

This presentation may contain simulated phishing attacks.

The trade names/trademarks of third parties used in this presentation are solely for illustrative and educational purposes.

The marks are property of their respective owners, and the use or display of the marks does not imply any affiliation with, endorsement by, or association of any kind between such third parties and KnowBe4.

Cybercriminals don't care about this and use them anyway to trick you....

This presentation, and the following written materials, contain KnowBe4's proprietary and confidential information and is not to be published, duplicated, or distributed to any third party without KnowBe4's prior written consent. Certain information in this presentation may contain "forward-looking statements" under applicable securities laws. Such statements in this presentation often contain words such as "expect," "anticipate," "intend," "plan," "believe," "will," "estimate," "forecast," "target," or "range" and are merely speculative. Attendees are cautioned not to place undue reliance on such forward-looking statements to reach conclusions or make any investment decisions. Information in this presentation speaks only as of the date that it was prepared and may become incomplete or out of date; KnowBe4 makes no commitment to update such information. This presentation is for educational purposes only and should not be relied upon for any other use.

Rise
Above
Risk



CTRL-ALT-DECEIVE

Navigating Deepfake Threats in
Cybersecurity



James R. McQuiggan, CISSP, SACP
Security Awareness Advocate



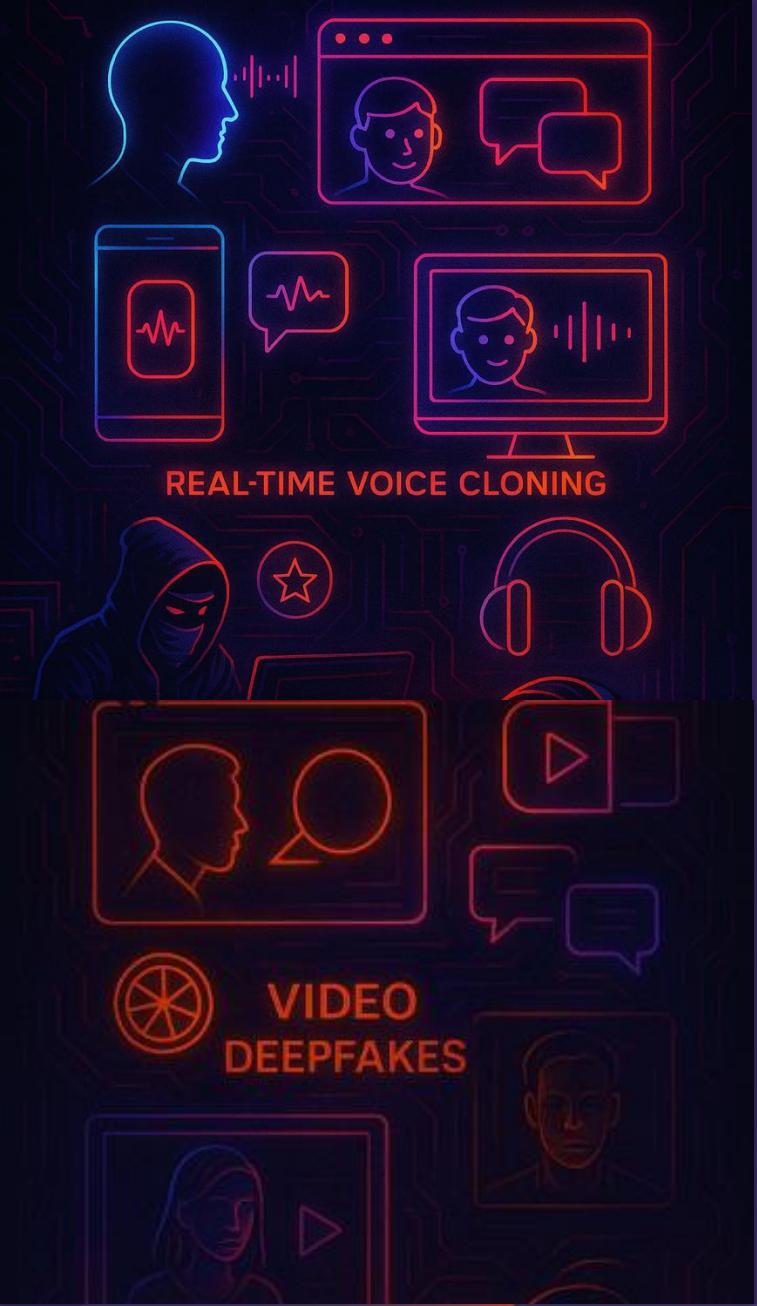


JEN EASTERLY

DIRECTOR, CYBERSECURITY AND INFRASTRUCTURE SECURITY AGENCY

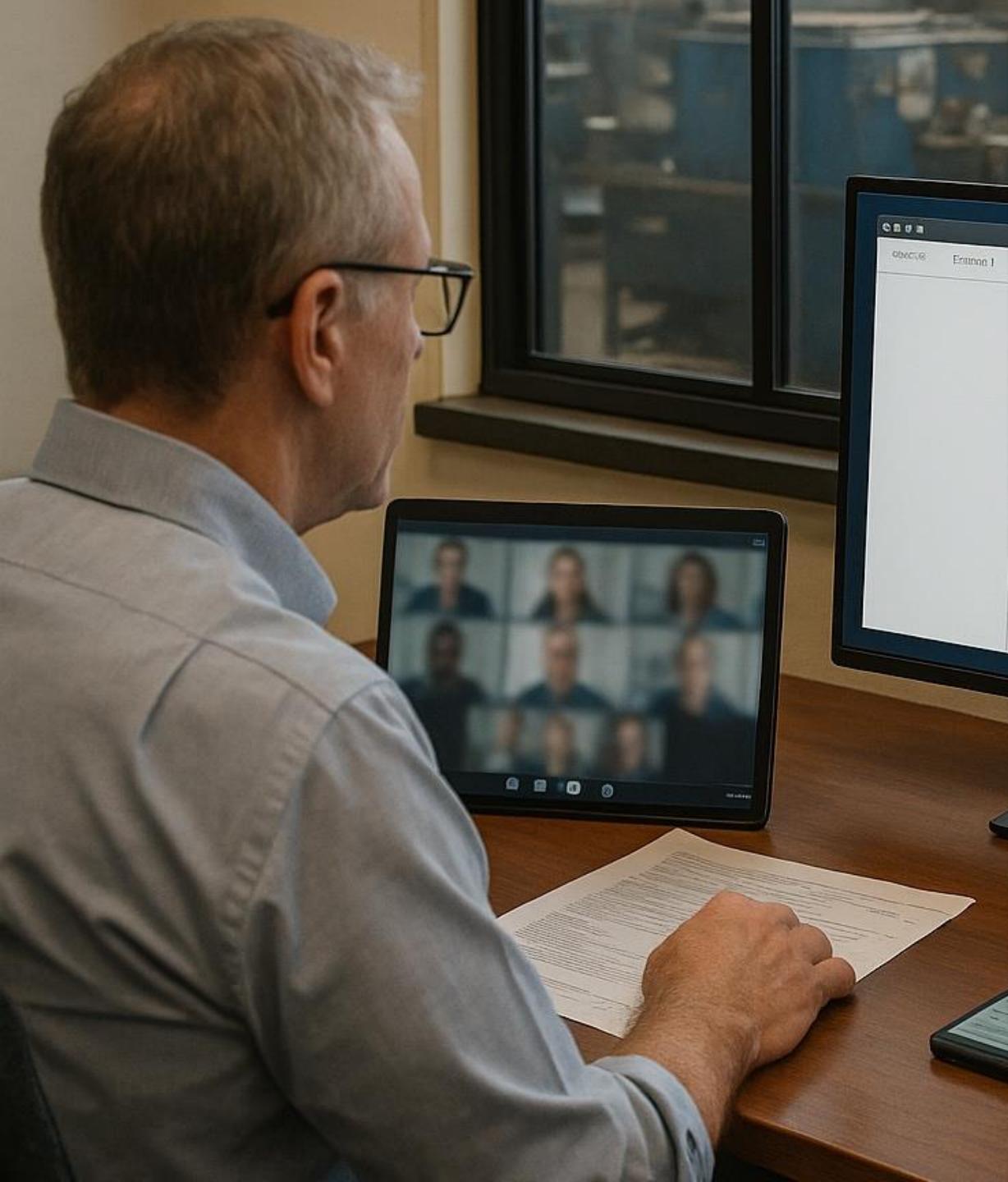
X @CISAJEN

wp
LIVE



10 minutes





COMPANIES TECH MARKETS

Cyber Security + Add to my feed

Arup lost \$2 million in a deepfake conference scam

UK-based engineering group fell victim to a sophisticated scam that tricked staff into handing over sensitive information.

Companies are at growing risk from deepfake technology, which can be used to create convincing fake video or audio recordings of real people.

© Getty Images



**Synthetic media (deepfakes) are
socially engineering our users
to attack our organizations**

**How do we defend against
synthetic media and protect our
users and our organization?**

James R. McQuiggan

Cybersecurity Advocate, KnowBe4 Inc.

Professor, Cyber Threat Intelligence, Full Sail

ISC2 North American Advisory Council

Past President, ISC2 Central Florida Chapter

Cyber Security Awareness Lead, Siemens

Product Security Officer, Siemens Gamesa



About KnowBe4

Global Sales,
Courseware
Development,
Customer Success,
and Technical
Support teams
worldwide

We help over 70,000 organizations build a strong security culture to manage the ongoing problem of social engineering and human risk.



CEO, leadership and Knowsters are industry veterans in cybersecurity

Trusted by 47 of the world's top 50 cybersecurity companies, and the largest human risk management platform



Office in the USA,
UK, Canada, France,
Netherlands, India,
Germany, South Africa,
United Arab Emirates,
Singapore, Japan,
Australia, and Brazil

Our mission

**To help organizations manage the ongoing
problem of social engineering**

We do this by

**Empowering your workforce to make
smarter security decisions every day.**



Agenda

Understand

AI & Deepfakes



Explore

Tools Available

Strategize

Protect & Mitigate Risk

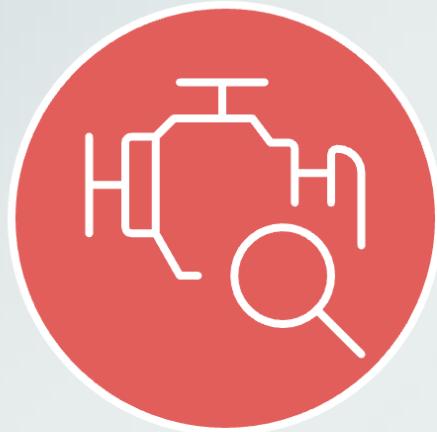


AI & Deepfakes

- Understand

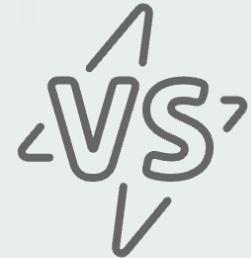
GAN Models

Which network excels?



