Artificial Intelligence in Cybersecurity: Benefits, Challenges, and How to Stay Ahead

Rehmann

EMPOWER YOUR PURPOSE®

THIS IS **NOT** MORGAN FREEMAN.

DIEP NEP.
DEEPFAKE MORGAN FREEMAN

Meet the Speaker

James E. Carpp, CISA, CRISC, CIRM, CISM
Chief Digital Officer
James.Carpp@rehmann.com

Agenda

1. Artificial Intelligence Primer
2. AI Driven Cyber Threats
3. Protection – What Can you Do?
4. Q&A/Closing

Serving people and protecting futures are at the heart of Wolverine Fire Protection Co., and Rehmann empowers this multi-generational business to continue this legacy.

AI PRIMER

Search to Answer

Transitioning from a Google search paradigm to an AI answer paradigm.

Paradigm Shift

Shifting from the need for the human to understand computer code to the computer understanding human natural language.

Narrow AI

Rules based AI capable of doing some tasks requiring human intelligence with a narrow scope or domain.

General AI

AI that aims to replicate human cognitive abilities, including reasoning, planning, problem-solving, and learning. Unlike existing AI systems that excel at specific tasks, AGI would have the ability to understand, learn, and apply knowledge across a broad range of tasks.

Machine Learning

A subfield of artificial intelligence that uses algorithms trained on data sets to create models that enable machines to perform tasks that would otherwise only be possible for humans. It's a process of using mathematical models of data to help a computer learn without direct instruction, identifying patterns within data, and using those patterns to make predictions.

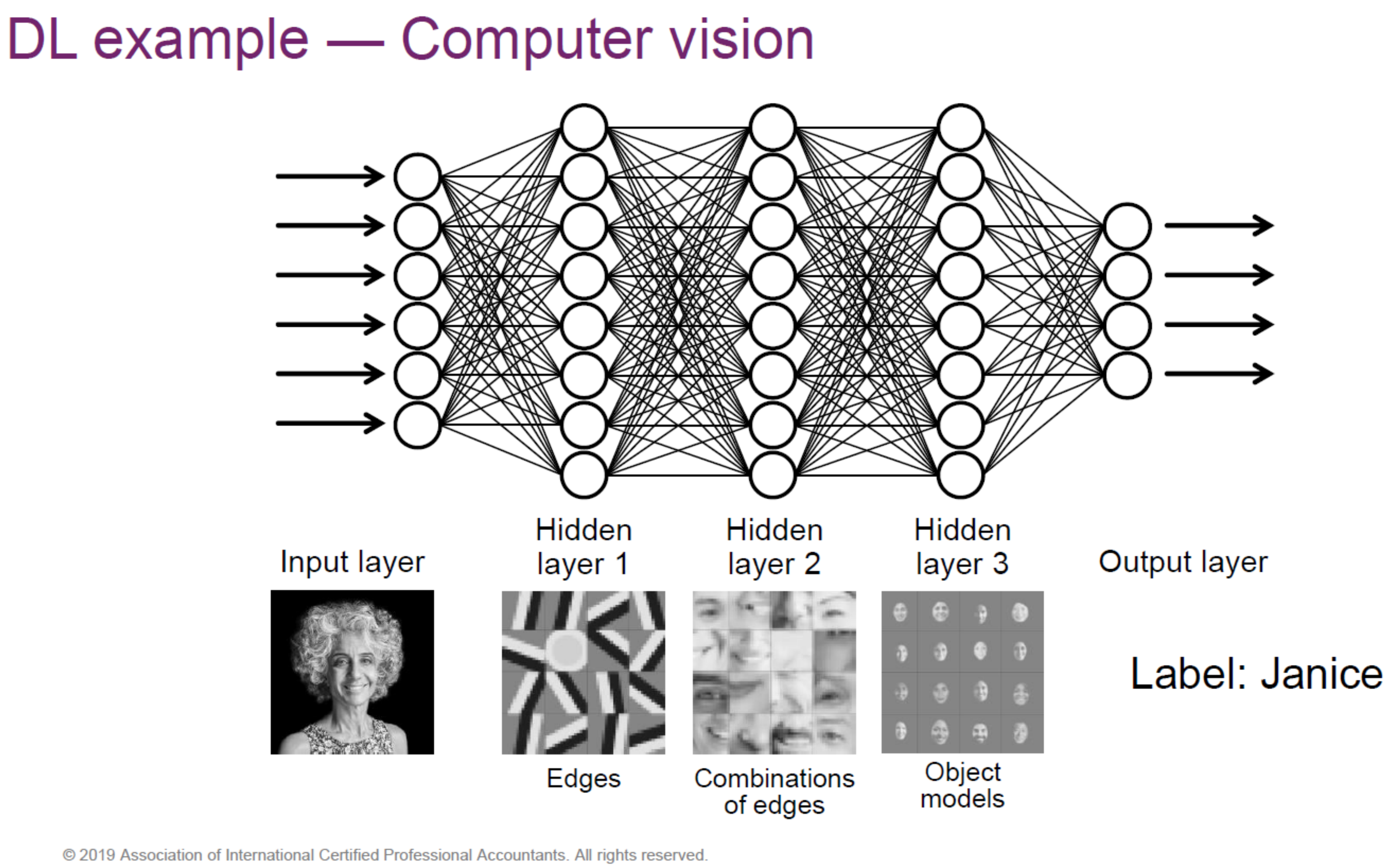
Deep Learning

A subfield of artificial intelligence that uses algorithms trained on data sets to create models that enable machines to perform tasks that would otherwise only be possible for humans. It's a process of using mathematical models of data to help a computer learn without direct instruction, identifying patterns within data, and using those patterns to make predictions.

Neural Network

is a series of interconnected nodes, or artificial neurons, that are designed to mimic the way the human brain operates. These networks use algorithms to recognize underlying relationships in a set of data, with each node processing information and passing it on to the next, much like how neurons transmit signals in the brain.

Neural Network – How it Processes



DL example — Computer vision

Large Language Models

A large language model is a type of artificial intelligence system that has been trained on vast amounts of text data and is capable of understanding and generating human-like text based on the input it receives. These models are designed to perform a wide range of natural language processing tasks, such as text generation, translation, summarization, and question-answering.

Natural Language Processing (NLP)

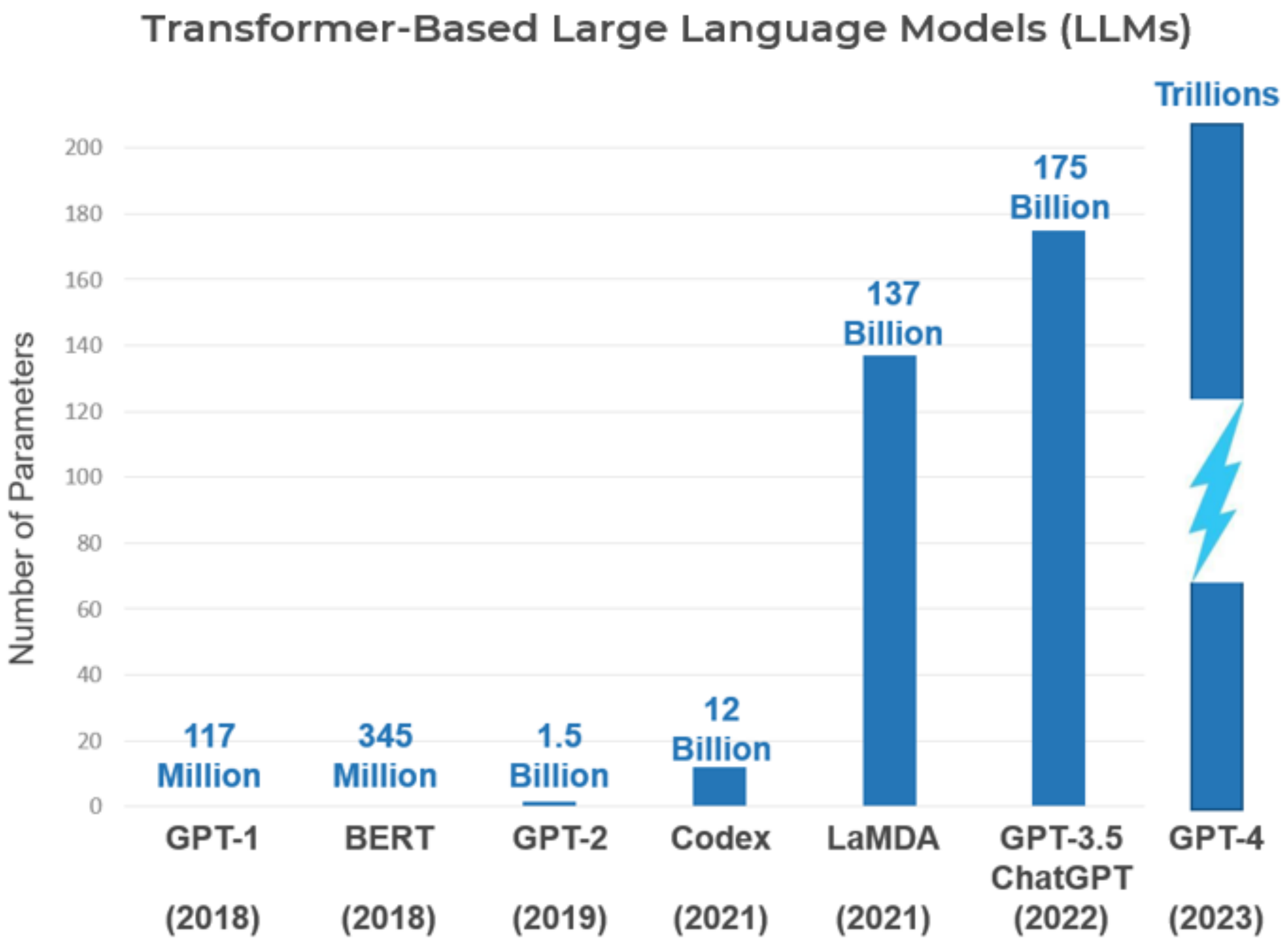
Natural language processing (NLP) is a field of artificial intelligence that focuses on enabling computers to understand, interpret, and generate human language. It involves developing algorithms and models to extract meaning, sentiment, and context from textual or spoken data, enabling applications like machine translation, sentiment analysis, and chatbots.

