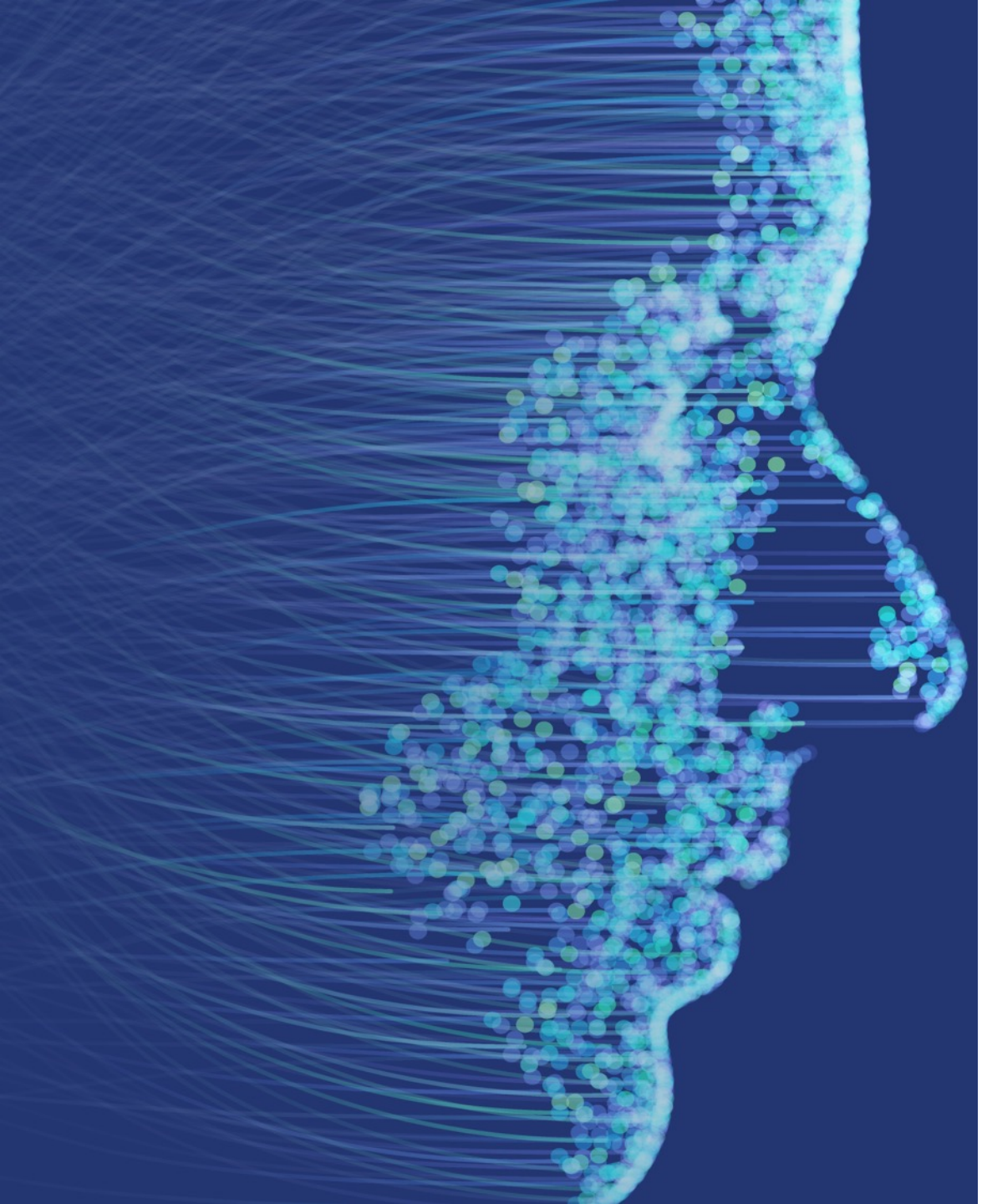
Leben in einer Welt mit Deepfakes

it-sa | IT Security Talks | 16. - 17. April 2024

KnowBe4



About me

Dr. Martin J. Kraemer

Security Awareness Advocate
KnowBe4

Twitter: *@markraemer*
LinkedIn: */in/martinkra/*

About Me

- Thought leadership and PR at KnowBe4 (DACH/EMEA)
- Researcher and lecturer in human-centered security