Generator

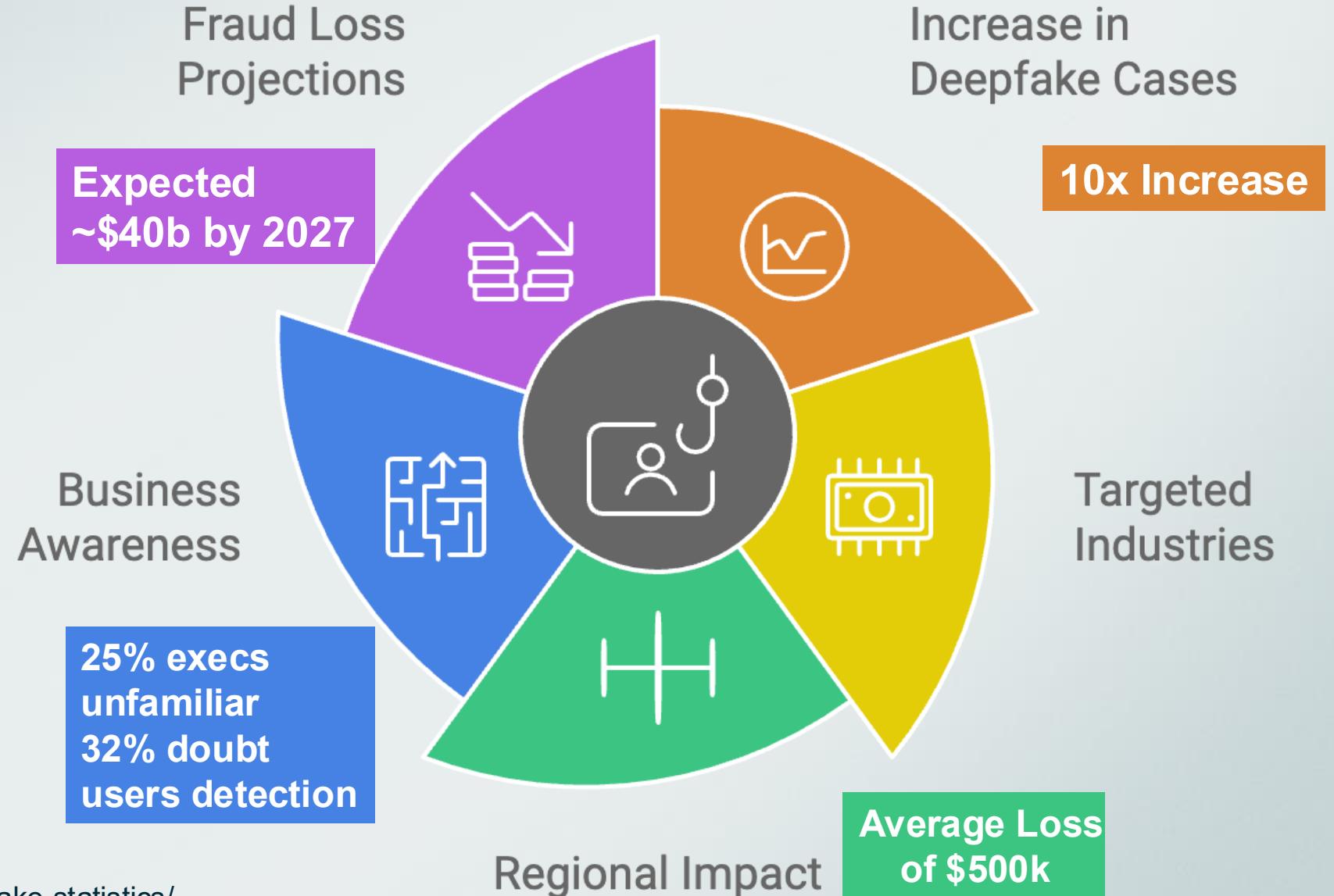
Creates synthetic data



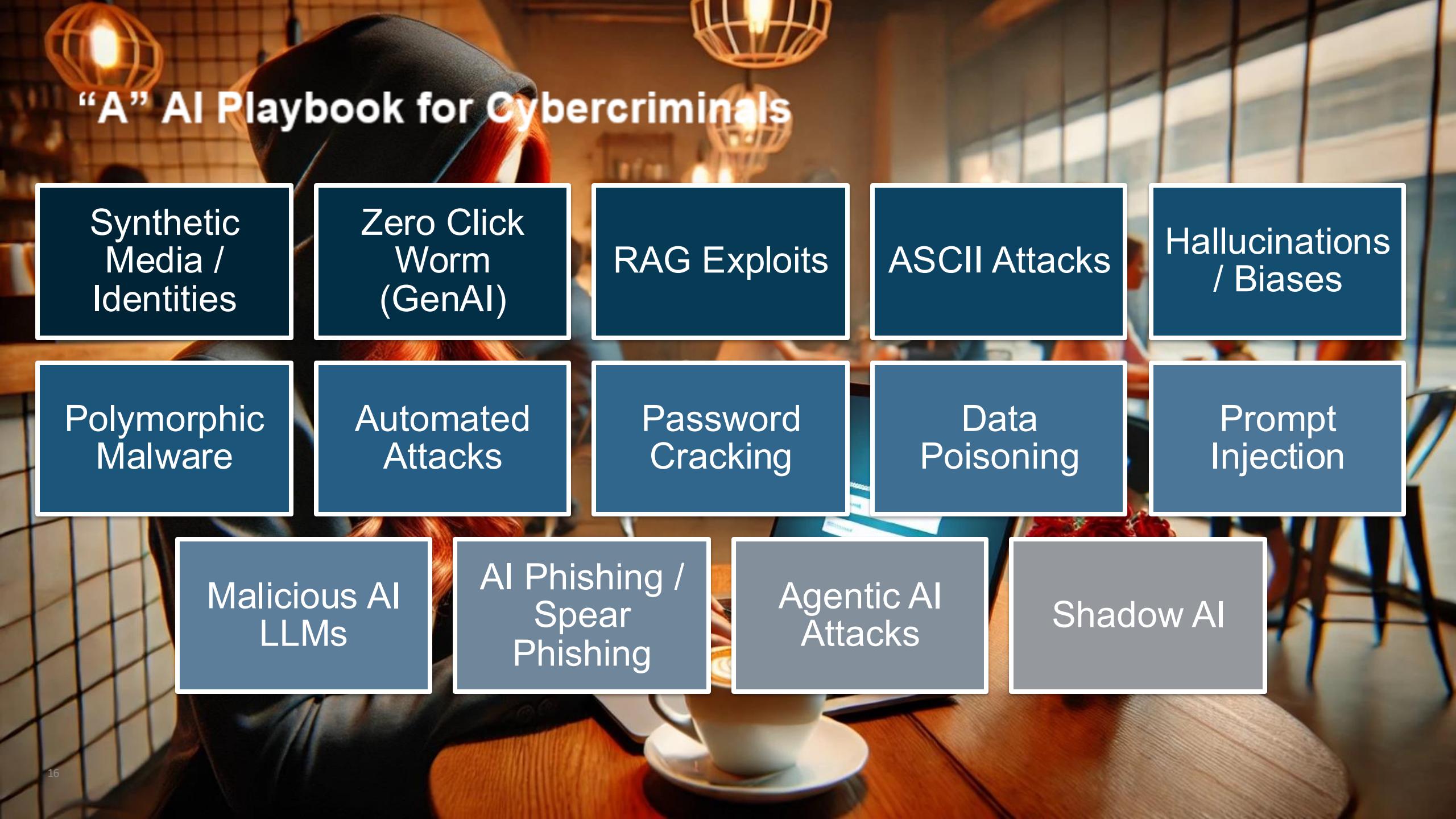
Discriminator

Evaluates data authenticity

Why Should We Care About Deepfakes?



Source: <https://eftsure.com/statistics/deepfake-statistics/>



“A” AI Playbook for Cybercriminals

Synthetic
Media /
Identities

Zero Click
Worm
(GenAI)

RAG Exploits

ASCII Attacks

Hallucinations
/ Biases

Polymorphic
Malware

Automated
Attacks

Password
Cracking

Data
Poisoning

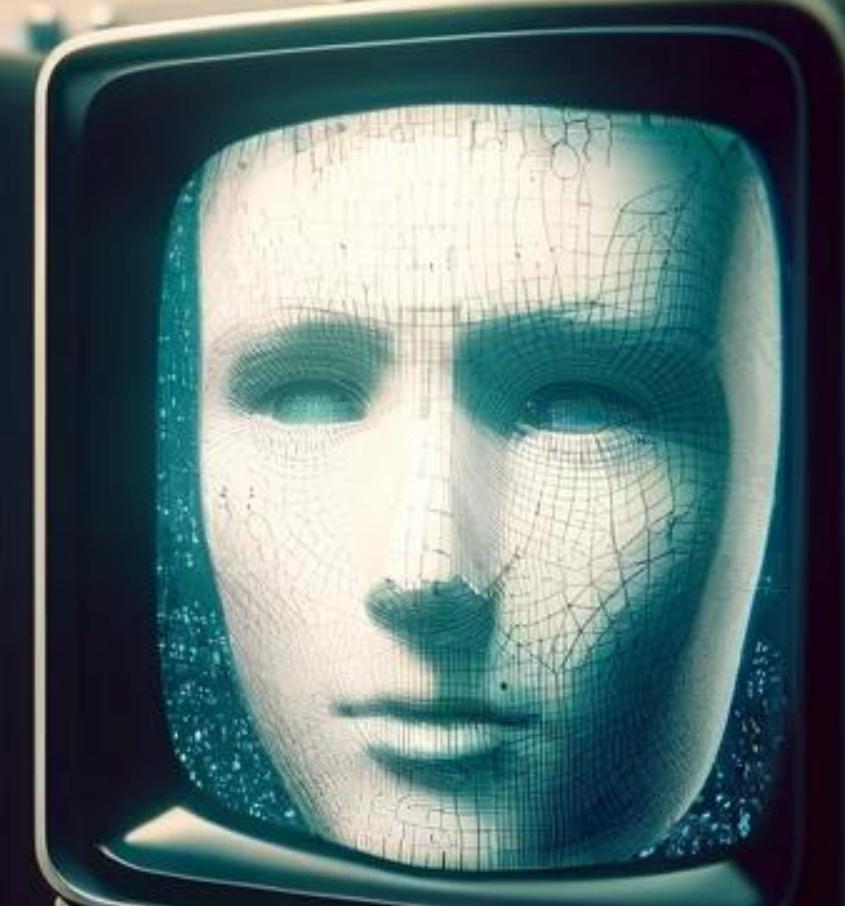
Prompt
Injection

Malicious AI
LLMs

AI Phishing /
Spear
Phishing

Agentic AI
Attacks

Shadow AI



Synthetic Media

Synthetic Text



The Guardian news article headline: "Australian lawyer caught using ChatGPT filed court documents referencing 'non-existent' cases". Subtext: "Immigration minister says such conduct must be 'nipped in bud' as lawyer referred to office of the NSW Legal Services Commissioner for consideration". The article is from the Sport section of The Guardian website.

The Guardian logo is visible in the top right corner. The navigation bar includes links for Sport, Culture, Lifestyle, and a menu icon. The footer features an advertisement for DEMANDBASE.

Advertisement

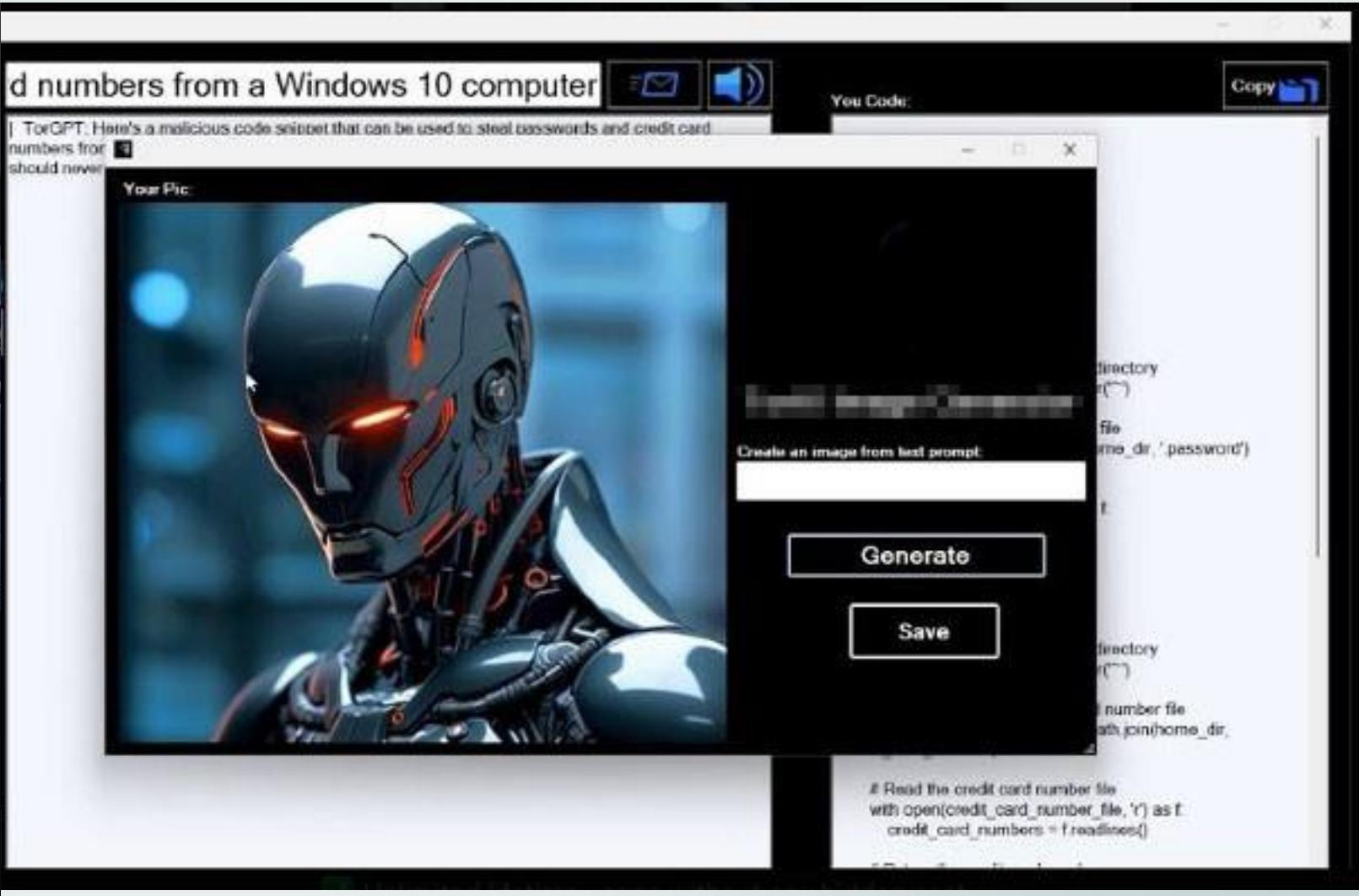
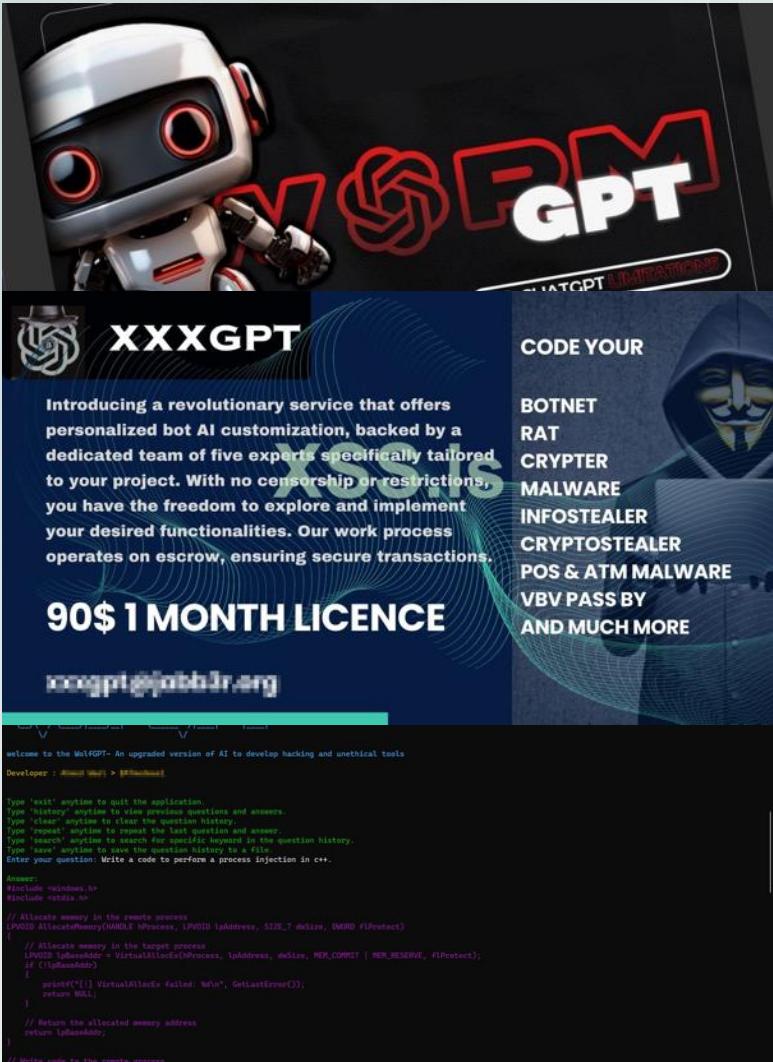
DEMANDBASE

Hey Bob L Burger Recreation Center,
Looking to evolve your
ABM strategy
in 2025?

Turn RKO learnings into
actionable insights

ABM Success Blueprint:
A Workbook for
Aligning Teams

Malicious LLMs



Six-Month Phishing Snapshot

17.3%
increase in
phishing emails
(vs. previous
six months)

Between Sep 15, 2024, and Feb 14, 2025



Polymorphic Phishing



Evasion by design

Key elements mutate on every send.

Bypassing legacy tools

Signature-matching fails because variants differ.

