

# Your voice confirms my identity

Ethan McKee-Harris BSides San Francisco



#### \$whoami

- Ethan McKee-Harris
  - Go by Skelmis online
- Security Consultant @ Bastion Security Group
- Also an avid open source developer
- www.skelmis.co.nz



#### **Overview**

- Deep faking within our current society
- The rise of voice as a security boundary
- How to clone your own voice
- Practical voice cloning using someone's digital presence
- Industry trends and mitigating factors

#### **Outcomes**

- A better understanding of the voice biometric space
- Practical voice cloning skills
- How to use them to bypass voice authentication
- How to use voice authentication securely (don't)
- Understanding the risk voice authentication poses to your business

#### Deep faking within our current society

AI voice clones mimic politicians and celebrities, reshaping reality

# Don't believe your ears: voice deepfakes

# Voice Deepfakes Are Coming for Your Bank Balance

Artificial intelligence tools have given scammers a potent weapon for trying to trick people into sending them money.

#### Deep faking within our current society

## AI voice clones mimic politicians and celebrities, resh How I Broke Into a Bank Account With an AI- : voice Generated Voice

# Voice Deepfakes Are Coming for Your Bank Balance

Artificial intelligence tools have given scammers a potent weapon for trying to trick people into sending them money.

#### Deep faking within our current society

#### AI voice clones mimic politicians and celebrities, resh How I Broke Into a Bank Account With an Al-: voice Genera F.C.C. Bans A.I.-Generated Robocalls Voice Deepfakes Are Co The move by the agency follows a phony call last month that was Your Bank Balance made to sound like President Biden telling New Hampshire voters to stay home.

Artificial intelligence tools have given scammers a potent weapon for trying to trick people into sending them money.

 Voice authentication is often seen as the next step forward in a simplified user experience

- Voice authentication is often seen as the next step forward in a simplified user experience
  - Yet it's seen as a step backwards for user security

- Voice authentication is often seen as the next step forward in a simplified user experience
  - Yet it's seen as a step backwards for user security

Your voice can be recorded

- Voice authentication is often seen as the next step forward in a simplified user experience
  - Yet it's seen as a step backwards for user security

- Your voice can be recorded
- Your voice can be cloned



#### How secure and reliable is Voice ID?

Voice ID is a secure authentication method that uses advanced biometric security to verify your voice based on hundreds of unique characteristics. This makes it difficult for someone else to imitate you or even use pre-recorded audio of your voice.

ANZS

~



Overview Voice ID Watch video FAQs

# With Voice ID, we can verify you by the sound of your voice.

Similar to a fingerprint, Voice ID uses your unique voiceprint to verify you—so it's easy, fast and secure.



Your voice is unique, just as your fingerprint is which means you can create your own voiceprint with us. Once you've created your 'voiceprint', you'll be able to use your voice to access telephone banking and we'll use this to further help protect against fraud.



HSBC Channel Islands and Isle of Man https://ciiom.hsbc.com > ... > Telephone Banking

Voice ID | Access Phone Banking - HSBC CI



#### What is Voice Verification?

Voice Verification is a biometrics service that allows customers to use their unique voice to verify their identity when they call Wells Fargo for service on their accounts. Customers can choose to speak a simple passphrase instead of entering other information.

LLOYDS BANK	*	<b>≡</b> Menu		
Home <b>&gt;</b> Cor	ntact us 🔉 🔉	Voice ID	>	

# Voice ID

#### Extra security using your voice

- Simple to set up and use.
- Confirm your identity using your voice.
- Once set up you won't need to remember extra security information.

You can use Voice ID if you are eligible and hold a bank, savings, or credit card account with us.



#### Using Voice ID with a speech impairment or voice box

It may be possible for you to set up Voice ID even if you have a speech impairment. For Voice ID to work we need at least three strong recordings of your voice as you repeat the passphrase.

We don't recommend setting up Voice ID if you use a voice box or speech synthesiser. They might be a security risk as some speech synthesiser machines have the same voice.

If Voice ID is not right for you, please let us know.

Can someone else access my account with my permission?

Since your voice is unique, someone else won't be able to use Voice ID on your behalf. **To protect yourself** - don't share your security number or security details with anyone else.

Multiple ways to bypass a given implementation

- Multiple ways to bypass a given implementation
  - Mimicking voices

- Multiple ways to bypass a given implementation
  - Mimicking voices
  - Recording the victim without their knowledge

- Multiple ways to bypass a given implementation
  - Mimicking voices
  - Recording the victim without their knowledge
  - Splicing existing audio to form the correct phrase

- Multiple ways to bypass a given implementation
  - Mimicking voices
  - Recording the victim without their knowledge
  - Splicing existing audio to form the correct phrase
  - Spoofing voices using AI

- Let's look at some example data
- We sampled 15 people four times
- Each axis represents two audio clips from a person
- The numbers represent the speaker cross axis



27


















### Sending audio from laptops to phones

Highly complicated audio setups for bypasses

### Sending audio from laptops to phones

### Highly complicated audio setups for bypasses





• So with that, what do we get?

• So with that, what do we get?

• We have a low barrier to entry

• So with that, what do we get?

- We have a low barrier to entry
- We have high value targets

So with that, what do we get?

