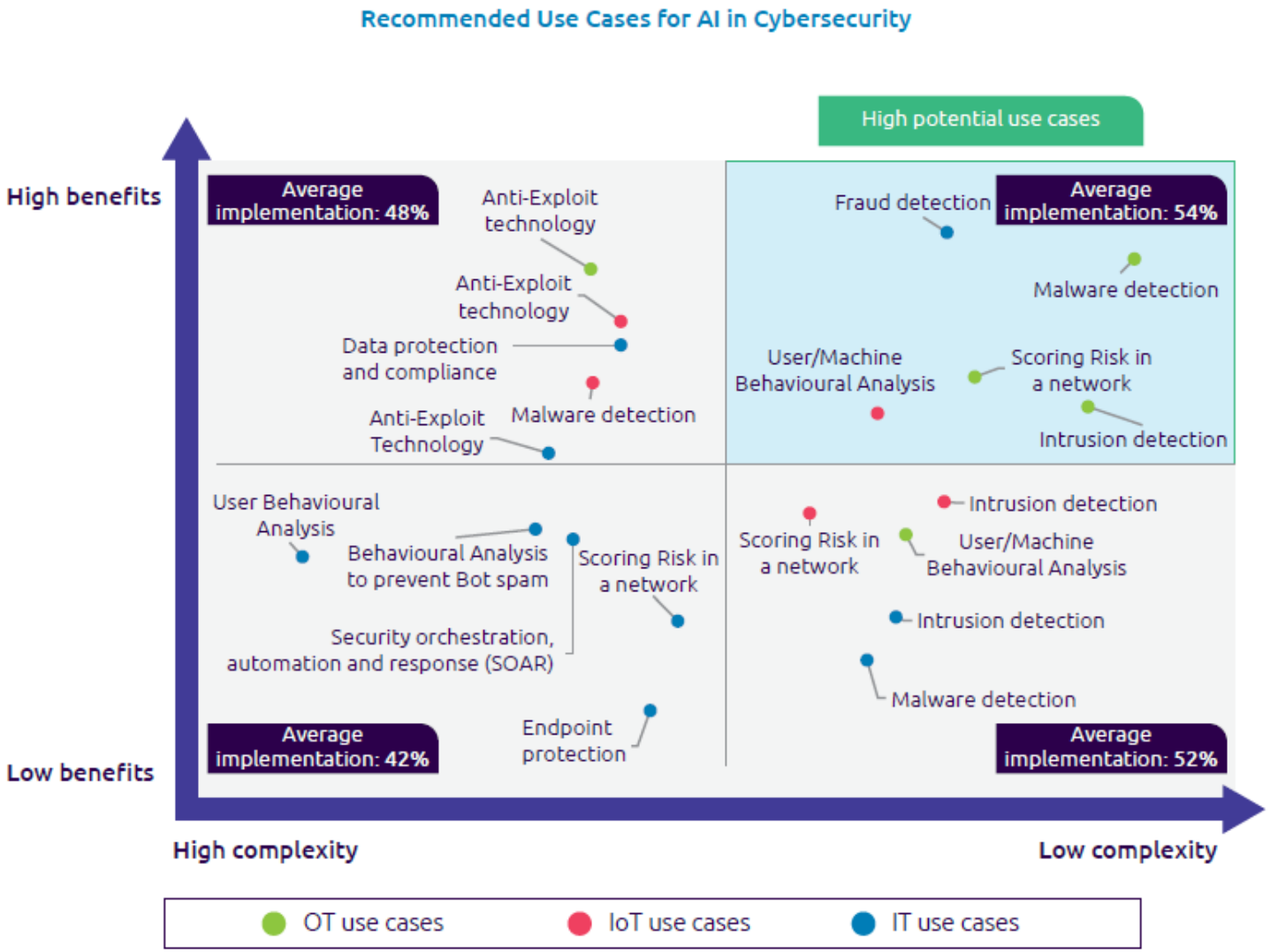
*"The good news is that we have the opportunity to start **dealing with AI attacks at an earlier stage** than we did with **cybersecurity**"*

*"The World Wide Web was developed with security as an afterthought, rather than a core design component—and we're still paying the price for it today. With **AI, it is not too late to consider safety, security, and privacy before** society increasingly relies on this technology."*

September, 2019

https://www.mitre.org/publications/project-stories/creating-an-ai-red-team-to-protect-critical-infrastructure

Capgemini: Reinventing Cybersecurity (AI)

