BASTION
SECURITY GROUP

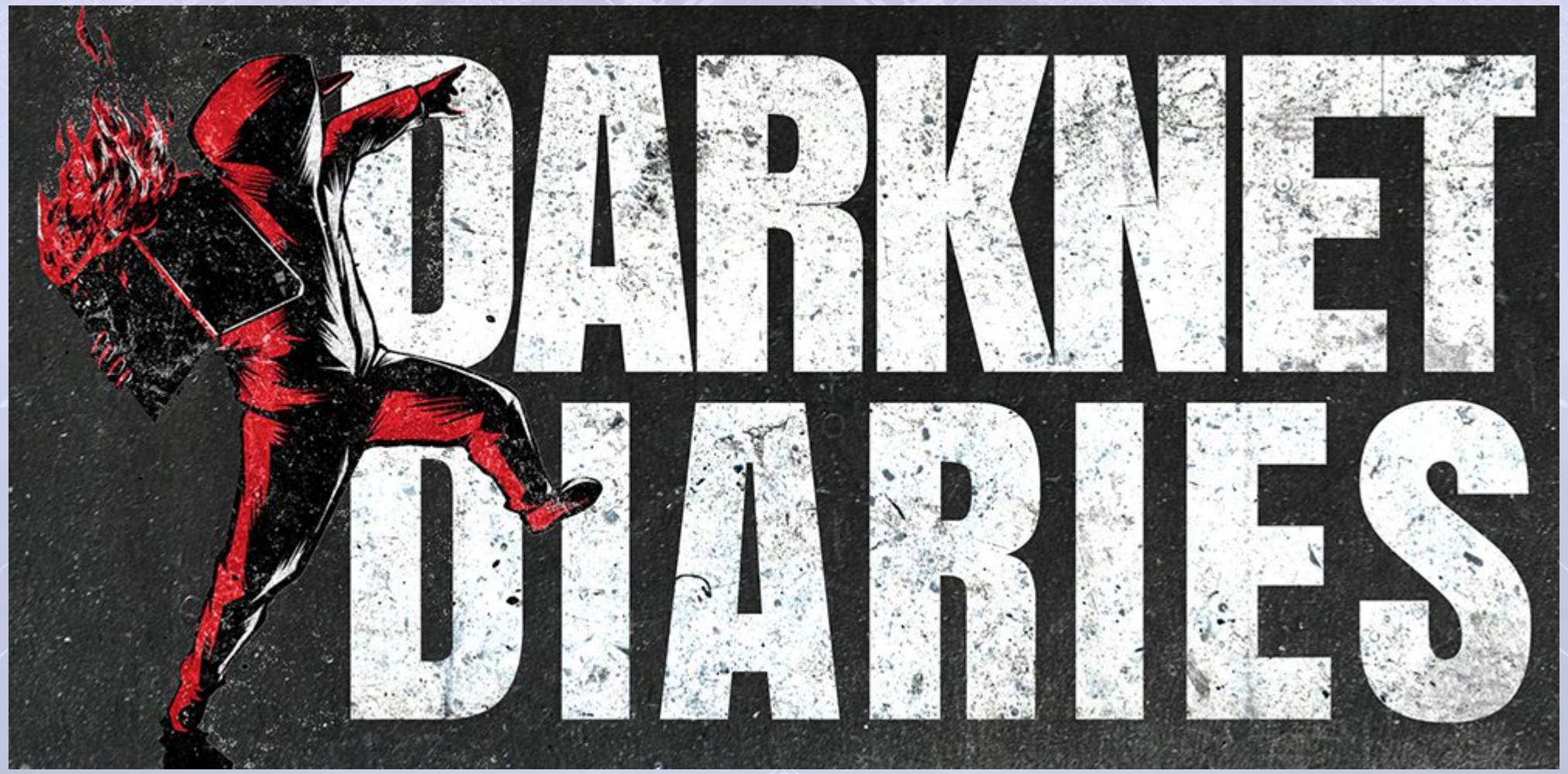
Your voice confirms my identity

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


DARKNET
DIARIES

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-Generated Voice

: 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 AI-

: voice

Genera

F.C.C. Bans A.I.-Generated Robocalls

The move by the agency follows a phony call last month that was made to sound like President Biden telling New Hampshire voters to stay home.