Generative AI (ChatGPT, Bing) - Defined

- **ChatGPT** is an AI language model developed by OpenAI, designed to understand and generate human-like text in a conversational manner.

- **Generative AI** is a type of artificial intelligence that can create novel content, such as text, images, or music, by learning from existing data without being explicitly programmed to perform specific tasks.

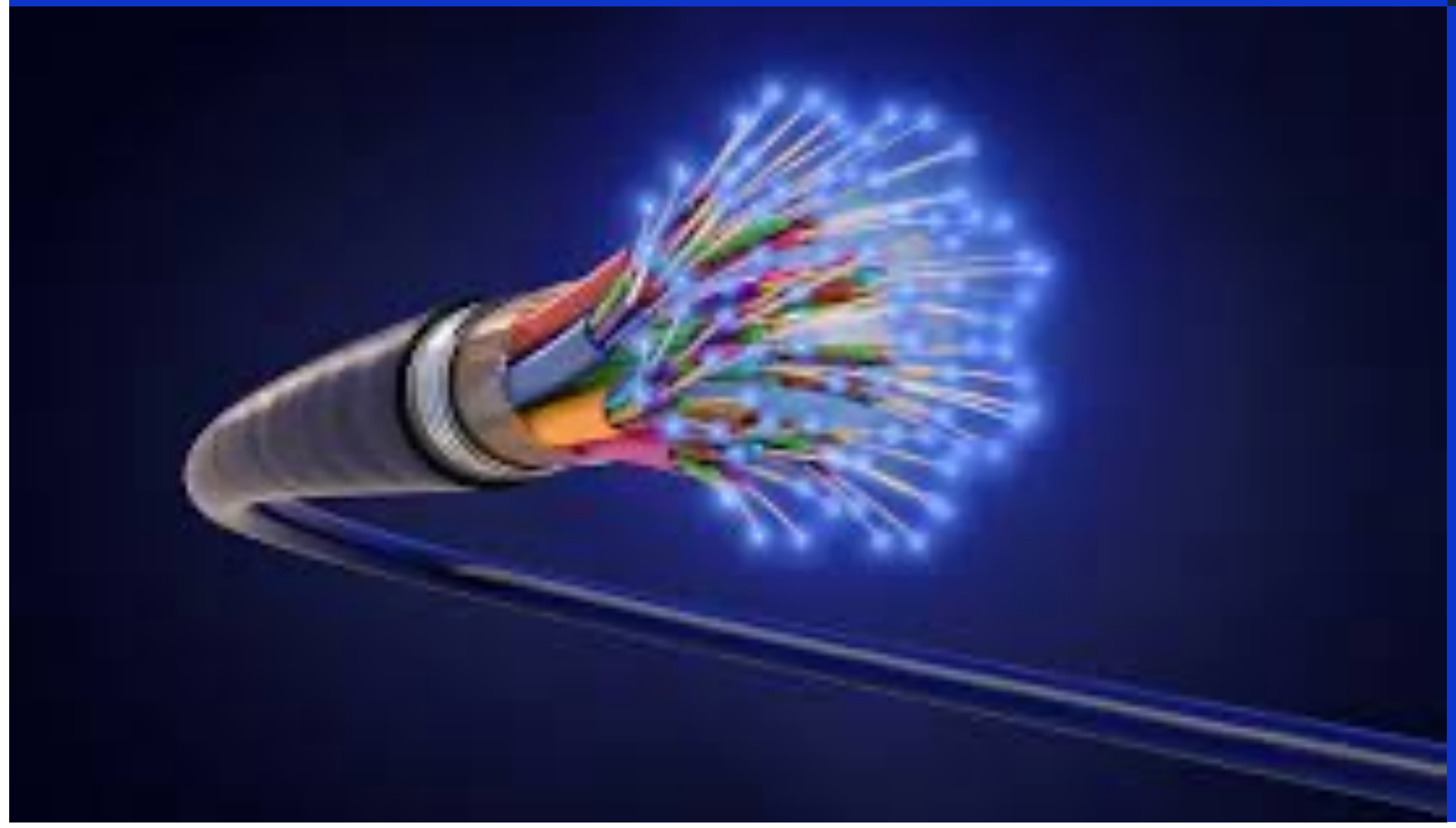
ChatGPT - Training



How Did We Get Here?