Areas of Interest

- Human-centered security (awareness, behavior, and culture) and security-by-design
- Data protection, privacy, and privacy-by-design
- Ethics of AI and responsible research and innovation

HONG KONG

Multinational loses HK$200 million to deepfake video conference scam, Hong Kong police say

Police received a report of the incident on January 29, at which point some HK$200 million (US$26 million) had already been lost via 15 transfers.

NEWS SERVICE by AFP
12:09, 5 FEBRUARY 2024

Listen to this article

Scammers tricked a multinational firm out of some US$26 million by impersonating senior executives using deepfake technology, Hong Kong police said Sunday, in one of the first cases of its kind in the city.

https://hongkongfp.com/2024/02/05/multinational-loses-hk200-million-to-deepfake-video-conference-scam-hong-kong-police-say/

PRO CYBER NEWS

Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case

Scams using artificial intelligence are a new challenge for companies

PHOTO: SIMON DAWSON/BLOOMBERG NEWS

By Catherine Stupp
Updated Aug. 30, 2019 12:52 pm ET

Criminals used artificial intelligence-based software to impersonate a chief executive's voice and demand a fraudulent transfer of €220,000 ($243,000) in March in what

Commercial Generative AI Today

Imagery

Video Avatars

Voice & Sound

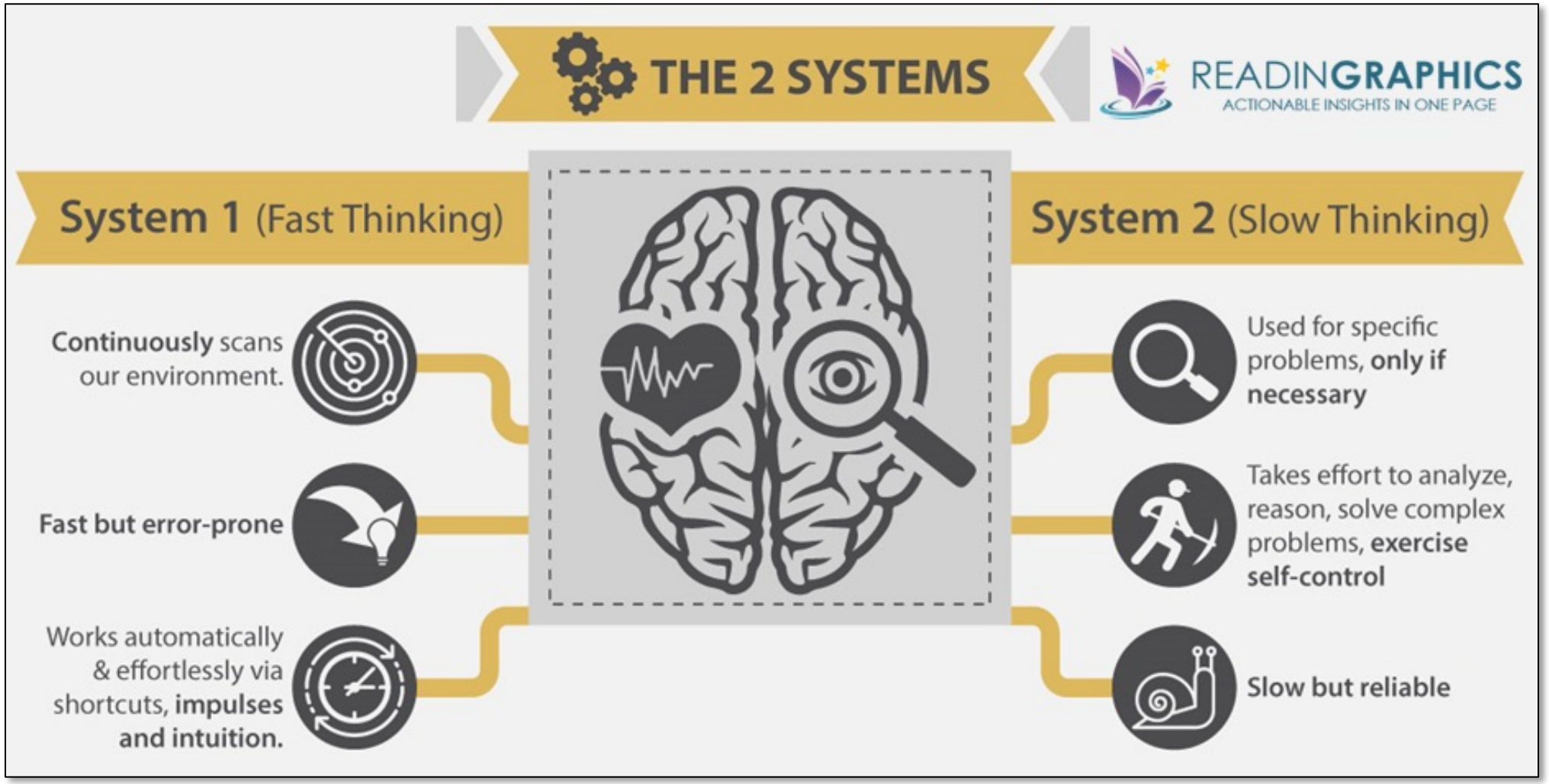
Realistic and imaginative scenes

The Number One Defense Against AI-Related Attacks

- The best way to defend against advanced deepfake and AI-generated attacks is to do what are already doing, but with some refinement
- AI-generated attacks will push technical defenses to their limits, however, AI still can’t match the power of the human mind
- Asking ourselves if something ‘feels’ odd or if a request is abnormal will go a long way toward countering these attacks

The Ideal Situation for a Social Engineer is to Create a Knee-jerk Action That Effectively Bypasses Critical Thinking and Results in the Attacker's Intended Action

Thinking, Fast & Slow (Daniel Kahneman)



Be Aware of Our Emotions

Getting to the Bottom of Deepfakes

Deepfakes are popping up more and more on social media.

What is a deepfake?

Deepfakes are digital versions of people, manipulated to make it seem like they are saying things they're not actually saying.

A bit of fun right?

Wrong.

The consequences of deepfakes can be really serious.

> **For example: If people are exposed to a deepfake of a politician spreading false claims it can have a dramatic impact on public perception.**

It is therefore critical, that when you see your favourite celebrity endorsing something, or a politician make a statement on a social media channel, you discern whether the content is indeed real or fake.

Some tips on how to spot deepfakes:

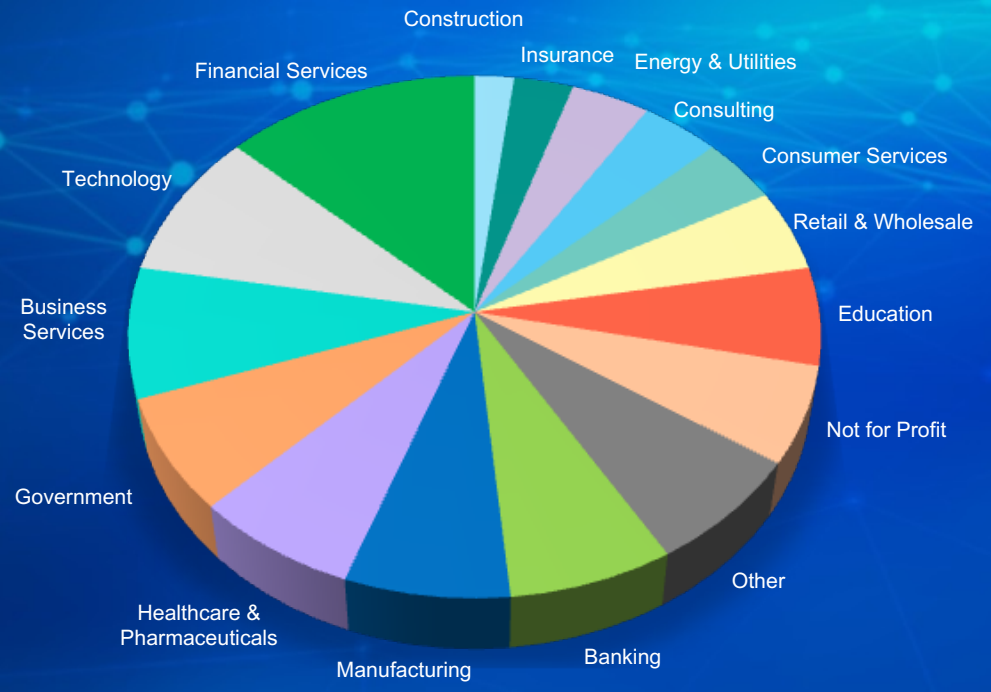
- Look for discolouration.
- Lighting inconsistencies.
- Synchronisation issues - people's eyes often look strange.
- Blurred lines, fuzzy hairlines, the hair and teeth can look unnatural.
- Proportions could be out of whack.
- Watch the video on a bigger screen. The inconsistencies are easier to spot when magnified.

KnowBe4 Human error. Conquered. | Popcorn Training a KnowBe4 company

Over

65,000

Customers



About Us

- The world's largest integrated Security Awareness Training and Simulated Phishing platform
- We help tens of thousands of organizations manage the ongoing problem of social engineering
- CEO & employees are industry veterans in IT Security
- Global Sales, Courseware Development, Customer Success, and Technical Support teams worldwide
- Offices in the USA, UK, Netherlands, India, Germany, South Africa, United Arab Emirates, Singapore, Japan, Australia, and Brazil

FORRESTER® WAVE LEADER 2022 Security Awareness And Training Solutions

AMERICA'S FASTEST-GROWING PRIVATE COMPANIES Inc. 500

Gartner peerinsights customers' choice 2021

Get started with KnowBe4 today!

Think before you click!

Raise Awareness

Red Flags!

Red flags are signs of danger or a problem. Protect yourself and your organization from cybercriminals by being aware of these warning signs and knowing actions to stay safe.

Common Red Flags	Actions to Stay Safe
Someone you don't know following you or your co-workers inside the office.	Contact security about unknown individuals.
Someone looking at your screen or watching what you type.	Pay attention to your surroundings and safeguard organizational information.
Someone you don't recognize looking through a desk.	Keep confidential information and devices locked-up/secured when not in use.
Social media connection requests from someone you don't recognize.	Don't accept unsolicited requests; report them to the service.
Receiving an unusual request from someone you know.	Contact the person directly to verify it's legitimate.
Requests that offer you something in exchange for private organizational information.	Be cautious before sharing any personal or organizational information.

BEC, Vishing, VEC, GenAI red flags

Defend Against Social Engineering

Create and sustain engagement

Foster Cybersecurity Mindset

AI Chatbots: Understanding Their Use, Risks, and Limitations in the Workplace

What's This AI Everyone Is Talking About?

What you'll learn:

- What an AI (artificial intelligence) chatbot is
- When and how to appropriately use AI chatbots
- The security and privacy risks of using AI chatbots
- How to identify when an AI chatbot provides biased or inaccurate information

Critical thinking

Build Competencies

KnowBe4
Human error. Conquered.

Dr. Martin J. Kraemer
martink@knowbe4.com
@markraemer
linkedin.com/in/martinkra/

NEWSLETTER

Securing Reality

AI, Humans, Social Engineering, Disinformation and the Battle For Trust