Voice Deepfakes Are Co
Your Bank Balance

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

The rise of voice as a security boundary

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

The rise of voice as a security boundary

- 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

The rise of voice as a security boundary

- 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

The rise of voice as a security boundary

- 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

Voice biometric usage

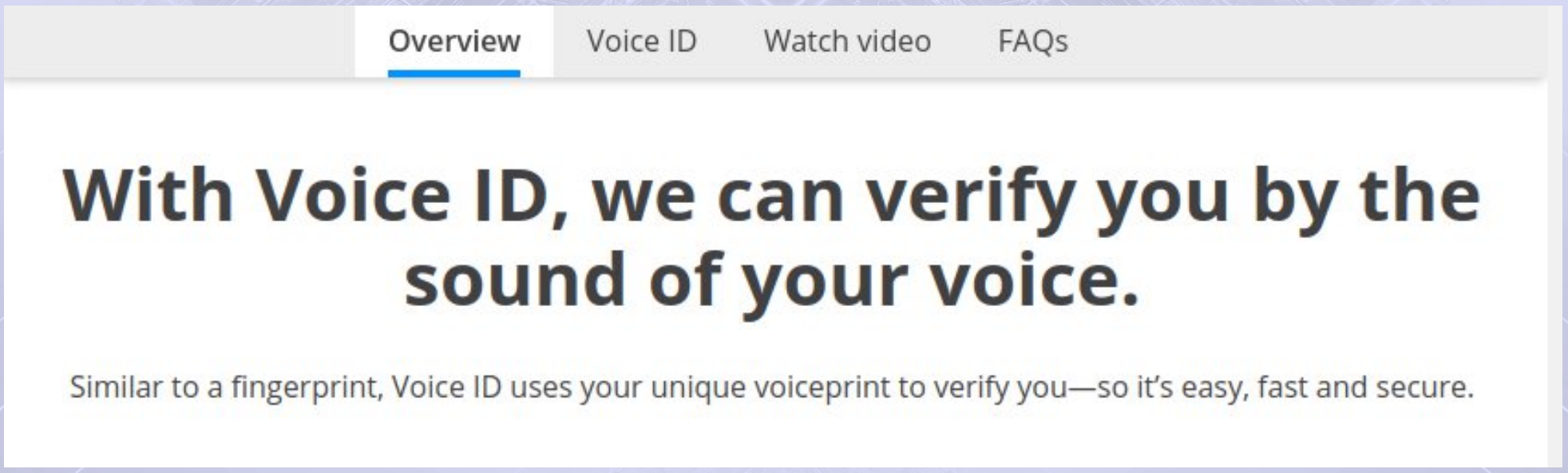
Voice biometric usage

Voice biometric usage

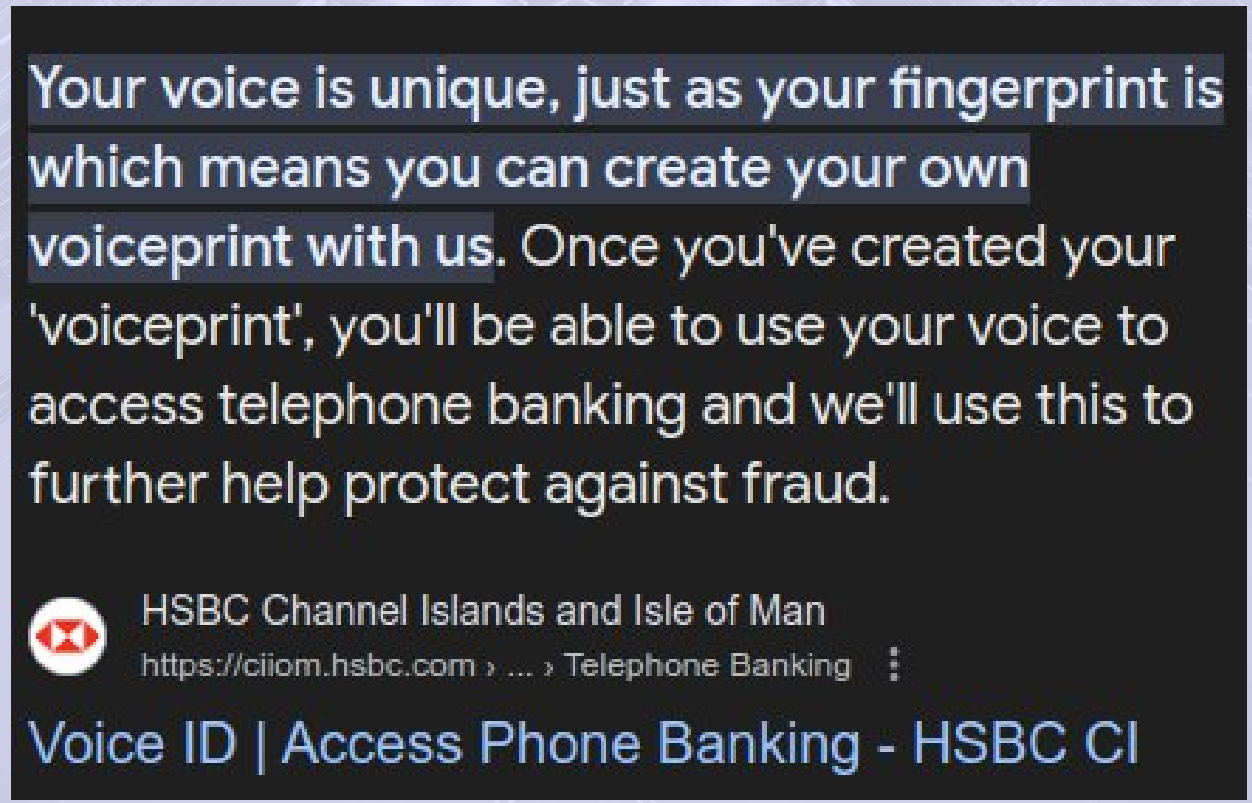
How secure and reliable is Voice ID?

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.

Voice biometric usage



Voice biometric usage



Voice biometric usage

WELLS FARGO

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.