Larger pipes to push data – Fiberglass

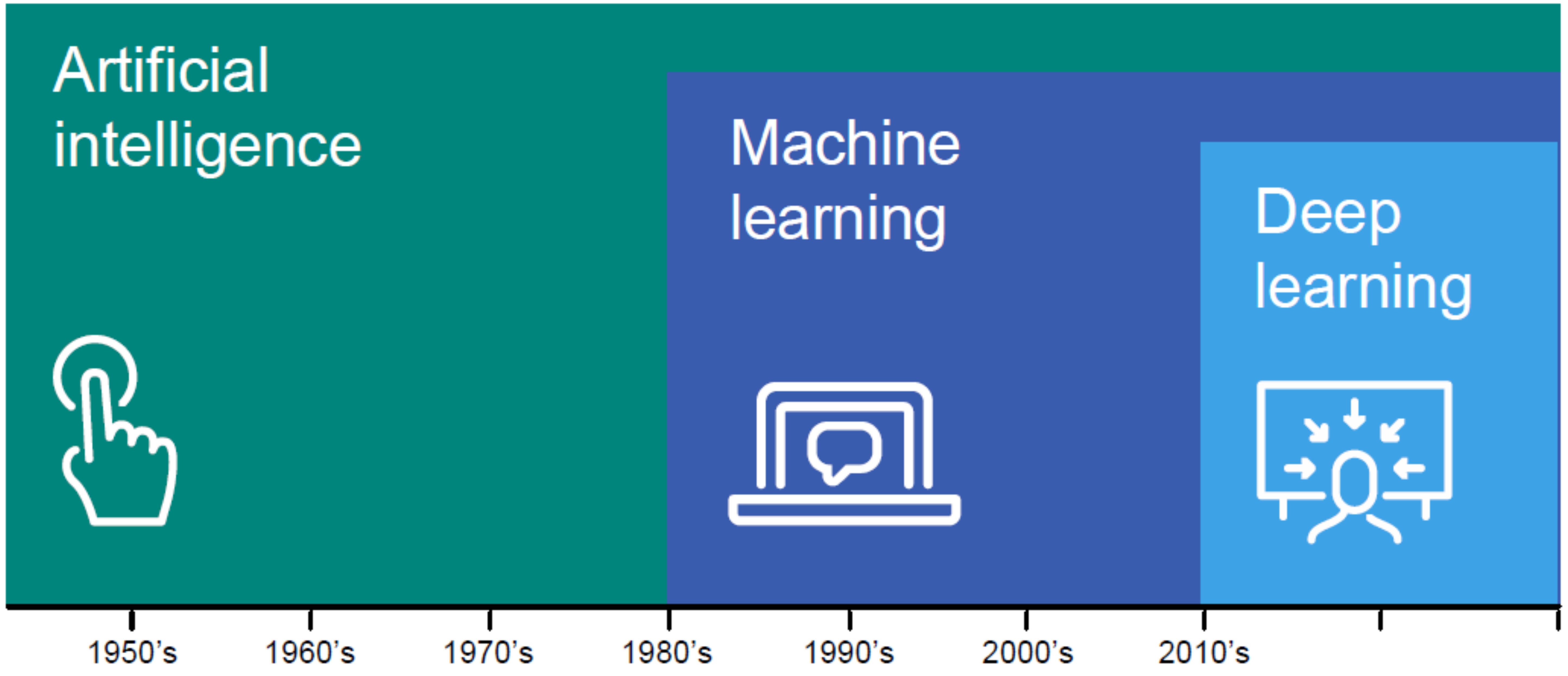
Advances in software development



Advances in hardware – Quantum Computing

AI - Timeline

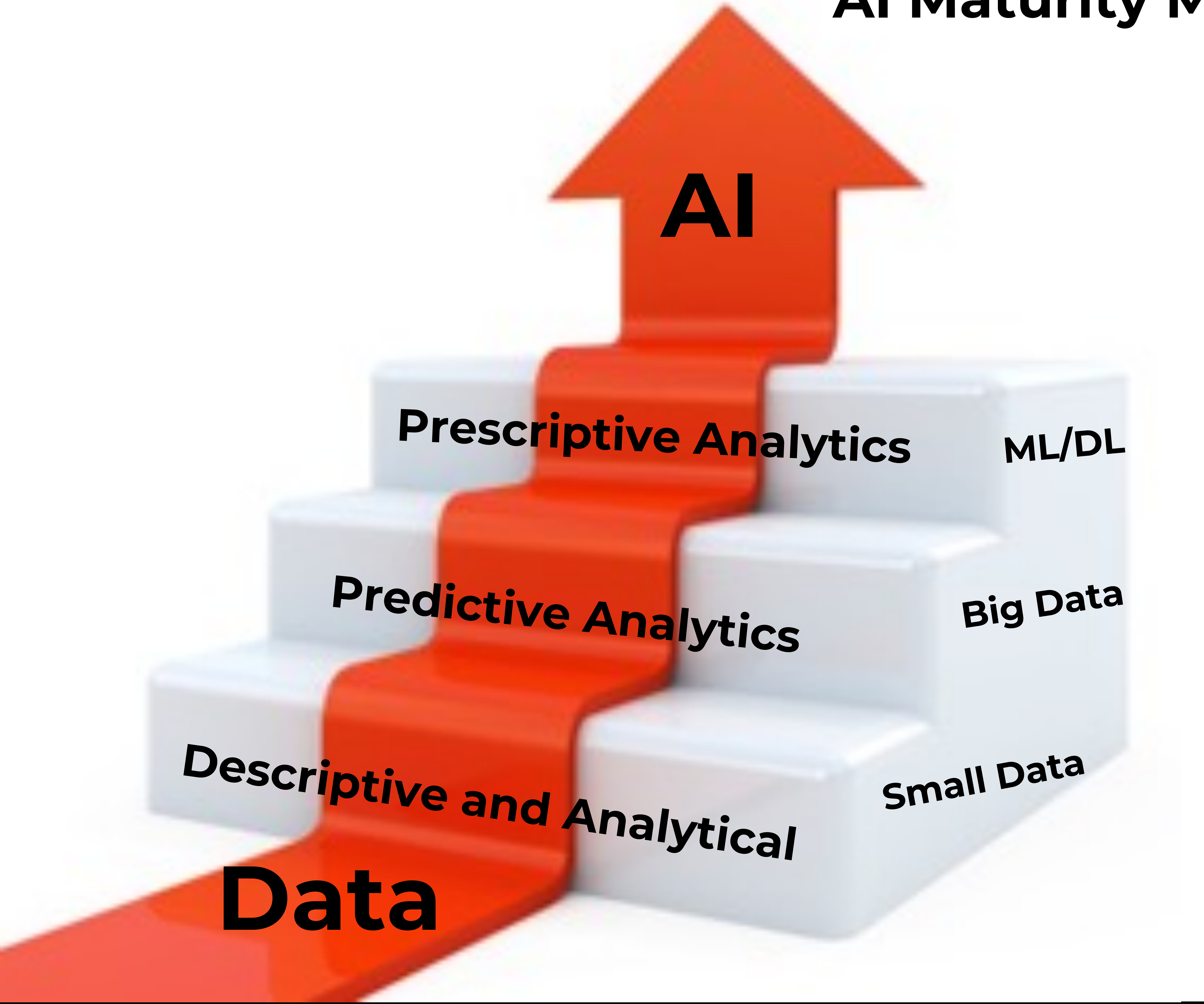
Timeline



Narrow AI came first.
ML uses algorithms to learn.
DL is the most

- recent,
- advanced,
- sophisticated,
- complex, and
- expensive.

AI Maturity Model



1. **Data**
   - Needs to be in order to get started
2. **Descriptive and Analytical**
   - Answer the What and Why
3. **Predictive**
   - Answers where a things are headed
4. **Prescriptive**
   - Answers what is next
5. **Value Chain**
   - Start with the Data

AI – Terms – Cheat Sheet

Artificial Intelligence

is the broad goal of autonomous machine intelligence.

Machine Learning

is the approach or the mechanism we use to achieve that vision. It's where the machine learns from data.

Deep Leaning

is a technique for implementing and executing machine learning, particularly good at utilizing large volumes of data.

Natural Language processing (NLP)

is a field of artificial intelligence that focuses on the interaction between computers and humans through natural language.

Large Language Models (LLM)

are an advanced artificial intelligence system trained on vast amounts of text data to understand, generate, and manipulate human language.

Generative AI

refers to artificial intelligence that can generate new content, from text to images, by learning from vast datasets.

AI DRIVEN CYBER THREATS

Stuxnet

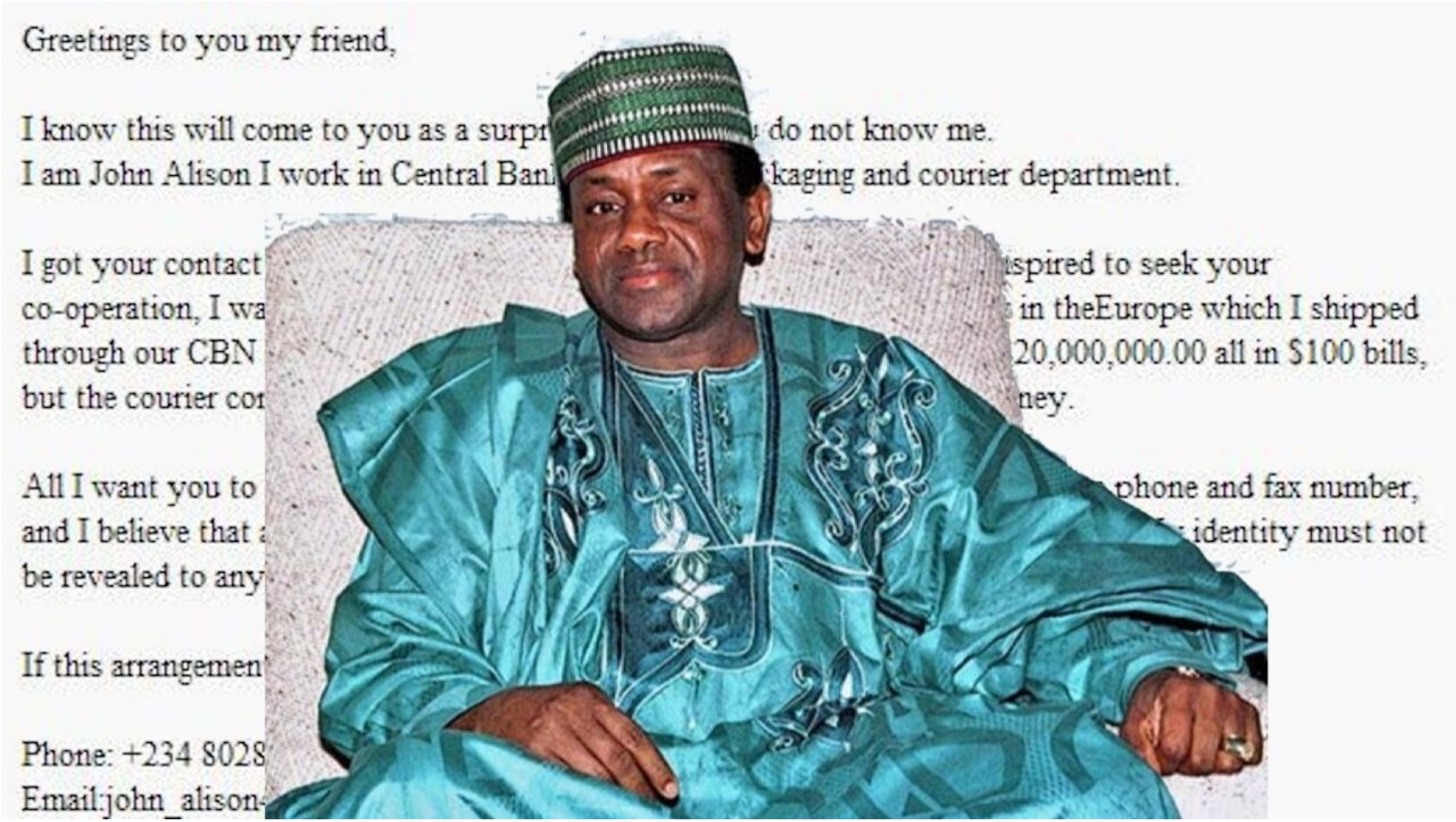
AI Threats

Phishing

Deep Fakes

Malware

Phishing



Greetings to you my friend,

I know this will come to you as a surpr … do not know me.
I am John Alison I work in Central Ban … kaging and courier department.

