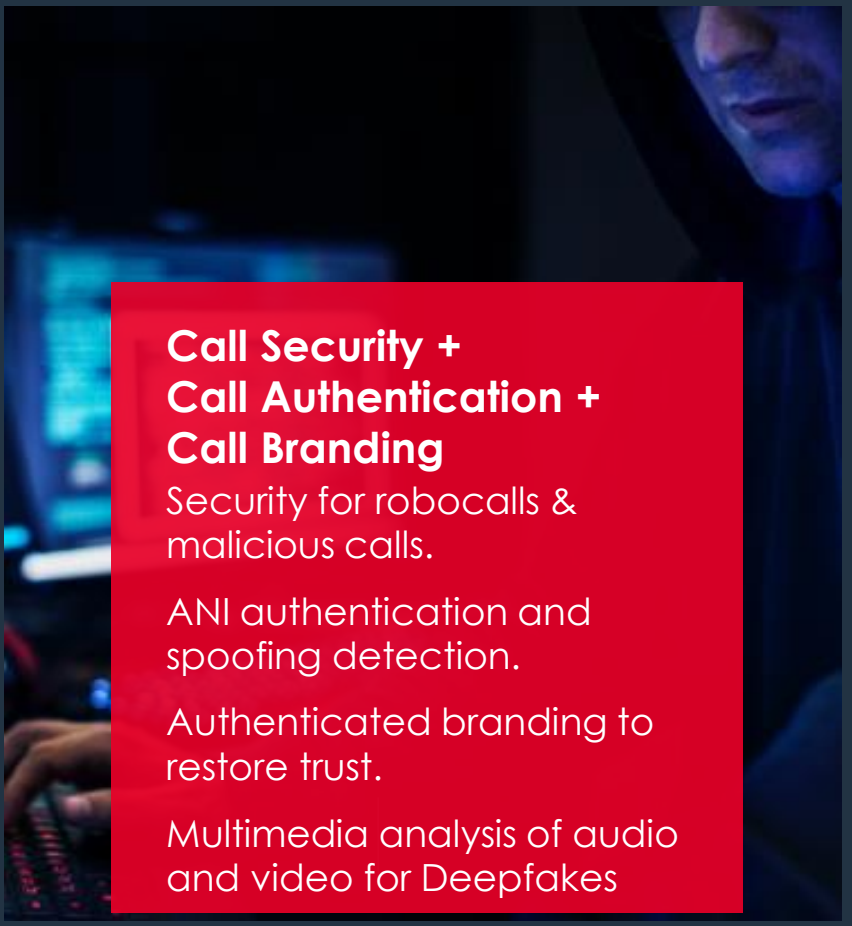
Audio Deepfakes Discussion

Mark Collier
Chief Technology Officer

Call Security & Trust



Call Security + Call Authentication + Call Branding

Security for robocalls & malicious calls.

ANI authentication and spoofing detection.

Authenticated branding to restore trust.

Multimedia analysis of audio and video for Deepfakes

Industry Leading Expertise

Professional services team has 20+ years of experience securing billions of calls.

Partnerships, Research & Innovation

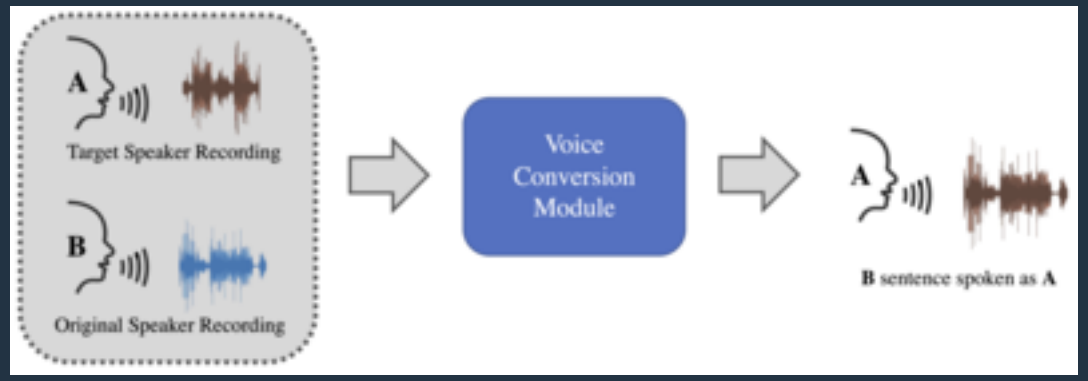
Strategic relationship with Verizon, AT&T, and T-Mobile. PhD-led R&D work with U.S. DHS and DoD.