Voice biometric usage

Voice biometric usage

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.**

Bypassing voice biometric solutions

- Multiple ways to bypass a given implementation

Bypassing voice biometric solutions

- Multiple ways to bypass a given implementation
    - Mimicking voices

Bypassing voice biometric solutions

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

Bypassing voice biometric solutions

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

Bypassing voice biometric solutions

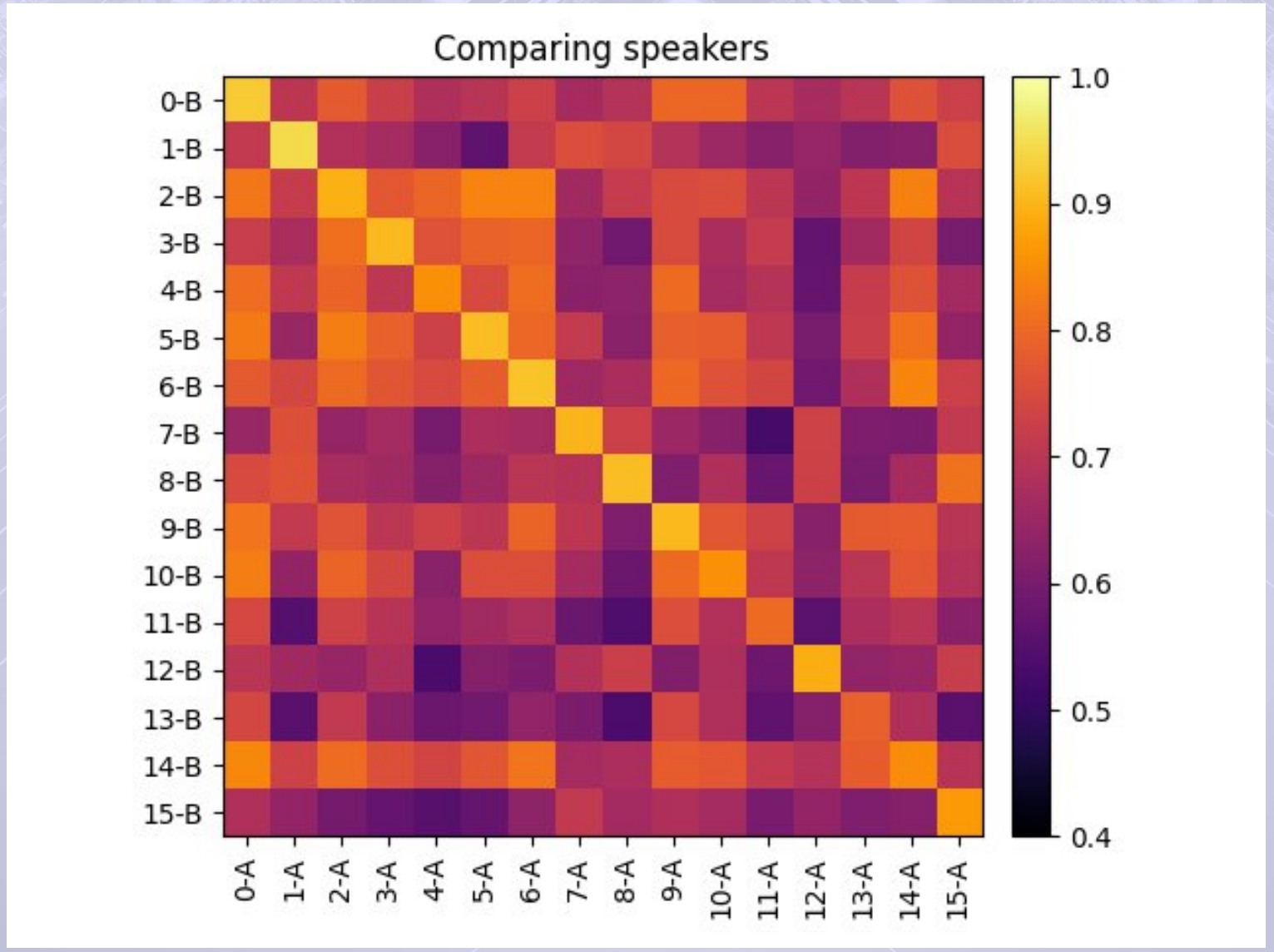
- Multiple ways to bypass a given implementation
  - Mimicking voices
  - Recording the victim without their knowledge
  - Splicing existing audio to form the correct phrase
  - Spoofing voices using AI