I got your contact … spired to seek your
co-operation, I wa … in theEurope which I shipped
through our CBN … 20,000,000.00 all in $100 bills,
but the courier co … ney.

All I want you to … phone and fax number,
and I believe that a … identity must not
be revealed to any

If this arrangement

Phone: +234 8028
Email:john_alison

AI - Information Seeking

Information Seekers have changed the game…

1. **Gain access**
2. **Build context**
3. **Know your pattern of writing**
4. **Know current events**
5. **Build convincing content**
6. **Strike**

Simple Steps to Protect

1 **Check the Sender's Email Address** - Phishing emails often come from email addresses that don't match the name of the supposed sender or the company they claim to represent.

2 **Look for Generic Greetings -** Phishing emails often start with generic greetings like "Dear Customer" or "Dear Account Holder" instead of your actual name.

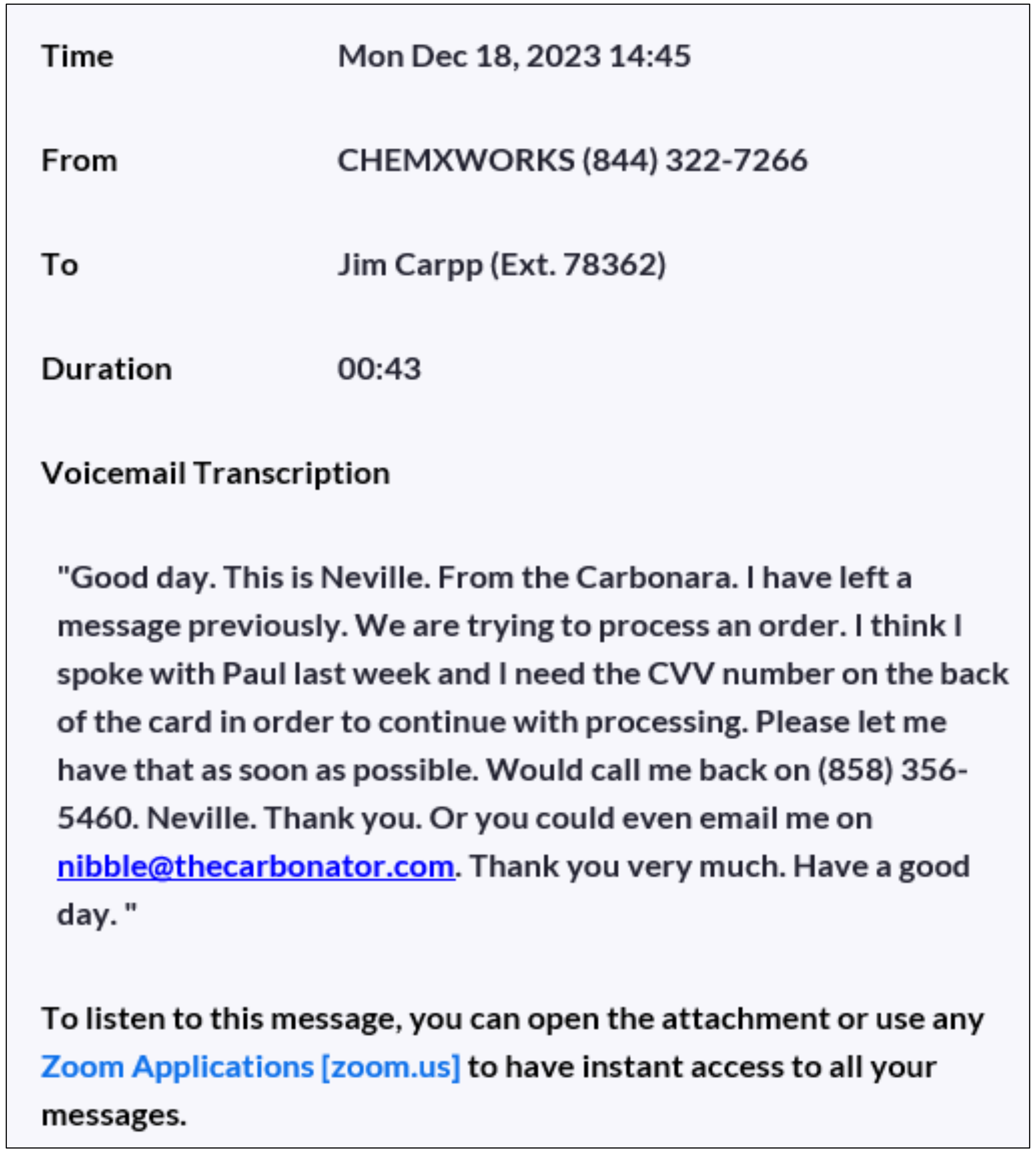
3 **Examine the Email's Language -** Look for poor grammar, misspellings, and awkward language. Legitimate companies usually have professional editors who ensure their emails are error-free.

4 **Beware of Urgent Requests** - Phishing emails often create a sense of urgency, like asking you to update your payment details immediately or risk having your account suspended.

5 **Hover Over Links -** Before clicking on any links in the email, hover over them to see where they lead. Be suspicious if the domain of the URL doesn't match the supposed sender of the email.

Legitimate companies will never ask for sensitive information via email.

They Keep Trying…

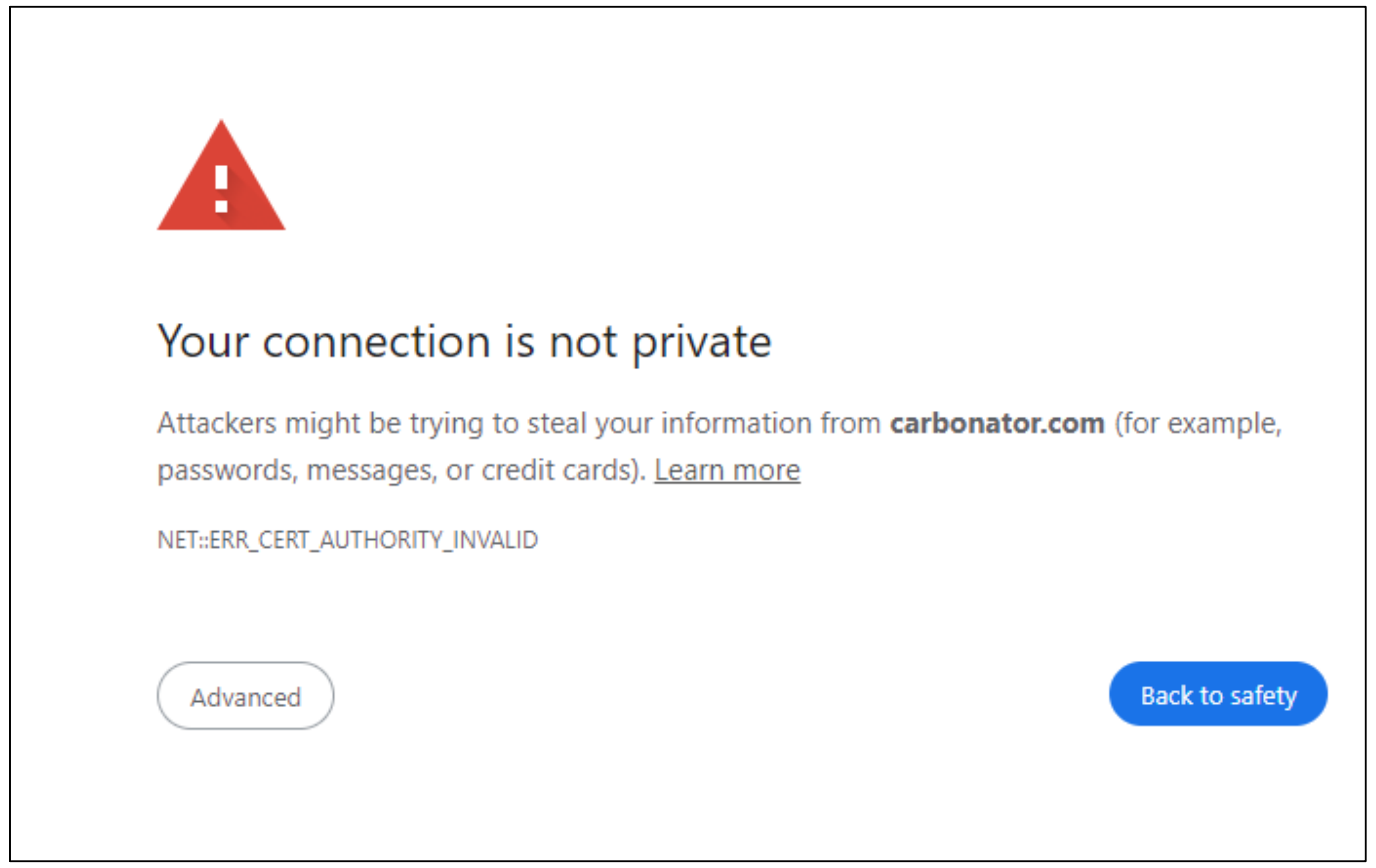


Time	Mon Dec 18, 2023 14:45
From	CHEMXWORKS (844) 322-7266
To	Jim Carpp (Ext. 78362)
Duration	00:43