Definition

- A Deepfake uses multimedia content to impersonate a person
- Includes video and messaging (video is more difficult than audio)
- Focused today on audio/voice Deepfakes
- Has the potential to make impersonation attacks more damaging

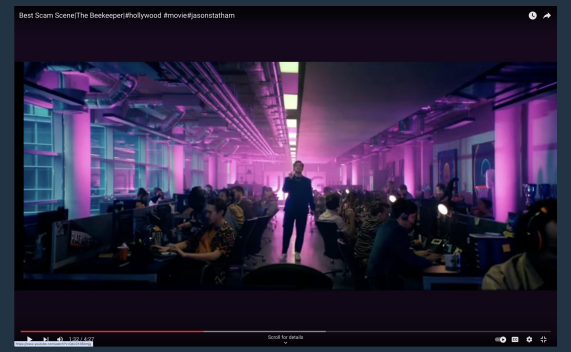
Threat Overview

- A Deepfake uses generative Artificial Intelligence (GenAI)
- Can impersonate a person or specific “important” person
- Builds upon impersonation based on spoofed calling number/identity
- Consider how effective SCAMs and Account Take Over (ATO) are now
- A Deepfake has the potential to make attacks far more convincing
- AI for attacks will outpace AI for defense (at least a year)



Types of Audio Deepfakes

- Asynchronous versus synchronous conversation
- Impersonate a customer calling into retail contact center
- Impersonate person for SWATting, active shooter hoaxes, and bomb threats
- Impersonate a person for robocall SPAM and SCAMs
- Impersonate “important” person for social engineering



Retail Contact Center

SecureLogix®

- Threat – now to soon – involves many calls
- Moderate loss of funds or personal information
- Not impersonating a known person to a random agent
- Goal is to fool voice authentication and biometrics
- Lively debate from voice authentication companies
- Will probably be effective
- Will increase the need for multi-factor authentication
- Increases emphasis on calling number authentication/identity
- Relatively easy to deploy countermeasures on contact center voice

SWATting

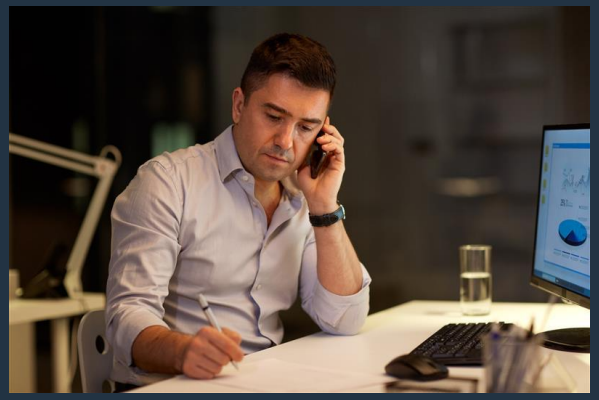
- Threat – now – affects a small number of calls – potentially very damaging
- Focused on disruption or misdirection – not money
- Somewhat unique to public safety and education
- SWATting, active shooter hoaxes, bomb threat, etc.
- SWATting as a Service is available
- Can be used to generate many calls simulating a real event
- Must be taken seriously
- Increases emphasis on calling number authentication/identity
- Difficult to deploy countermeasures at many different sites

Robocall SPAM and SCAMs

- Threat – soon – affects enormous number of calls – very damaging
- Robocalls are already very effective
- Maybe Deepfakes are not needed for simple robocalls
- Could make scams more effective
- Recon and grooming targets
- More believable audio – Not “Roger” with thick accent
- Possibly useful for voice mails
- Increases emphasis on calling number authentication/identity
- Targets are consumers with mobile or land lines and hard to protect