Bypassing voice biometric solutions

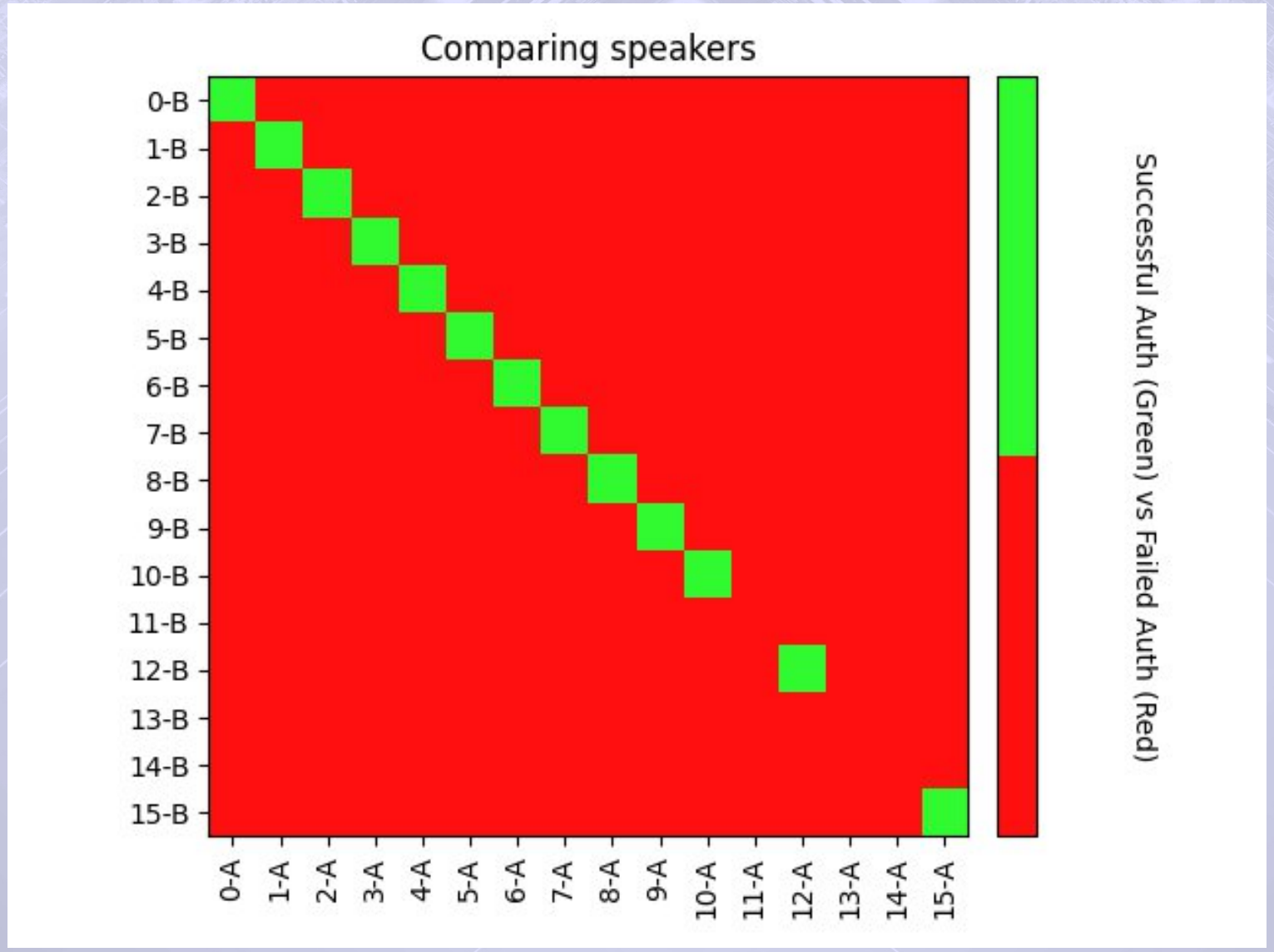
- Let’s look at some example data

- We sampled 15 people four times
- Each axis represents two audio clips from a person
- The numbers represent the speaker cross axis