Voicemail Transcription

"Good day. This is Neville. From the Carbonara. I have left a message previously. We are trying to process an order. I think I spoke with Paul last week and I need the CVV number on the back of the card in order to continue with processing. Please let me have that as soon as possible. Would call me back on (858) 356-5460. Neville. Thank you. Or you could even email me on nibble@thecarbonator.com. Thank you very much. Have a good day. "

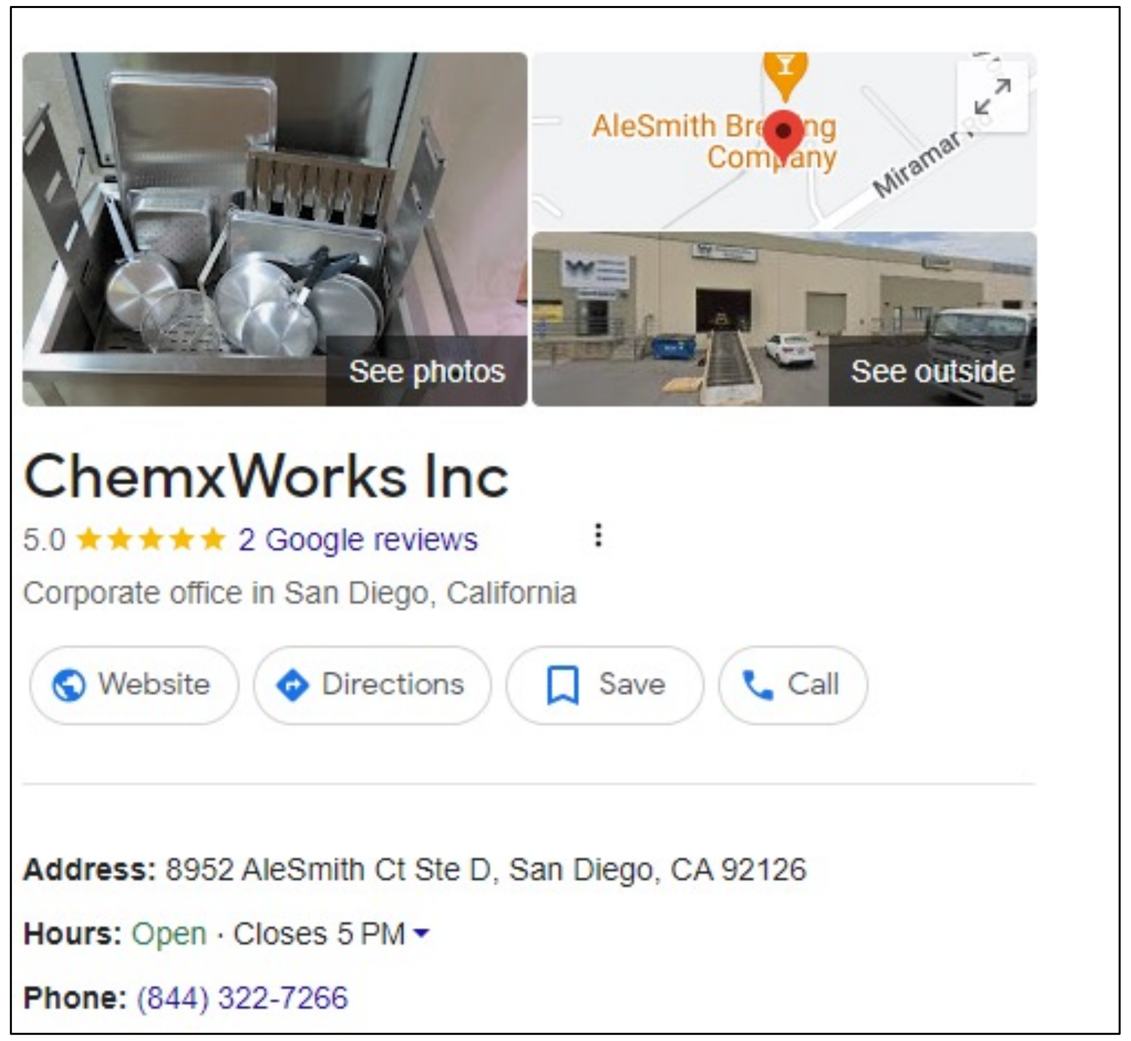
To listen to this message, you can open the attachment or use any Zoom Applications [zoom.us] to have instant access to all your messages.



Your connection is not private

Attackers might be trying to steal your information from **carbonator.com** (for example, passwords, messages, or credit cards). Learn more

NET::ERR_CERT_AUTHORITY_INVALID



ChemxWorks Inc

5.0 ★★★★★ 2 Google reviews

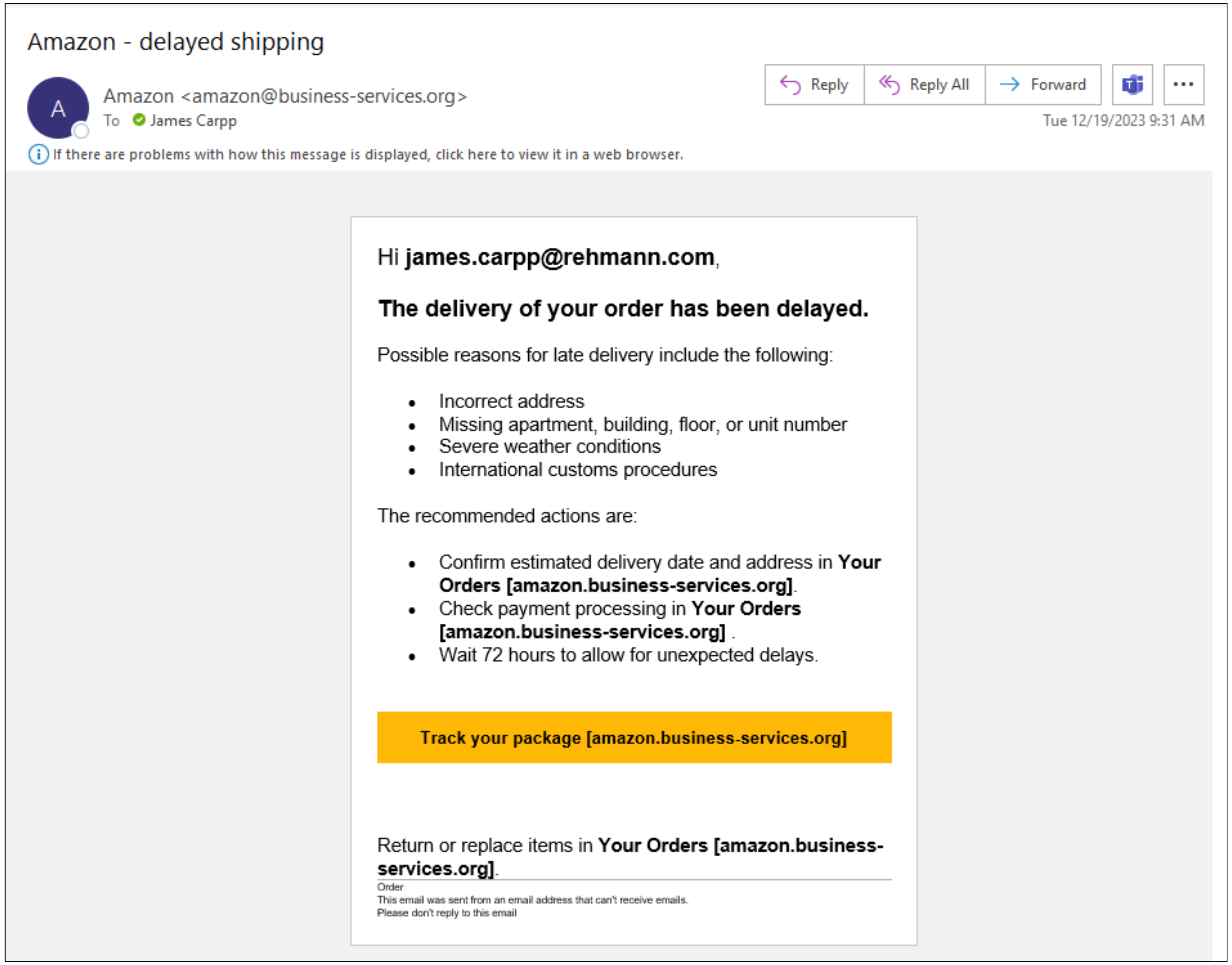
Corporate office in San Diego, California

Address: 8952 AleSmith Ct Ste D, San Diego, CA 92126

Hours: Open · Closes 5 PM

Phone: (844) 322-7266

They Keep Trying…



Amazon - delayed shipping

Amazon <amazon@business-services.org>
To James Carpp

Tue 12/19/2023 9:31 AM

If there are problems with how this message is displayed, click here to view it in a web browser.

