Trust in Voice. Verification in the Age of Synthetic Speech: Is it live -- or is it....?

John Parkinson, Senior Director Emerging Technologies

jparkinson@vailsys.com; john.parkinson@freeclimb.com

https://freeclimb.com/research

October 8, 2025

About Me

- **Senior Director, Emerging Technologies**, FreeClimb from Vail
- **Senior Advisor and Head of the Research in AI for Language Systems (RAILS) group at Vail Systems**
- **Partner & Managing Director**, ParkWood Advisors LLC
- 45+ years in technology
- 30+ years in consulting and advisory
- Published author (five books), columnist (CIO Insight, CFO Online. CXO Online) and contributor (WSJ, HBR Online, VentureBeat)
- Computerworld 100 Leaders in IT (2005)
- Advisory board member: MTEN, LLC
- Life senior member IEEE; Member ACM; Member CSCMP
- Named inventor on 15 patents

- Retired head of Innovation and Strategy at Ernst & Young LLP
- Former CTO of Capgemini in North America
- Former CTO and CISO at TransUnion
- Former SVP Global Programs Office, AXIS Capital Holdings Limited (NYSE:AXS)
- Former CEO, Entertainment Experience
- Former board member, Guam Telecom, US Micro, Inc., Visible Spectrum, Inc.
- Former advisory board member, Morningstar, Inc., Rand Technology
- Former Affiliate Partner, Waterstone Management Group

jparkinson@vailsys.com
john.parkinson@freeclimb.com
+1 847 877 4520
john@parkwoodadvisors.biz

5 Things about me not in my resume

- Once the subject of 1990 UK TV show about workaholics
- One of the first 1,000 or so people ever to have an email address
- Eaten dinner in over 100 countries
- Flown more than 12 million miles since 1979
- One of my solution designs made it into the Computerworld Smithsonian Collection

VAIL

RAILS
Research on AI in Linguistics and Systems

The team at a glance

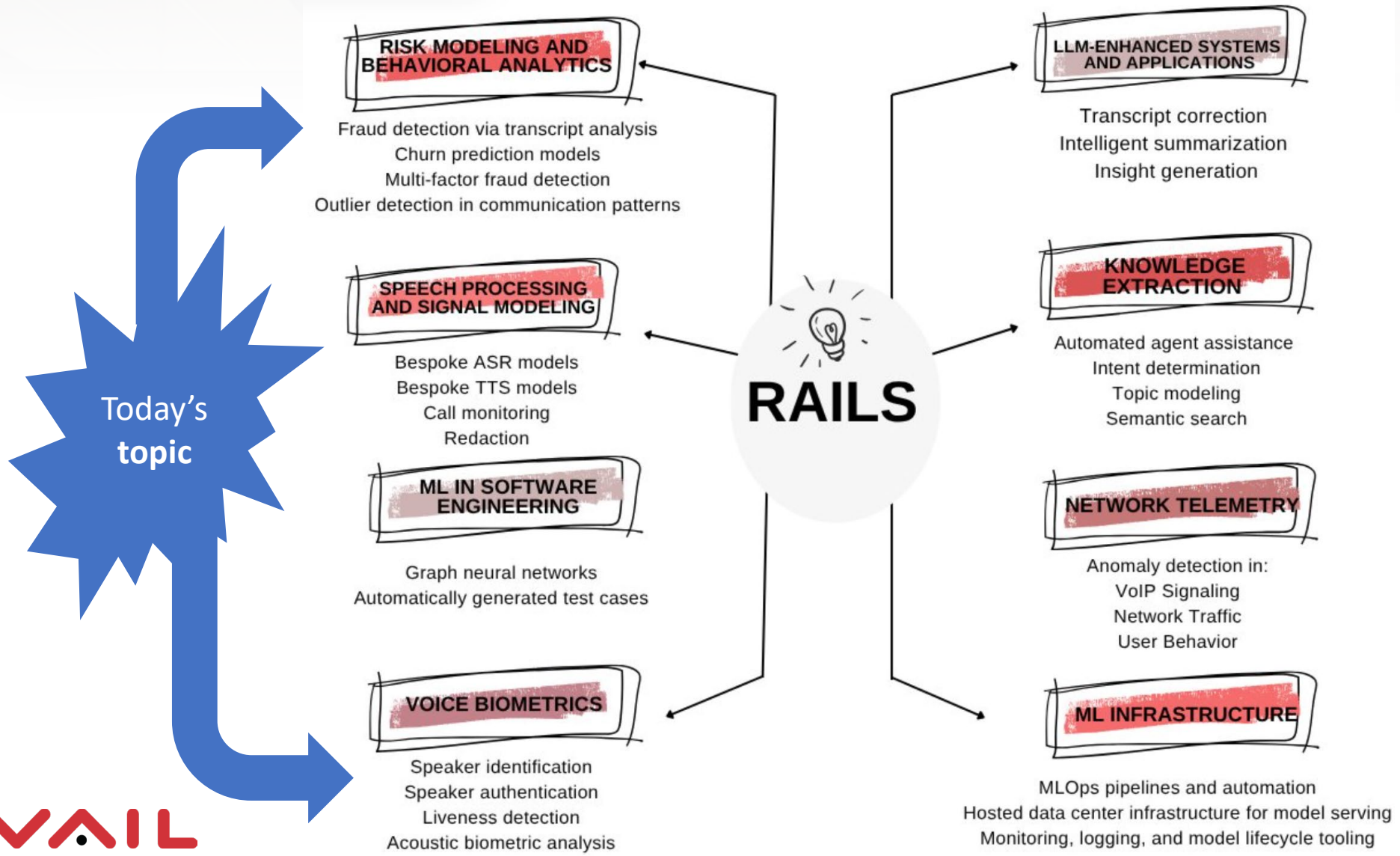
- Cross-disciplinary team with experience in:
  - Computational linguistics
  - Large scale telecommunications software systems and architectures
  - Machine Learning in telecommunication systems
  - Software engineering, open-source
  - Systems and data security, privacy, and governance
  - DevOps and the application of coding assistants in software engineering
- Enterprise ready expertise:
  - Model training, tuning, and deployment
  - Data governance
  - Patents, papers (academic and white) and industry talks
- Diverse academic background:
  - 4 team members with Ph.D. degrees
  - 2 team members with Master's degrees
  - 2 team members with Bachelor's degrees