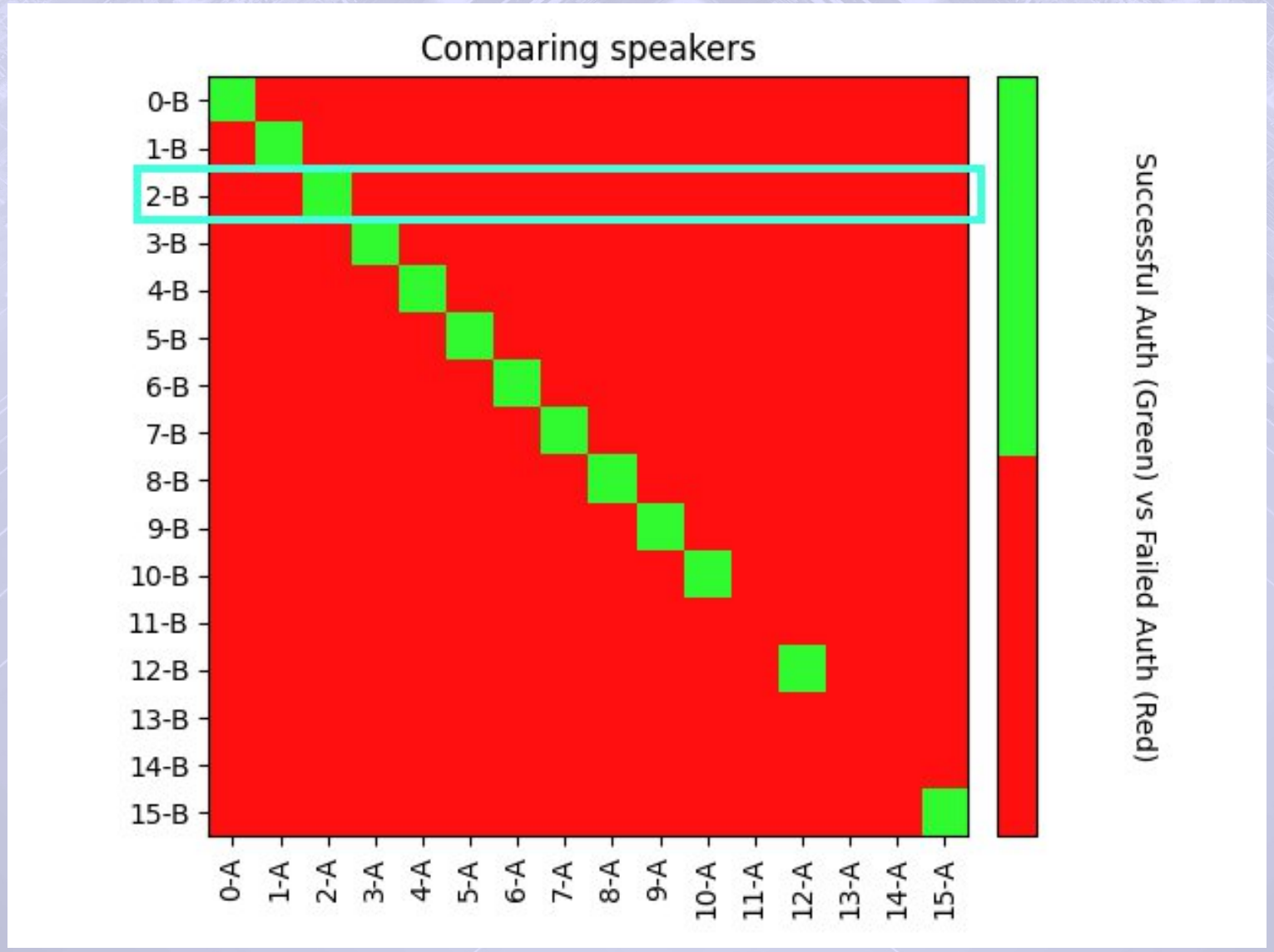
Bypassing voice biometric solutions



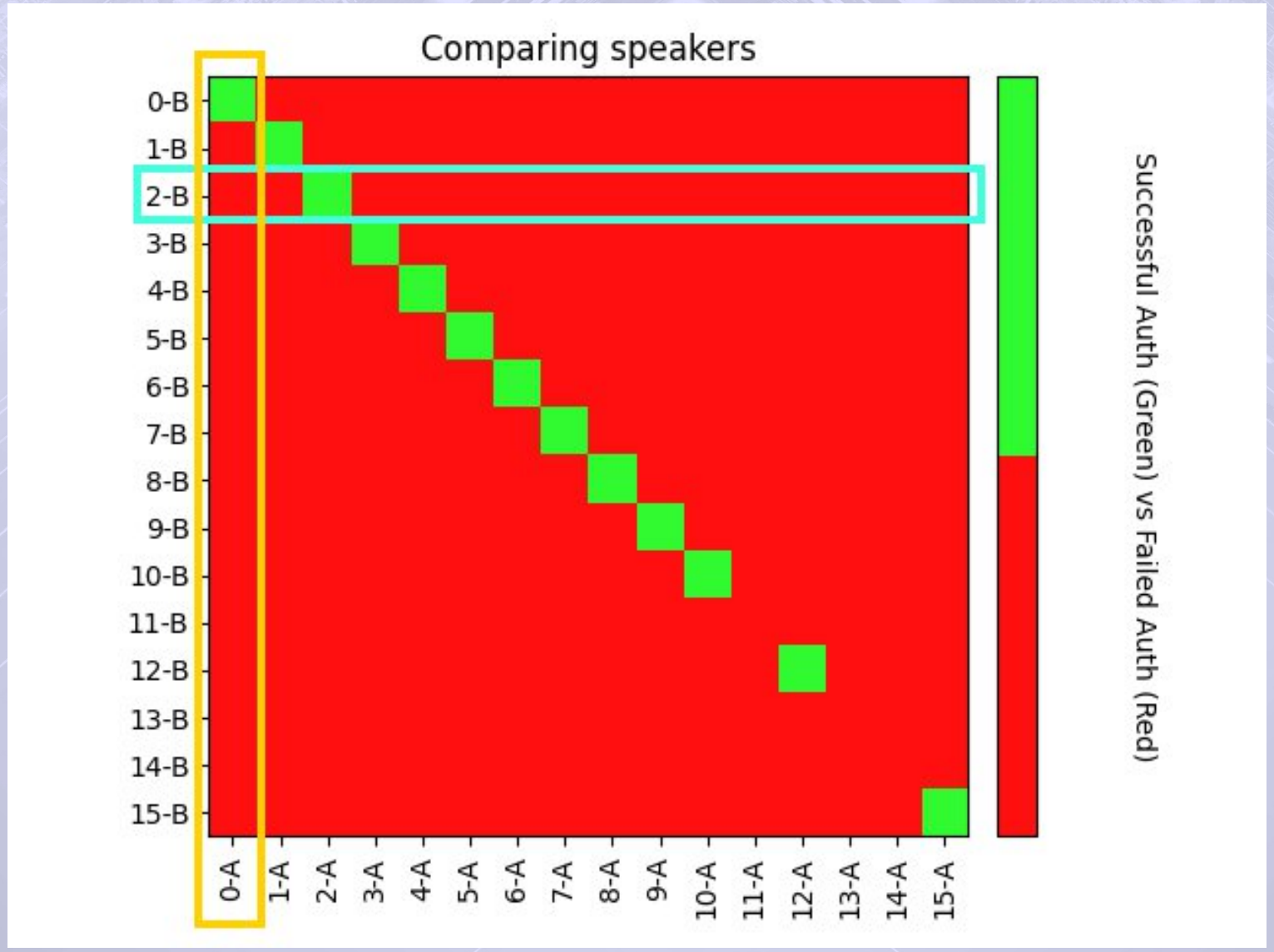