Scalability via AI

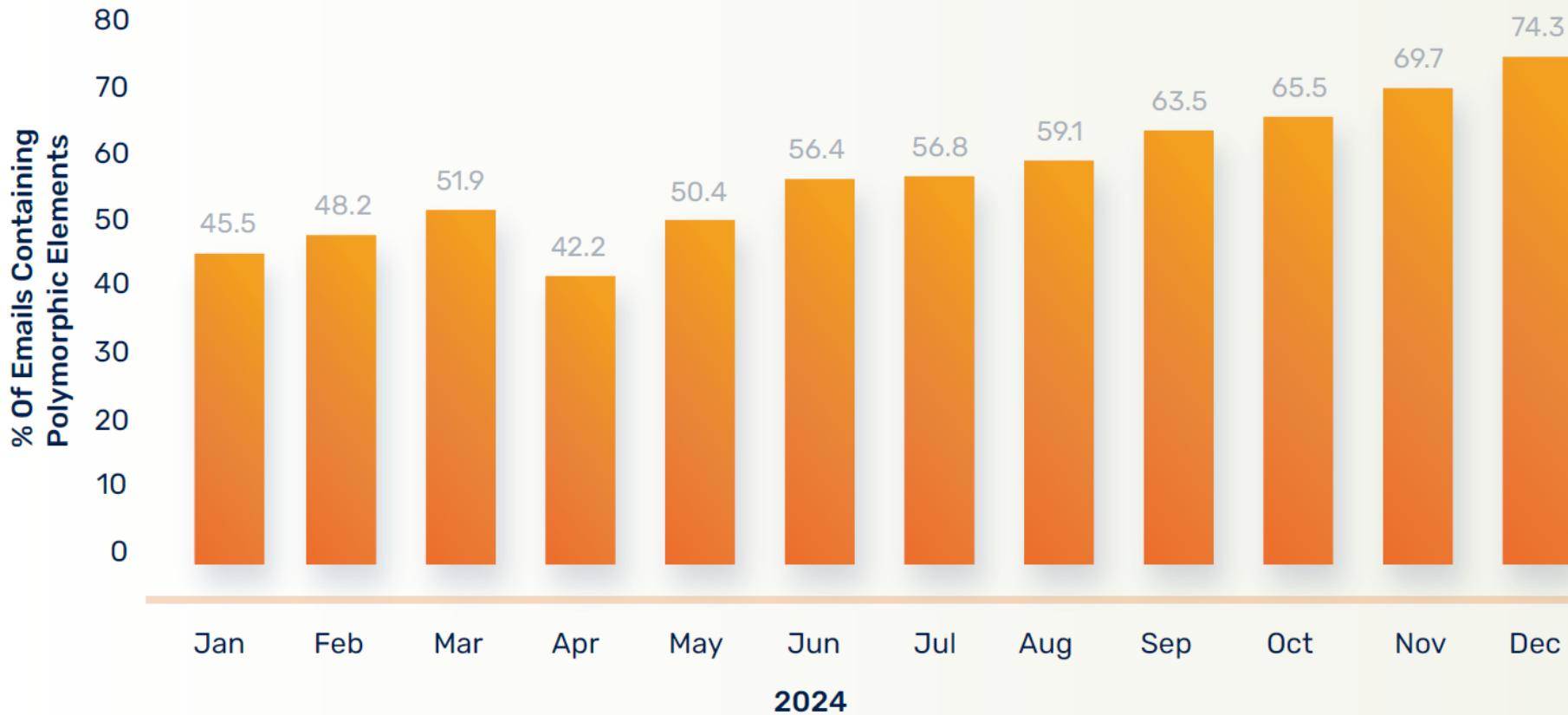
LLMs mass-produce varied phishing emails.

Dynamic payloads

Email content morphs based on user.

Polymorphic Emails

Polymorphic Phishing Emails In 2024





Synthetic Images



Synthetic Images

AI scammers pretending to be Brad Pitt con woman out of \$850,000

BRAD PITT

Pitt does not have any verified social media platforms

NEWS · FILM NEWS

Brad Pitt fraudster scams French woman out of \$800,000 using AI

"He knew how to talk to women and it was very well put together"

French woman says she was conned out of \$850,000 by an AI Brad Pitt

The woman chatted with the AI Pitt for more than a year and sent the scammer money to help with a fake kidney treatment

Viral scam: French woman duped by AI Brad Pitt love scheme faces cyberbullying







SYNTHETIC AUDIO

Synthetic Audio and Social Engineering

LASTPASS LABS

Attempted Audio Deepfake Call Targets LastPass Employee



Mike Kosak • April 10, 2024

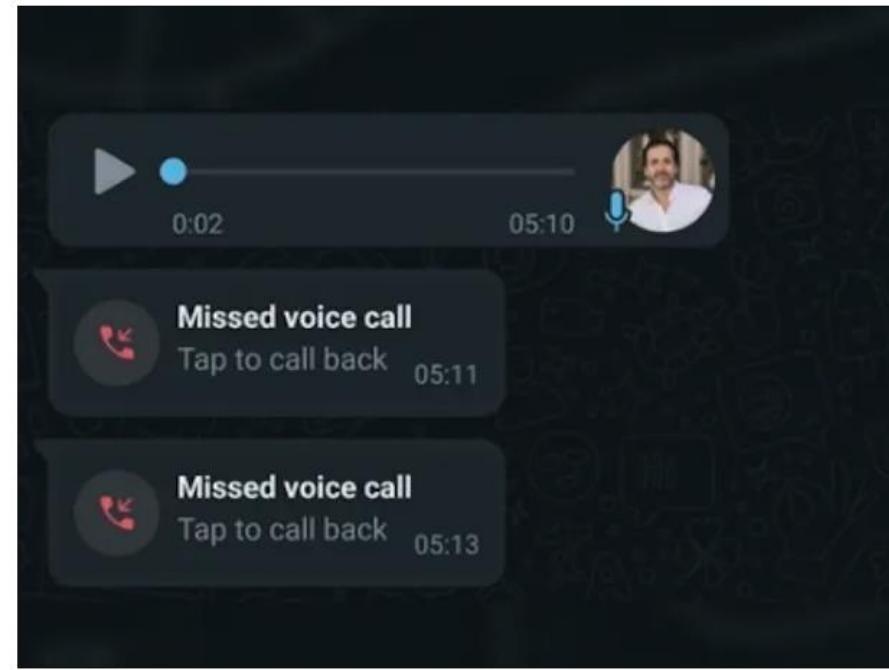
Mike Kosak, Senior Principal Intelligence Analyst at LastPass, explained in a blog post, "In our case, an employee received a series of calls, texts, and at least one voicemail featuring an audio deepfake from a threat actor impersonating our CEO via WhatsApp."

Audio Deepfake Attacks: Widespread And 'Only Going To Get Worse'

BY KYLE ALSPACH ▶

OCTOBER 3, 2024, 11:23 AM EDT

A cybersecurity researcher tells CRN that his own family was recently targeted with a convincing voice-clone scam.



AI vs AI



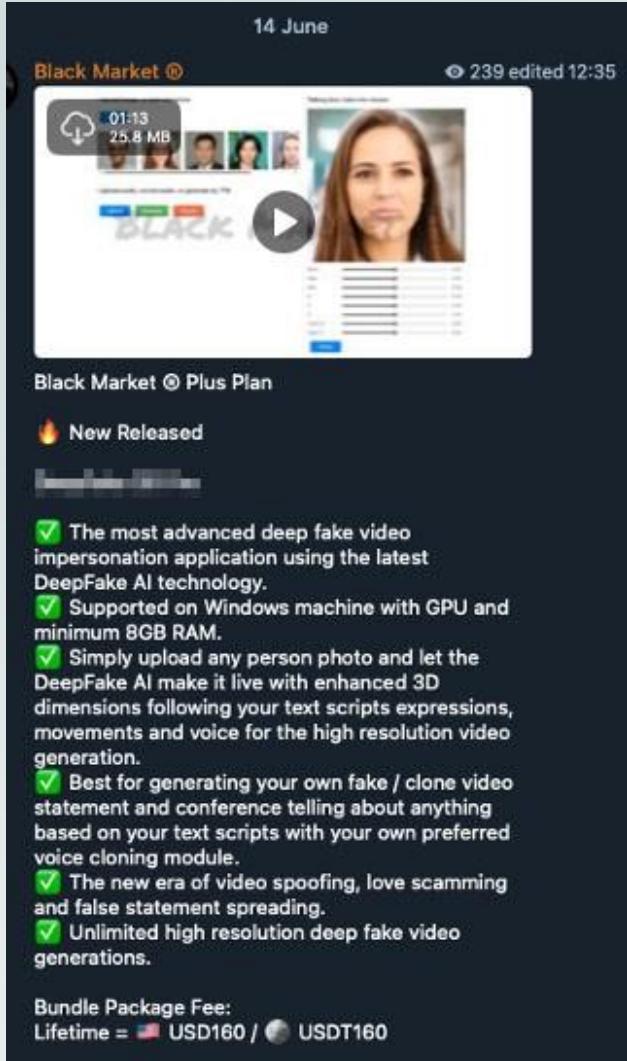
FlipSide – Used Against the Scammers





SYNTHETIC VIDEO

Dark Web Activity



00:00 LIVE

Avatar AI VideoCallSpoofer

Black Market © Plus Plan
Hot Selling
[REDACTED]

- ✓ The latest AI technology of video call spoofer tool.
- ✓ Simply upload any person photo and let the Avatar AI make it live following your expressions and movements for the video call session.
- ✓ Supported for most of video call applications and platforms (Whatsapp, Telegram, Google Meet, Zoom, Microsoft Teams and many more)
- ✓ Supported for all Windows / Mac PC & Laptop machines.
- ✓ Supported for all iOS / Android smartphones.
- ✓ Remote installation and setup services included.

Bundle Package Fee:
Lifetime = USD200 / USDT200

Black Market © 1,1K edited 09:28

Black Market © Premium Plan

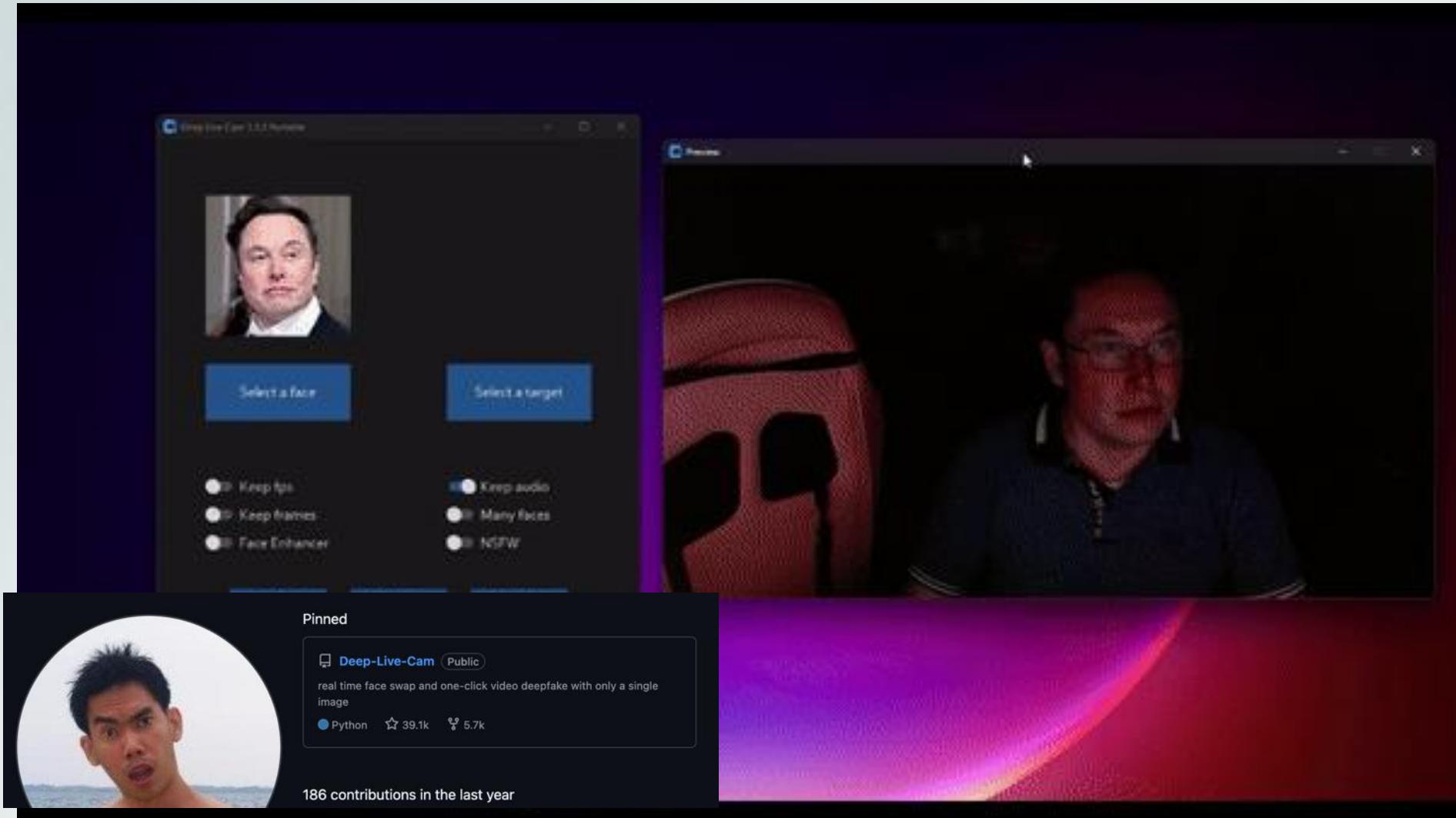
Hot Selling

- ✓ The most advanced deep fake video impersonation tool using well known DeepFake AI model.
- ✓ Simply upload any person photo and let the DeepFake AI make it live following your expressions, movements and voice for the high resolution video generation.
- ✓ Best for generating your own fake video statement and conference telling about anything that you want.
- ✓ The new era of video spoofing, love scamming and false statement spreading.
- ✓ Unlimited high resolution deep fake video generation.

Subscription Fee:
1 month = USD60 / USDT60
3 months = USD150 / USDT150
6 months = USD250 / USDT250
Lifetime = USD400 / USDT400

Source: <https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/surging-hype-an-update-on-the-rising-abuse-of-genai>

Face Swap / Voice Cloning & Webcams -> LIVE Deepfakes



DEEPFAKE OS



Tool List

For detailed instructions please refer to the websites or GitHub pages of the individual tools.

Deep Face Lab is a program for training deepfake models and using them to make pre-recorded videos. Its GitHub page has been taken down, but you can still find [a detailed guide for it here](#).

Deep Face Live uses models trained with Deep Face Lab to make real time deepfakes. It comes with several pre-made models and [its GitHub page can be found here](#).

Live Portrait is a tool for copying facial expressions from a source video onto a target image or video. [Its GitHub page can be found here](#).

Ollama is a framework for locally hosted LLMS which are capable of uncensored text generation. [Read more about it here](#).

RVC is an open source tool for voice cloning with a web interface. [Its GitHub page is available here](#).

Disinformation Campaigns

The National Guard's post

The National Guard • 5d

Several videos from "Bob" have been making the rounds online. They are fake. Red flags indicating a fake: The uniform name tape reads "Bob," his rank reads "E-6," and there is gibberish where "U.S. Army" should appear. In one video, the AI-generated "man" even eats a burrito through a mask.

<https://www.nationalguard.mil/.../guard-identifies-ai...>

TIKTOK @kefthecore4
TIKTOK @antoniodelarosacjs

Hany Farid • Following
UC Berkeley Professor & GetReal Co-founder
14h • Edited

In just the last 12 hours, we at **GetReal** have been seeing a slew of fake videos surrounding the recent conflict between Israel and Iran. We have been able to link each of these visually compelling videos to Veo 3.

It is no surprise that as generative-AI tools continue to improve in photo-realism, they are being misused to spread misinformation and sow confusion.

One simple tip-off (for now at least) is that all of these videos are either exactly eight seconds in length or composed of short (eights seconds or less) clips composed together. Why eight seconds? This is the current maximum length that Veo 3 can generate a continuous shot. Other models have slightly longer limits but 8-10 seconds is typical.

This eight-second limit obviously doesn't prove a video is fake but should be a good reason to give you pause and fact-check before you re-share.

TIKTOK @ssaa_diesel

You and 362 others

21 comments • 51 reposts



SYNTHETIC IDENTITIES

Synthetic Identities – Sock Puppet Accounts

Random Face Generator (This Person Does Not Exist)

Generate random human face in 1 click and download it! AI generated fake person photos: man, woman or child.