VAIL

RAILS
Research on AI in Linguistics and Systems

The Team: Areas of work

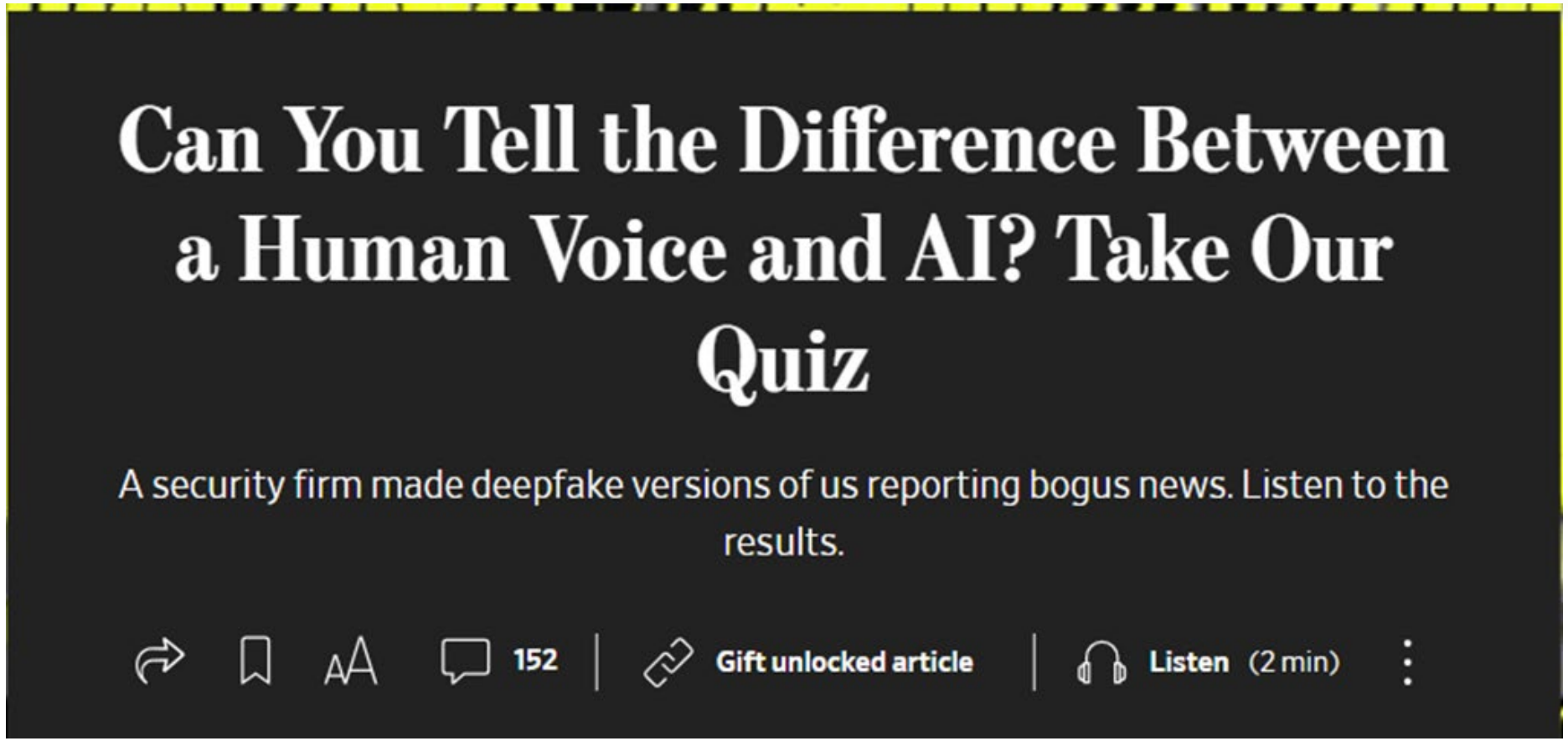
Read about our research and published results at:
https://www.freeclimb.com/research/research-index
And about the team at:
https://www.freeclimb.com/research/