Bypassing voice biometric solutions



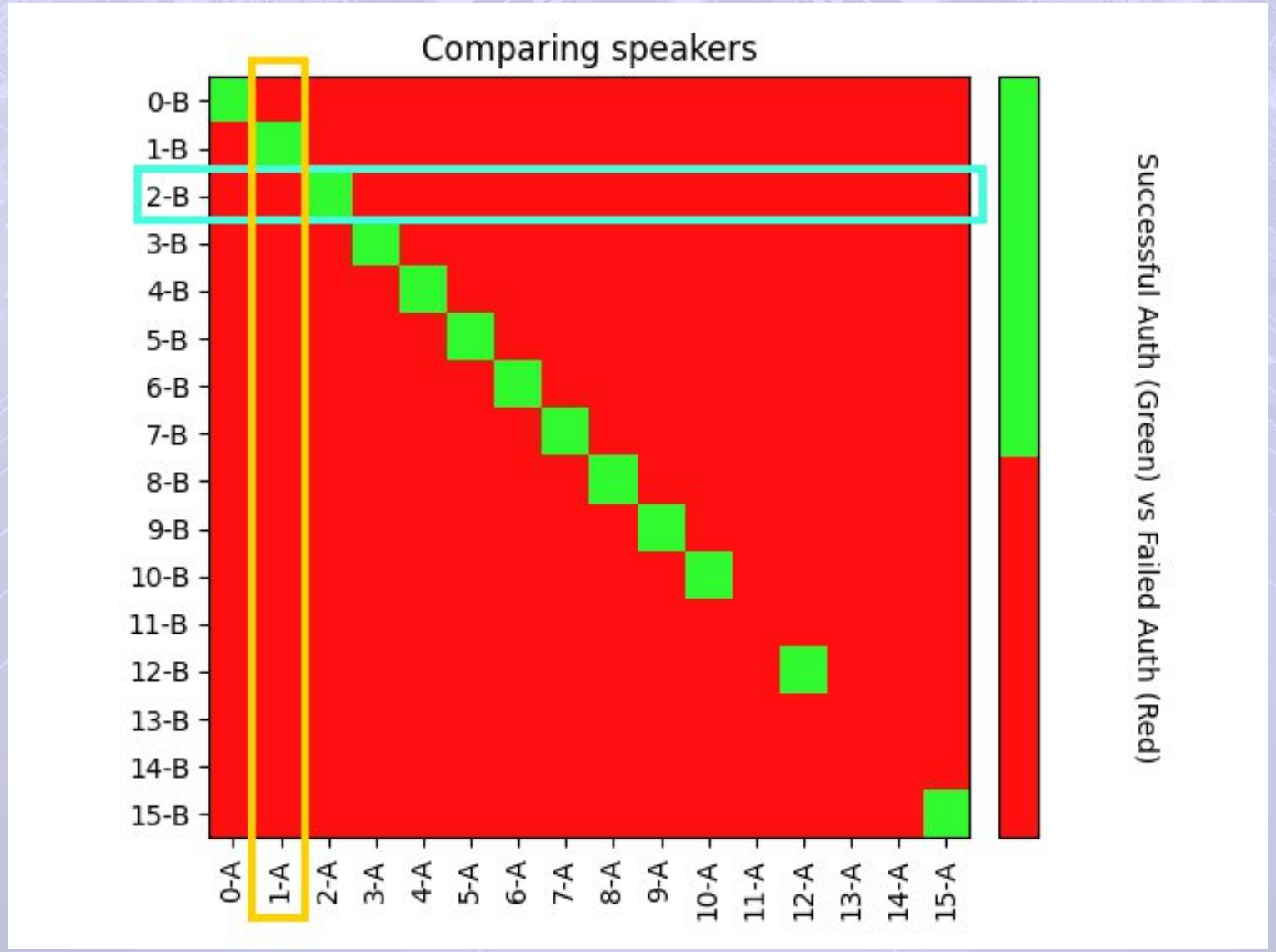
Bypassing voice biometric solutions