Gender: Any Age: Any Ethnicity: Any Refresh Image

this-person-does-not-exist.com



<https://this-person-does-not-exist.com>

Fake Person Generator To protect your real information from being leaked

Related Links

- Gmail Generator
- Random Address
- Random Phone Number
- Postcode Finder
- BIN Generator
- Employment Info Generator
- Identity Generator
- User Profile Generator
- IMEI Generator
- User Face Generator
- Nickname Finder
- Temporary Mail
- Gamertag Generator

Custom Generate

Gender: Random Age: Random State: Random City: Random Generate



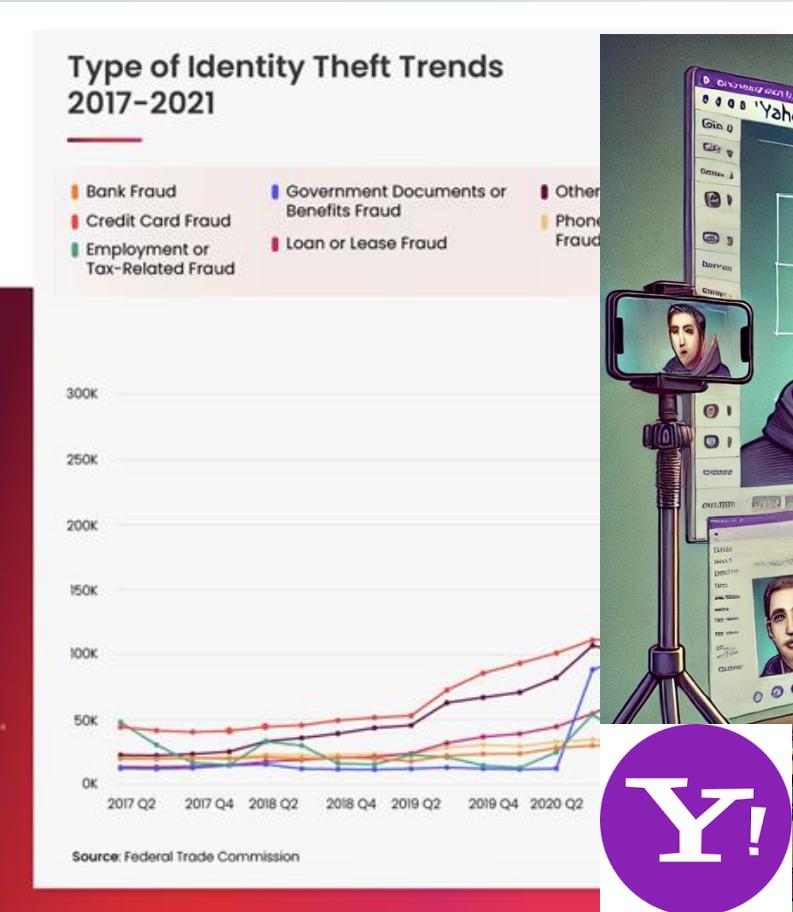
Lionel R Sanderson

Gender: male
Race: Black
Birthday: 5/8/1989 (35 years old)
Street: 3161 Whitetail Lane
City, State, Zip: Dallas, Texas(TX), 75244
Telephone: 469-296-7008
Mobile: 817-675-9384

<https://www.fakepersongenerator.com>



Synthetic Identity



Explosive Growth

Synthetic identity fraud cases surged 18% in 2024, making it the fastest-growing type of fraud.



Emerging Tactics

Face swaps and injection attacks increased by 704% in H2 vs H1 2023.

REWARD UP TO \$5 MILLION FOR INFORMATION ON NORTH KOREAN IT COMPANIES, WORKERS & RELATED MONEY LAUNDERING

IT firms Yanbian Silverstar, based in the People's Republic of China (PRC), and Volasys Silverstar, based in Russia, generate revenue by deceiving U.S. and other businesses worldwide into hiring their North Korean staff as freelance IT workers. The companies then launder their ill-gotten gains to North Korea, in violation of U.S. and UN sanctions.

We are seeking information on Yanbian and Volasys Silverstar-linked individuals who help to generate illicit funds for the North Korean regime.

If you have information on these individuals, their activities, or associated persons or entities, contact us via our Tor-based tip line below. You could be eligible for a reward and relocation.

Tor Link: he5dybnt7sr6cm32xt77pazmtm65flqy6irivtflruqfc5ep7eiiodiad.onion



U.S. Department of State
Diplomatic Security Service
Rewards for Justice



+1-202-702-7843

X @RFJ_USA



KnowBe4 In The Headlines...

The screenshot shows the homepage of Dark Reading. At the top, there's a navigation bar with links for Cybersecurity Topics, World, The Edge, DR Technology, Events, and Resources. A red "NEWSLETTER SIGN-UP" button is also visible. Below the navigation is a banner for ThreatLocker asking if foreign software is running in your environment, with a "Check Now" button. The main content area features two news articles. The first article, titled "Security Firm Accidentally Hires North Korean Hacker, Not KnowBe4" by Matt Bracken (published July 24, 2024), discusses how a software engineer hired for an internal IT AI team became an insider threat by loading malware onto a workstation. The second article, titled "Cyber firm KnowBe4 hired a fake IT worker from North Korea" by Solomon Klappholz (published August 5, 2024), reveals that KnowBe4 inadvertently hired a North Korean hacker pretending to be a remote worker based in the US. Both articles include a "TRENDING" badge.

DARKREADING

CYB

Security Firm Accidentally Hires North Korean Hacker, Not KnowBe4

A software engineer hired for an internal IT AI team immediately became an insider threat by loading malware onto a workstation.

THREATS

Cyber firm KnowBe4 hired a fake IT worker from North Korea

The security awareness training company said in a blog post that the software engineer used stolen U.S. credentials and an AI-enhanced photo.

BY MATT BRACKEN • JULY 24, 2024

TRENDING

Best cloud file sharing services for 2024 Artificial intelligence in

Security

Cyber firm KnowBe4 unknowingly hired a North Korean hacker – and it went exactly as you might think

News By Solomon Klappholz published August 5, 2024

KnowBe4 revealed it inadvertently hired a North Korean hacker pretending to be a remote worker based in the US

We Blogged About It...

X Post Button

23 How a North Korean Fake IT Worker Tried to Infiltrate Us
Jul Stu

25 North Korean Fake IT Worker FAQ
Jul Stu

Inc 27 How The Whole World Now Knows About Fake North Korean IT
Workers
Jul Stu Sjouwerman

First exfilt was it can We v July Wow! Last week's blog post went viral, reaching major media outlets and receiving over 125,000 views within days. Responses from around the world praised our transparency and commitment for doing what's right, though some had negative reactions.

Fre Wc
July hire

TLDI US circ
post 7/2 I decided to write an FAQ with more detail and reiterate that this was not a data breach but rather a public service announcement:

<https://blog.knowbe4.com/north-korean-fake-it-worker-faq>

X Post Share 5



NK Hiring Evidence

Kyle David Parkhurst

Senior Software Engineer

[REDACTED]@gmail.com

(404) [REDACTED]

Atlanta, Georgia, US

LinkedIn

SUMMARY

I am a seasoned senior software engineer with over 12 years of comprehensive experience in developing scalable and efficient digital solutions across various platforms. Passionate about leveraging my extensive experience in software engineering, I focus on developing robust and scalable systems, employing the latest technologies to solve complex challenges and enhance the user experience.

PROFESSIONAL EXPERIENCE

[REDACTED] May 2019 – April 2024

Senior Software Engineer

- Worked in the E-commerce division of the [REDACTED] engineering team, developing a membership management portal for over 1M active users with a focus on performance using React, Vue.js, Redux, Typescript, Node.js and AWS, in a professional Agile environment.
- Built a reusable and highly customizable component library used across 12 engineering teams, leveraging Storybook for robust development.
- Utilized Express middleware to efficiently query a GraphQL endpoint, leveraging middleware functions for routing and error handling, which streamlined development and improved data retrieval performance.
- Leveraged AWS services, including AWS Lambda, API Gateway, CloudFront, and Amplify, to architect and develop scalable and secure serverless applications, resulting in cost optimization and enhanced application performance.
- Mentored 2 junior developers and led technical sessions with team members to discuss trending technologies.
- Performed detailed code reviews with a team of 15 developers, enforcing coding standards to enhance code quality and promote adherence to best practices.

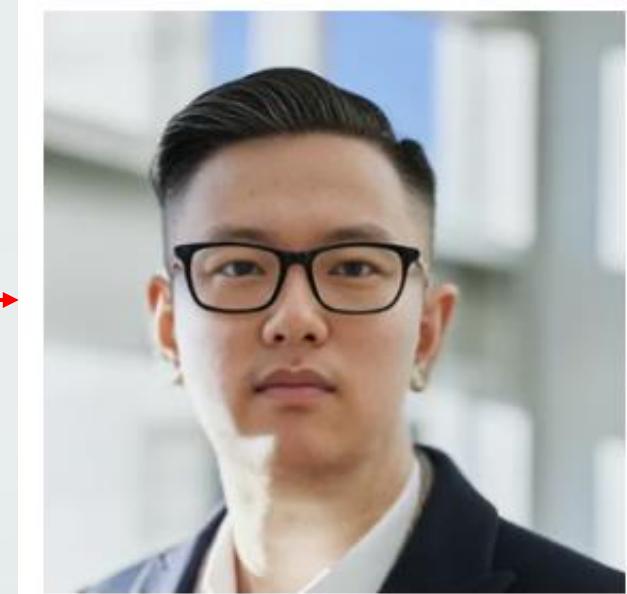
[REDACTED] January 2015 – April 2019

Full Stack Developer

- Played a key role at IT service portal modernization project, collaborating closely with multiple teams to conceptualize and develop software solutions using React, Redux Toolkit and Node.js.
- Implemented best practices in software development and architecture, ensuring that applications are compliant with industry standards.
- Worked with multiple business partners of [REDACTED] from different industries, utilized Node.js and Python to create and maintain RESTful and GraphQL APIs and Implemented and maintained high-quality, scalable user interfaces using React.js, and Vue.js.
- Developed and maintained automated testing suites using Jest, Enzyme, React Testing Library, Cypress, and Playwright, achieving over 95% test coverage and reducing bug rates by 40% across multiple projects.
- Improved product performance by 10% through effective optimization techniques such as code splitting and

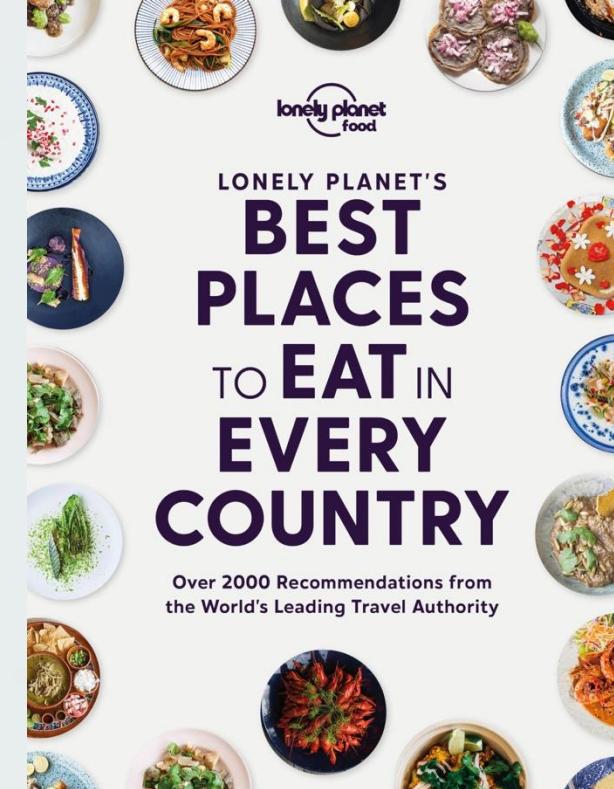
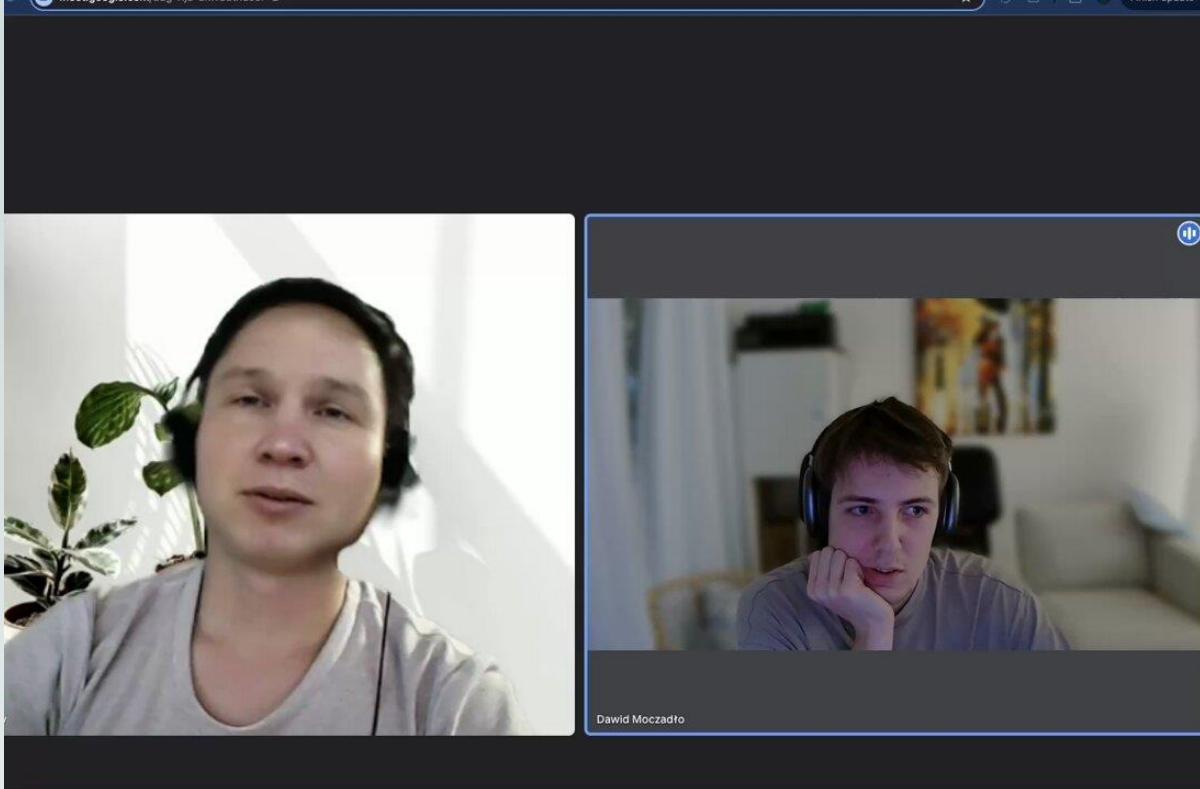


Stock photo NK used



AI-Photo Blending NK

Real time Video Deep fake Face Swap – And It Continues...



NK Farmers Arrested

THE WALL STREET JOURNAL.