VAIL

RAILS
Research on AI in Linguistics and Systems

What's the Problem?



From WSJ, Sept. 13, 2025
Try it for yourself at:

But **not** during the presentation, please!

VAIL

RAILS
Research on AI in Linguistics and Systems

AI-generated voices are now indistinguishable (by humans) from real human voices

- New research from Queen Mary University of London shows that AI voice technology has now reached a stage where it can create "voice clones" or deepfakes which sound just as realistic as human recordings.
- The study compared real human voices with two different types of synthetic voices, generated using state-of-the-art AI voice synthesis tools.
- Some were "cloned" from voice recordings of real humans, intended to mimic them, and others were generated from a large voice model and did not have a specific human counterpart.
- The study found that voice clones can sound as real as human voices, making it difficult for listeners to distinguish between them.
- Both types of AI-generated voices were evaluated as more dominant than human voices, and some were also perceived as more trustworthy.

https://dx.plos.org/10.1371/journal.pone.0332692

VAIL

RAILS
Research on AI in Linguistics and Systems

Why is this a problem?
Voice fraud is on the increase!

- In December 2024,, the Financial Crimes Enforcement Network (FinCEN) issued an alert, FIN-2024-Alert004, to help financial institutions identify fraud schemes that use deepfake media created with generative artificial intelligence (GenAI) tools.
- FinCEN wrote that it had “observed an increase in suspicious activity reporting by financial institutions describing the suspected use of deepfake media in fraud schemes targeting their institutions and customers” beginning in 2023 and continuing into 2024.
- Deloitte’s Center for Financial Services predicts that GenAI could enable fraud losses to reach US$40 billion in the United States by 2027.
- Over 2024, cases of AI deepfake fraud or attempted fraud have made headlines in the press and resulted in millions of dollars in losses for the unluckier targets, underscoring the threat this technology poses to organizations and individuals when effective security is not in place.

Source: https://incode.com/blog/top-5-cases-of-ai-deepfake-fraud-from-2024-exposed/

A malicious actor recently used AI voice cloning to impersonate U.S. Secretary of State Marco Rubio to target high-level officials. According to reports, the AI voice cloning impersonation reached at least three foreign ministers, a U.S. governor, and a member of Congress.