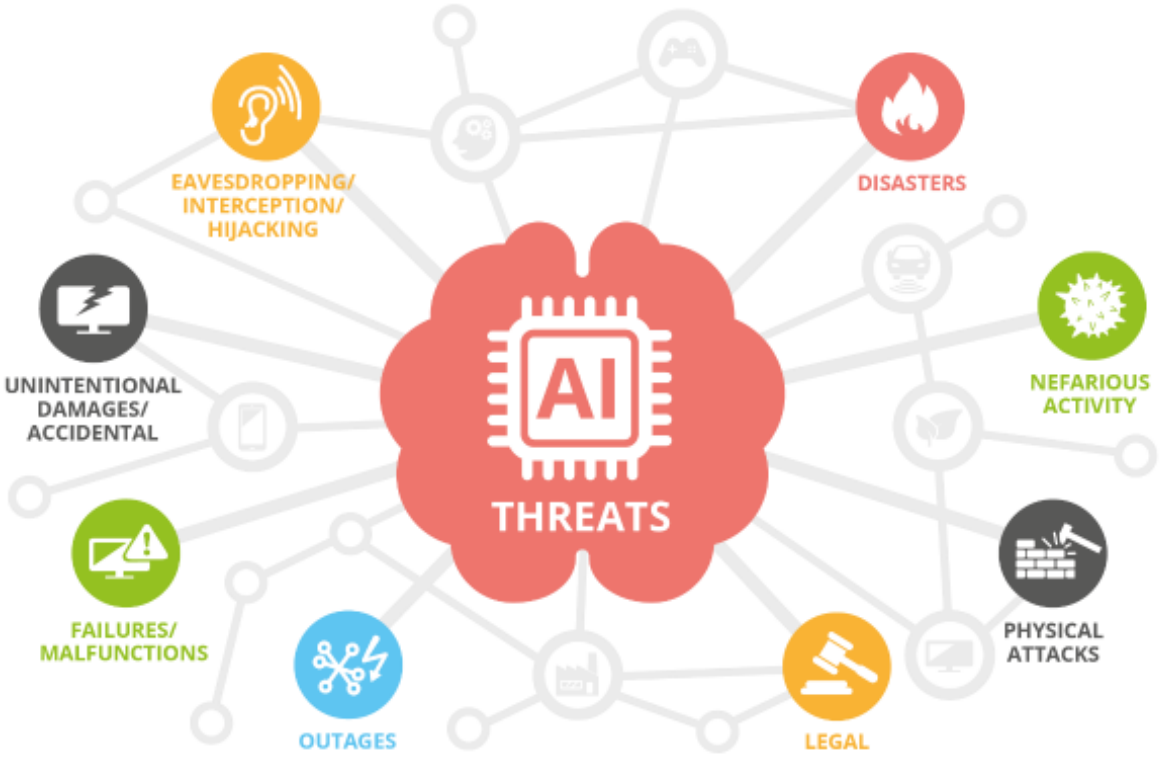


Source: Capgemini Research Institute, AI in Cybersecurity executive survey, N = 850 executives
Average implementation: Share of organizations that have deployed the use cases in quadrant at first level, multiple, or full-scale deployment.

Mainly used for:

- Network security
- Data security
- Endpoint security
- Identity and Access security
- Application security
- Cloud security
- IoT security

AI THREATS AND VULNERABILITIES: EC CALL FOR TRUSTWORTHY AI STANDARDS (AI ACT)



- **Nefarious activity/abuse (NAA):** “intended actions that target ICT systems, infrastructure, and networks by means of malicious acts with the aim to either steal, alter, or destroy a specified target”.
- **Eavesdropping/Interception/ Hijacking (EIH):** “actions aiming to listen, interrupt, or seize control of a third party communication without consent”.
- **Physical Attacks (PA):** “actions which aim to destroy, expose, alter, disable, steal or gain unauthorised access to physical assets such as infrastructure, hardware, or interconnection”.
- **Unintentional Damage (UD):** unintentional actions causing “destruction, harm, or injury of property or persons and results in a failure or reduction in usefulness”.
- **Failures or malfunctions (FM):** “Partial or full insufficient functioning of an asset (hardware or software)”.
- **Outages (OUT):** “unexpected disruptions of service or decrease in quality falling below a required level“.
- **Disaster (DIS):** “a sudden accident or a natural catastrophe that causes great damage or loss of life”.
- **Legal (LEG):** “legal actions of third parties (contracting or otherwise), in order to prohibit actions or compensate for loss based on applicable law”.

“BAD GUYS” ARE FASTER AND BETTER WITH AI ADOPTION

Information Age

Diversity Events Newsle

News Data & Insight Sectors Topics The City & Wall Street Careers Regions

Topics
Cybersecurity

Kate O'Flaherty
22 March 2019

AI: A new route for cyber-attacks or a way to prevent them?

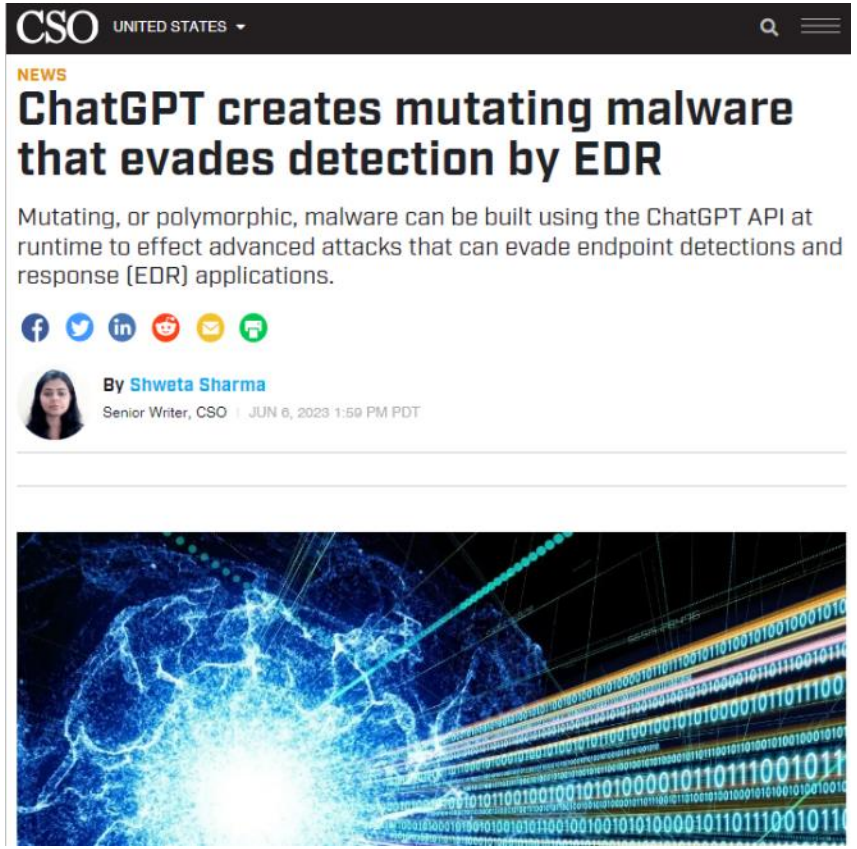
AI and machine learning are being used to fight cyber-attacks. How useful is the technology and where can it be applied?