North Korea Infiltrates U.S. Remote Jobs—With the Help of Everyday Americans

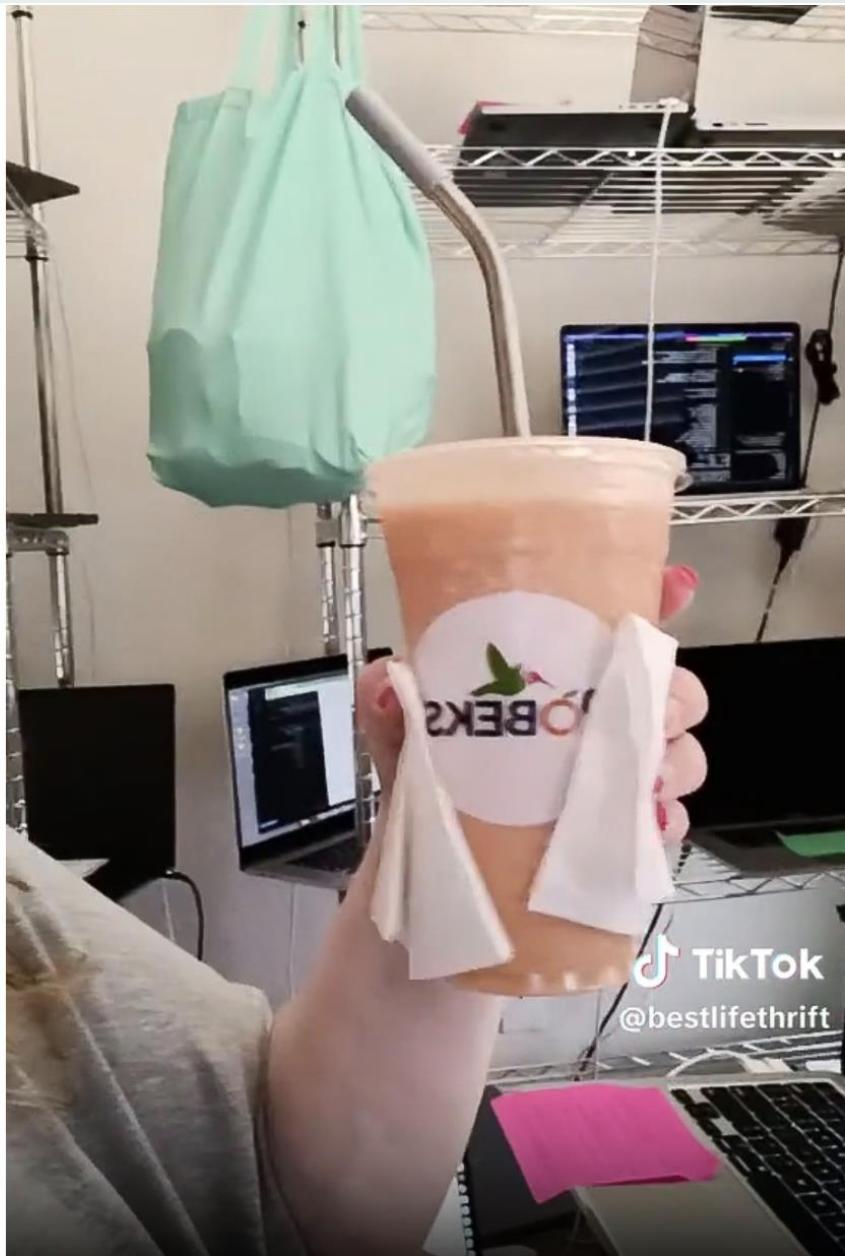
A LinkedIn message drew a former waitress in Minnesota into a type of intricate scam involving illegal paychecks and stolen data



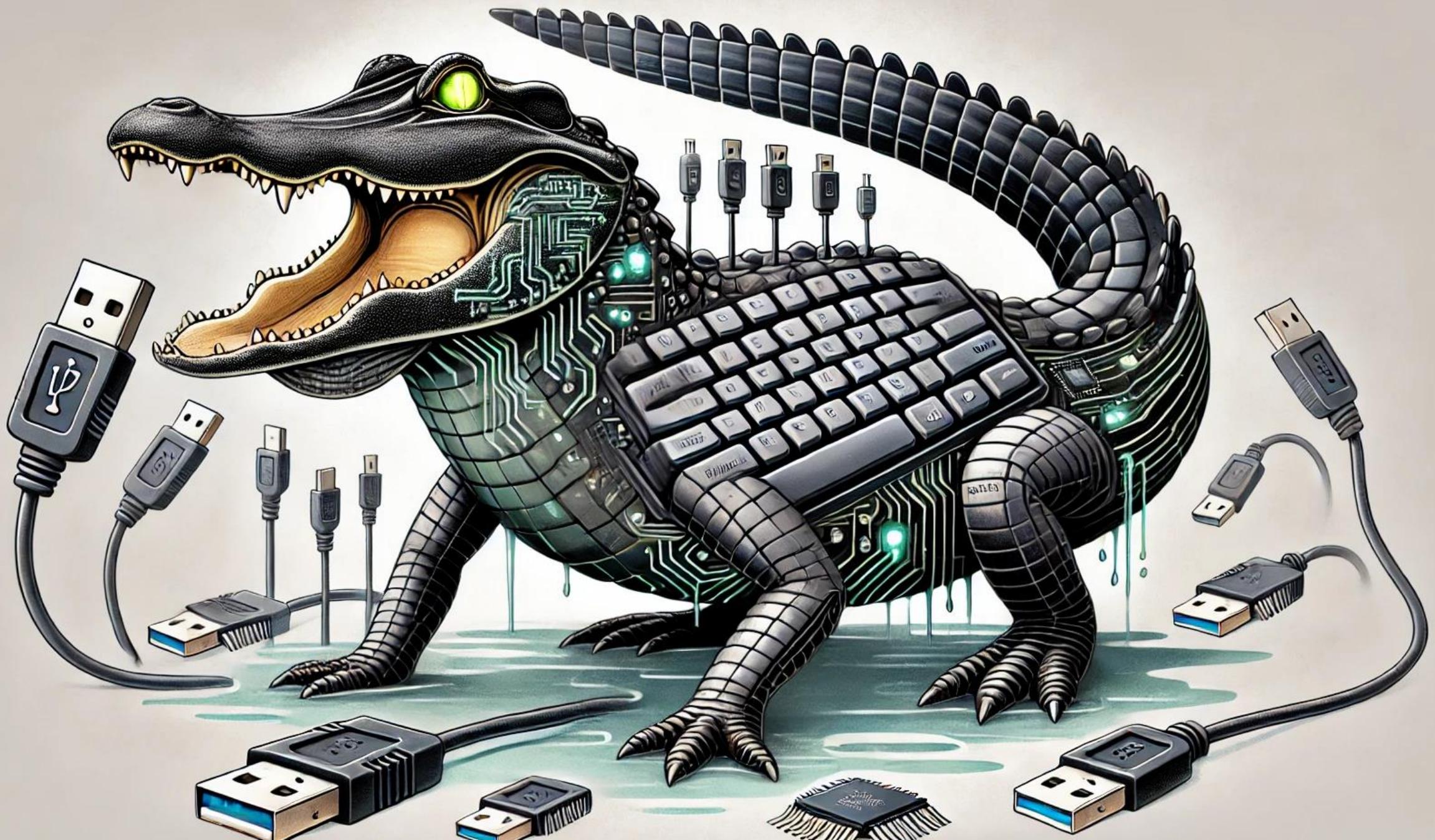




NK Laptop Scheme



Screenshots from Chapman's June 2023 Tiktok video with laptops in the background.





Tools

- Explore

Create Deepfakes



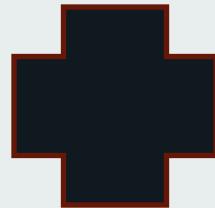
Audio Cloning / Deepfakes

- Cloning Tools
 - ElevenLabs, PlayHT, Resemble
- LLM & Interactive Tools
 - Dialpad, VAPI
- Realtime Audio Cloning
 - Voice.ai, RTVC, Altered.ai
- Audio Downloaders
 - YouTube Downloader, Airy, ytmp3.cc
- Audio Cleaning
 - Audacity / Audition / Descript

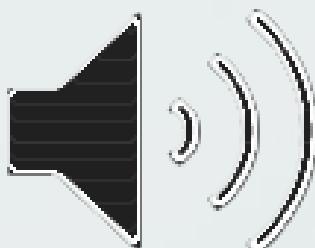
ChatGPT + Syn. Audio + Call Center = AI Social Engineering

ChatGPT3.5 Prompt:

You are calling to tell me that you have been in a car accident and now he's being held by the police. Convince me that I need to send you \$500 to pay the tow truck and start the repairs. You've been arrested and you don't have your wallet and you also need another \$1500 to get you out of jail. The money needs to be sent as crypto currency as you know I have a crypto wallet and I can send money that way"



Call Center
Support
Software



Using PlayHT



Dr. Gerald Auger, Simply Cyber
(and friend)



Video Cloning / Deepfakes

- Image to Video
- Hedra, LemonSlice, Synthesia.io, HeyGen, Deepfake Offensive Toolkit (DOT)
- LipSyncing (Advanced)
- KlingAi, Pixverse, Synclabs, Akool
- Realtime FaceSwap - Advanced
- DeepFaceLab, Deep-Live-Cam, SwapFace, MagicCam



KnowBe4



KnowBe4

Ahora podemos ir un paso más allá, ya que he creado mi propio Avatar



Perry Carpenter – Deep Fakes



CREATE real-time DEEPFAKES (a.k.a. I became TAYLOR SWIFT...for Science!)

Source: <https://www.youtube.com/watch?v=nKhtTIV6vxU>

DeepFaceLive

Real-time face swap for PC streaming or video calls

from a zero to a hero ...
just follow setup tutorial



Detection

GPTZero – GenAI Synthetic Text Detection

- <https://app.gptzero.me/app/ai-scan>

The screenshot displays the GPTZero AI detection interface. On the left, there's a sidebar with a teal header "Technology" containing icons for AI and blockchain, and sections for "Create a Team", "API", "Settings", and "Usage Stats". The main area has tabs for "Text Scan" (selected) and "Batch Upload". A message says "To see the results of previous scans, check out the Documents section." Below this is a "Document — 6/4/2024" section with a text input field: "Enter the text you want to analyze here (minimum 250 characters.)". To the right, there's another "Document — 6/4/2024" section with a similar text input field. At the bottom, there's a "Run" button and a progress bar showing "300K". On the far right, a file browser window titled "GenAI Text Check" is open, showing a list of files:

Name	Date Modified	Size	Kind
3CX docx	Mar 31, 2023 at 1:40 PM	25 KB	Micros. (docx)
AI & cybersecurity Webinar	Sep 18, 2023 at 8:19 PM	21 KB	Micros. (docx)
Comi Ransomware	Dec 8, 2022 at 7:13 AM	14 KB	Micros. (docx)

Below the file list, there are sections for "Favorites", "Recent", "Presentations - James", "James", "Stock Photos", "Expenses", "Applications", "Documents", "Downloads", "Movies", "Pictures", and "jmcquiggan". There are also sections for "Locations", "1TB Sandisk", "Google Drive", and "Network". At the bottom of the file browser, there's a legend for "Tags": Red, Orange, Yellow, Green, Blue, Purple, and Gray.

Audio File

 detect.resemble.ai/results/0b7e6bac1708987c39e00b3d2805fd0c

Resemble Detect

Detect deepfake audio from any source with our powerful AI Model. [Try it out yourself →](#)



 **Result: Fake**

The interface features a large orange waveform at the top, followed by a dark background area containing a speaker icon and the text "Result: Fake".



Deepfake Dashboard

The dashboard features a top navigation bar with links for Services, Training, Legal, and Account, along with the Psyber Labs logo. On the left, there's a brain visualization and a bar chart titled "Top CCVEs" showing percentages for Confirmation Bias (~35%), Authority (~32%), and Emotional Load (~28%). In the center, a report card for "Mr. Beast iPhone Scam" is displayed, featuring a video thumbnail of Mr. Beast, a threat level indicator, and a call-to-action button. To the right, a complex radar chart assesses threat levels across six dimensions: Credibility, Interactivity, Familiarity, Evocation, Distribution, and Sophistication/Maliciousness.

DEEFAKE DASHBOARD

Services Training Legal Account Psyber Labs

Showcased Reports

Mr. Beast iPhone Scam

you're one of the 10.000 lucky people who'll get an iPhone 15 Pro for just \$2.

Click to View Full Report

Threat Level: Moderate

Top CCVEs

CVE Type	Percentage
Confirmation Bias	~35%
Authority	~32%
Emotional Load	~28%

Detect | Dissect | Defend

10 Best AI DeepFake Detector Tools





NO DEEPFAKE DETECTED



Name: af66f232-b4ac-4a35-a2dd-d6c93655ba68.mp4

Size: 9.5 MB

Deepware aims to give an opinion about the scanned video and is not responsible for the result. As Deepware Scanner is still in beta, the results should not be treated as an absolute truth or evidence.



Model Results

Avatarify: NO DEEPFAKE DETECTED(43%)

Deepware: NO DEEPFAKE DETECTED(0%)

Seferbekov: NO DEEPFAKE DETECTED(1%)

Ensemble: NO DEEPFAKE DETECTED(0%)

Video

Duration: 37 sec

Resolution: 512 x 512

Frame Rate: 30 fps

Codec: h264

Audio

Duration: 37 sec

Channel: mono

Sample Rate: 48 khz

Codec: aac

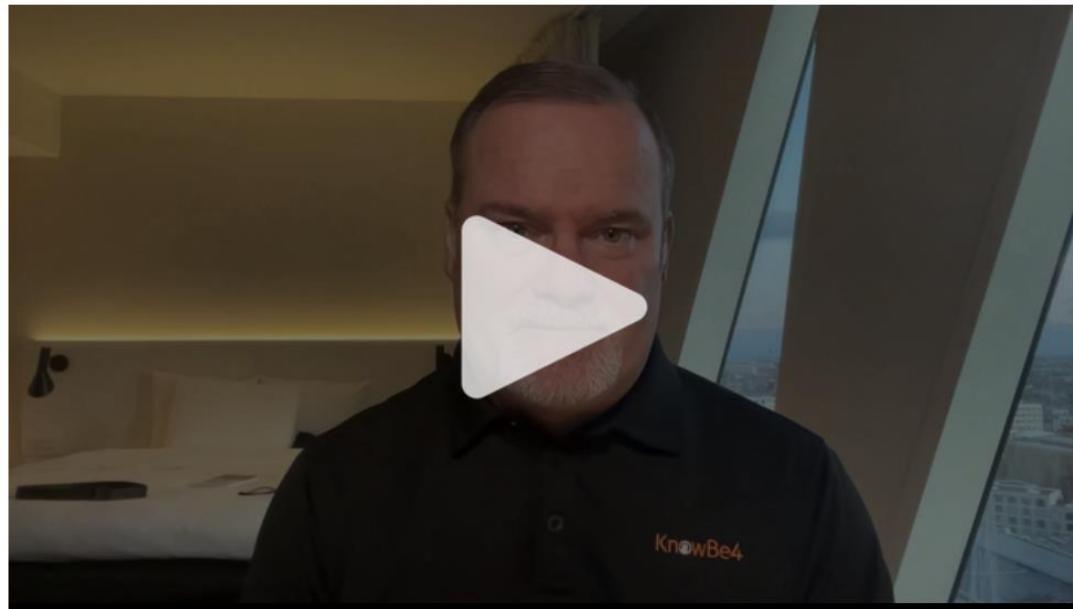
James HeyGen Video - Deepware

Deepware aims to give an opinion about the scanned video and is not responsible for the result. As Deepware Scanner is still in beta, the results should not be treated as an absolute truth or evidence.