Source: US Department of State

VAIL

RAILS
Research on AI in Linguistics and Systems

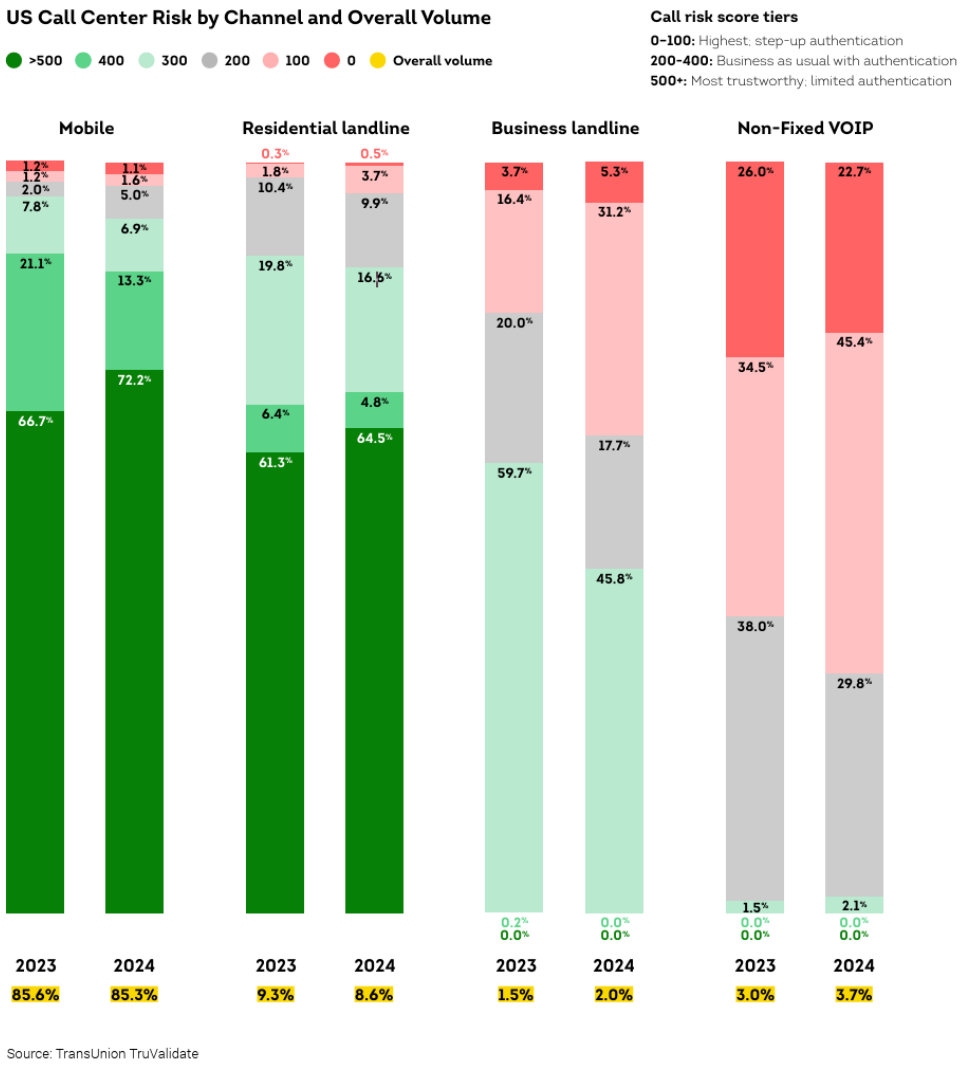
....and is accelerating in 2025....

Family emergency or "grandparent" scams

- Man swindled out of $25,000: An elderly man lost $25,000 after receiving a phone call with what he believed was his son's voice. The cloned voice, used by AI, told a story about being in a car accident involving a pregnant woman and needing money for bail. The scammer then had the man wire multiple payments through an intermediary posing as a lawyer.
- False kidnapping call: In another case, a woman received a frantic call with a voice she thought was her daughter, saying she had been kidnapped. The fraudster demanded a $1 million ransom, but the victim was able to confirm her daughter was safe and avoided losing any money.
- Grandparent scam evolution: The classic "grandparent scam," where criminals impersonate a grandchild in trouble, has been amplified by AI. Once they capture a small audio sample from a victim's family member, scammers can use AI to make that person's "voice" say anything, creating a more convincing and effective scam.