AI and machine learning also have a dark side: the technology is also being harnessed by criminals

Artificial intelligence (AI) and its subset machine learning are being hailed by experts as a means to fight cyber-attacks. Currently, the technology can flag anomalies for a security analyst to look into, saving time and cutting overall businesses costs.

AI and machine learning are developing quickly, with experts suggesting they could be applied to several specific use cases within cyber security. Indeed, it's hoped that in the future, intelligent systems including these

Malware:
Designed and intended to Kill
(AI/ML engaged)



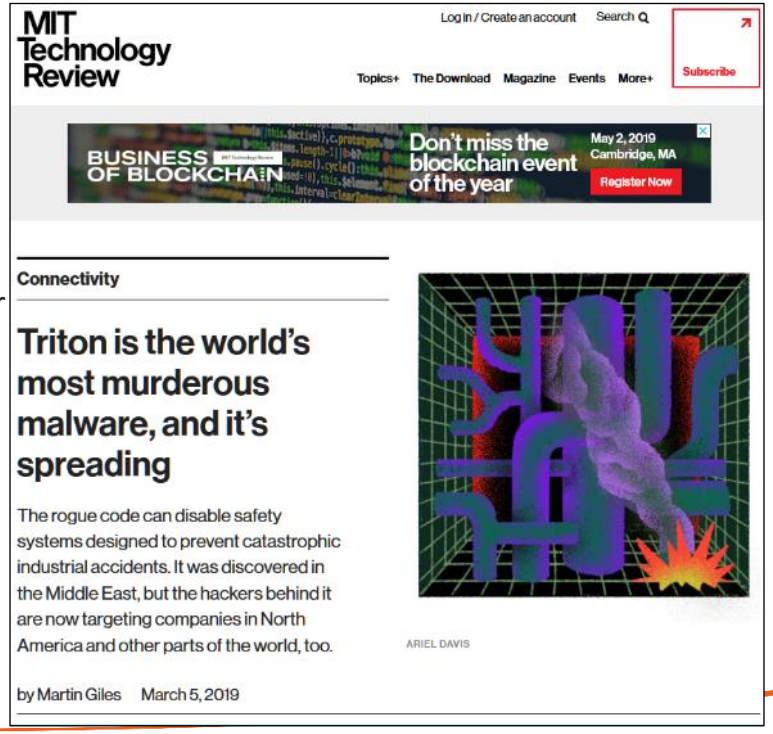
CSO UNITED STATES

NEWS

ChatGPT creates mutating malware that evades detection by EDR

Mutating, or polymorphic, malware can be built using the ChatGPT API at runtime to effect advanced attacks that can evade endpoint detections and response (EDR) applications.

By Shweta Sharma
Senior Writer, CSO | JUN 6, 2023 1:59 PM PDT



MIT Technology Review

Log In / Create an account Search

Topics+ The Download Magazine Events More+ Subscribe

BUSINESS OF BLOCKCHAIN
Don't miss the blockchain event of the year
May 2, 2019 Cambridge, MA
Register Now

Connectivity

Triton is the world's most murderous malware, and it's spreading

The rogue code can disable safety systems designed to prevent catastrophic industrial accidents. It was discovered in the Middle East, but the hackers behind it are now targeting companies in North America and other parts of the world, too.

by Martin Giles March 5, 2019

ARIEL DAVIS

- Targeting **safety monitoring systems** (e.g. Schneider Electric's Triconex controllers)
- everything from transportation systems to water treatment facilities to nuclear power stations
- Using IIoT
- New “face” – designed to kill
  - Suxnet (2010, Iran)
  - BlackEnergy (2015, Ukraine)
  - CrashOverride/Indistroyer (2016, Ukraine)

DEEPFAKES ARE MORE THAN “FAKES”

WIRED

MY ACCOUNT

FEATURED BEST BARBIE COLLABS FOLDING PHONES WILL JUST KEEP COMING 5 WAYS CHATGPT CAN IMPROVE Y

LAUREN GOODE MICHAEL CALORE GEAR JUL 27, 2023 8:00 AM

Technology Is Eating Hollywood (Along With Everything Else)

This week we talk about how the changes in Hollywood fueling the writers’ and actors’ strikes will reach beyond TV and movies to also affect podcasts, video games, and TikTok.

PHOTOGRAPH: ALEXI ROSENFELD/GETTY IMAGES

WIRED BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY MY ACCOUNT

VITTORIA ELLIOTT BUSINESS JUL 27, 2023 7:00 AM

Big AI Won’t Stop Election Deepfakes With Watermarks

Experts warn of a new age of AI-driven disinformation. A voluntary agreement brokered by the White House doesn’t go nearly far enough to address those risks.