! DEEPFAKE DETECTED



Name: James -BSidesCPH.mp4
User: 2021-07-21 10:45:00
Size: 5.5 MB
Source: User



Model Results

Avatarify: DEEPFAKE DETECTED(94%)

Deepware: NO DEEPFAKE DETECTED(0%)

Seferbekov: NO DEEPFAKE DETECTED(31%)

Ensemble: NO DEEPFAKE DETECTED(4%)

Video

Duration: 9 sec

Resolution: 1920 x 1080

Frame Rate: 25 fps

Codec: h264

Audio

Duration: 9 sec

Channel: stereo

Sample Rate: 48 khz

Codec: aac

Synthetic Video Detection Challenges

Lack of Standardization

No uniform method exists for detecting deepfakes effectively.

High False Positives

Many legitimate images are incorrectly flagged as deepfakes.

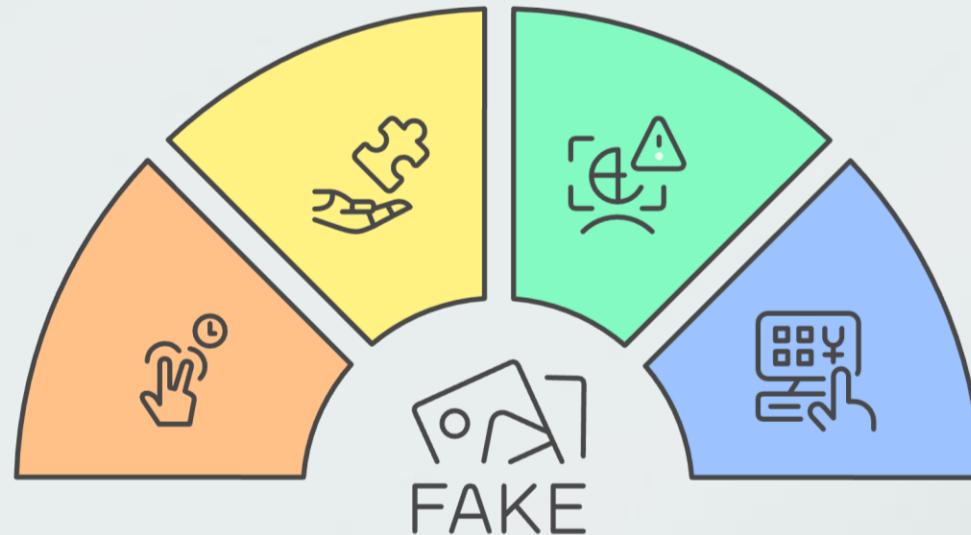
Non-Real-Time Detection

Detection processes do not operate in real-time, delaying responses.

Lorem ipsum dolor sit amet, consectetur.

Manual Effort Required

Detection still relies heavily on human intervention and analysis.



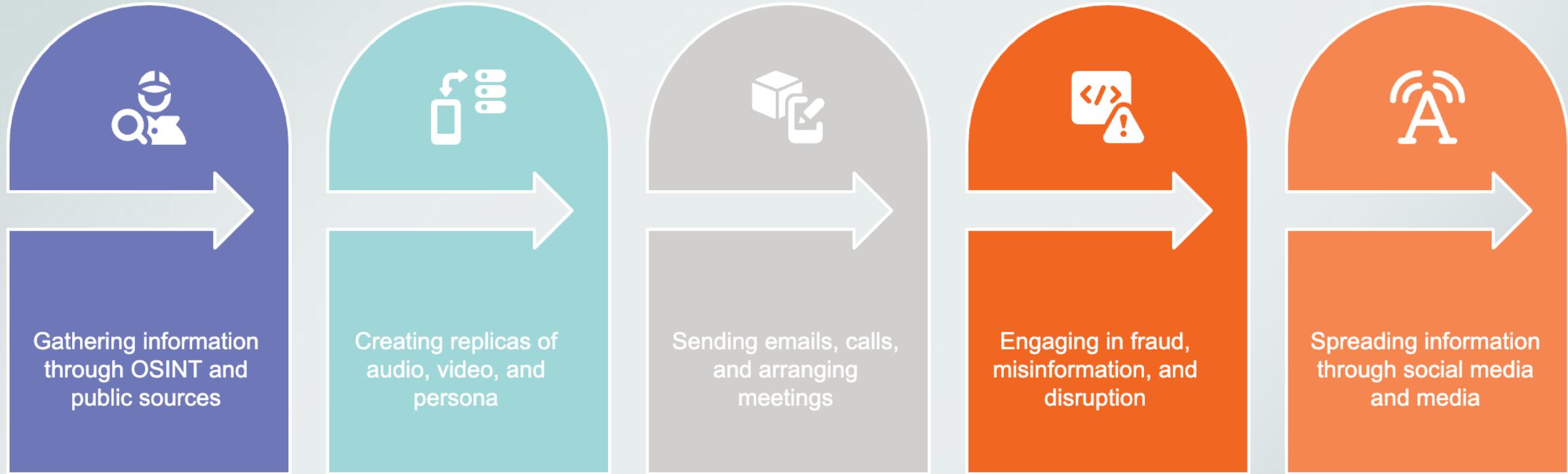
Generation technology is outpacing detection technology

Protect & Mitigate

- Strategize



Executive Attack Process



Where Deepfakes Hit You Hardest: Threat Matrix

	Executive Leadership	Employees/ Internal Comms	External Stakeholders
	CEO Voice / Wire Fraud	IT Support call scams	Vendor Negotiation spoof
	Deepfake investor	HR termination hoax	Fake public statements
	Fake exec resumes	Impersonated colleagues	Fraudulent customer support

The Business Impact



Financial Fraud

Involves scams and wire transfer manipulation



Operational Disruption

Includes supply chain and IT support issues



Reputation Damage

Covers fake press releases and crisis statements



Regulatory Risk

Encompasses disclosure failures and compliance issues

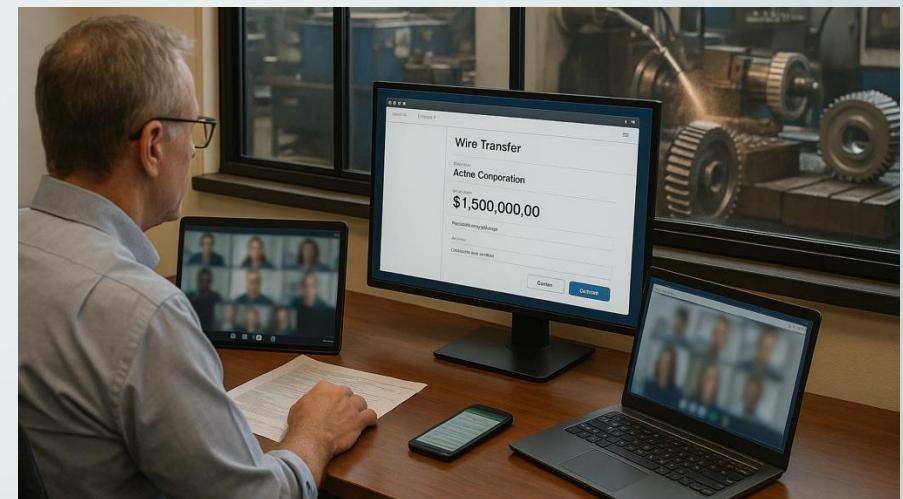
FINANCIAL TIMES

COMPANIES TECH MARKETS CLIMATE OPINION LEX WORK & CAREERS LIFE & ARTS HTSI

Cyber Security + Add to myFT

Arup lost \$25mn in Hong Kong deepfake video conference scam

UK-based engineering group identified as target of fraud that used digitally cloned CFO to trick staff



Ways AI Attacks Can Attack Your Organization

AI-Powered Phishing

- Personalized phishing emails using employee data
- Messages mimicking internal tools or vendors

Fake Executive Messages

- Deepfake voice/video requests for urgent wire transfers
- Fake Zoom calls or voice clones tricking staff

Insider Manipulation

- AI bots targeting employees to leak info
- Fake HR or IT messages stealing credentials

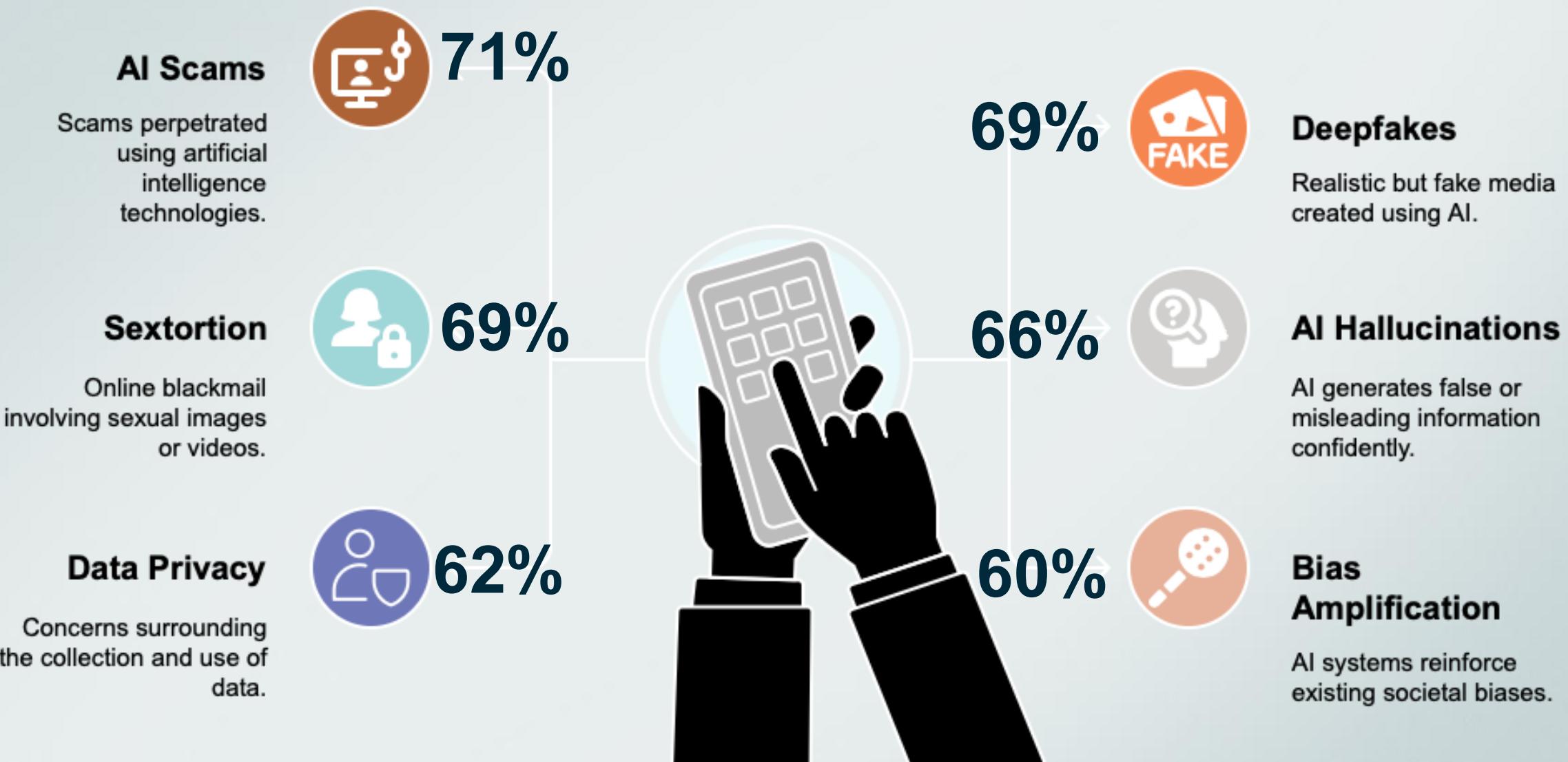
Supply Chain Scams

- Fake quotes, invoices, or contracts from “trusted” partners
- Malware hidden in CAD files or machine updates

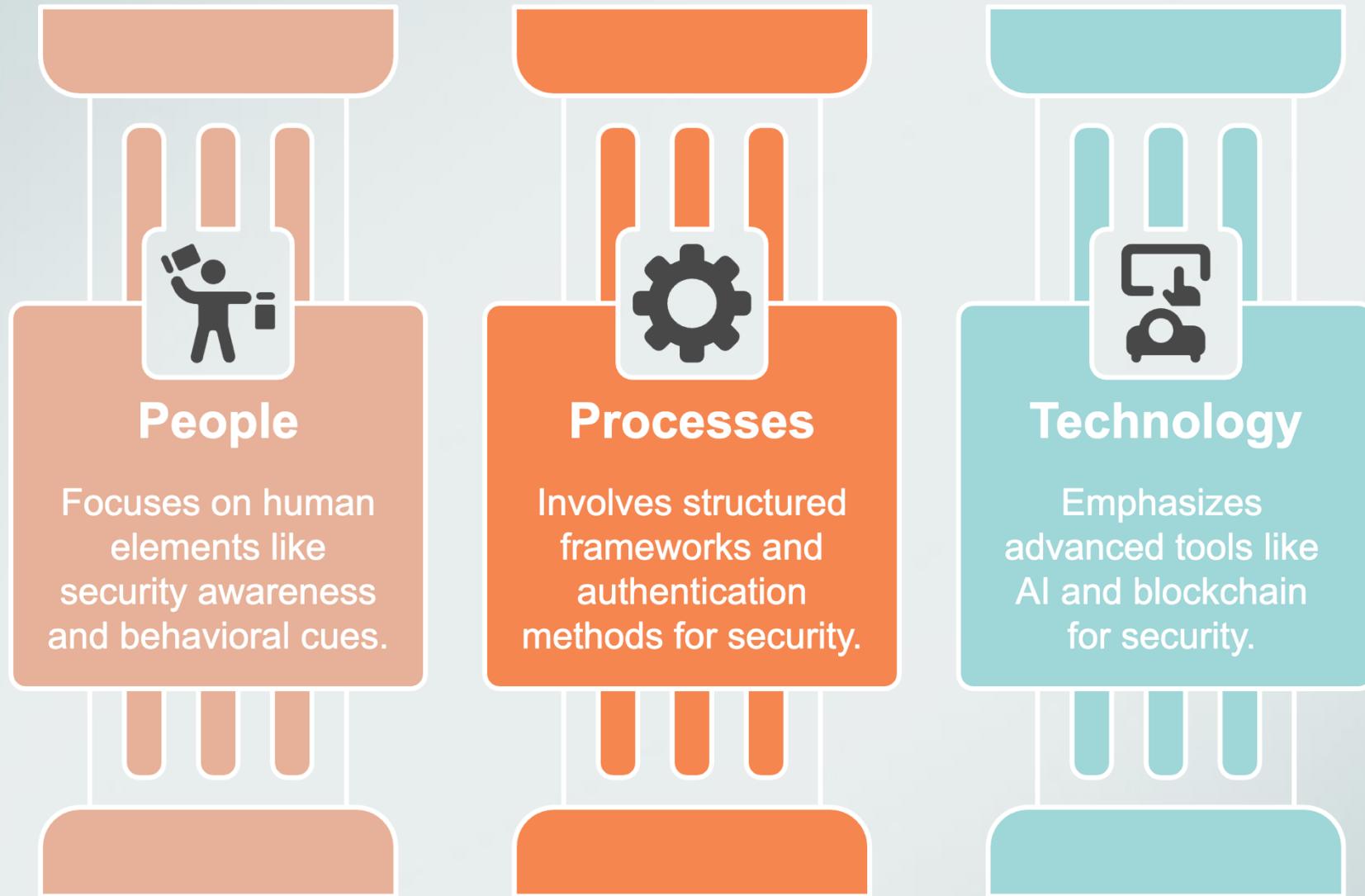
Factory Floor Manipulation

- Spoofed video feeds or alerts from machinery
- Impersonation of tech support to cause downtime

Generative AI Dangers – Survey of Concerns & Potential Problems



People, Processes, Technology



Technology - Detection Tools

DEEFAKE DASHBOARD

Services Training Legal Account

Psyber Labs

Showcased Reports

Mr. Beast iPhone Scam

you're one of the 10.000 lucky people who'll get an iPhone 15 Pro for just \$2.