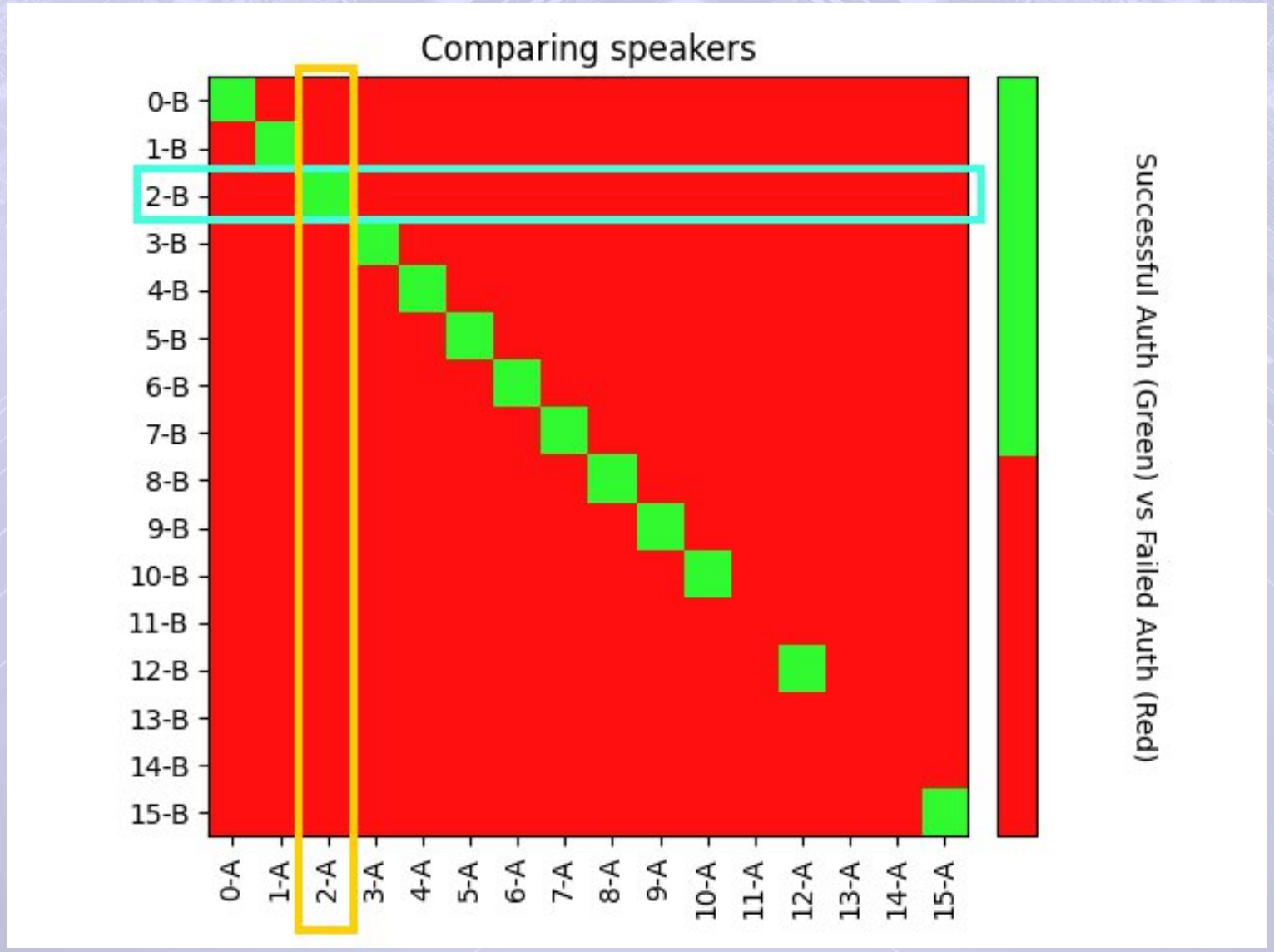
Bypassing voice biometric solutions



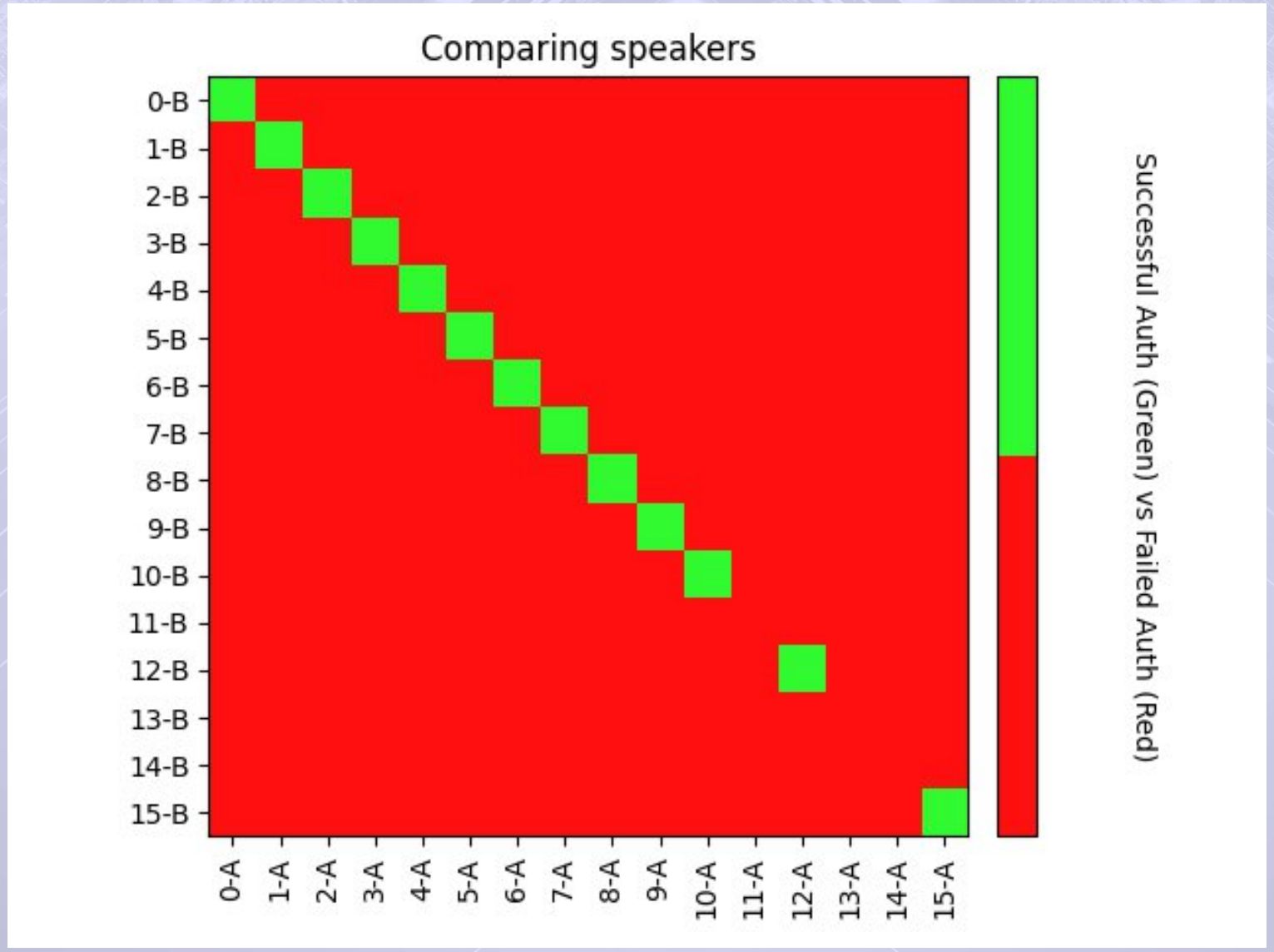