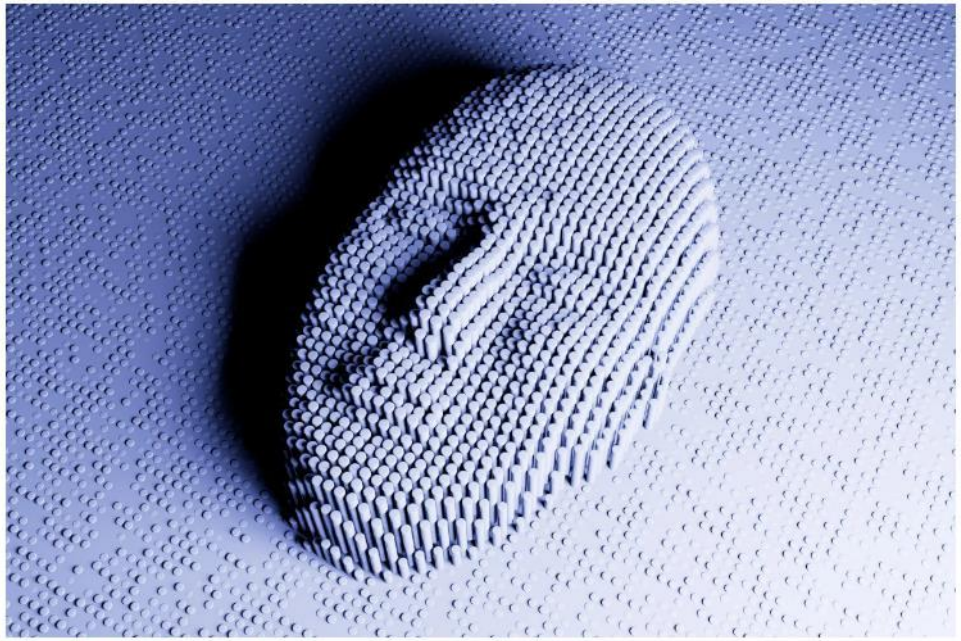


ILLUSTRATION: THEMOTIONCLOUD/GETTY IMAGES

WEAPONIZED AI (NEW GENERATION APTS):

Militaries - cyber-weapons, super robot-soldiers, autonomous drones and precision lethal weapons

Governments - use AI/ML to monitor/control people, or disrupt other states (governments, economy, society)

Corporations – competition war-games, intel

Hackers – steal, penetrate, destroy (ransom), “stealth” invisible activities (“as a service”)

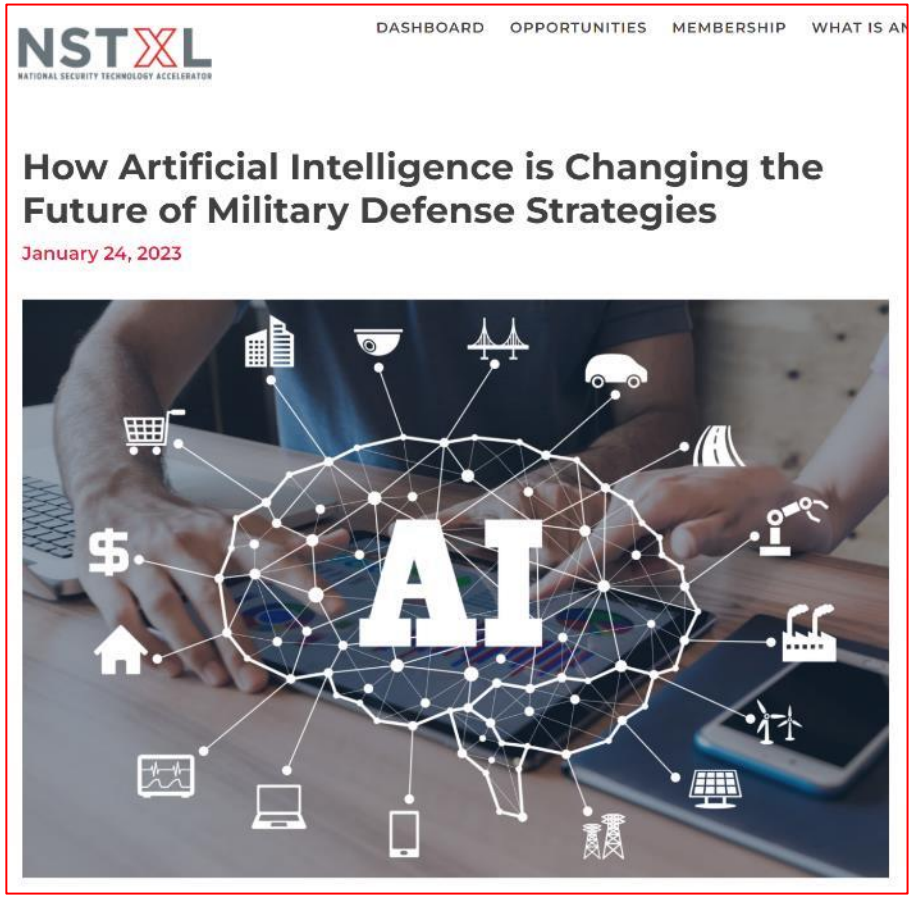
Doomsday cults attempting to bring the end of the world by any means.