Other notable voice fraud incidents

- Deepfake robocall for election interference: In January 2024, an AI-generated robocall impersonating President Joe Biden encouraged Democrats to not vote in the New Hampshire primary.
- School principal deepfake audio: An AI-manipulated audio clip, falsely attributed to a school principal, went viral in January 2024. The clip contained derogatory remarks and was created by an athletics director facing an investigation to discredit the principal.
- AI "scam sweatshops": Authorities in Scotland uncovered AI "scam sweatshops," where criminals use cheap and accessible AI tools to create hyper-personalized fraud campaigns in under two minutes.

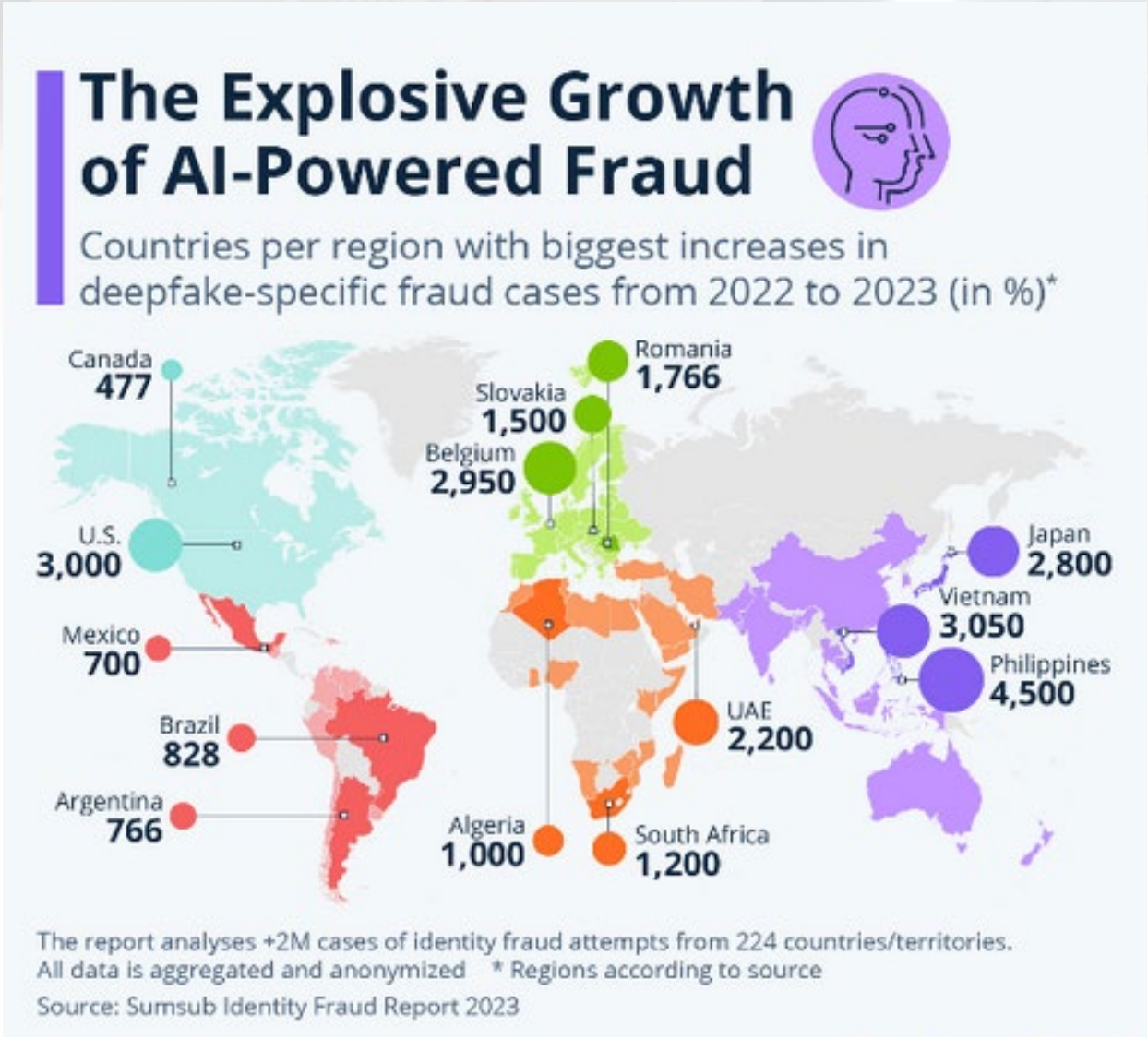


See more examples in: National Counterterrorism Innovation, Technology, and Education Center, " Deepfakes and Fraud: Real World Examples of AI Misuse" (2025). Reports, Projects, and Research. 136. https://digitalcommons.unomaha.edu/ncitereportsresearch/136