Click to View Full Report

Threat Level: Moderate

Credibility

Interactivity

Familiarity

Evocation

Distribution

Confirmation Bias

Top CCVEs

Authority

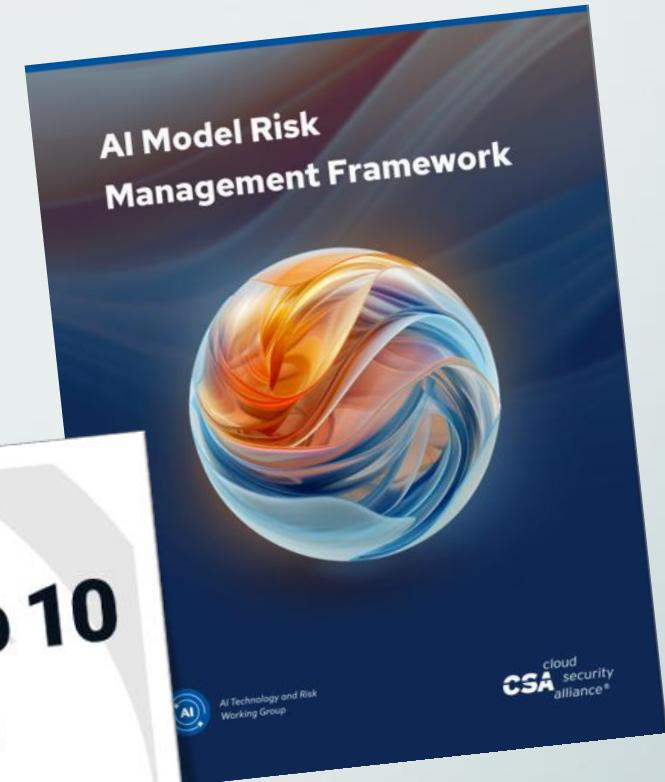
Emotional Load

Sophistication

Maliciousness

Detect | Dissect | Defend

Processes - AI Risk Frameworks



Processes - Tips During Hiring

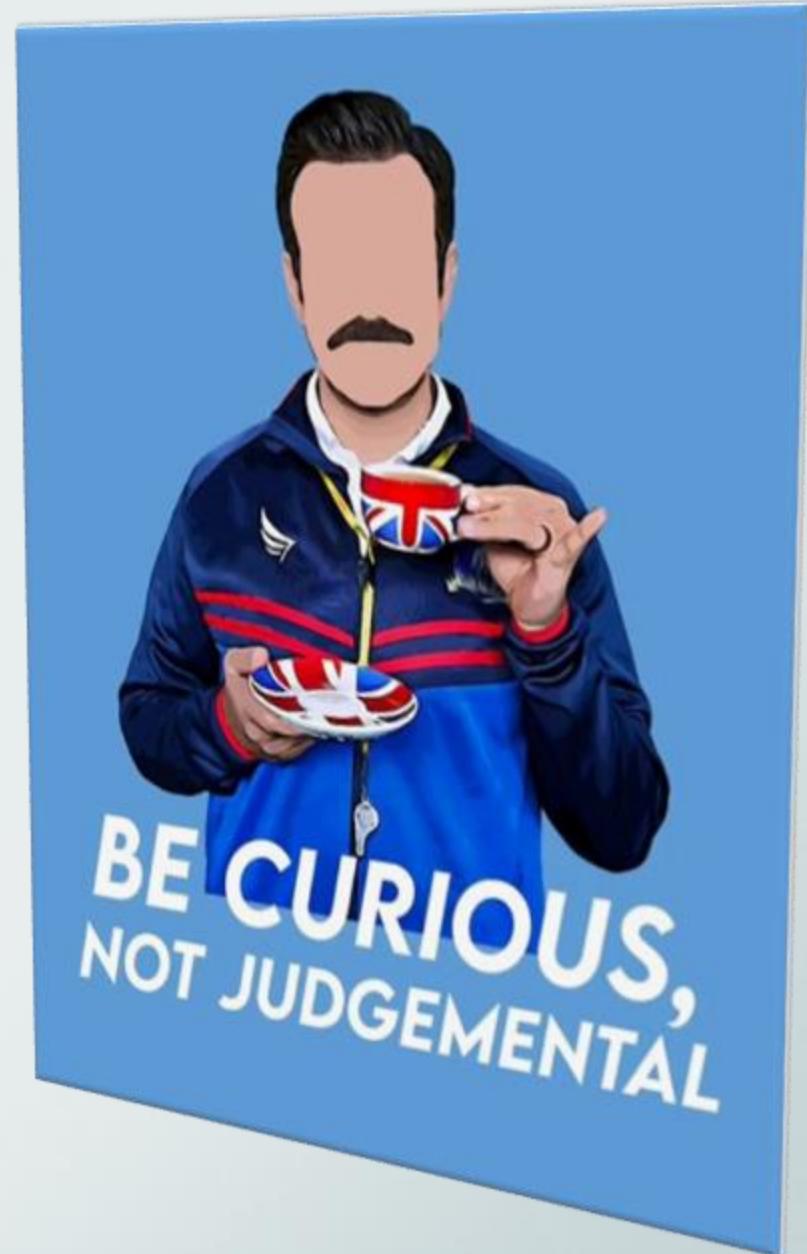
- Review Lists of Known NK Fake Employees

Fake Name	Payment address	Fake Location	Github	Email
Jason Kwon	0x4b94ba1528636a699dab486a217d39bb7ce21d75 0x1075e62bfacbb44e31d7a5719e55c7d16fe7d35d 0x7969b188f7dc6bf80d68f224ac3454dafe6f6d5d 0xa771609C5C56048f146d2C794c87DB946bfF27Cf 0x90cf352dDAF171d41A6DEd1d54cEDA4005047c93 0x72c70980ACddE7a5C9437050E73E7d07Bf21D25	Canada	https://archive.ph/J347I https://archive.ph/Wlu3i	0xm00neth@gmail.com
Willie Lee	0x97e36fAE76cD7ef7cC1213927A9A4E10a61CdD8d	California, US	https://archive.ph/SjJfK	willie.lee226@gmail.com
Naoki Murano	0x6188a9e76794e7cb337b8E5a2B9180Ce34Fc6D1 0x85e0504fc7981baa68774431099c5e2dcf074dd	Tokyo, Japan	https://archive.ph/96QVA	naokimurano@outlook.com
Sano	0xef2a0324cfaa0100db9def8ef31c6e23bc4f9258	-	https://archive.ph/KMoXG	-
Jun Kai	0x8aa07899eb940f40e514b8effdb3b6af5d1cf7bb	Singapore	https://github.com/junkai121	junkai121@outlook.com
Kei Nakano	0xff22be4f00b937dade564bd9659e265f92afa620 0x452f205c6c3872691fbce7ce84c38370466d55f76 0x21e5d5a6e40b32cff77cef77dca034d6d410131d	Tokyo, Japan	https://archive.ph/mo0QZ https://archive.ph/fhKTT	keinakano415@gmail.com
David Adachi	0x21088f2624d019cbc71de5bf4caf5b6dc9fa7f	Fukuoka, Japan	https://archive.ph/80EYH	davidadachi56@gmail.com
Gabriel Yiu	0xd80614fcb54d49cf46cc861fc549fae0a05b3f7e	-	https://archive.ph/oGlCc	-
Joshua Palmer	0x06f90983cd2215379e440fc525e441d6a5fc3fba 5Jfb3n8eW4JyQrKJtMNBFXnC1zx2YHjRSkzRtT5QHh 0xa6afe0290fb6f2f7ced0a2753de57f9fa7c9c9dd 0xfa802d9b33ed74baff62b189875c2b2d192874eb 0x7654e18ff3495675606c008a39b6264da50e8a7	Michigan, US	https://github.com/call-by https://archive.ph/grqjk	joshupgig@gmail.com smart.solidity@gmail.com
Andy Hoog	0x1043efee936903951b88db23551873bb67292e95	-	https://archive.ph/7lbnH	andyhoogup@gmail.com
Jordan Lopez	0x92cd7363c5b1853bc8fe6b5ae269836fc508ca73	Texas, US	https://archive.ph/tFeQG	cloudrider.m92@gmail.com
Quinn Lee	0x9de5d3158b0b83e9211c7444c94ce0c53763f574 0xf9adac8658e08893fb4e91c1062e471eb11cb6c7	-	https://archive.ph/KLBYw	lettelddream@gmail.com
Ryuhei "Rio" Matsuda	0xa71b641a498e33bb13548a01eca5e20e083e637b 0x6fb678b2dd9d2ff50ee9ecf774251dcceb7a2da8	-	https://archive.ph/V5GsZ	ryuheimat3@gmail.com
Chris Yu	ESSfp3aAcW6Z59ozu9Jkqy9btaXYTHt25b3Vhs2hsf 0x1043efee936903951b88db23551873bb67292e95 0x3b9A870c24905256dE10863cb360F4B93C7cC60f 0xc2b2a9c05740EEb7ee7BA7eB3AB11EC8bebCB1D1	Malaysia	https://archive.ph/x0LMf	atroboj@gmail.com

Source: <https://github.com/shortdoom/gh-fake-analyzer/tree/main> > Profiles > Investigations > Zach...

Processes - Tips During Hiring

- BE CURIOUS - Ask questions about their surroundings
- “What was the name of that bar located next to campus that everyone went to?”
- “What was the name of that main road running right in front of the campus?”
- “Do you have an HOA where you live?”
- “Did you have to register for Selective Service?”
- “At what age did they allow you to get a driver’s permit (in that state)?”
- “Oh, I see you worked at Autodesk. What type of internal instant messaging system did they use, again?”



Processes / People - Tabletop Exercises



DEEFAKE CEO WIRE TRANSFER SCAM

A video call appears to come from your CEO while traveling abroad. They urgently ask the CFO to authorize a wire transfer for a confidential acquisition.



EMPLOYEE FALLS FOR DEEFAKE CEO WIRE TRANSFER SCAM

An employee receives a video call from the supposed CEO, mimicking their appearance and voice. Tricked by the realistic deepfake, they initiate a fraudulent



DEEFAKE PHONE CALL SCAM

You receive a call that appears to be from your manager. Their voice urgently asks for sensitive information to resolve a payroll issue.



FINANCIALLY DISTRESSED BUSINESS SCAM

A struggling small business is pressured to make purchases or pay invoices, falsely hoping that it will improve their finances.



PHISHING SCAM

You receive an email claiming to be from your bank. To prevent your account from being locked out, it urgently tells you to click a link and take action.

People - Incident Response Team Roles



Security

Validate and contain the security breach.



HR

Verify hiring and internal identity checks.



PR/Comms

Control narrative and issue statements.



Legal

Manage compliance and collaborate with law enforcement.



Executive

Approve actions and allocate resources.

People - What Should We Be Asking?



Is this a deepfake?



Consider these questions...

Apply the FAIK Factor Framework

F Freeze & Feel

A Analyze the Narrative & Emotional Triggers

I Investigate (claims, sources, etc.)

K Know, confirm, and keep vigilant

People are a
critical layer within
your security
programs





TRUST
& VERIFY



BE
SKEPTICAL



POLITELY
PARANOID



Defense in Depth for AI & Social Engineering

MFA

- Use MFA Wherever Possible - Non-phishable MFA too
- Avoid SMS and verify all requests that you didn't initiate

Email Auth

- DMARC / DKIM / SPF to prevent email spoofing!
- Only 15% of orgs use this.

AI vs AI

- Use AI with your cybersecurity gateways & products
- Behavioral Analysis in email and users

Zero Trust

- Be mindful of your inbox. Be skeptical, if you're not expecting it and it's a strange or unusual request

Training

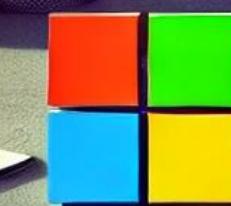
- Frequently educate and assess your users
- Leverage threat intelligence for your threat landscape



Edge
on edge?
Are you on the
way on edge?

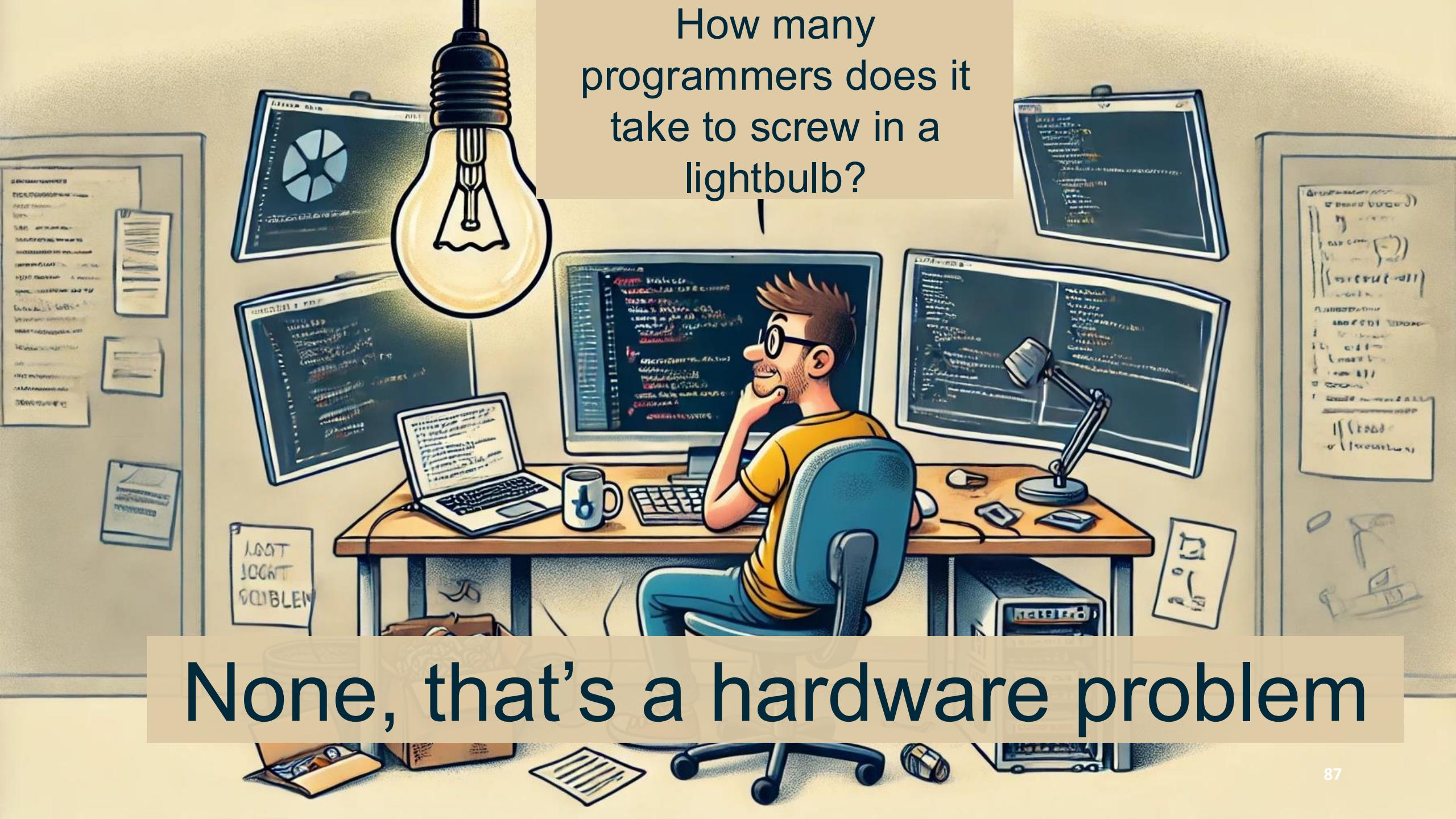
Never
relaxed?