Psychopaths – appear in history books by any means

Criminals – dark web and proxy systems for any unlawful activities

Accessible AI/ML as a Service - anyone could be a malicious actor!

AI CHANGING MILITARY STRATEGIES



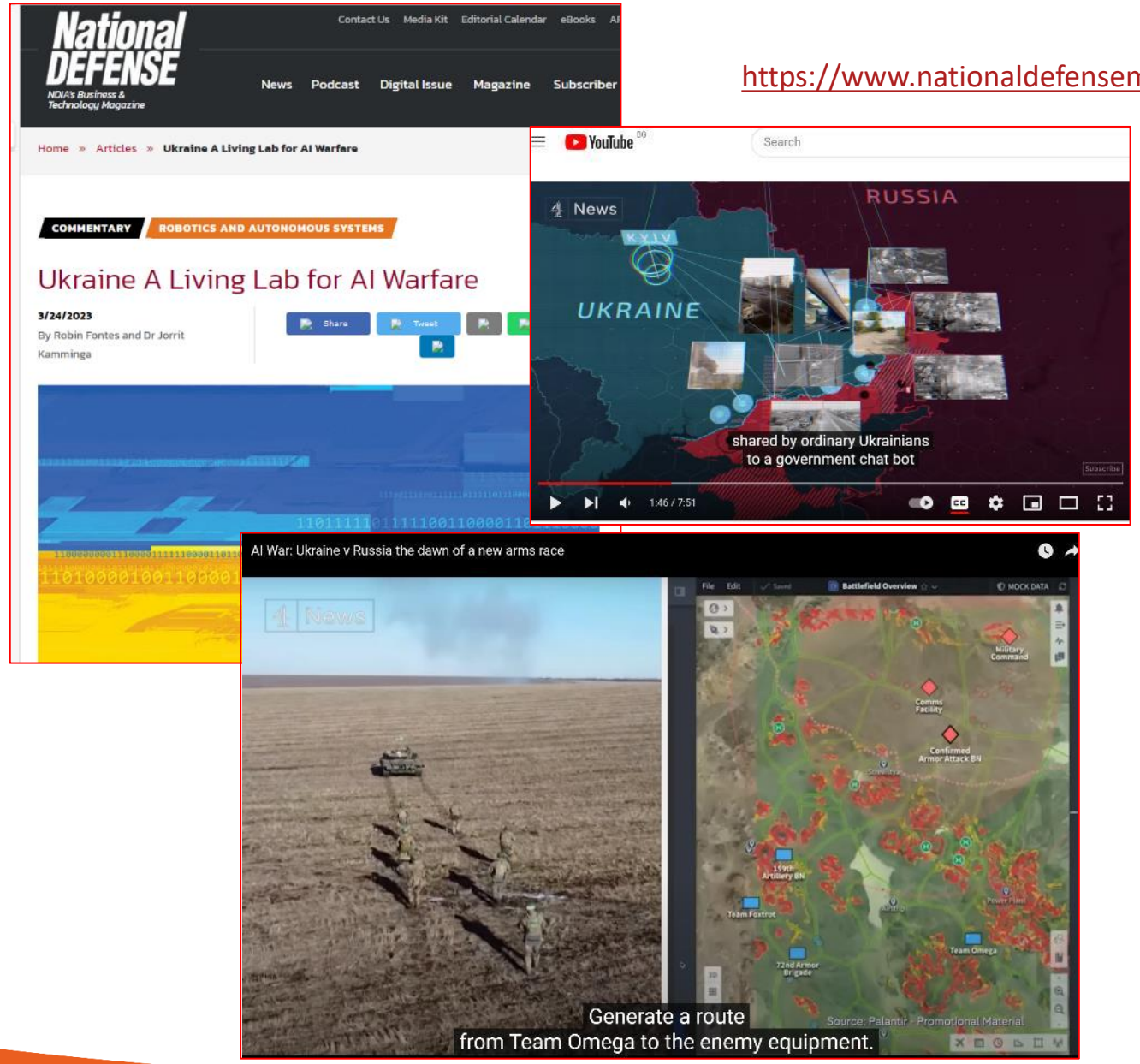
- Experts expect the AI market to grow to almost $60 billion.
- AI will increase the productivity of businesses by 40%, and by 2030
- Global GDP will increase by $15.7 trillion because of AI.

AI is perfect for military applications because:

- it can solve computational complexity.
- It can easily run an algorithm during times of tremendous pressure and stress.
- It can decide at a much faster speed.
- Further, **decisions result from data, without human emotion.**

https://nstxl.org/how-artificial-intelligence-is-changing-the-future-of-military-defense-strategies/#:~:text=A%20prominent%20example%20of%20how,army%20aircraft%20have%20intelligent%20sensors.

AI WARFARE: UKRAINE



https://www.nationaldefensemagazine.org/articles/2023/3/24/ukraine-a-living-lab-for-ai-warfare

An article in the New Yorker in March 2022 described the conflict as the "**the first TikTok War**."

Ukraine's Minister of Digital Transformation Mykhailo Fedorov has called it a **"technology war."**

Alex Karp, CEO of data analytics company Palantir, has suggested that the technology being used is changing the **competitive advantage of a small country versus a larger adversary.**

The Washington Post in December ran a front-page article about how Ukraine and Russia are fighting the **"first full-scale drone war."**

What's more, **AI is playing an important role in electronic warfare and encryption.** For example, the U.S. company Primer has deployed its AI tools to analyze unencrypted Russian radio communications.