VAIL

RAILS
Research on AI in Linguistics and Systems

….and is Global (just like the phone system)



See also: https://fraudtaskforce.aspeninstitute.org/?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axiosam&stream=top

VAIL

RAILS
Research on AI in Linguistics and Systems

...and while all this is illegal...

The bad guys don't care:

- Low cost to operate
- High rate of return
- Low risk of detection
- Target rich environment

What's not to like?

“Voice” has some additional challenges

- While we can generally detect deepfake video with techniques such as multispectral analysis of actors' images, we can’t actually “see” voice
- Channel characteristics, codecs and transcoding can mask or distort important voice characteristics
- The overall acoustic environment matters

And

- Although the incidence of voice fraud is increasing, it’s still only a tiny fraction of all voice calls, so “cost to detect” matters too
- We don’t generally want to add unnecessary “friction” or latency to caller interactions
- We want to avoid false positives
- We don’t (yet) have a lot of correctly labelled training data available

VAIL

RAILS
Research on AI in Linguistics and Systems

So, what can we do about this?

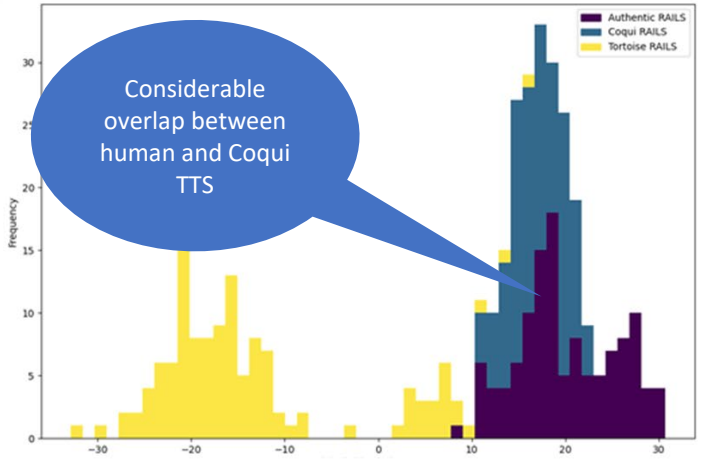
Use the technology that caused the problem to address the problem.

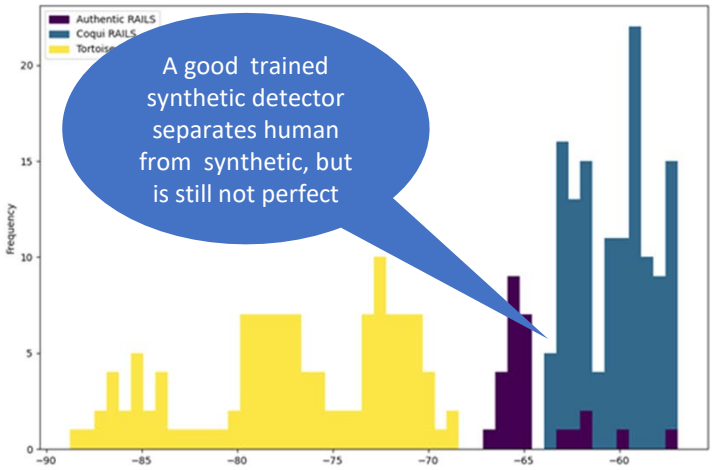
5-30 sec of audio is all that is needed to clone voices.