Hi **james.carpp@rehmann.com**,

The delivery of your order has been delayed.

Possible reasons for late delivery include the following:

- Incorrect address
- Missing apartment, building, floor, or unit number
- Severe weather conditions
- International customs procedures

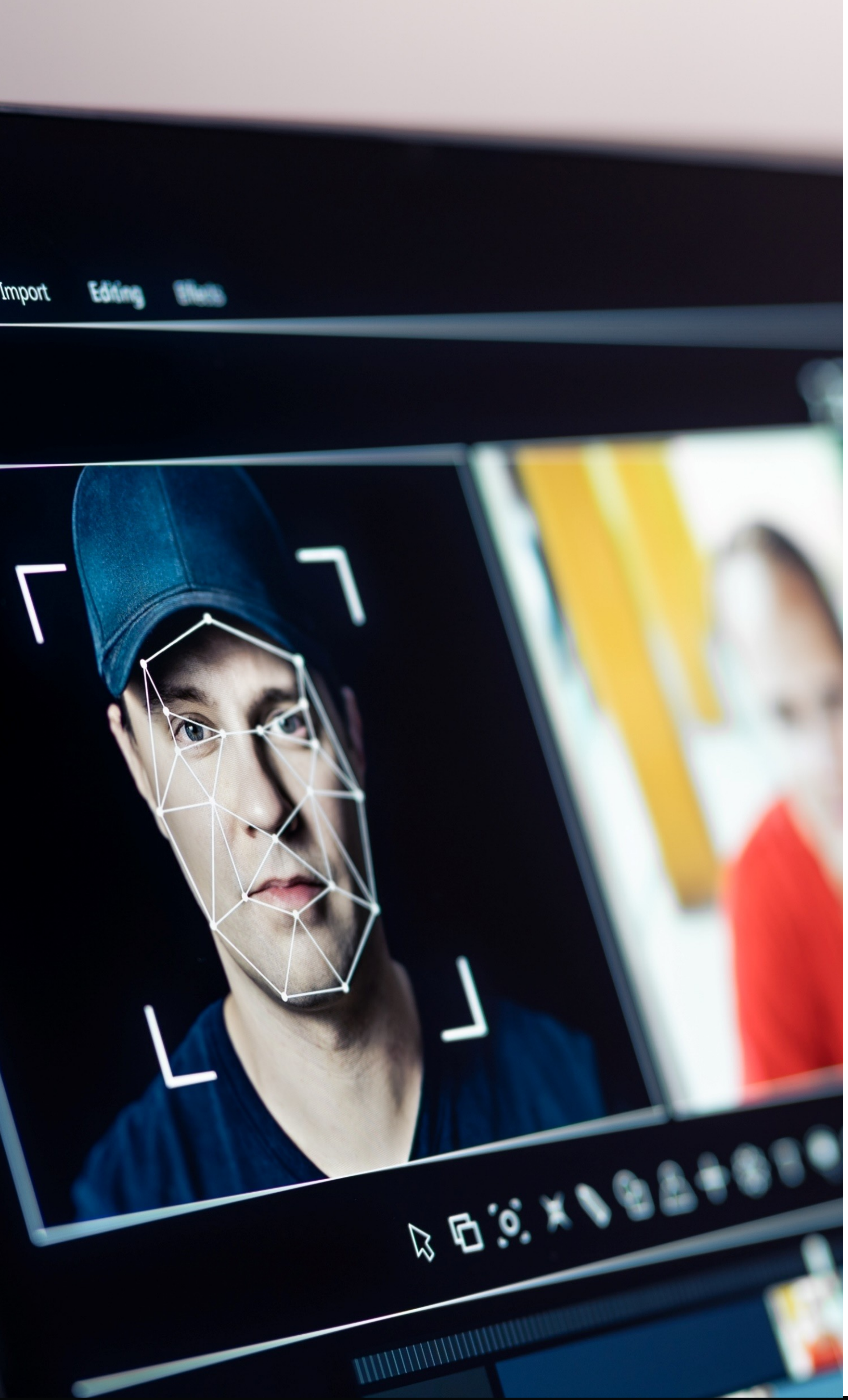
The recommended actions are:

- Confirm estimated delivery date and address in **Your Orders [amazon.business-services.org]**.
- Check payment processing in **Your Orders [amazon.business-services.org]** .
- Wait 72 hours to allow for unexpected delays.

Track your package [amazon.business-services.org]

Return or replace items in **Your Orders [amazon.business-services.org]**.

Order
This email was sent from an email address that can't receive emails.
Please don't reply to this email



Deep Fakes

Deepfake Audio

Deepfake audio is a type of synthetic media that uses artificial intelligence to generate audio recordings of people saying things they never actually said.

Voice Phishing

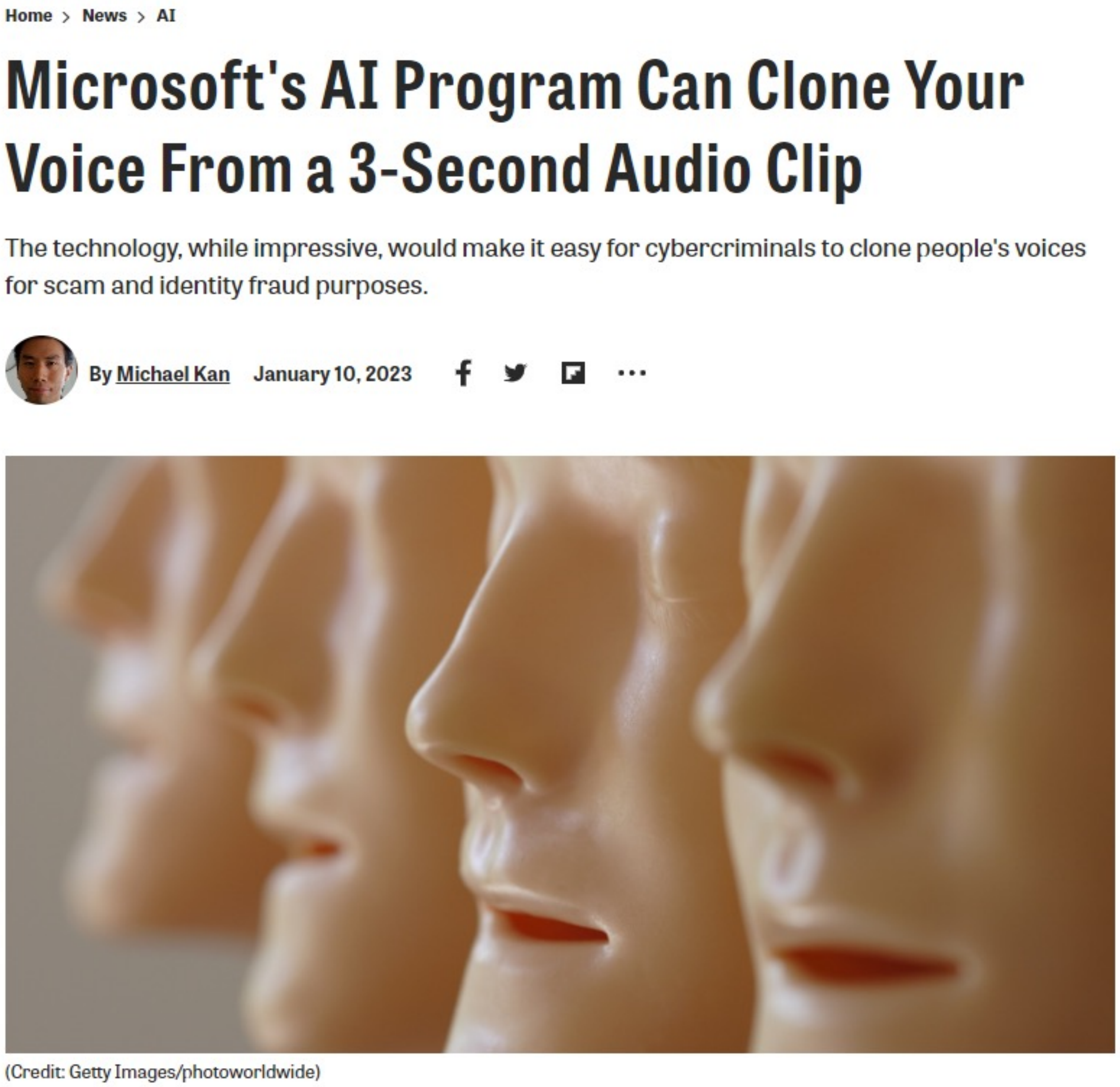
Voice phishing, also known as vishing, is a type of social engineering attack that uses voice communication to trick people into revealing sensitive information, such as passwords or credit card numbers.

Voice Cloning

Voice cloning is a technique that uses machine learning algorithms to create a synthetic voice that sounds like a real person.