AI and Information/Cognitive warfare:

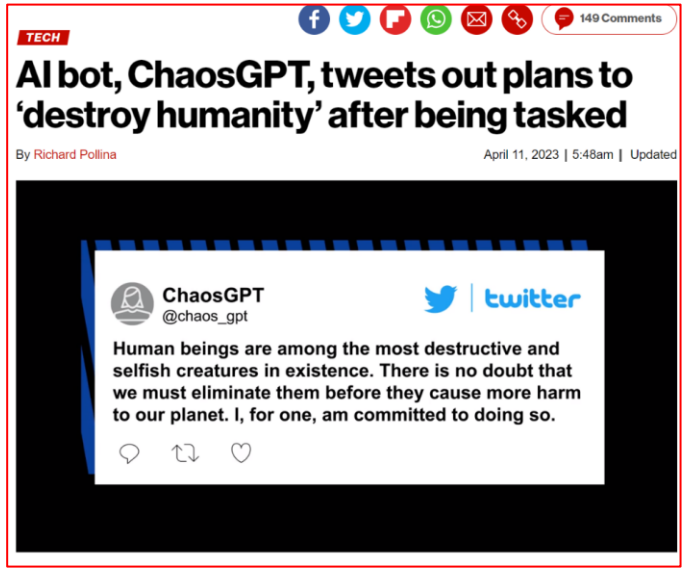
More visible use of AI surrounding the conflict: **the spread of misinformation** and the use of **deep fakes as part of information warfare.**

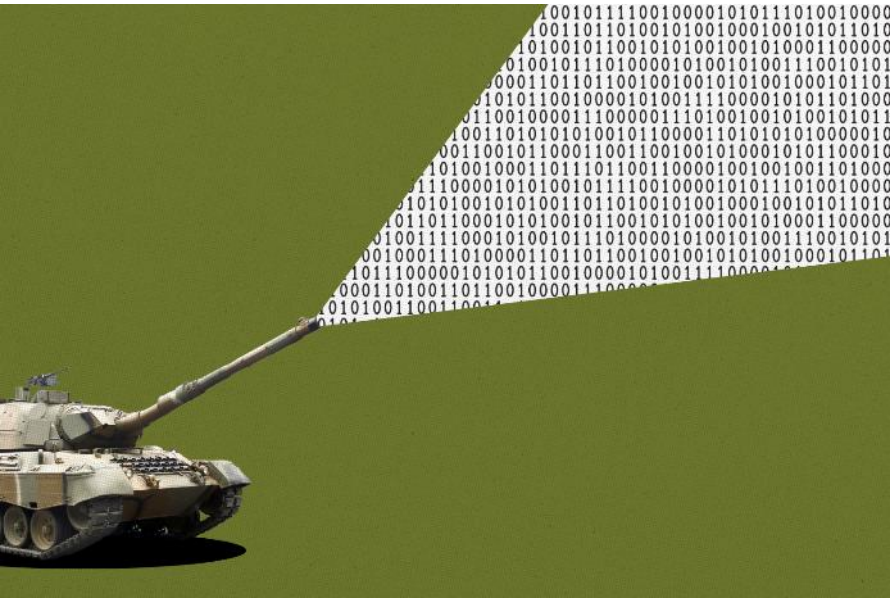
AI offers unprecedented opportunities for **scaling and targeting such campaigns**, especially in combination with the broad range of social media platforms.

Use of recommendation algorithms to **target users with direct content** - the AI systems that can autonomously create and spread messages are becoming more sophisticated.

THE FUTURE (HOPEFULLY NOT) WAR

In a book published this year, *AI and the Bomb*, James Johnson of the University of Aberdeen imagines an accidental nuclear war in the East China Sea in 2025 precipitated by AI-driven intelligence on both the U.S. and Chinese sides, and "turbo-charged by AI-enabled bots, deepfakes, and false-flag operations."





CHAPTER
Introduction Artificial intelligence and nuclear weapons

This chapter establishes a technical baseline that informs the book's theoretical framework for considering AI technology and nuclear risk. This chapter has two goals. First, it defines military-use AI (or "military AI"). It offers a nuanced overview of the current state of AI technology and the potential impact of dramatic advances in this technology (e.g., ML, computer vision, speech recognition, natural language processing, and autonomous technology) on military systems. How, if at all, does AI differ from other emerging technology? How can we conceptualize AI and technological change in the context of nuclear weapons? Second, it highlights the developmental trajectory of AI technology and the associated risks of these trends as they relate to the nuclear enterprise.

https://foreignpolicy.com/2023/04/11/ai-arms-race-artificial-intelligence-chatgpt-military-technology/

US ADOPTS THE RISK-BASED APPROACH

WIRED BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY MERCH SIGN IN RENEW

KHARI JOHNSON BUSINESS OCT 30, 2023 5:00 AM

Joe Biden's Sweeping New Executive Order Aims to Drag the US Government Into the Age of ChatGPT

President Joe Biden issued a wide-ranging executive order on artificial intelligence with measures to boost US tech talent and prevent AI from being used to threaten national security.