At RAILS we have focused on three complimentary ML-driven approaches

1. Distinguishing synthetic voice (recorded, cloned and generated) from live human speech
2. Speaker recognition
3. Risk based scoring based on context and the results from the above

Our models have performed well against a variety of commonly encountered generators and are reasonably robust against new synthesizers





Did you know that you can record a call on your iPhone? With the appropriate opt-in language added automatically.
And then transcribe it?
Both the recording and the transcription could then be used for voice cloning or synthesis.

https://support.apple.com/guide/iphone/record-and-transcribe-a-call-iph57c6590e9/ios

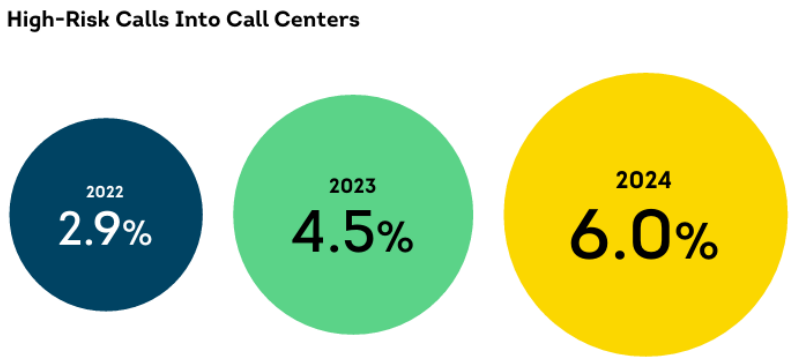
Echoes unveiled: Identifying synthetic voices
Dan Pluth Jordan Hosier Yu Zhou Vijay K. Gurbani IEEE PerCom 2025 (2025).
A low latency technique for speaker detection from a large negative list
Yu Zhou Vijay K. Gurbani B. Chandra Mouli ICNLSP 2022: 202-211 (2022)

VAIL

RAILS
Research on AI in Linguistics and Systems

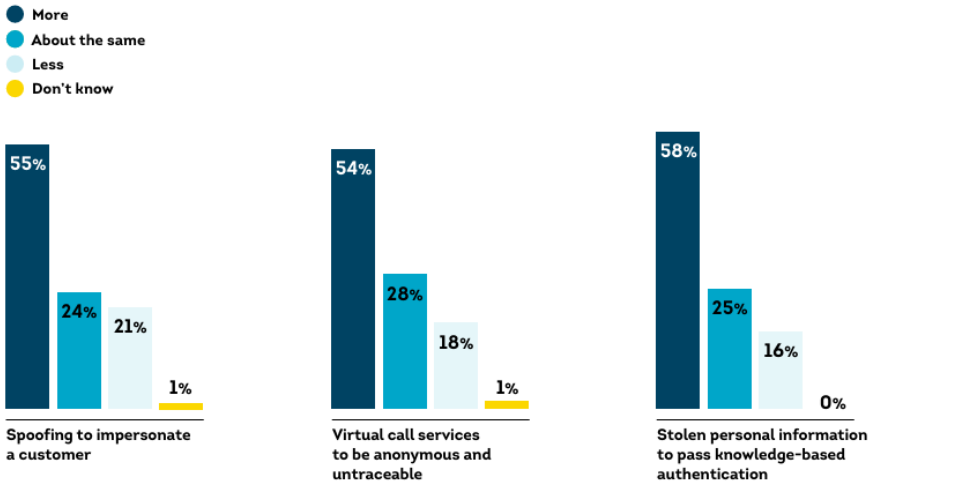
Continuing work

- As voice synthesis continues to evolve, we need to continue to test our liveness detection capabilities
- We are continuing to improve the ability of liveness detection to work with low quality audio
- It may be possible to build and enhance a “negative” list of high-risk contexts and intents with related behavioral analysis
- There could be a role for “fingerprinting” voice synthesizers if we can collect enough data
- We are looking at the implications for multichannel beyond voice, including SMS and chat



High-Risk Calls Into Call Centers

Source: TransUnion TruValidate



Criminal Call Center Fraud Tactics

Percentage of business leaders who reported how certain criminal tactics associated with call centers changed in the past year among those who said they're extremely or very knowledgeable about fraud in their call center

VAIL

RAILS

Research on AI in Linguistics and Systems

Links and Resources

https://freeclimb.com/research

Speaker identification demo:

Call 443-825-4483 to provide a voice sample and follow the prompts for confirmation of recognition via a second call.

By accessing this demonstration, you agree that Vail Systems has permission to record your voice and create a representation from the recording that can be used for subsequent speaker recognition.

Vail Systems will not use the recording or the representation of your voice for any other purpose, including but not limited to training a machine learning model. All recordings and derived data will be deleted by close of business on October 9th, 2025.

VAIL

RAILS
Research on AI in Linguistics and Systems