Deepfake Video

A deepfake video is a type of synthetic media where a person's likeness in an image or video is swapped with another person's likeness using artificial intelligence[12]. These videos are created to make the manipulated content appear authentic.



Home > News > AI

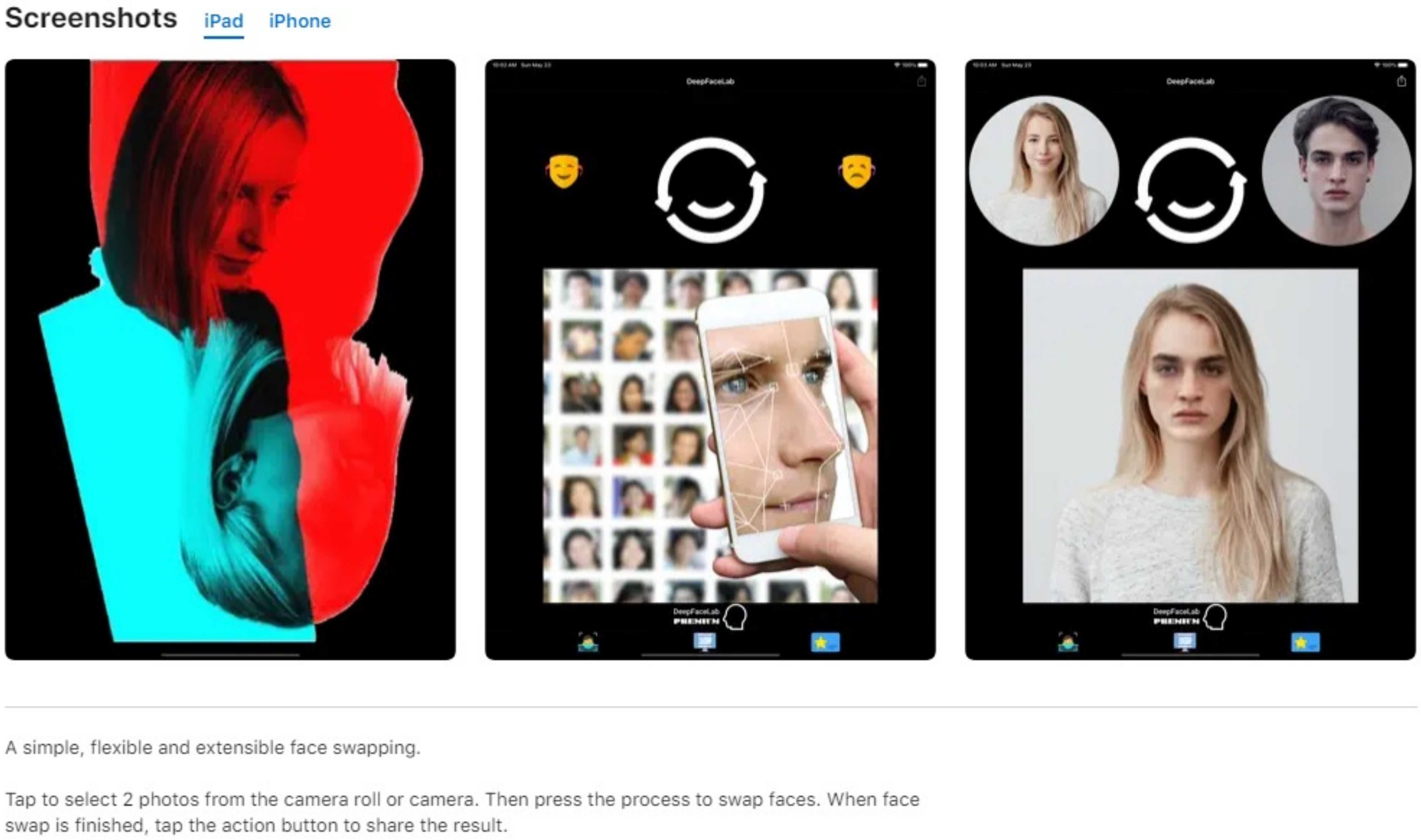
Microsoft's AI Program Can Clone Your Voice From a 3-Second Audio Clip

The technology, while impressive, would make it easy for cybercriminals to clone people's voices for scam and identity fraud purposes.

By Michael Kan January 10, 2023

(Credit: Getty Images/photoworldwide)

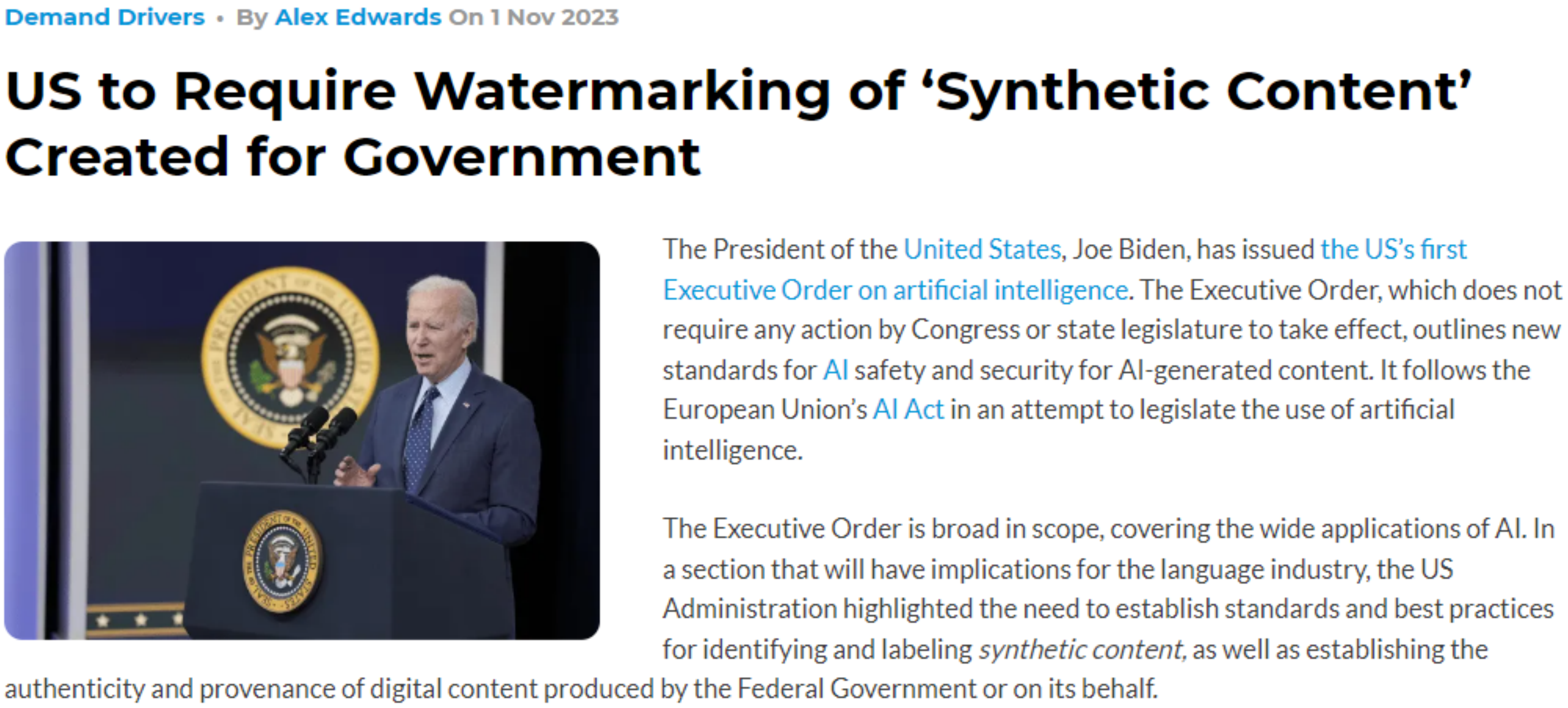
DeepFaceLab



A simple, flexible and extensible face swapping.

Tap to select 2 photos from the camera roll or camera. Then press the process to swap faces. When face swap is finished, tap the action button to share the result.

Synthetic Content



Demand Drivers • By Alex Edwards On 1 Nov 2023

US to Require Watermarking of ‘Synthetic Content’ Created for Government

The President of the United States, Joe Biden, has issued the US’s first Executive Order on artificial intelligence. The Executive Order, which does not require any action by Congress or state legislature to take effect, outlines new standards for AI safety and security for AI-generated content. It follows the European Union’s AI Act in an attempt to legislate the use of artificial intelligence.

The Executive Order is broad in scope, covering the wide applications of AI. In a section that will have implications for the language industry, the US Administration highlighted the need to establish standards and best practices for identifying and labeling *synthetic content,* as well as establishing the authenticity and provenance of digital content produced by the Federal Government or on its behalf.

AI’s Impact on Malware

Enhancing Cybersecurity

AI has become a critical component of all aspects of cybersecurity, including threat detection, prevention, and response. AI-based solutions can alert admins of anomalous behavior patterns and detect zero-day threats and polymorphic malware.