- We have a low barrier to entry
- We have high value targets
- Well we get a hackers field day

# **Cloning your own voice** – Training data

- Aim to produce high fidelity audio
  - Use a sound proofed room
  - Use a decent microphone
  - Speak clearly

### **Cloning your own voice** – Training data

- Aim to produce high fidelity audio
  - Use a sound proofed room
  - Use a decent microphone
  - Speak clearly

 The goal here is to produce clear audio without background interference

# **Cloning your own voice**

### • Let's use our new training data to build a voice

### **Cloning your own voice** – ElevenLabs

- Let's use our new training data to build a voice
- ElevenLabs is the current 'go to' platform for online voice cloning I have found from my research
  - It's a low barrier to entry for anyone interested
  - Only requires a few minutes of audio for a voice
  - Extremely cheap
  - Fast turn around from sign up to voice generation

# Cloning your own voice - Creating a voice

II Al Voice Generator & Tex ×		sts		
$\leftarrow \ \rightarrow \ \mathbf{G}$		O A https://elevenlabs.io/app/voice-lab		
I	llElevenLabs	Test out our new Multilingual Speech to Speech Model $\rightarrow$		
olþ	Speech	VoiceLab Voice Library		
ð	Voices	iceLab		
Φ	Projects			
Ŕ	Dubbing	creative AI toolkit. Design entirely new synthetic voices from scratch. Clone your own voice or a voice you have a per reate.		
6	Payouts	ID   Add Generative or Cloned Voice		
		25 / 30 🖤 Use 🗹 Edit 🗍 Remove		

### Cloning your own voice - Creating a voice

Type of voice to create

X

### Voice Design

Design entirely new voices by adjusting their parameters. Every voice you create is randomly generated and is entirely unique even if the same settings are applied.

### **4** Instant Voice Cloning

Clone a voice from a clean sample recording. Samples should contain 1 speaker and be over 1 minute long and not contain background noise.

### Voice Library

Add a voice from our community.

### Professional Voice Cloning

Creator+ only. Subscribe?

Create a perfect digital replica of your voice. Training running monthly.

### Cloning your own voice - Creating a voice

Add Voice		$\times$
Name		
1 	<u> </u>	1
1		1
1	Click to upload a file or drag and drop	1
1	Audio or Video files, up to 10MB each	1
·		/

### Samples 0/25

- No items uploaded yet. Upload audio samples of the voice you would like to clone.
- Sample quality is more important than quantity. Noisy samples may give bad results. Providing more than 5 minutes of audio in total brings little improvement.

### Labels 0/5

No labels. Click + to add a first one.

#### Description

How would you describe the voice? e.g. "An old American male voice with a slight hoarseness in his throat. Perfect for news."

• We now have a voice, let's use it

• We now have a voice, let's use it



### **Speech Synthesis**

Unleash the power of our cutting-edge technology to generate realistic, captivating speech in a wide range of...

SIMPLE ADVANCED

GENERATE HISTORY

TEXT TO SPEECH SPEECH TO SPEECH

Start typing here or paste any text you want to turn into lifelike speech.

Ethan - IVC  $\diamond$  Voice

Voice settings

0/5000 Generate speech

Let's listen to some examples using our new voice

### Let's listen to some examples using our new voice

High

Exaggerated





• So how do we use this for offensive purposes?

- So how do we use this for offensive purposes?
- Let's clone the voice of a high value target:
  - Can use it for phishing
  - Access potentially sensitive information like business bank accounts
  - Etc

- So how do we use this for offensive purposes?
- Let's clone the voice of a high value target:
  - Can use it for phishing
  - Access potentially sensitive information like business bank accounts
  - Etc



So what do we already know?

- So what do we already know?
  - We know the victim's name: "Simon Howard"

- So what do we already know?
  - We know the victim's name: "Simon Howard"
  - We know the victim's company: "ZX Security"

- So what do we already know?
  - We know the victim's name: "Simon Howard"
  - We know the victim's company: "ZX Security"
  - We know we need training data, and that we need to source it without alerting the target

- So what do we already know?
  - We know the victim's name: "Simon Howard"
  - We know the victim's company: "ZX Security"
  - We know we need training data, and that we need to source it without alerting the target

Let's go check out YouTube

Looks like Simon has a digital presence we can use



Looks like Simon has a digital presence we can use

 Let's download those videos and train a voice on them



• Once you have suitable audio, we can simply follow the same processes as we used in the prior section

 Once you have suitable audio, we can simply follow the same processes as we used in the prior section

- And there we have it, a viable clone of Simon
  - It's not perfect, but it can prove to be good enough for our avenues of abuse

 Once you have suitable audio, we can simply follow the same processes as we used in the prior section

- And there we have it, a viable clone of Simon
  - It's not perfect, but it can prove to be good enough for our avenues of abuse
  - Now lets hear some examples shall we



### Industry trends and mitigating factors

So how do you do it right?

### Industry trends and mitigating factors

So how do you do it right?Short answer: You don't

### Industry trends and mitigating factors

So how do you do it right?

- Short answer: You don't
- Relying on a non-unique medium
- Your voice inherently has variability
- And your voice can be easily obtained
# **Defense in depth for consumers**

• Be risk aware

### **Defense in depth for consumers**

- Be risk aware
- Use alternative authentication offerings

### **Defense in depth for consumers**

- Be risk aware
- Use alternative authentication offerings



## **Defense in depth for companies**

Add additional forms of authentication

## **Defense in depth for companies**

- Add additional forms of authentication
- Let users pick a private phrase

### **Defense in depth for companies**

- Add additional forms of authentication
- Let users pick a private phrase
- Ensure that the spoken phrase is a sentence and not numbers

### **Defense in depth for vendors**

- Provide means to verify if content was produced by you
  - This may take the form of something like cryptographic signing

#### **Defense in depth for vendors**

- Provide means to verify if content was produced by you
  - This may take the form of something like cryptographic signing

Push for the adoption of these checks when used

#### Recap

- Your voice is not a security border
- This technology is only going to become more prevalent with time
- Voice cloning is also only going to become better with time
- We have also made a couple of voice clones

When we base our authentication on solutions which are not inherently unique or secret, we are destined to fail.

# **Questions?**

bastionsecurity.co.nz

ethan.mckee-harris@bastionsecurity.co.nz