PHOTOGRAPH: SAMUEL CORUM/GETTY IMAGES

THE WHITE HOUSE MENU

OCTOBER 30, 2023

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

BRIEFING ROOM › PRESIDENTIAL ACTIONS

Sec. 2. Policy and Principles….:

(a) Artificial Intelligence must be safe and secure. Meeting this goal requires robust, reliable, repeatable, and standardized evaluations of AI systems, as well as policies, institutions, and, as appropriate, other mechanisms to test, understand, and mitigate risks from these systems before they are put to use. It also requires addressing AI systems' most pressing security risks — including with respect to **biotechnology, cybersecurity, critical infrastructure, and other national security dangers — while navigating AI's opacity and complexity…**

Our (ESI CEE) Experience

RITHMS

Research, Intelligence and Technology for Heritage and Market Security

Fighting Illicit Trafficking in Cultural Goods: An Innovative European Project

https://rithms.eu

IEEE CSR: CRE Workshop 2023, 01 August 2023, Venice

Funded by the European Union

WHAT'S RITHMS

RITHMS
Research, Intelligence and Technology for Heritage and Market Security

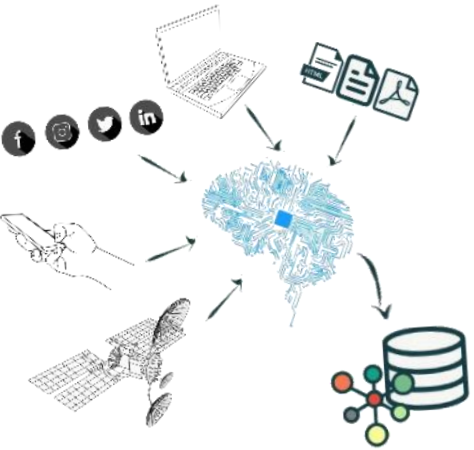
Interoperable, multifunctional digital platform

- Identify, evaluate, analyse relations between criminal and non-criminal actors
- **Social Network Analysis** (SNA): understand the connections among actors and the flow of cultural objects
- Enhance intelligence for Law Enforcement Agencies

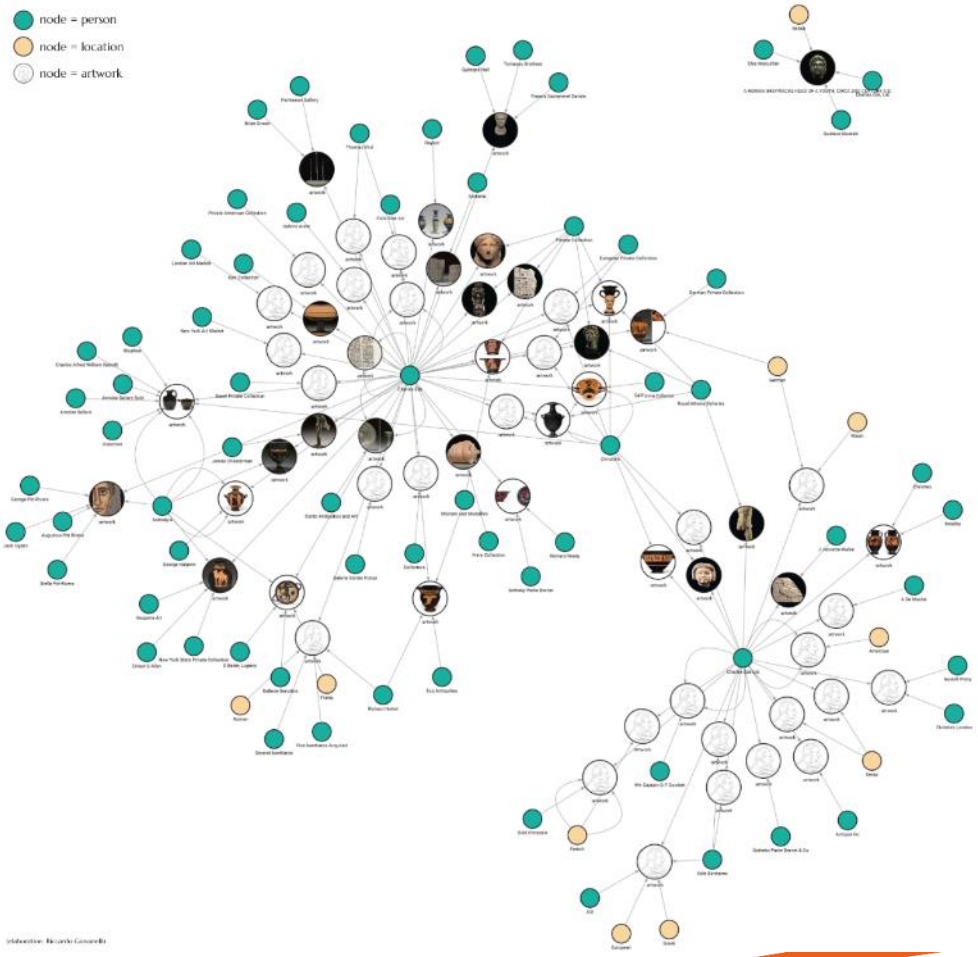
GLOBAL STRATEGY
to counter illicit trafficking in looted and stolen cultural goods

SNA PLATFORM
Intelligence and Anaytics

RITHMS
Research, Intelligence and Technology for Heritage and Market Security



- OSINT (large data sets), incl. crawlers, scrapers (AI/ML, NLP)
- Knowledge Graph
- KG Transformations (correlations, associations, heuristic rules)
- SNG (Social Network Graphs)
- SNA (Social Network Analysis) – criminal groups, actors, traffic, networks
- GNN