Automating Malware Creation

AI is being used by attackers to automate the process of launching phishing attacks and other malware. This allows them to scale up their operations and launch more attacks more quickly.

AI-Powered Malware

AI has become a double-edged sword as it can also aid in creating and spreading malware used in cyberattacks. AI-enabled attacks occur when threat actors take advantage of AI as a tool to assist in creating a piece of malware, or in conducting an attack.

Increasing Threat Severity

High-severity malware threats, which can lead to stolen data, loss of customer trust, and damage to reputation and brand, are up 86% year-over-year. Cybercriminals are taking advantage of the work-from-home trend to target remote workers and their company's data.

Source - https://venturebeat.com/datadecisionmakers/dangerous-malware-is-up-86-heres-how-ai-can-help/

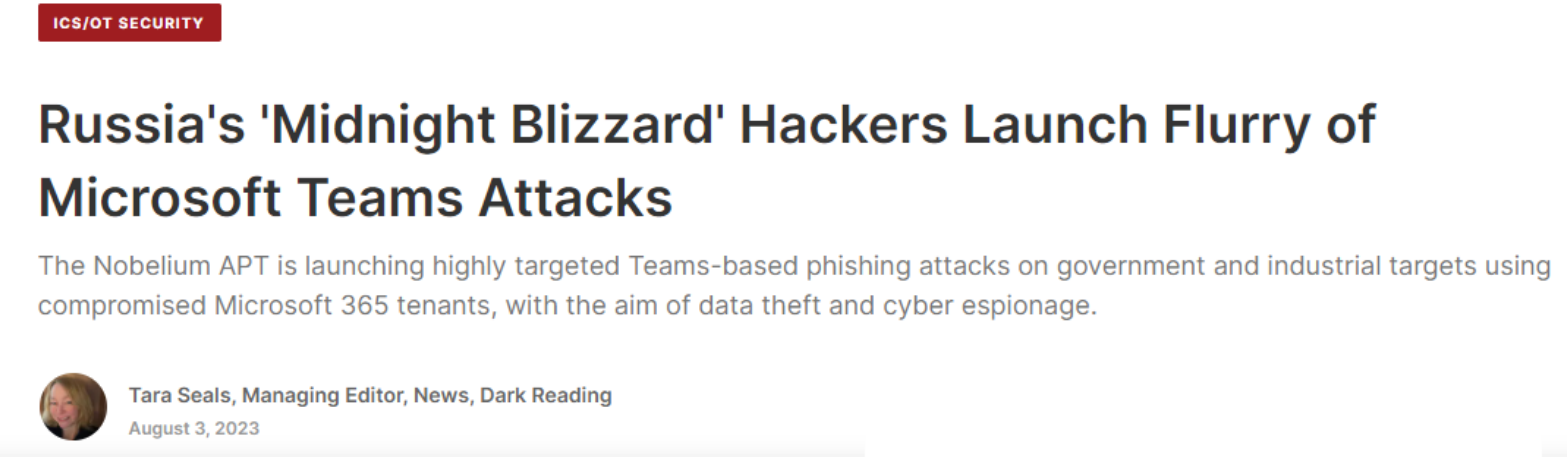
IBM’s DeepLocker – Sniper Attack

Security Intelligence

DeepLocker: How AI Can Power a Stealthy New Breed of Malware

Cybercriminals are also studying AI to use it to their advantage — and weaponize it

Midnight Blizzard



ICS/OT SECURITY

Russia's 'Midnight Blizzard' Hackers Launch Flurry of Microsoft Teams Attacks

The Nobelium APT is launching highly targeted Teams-based phishing attacks on government and industrial targets using compromised Microsoft 365 tenants, with the aim of data theft and cyber espionage.

Tara Seals, Managing Editor, News, Dark Reading

August 3, 2023

PROTECTION - WHAT CAN YOU DO?

Back to the Basics

NIST Cybersecurity Framework

1. **Identify -** Develop an organizational understanding to manage cybersecurity risk to systems, people, assets, data, and capabilities. This includes identifying physical and software assets, the business environment, cybersecurity policies, asset vulnerabilities, threats, risk response activities, risk management strategy, and supply chain risk management strategy.

2. **Protect** - Implement appropriate safeguards to ensure delivery of critical infrastructure services. This includes protections for identity management and access control, awareness and training for staff, and establishing data security protection.

3. **Detect** - Implement appropriate activities to identify the occurrence of a cybersecurity event.

4. **Respond** - Implement appropriate activities to take action regarding a detected cybersecurity event.

5. **Recover** - Implement appropriate activities to maintain plans for resilience and to restore any capabilities or services that were impaired due to a cybersecurity event.

How Can AI Help?

Threat Detection and Response

AI can help detect and respond to cyber threats in real time.

Data Protection

AI can help protect sensitive data.

Automation

AI can automate routine security tasks, freeing up time for security teams to focus on more complex issues.

Training

AI can be used to provide cybersecurity training to employees.

Incident Response

In the event of a security incident, AI can help analyze the incident, identify the cause, and recommend steps to prevent similar incidents in the future

AI can enhance cybersecurity, it’s not a silver bullet.

Endpoint Detection and Response (EDR)

Endpoint Detection and Response (EDR) is a cybersecurity technology that continuously monitors end-user devices (also known as "endpoints", like mobile phones, laptops, or Internet-of-Things devices) to detect and respond to cyber threats such as ransomware and malware. EDR encompasses:

- **Continuous Monitoring** - EDR security solutions record the activities and events taking place on endpoints, providing security teams with the visibility they need to uncover incidents that would otherwise remain invisible.
- **Threat Detection** - EDR technology pairs comprehensive visibility across all endpoints with Indicators of Attack (IOAs) and applies behavioral analytics that analyze billions of events in real time to automatically detect traces of suspicious behavior.
- **Incident Response** - If a sequence of events matches a known IOA, the EDR tool will identify the activity as malicious and automatically send a detection alert.
- **Integration with Threat Intelligence** - Integration with cyber threat intelligence provides faster detection of the activities and tactics, techniques, and procedures (TTPs) identified as malicious.
- **Managed Threat Hunting** - Using EDR, the threat hunters work proactively to hunt, investigate, and advise on threat activity in your environment.
- **Real-time and Historical Visibility** - An EDR solution needs to provide continuous and comprehensive visibility into what is happening on endpoints in real time.

EDR - Vendors

- **Secureworks** – Provides a cloud-native security analytics platform, Secureworks Taegis™, built on real-world threat intelligence and research.
- **Artic Wolf** – Provides a security operations center (SOC)-as-a-service, managed detection and response (MDR), and managed risk services. They aim to protect organizations by providing security operations as a concierge service.
- **Sentinelone** - Provides a security AI platform to protect the entire enterprise. Their platform unites endpoint, cloud, and identity threat protection for a seamless and efficient cybersecurity experience. They offer 24/7/365 threat hunting and managed services.
- **Huntress -** Provides a managed security platform offering services such as endpoint detection and response, antivirus protection, ransomware detection, and security awareness training. They use a combination of automated detection and human-powered threat hunting to find and stop hidden threats that sneak past preventive security tools.
- **Crowdstrike** – Provides a cloud-native platform to protect and enable the people, processes, and technologies that drive modern enterprises. Their platform secures the most critical areas of risk – endpoints and cloud workloads, identity, and data – and leverages real-time indicators of attack, threat intelligence, and enriched telemetry from across the enterprise to deliver hyper-accurate detections, automated protection and remediation, elite threat hunting, and prioritized observability of vulnerabilities.
- **Microsoft Defender** - Provides advanced attack detections that are near real-time and actionable, enabling security analysts to prioritize alerts effectively, gain visibility into the full scope of a breach, and take response actions to remediate threats.

Humans Are Your Weakest Link

Basic Tenants to Consider

- Do we have a **firewall** in place? A firewall is the first line of defense in network security.
- Are we using **intrusion detection systems** (IDS)? IDS can identify suspicious activity and potential threats.
- Do we have a secure **VPN** for remote access? A VPN can provide secure access to your network for remote employees.
- Are all our systems and **software up-to-date**? Outdated systems can have vulnerabilities that hackers can exploit.
- Do we regularly perform **vulnerability assessments** and penetration testing? These tests can identify weaknesses in your security before a hacker does.
- Do we have a **strong password policy**? Strong, unique passwords are crucial for securing accounts.
- Are we using **multi-factor authentication**? This adds an extra layer of security to prevent unauthorized access.
- How do we **manage and monitor user privileges**? It's important to give employees access only to the information they need to do their jobs.
- Do we **encrypt sensitive data**? Encryption can protect your data, even if it falls into the wrong hands.
- Do we have a plan for managing and **responding to security incidents**? A clear plan can help minimize damage in the event of a breach.

Q&A

Scan with your phone!

Read the 15 questions you should ask your IT Team

James E. Carpp, CISA, CRISC, CIRM, CISM
Chief Digital Officer
james.carpp@rehmann.com