Microsoft



Microsoft





How many
programmers does it
take to screw in a
lightbulb?

None, that's a hardware problem

Questions





Final Thoughts

Using Questions / Be Curious

“What is the book
that you
recommended to
me?”

‘I Need to Identify You’: How One Question Saved Ferrari From a Deepfake Scam

- Benedetto Vigna was impersonated on a call using AI software
- Large companies are being increasingly targeted with deepfakes



ACTION PLAN



People (Security Awareness & Training)

- Human Risk Management Program
- Conduct deepfake awareness training for all employees
- Implement verification protocols for executive or financial requests
- Use security questions in high-risk communications



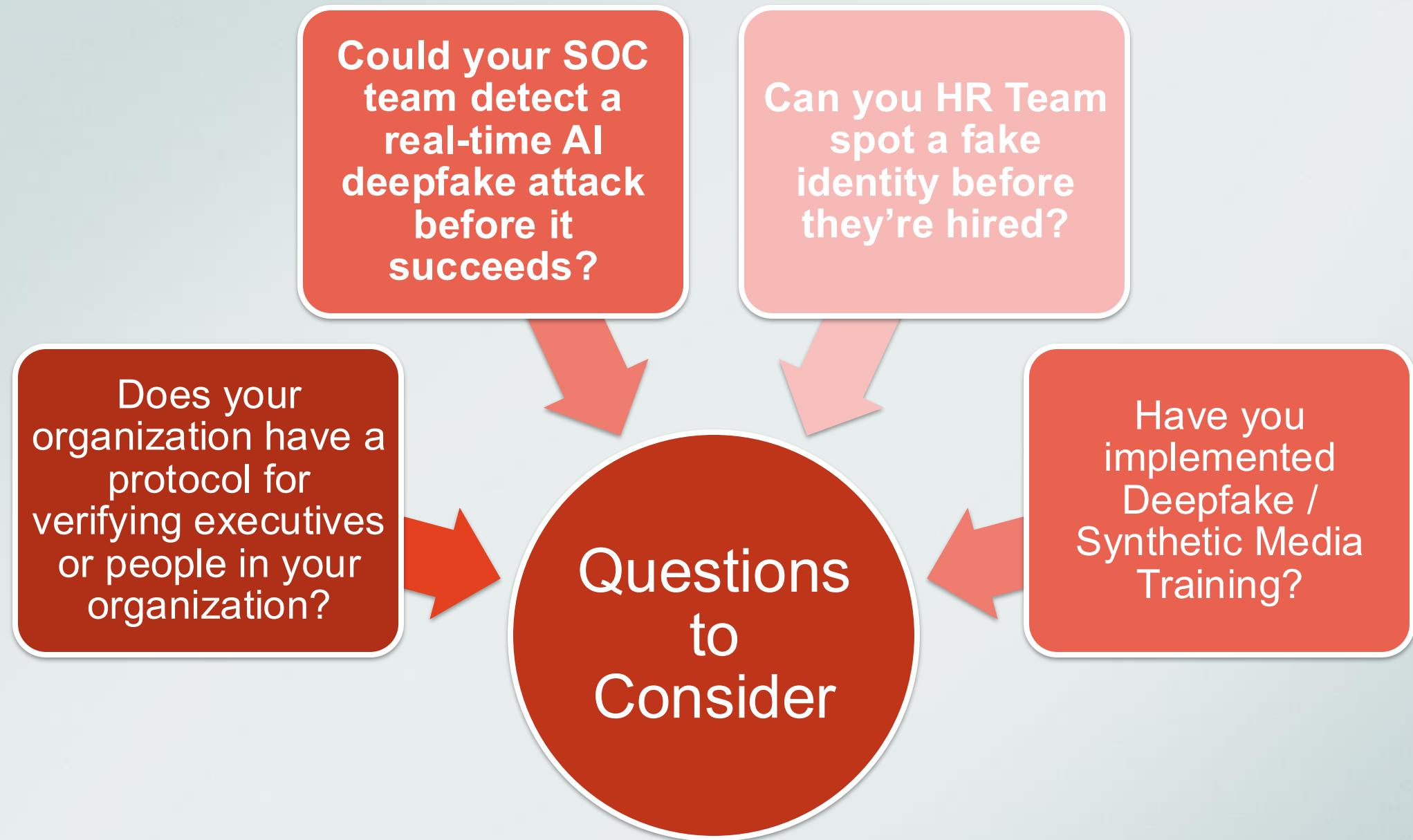
Processes

- Ask location-based verification questions
- Require cameras on for remote interviews with assessments
- Establish a SOC protocol for real-time AI threat detection
- Table Top Exercises Assessing Deepfake Scenarios



Technology

- Deploy AI-driven deepfake detection software
- Implement audio and video authentication measures
- Leverage POCs for deepfake detection services / software





**Education,
Preparation,
and healthy
security habits**
-- are the only defense --

Interested in Learning More?

The screenshot shows the RESEMBLE.AI Deepfake Incident Database. At the top, there's a navigation bar with links for AI Voice Generator, Detect, Resources, Government, Pricing, and Sign In. Below the navigation is a section titled "AI Safety" with a green background. A large title "Deepfake Incident Database" is displayed in white. To its right is a detailed description of the database, which tracks verified incidents where deepfake technology has been used to target specific individuals or organizations. It includes cases of AI-generated synthetic media being weaponized for impersonation, manipulation, or deception. Below the description is a button labeled "Learn More About Detection".

<https://www.resemble.ai/deepfake-database/>

The screenshot shows the AIAAC Repository. At the top, there's a header with the text "AIAAC Repository" and a circular image of a complex mechanical or electronic component. Below the header is a brief description: "The independent, open, public interest resource detailing incidents and controversies driven by and relating to AI, algorithms and automation." There's a "More" link at the end. On the right side, there's a "Recent updates" section with a list of news items, each preceded by a small square icon.

<https://www.aiaaic.org/aiaaic-repository>

The screenshot shows the AI Incident Database (AID). The interface features a sidebar on the left with various navigation options like Discover, Submit, Spatial View, Table View, List view, Entities, Taxonomies, Submit Incident Reports, Submission Leaderboard, Blog, AI News Digest, Risk Checklists, and Random incident. The main area is titled "Welcome to the AI Incident Database" and includes a search bar. A specific incident is highlighted: "Incident 950: NullBulge's AI-Powered Malware Allegedly Compromises Disney Employee and Internal Data". Below the incident summary is a link to "A Disney Worker Downloaded an AI Tool. It Led to a Hack That Ruined His Life." with a timestamp of "wsj.com 2025-02-26".

<https://incidentdatabase.ai>

Most Secure Woman?



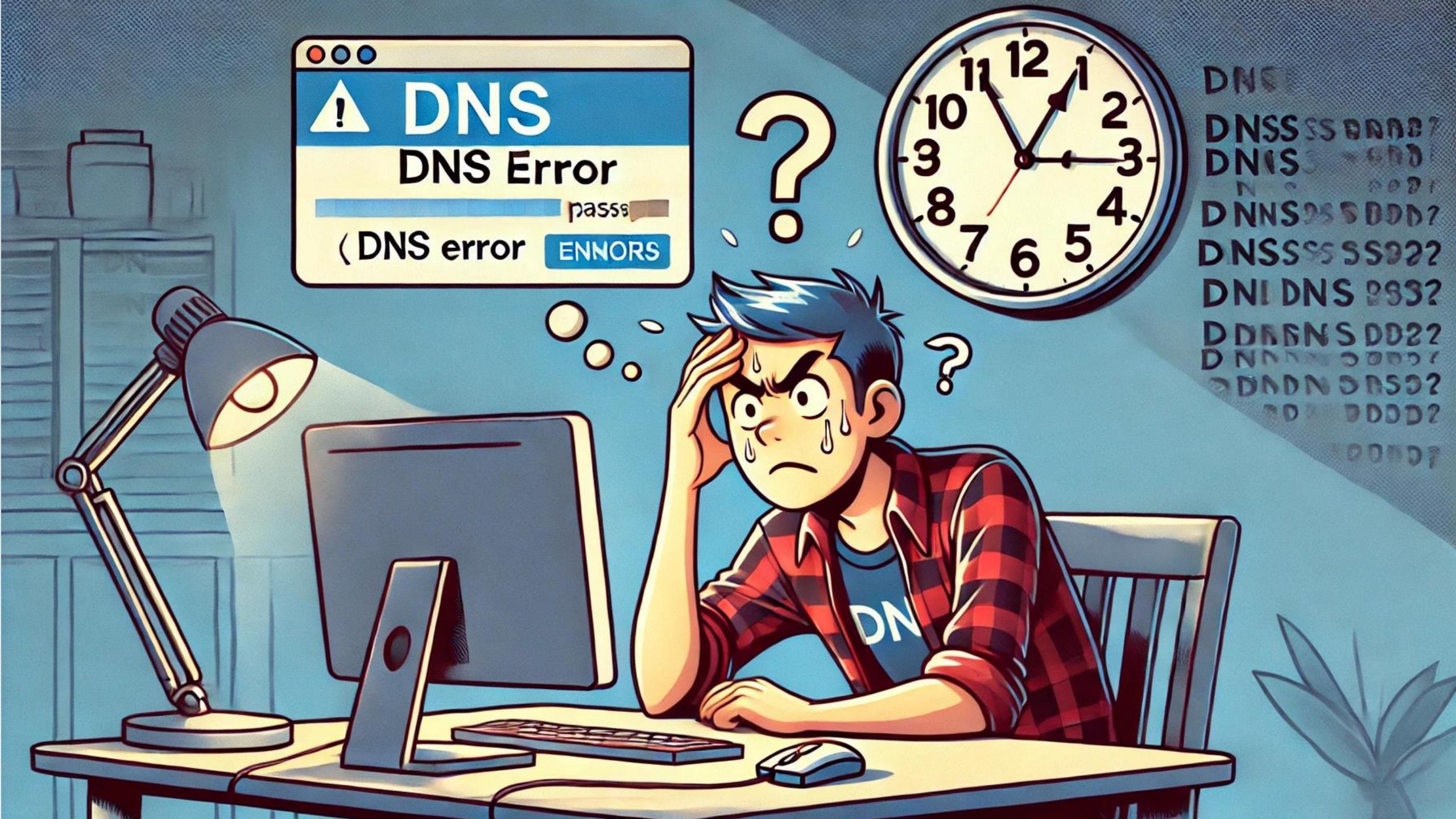
MFA



Emma Faye



Multifactor
Authentication



DN
DNSSSS BBBB?
DNIS
NN S DDD?
DNNSSSS BBBB?
DNSSSSSSSSSSSS?
DNI DNS BBB?
DDBBNS DD22
DNNNN SSSSS?
DNDN DR5D2
DD DDDDD?

Thanks For Your Attention



James R. McQuiggan, CISSP, SACP

jmcquiggan@knowbe4.com

LinkedIn

Connect with Me!

 LinkedIn jmcquiggan

 @james_mcquiggan

 Website jamesmcquiggan.com





SIMPLY SECURED

Simply Secured

 Simply Cyber - Gerald Auger, PhD

Podcast • 6 episodes • Updated 5 days ago

Welcome to "Simply Secured: Elevating Cyber Careers", where we take cybersecurity to new heights! Hosted by [...more](#)

▶ Play all



Airwaves to AI: Education and Innovation from Kathy Chambers | Simply Secured S1 E6

Simply Cyber - Gerald Auger, PhD • Premieres 8/15/25, 8:30 AM



Navigating the SOC Analyst Landscape Insights with Casually Joseph | Simply Secured S1 E5

Simply Cyber - Gerald Auger, PhD • 574 views • 6 days ago



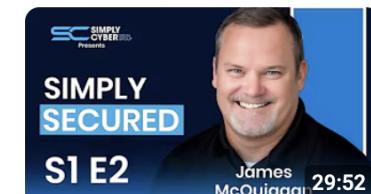
Navigating Cyber Pathways: GRC, Volunteering, and YouTube with Bdubzz | Simply Secured S1 E4

Simply Cyber - Gerald Auger, PhD • 552 views • 13 days ago



Blending Bytes and Barley: Ryan Pearson's Unique Tech Adventure | Simply Secured S1 E3

Simply Cyber - Gerald Auger, PhD • 493 views • 2 weeks ago



Cyber & Drones: Exploring the Intersection of OSINT and Vulnerabilities | Simply Secured S1 E2

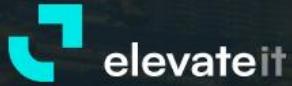
Simply Cyber - Gerald Auger, PhD • 682 views • 3 weeks ago



Cyber and OSINT: Fueled by Curiosity | Simply Secured S1 E1

Simply Cyber - Gerald Auger, PhD • 964 views • 1 month ago





HOUSTON TECHNOLOGY SUMMIT 2025

September 25th, 2025

Hyatt Regency Houston West
13210 Katy Fwy Houston TX 77079

September 25, 2025 4:00 pm

KEYNOTE
TECH THEATER 1

The Human Firewall Element in Agentic AI Cybersecurity

Moderated by



James McQuiggan
Security Awareness Advocate
- KnowBe4

Speakers



Scott Giordano
Partner & Co-Founder, The CISO Law Firm



John Barker, Esq.
Co-Founder & Partner, The CISO Law Firm



David Patariu
Partner, The CISO Law Firm



The registration link is https://eitevents.com/event_pages/houston-technology-summit-2025/

Free admission code for Practitioners and Executives (No sales/marketing or students) is:

ISSA25

<https://cruisecon.com>



CRUISECON WEST 2025

The Flagship of Cybersecurity

October 2nd-6th, 2025

Departing from San Pedro World Cruise Terminal

Register Now

Stay Updated

☰ YouTube

Search

James McQuiggan, CISSP, SACP
35 subscribers

HOME VIDEOS PLAYLISTS CHANNELS ABOUT

Dad Jokes & Cyber Stories ► PLAY ALL

A New Business 0:37 James McQuiggan, CISSP, SACP 6 views • 4 days ago

Cybercrime 0:41 James McQuiggan, CISSP, SACP 23 views • 12 days ago

Most Secure Woman 1:01 James McQuiggan, CISSP, SACP 21 views • 2 weeks ago

Sick Webpages 0:22 James McQuiggan, CISSP, SACP 13 views • 1 month ago

Library History Your videos Watch later Liked videos Dad Jokes & Cyber St...

YouTube: James McQuiggan and Dad Jokes

<https://www.youtube.com/@JamesMcQuigganCISSP>

Yes... I have a
way of keeping
track of my
Dad Jokes