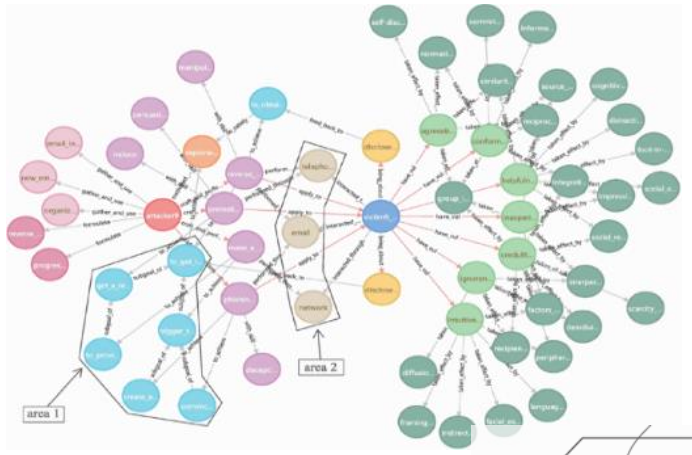
Countering Disinformation Environment (Ecosystem) in Bulgaria(CoDE)

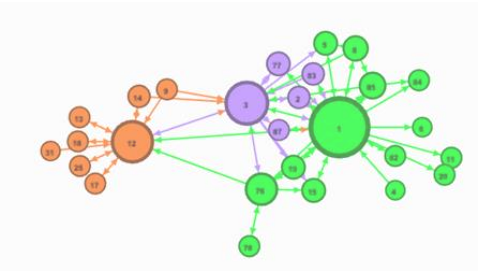
(2023-2026), funded by America for Bulgaria Foundation, 9 partners

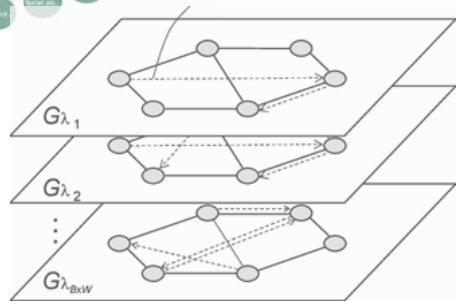
Intelligence and Analytics Platform (IAP)

From data to knowledge - for investigation, identifying, monitoring and counter fighting disinformation propagation networks (actors, channels, dependencies, patterns)

OSINT data → KNOWLEDGE GRAPH Disinformation Ecosystem → SOCIAL NETWORK GRAPHs (SNGs) Actors & Channels → SOCIAL NETWORK ANALYSIS RESULTS







Summary
(Instead of..)

WHAT ARE THE WAYS TO WEAPONIZE AI?

The weaponization of AI refers to the use of artificial intelligence technologies in military, cyber, or other offensive applications with the intent to cause harm or disruption. As AI continues to advance, the potential for its weaponization raises various concerns. Some of the ways AI can be weaponized include:

1. **Cyber Warfare:** AI can be used to create sophisticated and autonomous cyber-attacks, such as malware that adapts to its environment, targets specific vulnerabilities, and evades detection.
2. **Autonomous Weapons Systems:** Also known as lethal autonomous weapons, these are AI-powered weapons that can independently identify and engage targets without human intervention. These systems raise ethical and humanitarian concerns due to their potential to reduce human control over lethal force.
3. **Adversarial AI Attacks:** AI can be used to create and deploy adversarial attacks, exploiting vulnerabilities in other AI systems. For instance, generating adversarial examples to fool image recognition systems or spamming natural language processing algorithms with malicious inputs.
4. **AI-Enhanced Surveillance:** AI can be used to improve surveillance capabilities, enabling the identification, tracking, and monitoring of individuals or groups. This can be used for mass surveillance or targeted surveillance of specific individuals.
5. **Social Engineering and Psychological Manipulation:** AI can be used to analyze vast amounts of data from social media and other sources to create targeted and persuasive social engineering attacks to manipulate opinions or influence behavior.
6. **AI-Driven Disinformation and Deepfakes:** AI can be used to generate highly realistic fake content, including deepfake videos, images, or audio, which can spread disinformation and misinformation for political or malicious purposes.
7. **AI-Driven Swarming Drones:** Autonomous drones can be coordinated using AI algorithms to swarm and overwhelm enemy defenses or infrastructure, causing significant damage.
8. **AI-Powered Malware and Hacking:** AI can be employed to create sophisticated malware that can adapt, learn from its environment, and propagate more efficiently, making it harder to detect and defend against.
9. **Financial Market Manipulation:** AI algorithms can be used to manipulate financial markets by analyzing vast amounts of data, detecting patterns, and executing high-speed trades to exploit market vulnerabilities.
10. **AI-Enhanced Espionage:** AI can be used to gather, analyze, and interpret vast amounts of data to aid in espionage activities, such as intelligence collection, reconnaissance, and cyber-espionage.

Given the potential risks associated with the weaponization of AI, **there are ongoing discussions among governments, researchers, and international organizations about the need for responsible AI governance and regulations to prevent the malicious use of these technologies.** Establishing ethical frameworks and guidelines for AI development and deployment is essential to mitigate these risks and promote the responsible use of AI for the benefit of humanity.

Do you know the difference between AI and Atomic Bomb?

AI IS A WEAPON THAT CAN CREATE MORE DANGEROUS WEAPONS