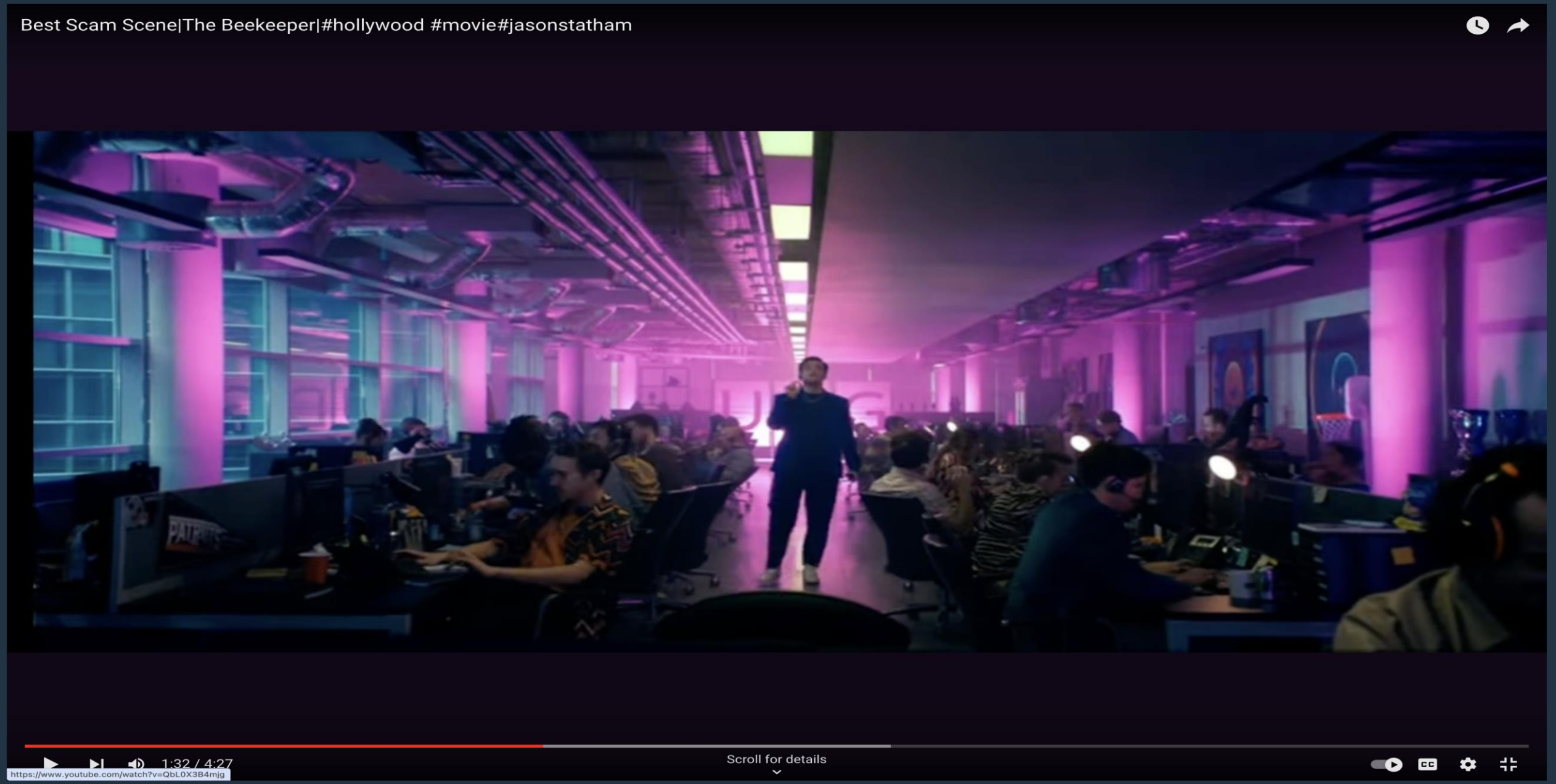
“Important” Person call

- Fewer targeted calls - Highly damaging
- CEO call to accounting to approve high value transfer (crypto…)
- High wealth person call to wealth management person
- Executive person to IT to gather/reset passwords
- Military/government/public safety examples
- Will require much more synchronous AI
- Increases emphasis on calling number authentication/identity
- Targets are consumers with mobile or land lines and hard to protect



AI-Based Scams – Can GenAI Do This?

SecureLogix®



Audio Deepfake Tools and Techniques

SecureLogix

1. Lyrebird (Descript)
   - https://www.descript.com/lyrebird
   - AI tool for creating realistic voice clones from minimal samples
2. Resemble.AI
   - https://www.resemble.ai
   - Custom voice synthesis platform offering real-time voice cloning
3. OpenAI's Jukebox
   - https://openai.com/research/jukebox
   - Neural network model for generating high-fidelity audio, including voice and music

Audio Deepfake Tools and Techniques *(contd.)*

4. WaveNet by Google DeepMind
   - https://deepmind.google/technologies/wavenet/
   - Deep neural network for generating realistic audio waveforms
5. DeepVoice by Baidu
   - https://research.baidu.com/Blog/index-view?id=91
   - Text-to-speech system that generates natural-sounding speech using deep learning
6. Speechify
   - https://speechify.com
   - Real-time text-to-speech platform that converts text into natural-sounding speech

Potential Countermeasures

1. **Advanced Voice Biometrics and Deepfake Detection in Contact Centers**
   - Analyze multiple voice characteristics and compare against known voice profiles
   - AI-based systems to detect anomalies and synthetic features in audio samples
2. **Enhanced Multi-Factor Authentication (MFA)**
   - Use additional biometrics layers combined along with the existing measures
   - Require contextual verification over alternative communication channels
3. **AI-Based Analysis of Call Content**
   - AI to analyze the context of conversations, language patterns, syntax, and semantics

     AI to monitor for deviations in behavioral patterns or speech anomalies during calls

Potential Countermeasures

4. **Detection of Audio Artifacts**
   - Detect artifacts (pauses, robotic, synthetic, or irregular content) in audio
   - Use forensic analysis techniques to examine the frequency spectrum and manipulation
5. **Audio Watermarking**
   - Embed inaudible audio watermarks in the content
   - Implement systems that can detect and decode watermarks
6. **Passphrases and Verification Codes**
   - Use dynamic passphrases exchanged between individuals
   - Use one-time verification codes or pre-established security questions

Potential Countermeasures *(contd.)*

SecureLogix®

7. **Deepfake Detection in Voice Networks and Consumer Devices**
   - Deploy deepfake detection tools within voice networks
   - Incorporate deepfake detection capabilities into consumer devices and applications
8. **Human Review and Verification**
   - Employ audio experts to manually review suspicious recordings for signs of deepfakes
   - Use crowdsourcing to gather multiple reviews and opinions on suspicious audio
9. **Improved Identity and Calling Number Authentication**
   - Implement improved identity solutions
   - Improve authentication of calling numbers
   - Implement robust end-to-end encryption for voice communications