Bypassing voice biometric solutions



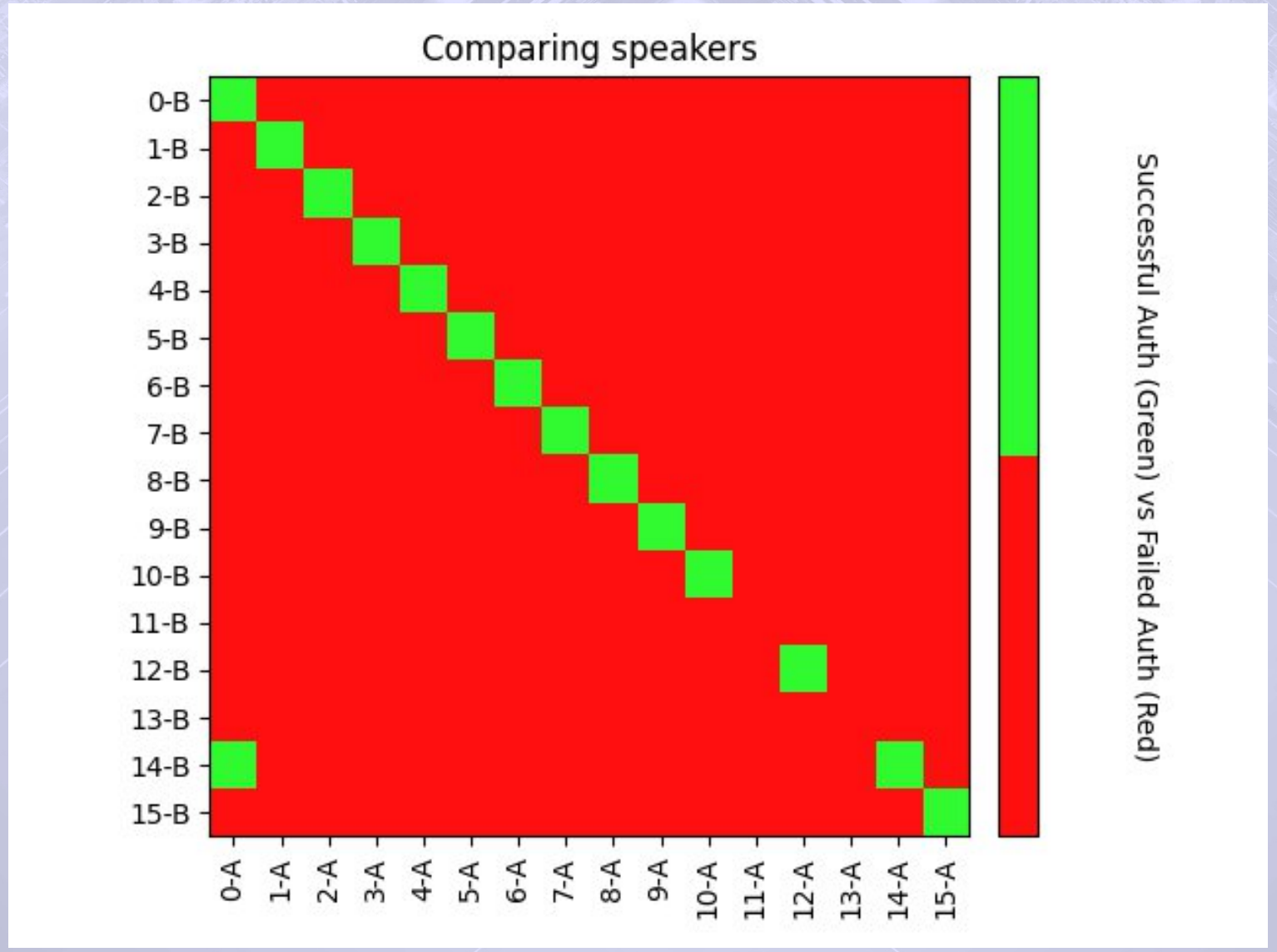
Bypassing voice biometric solutions



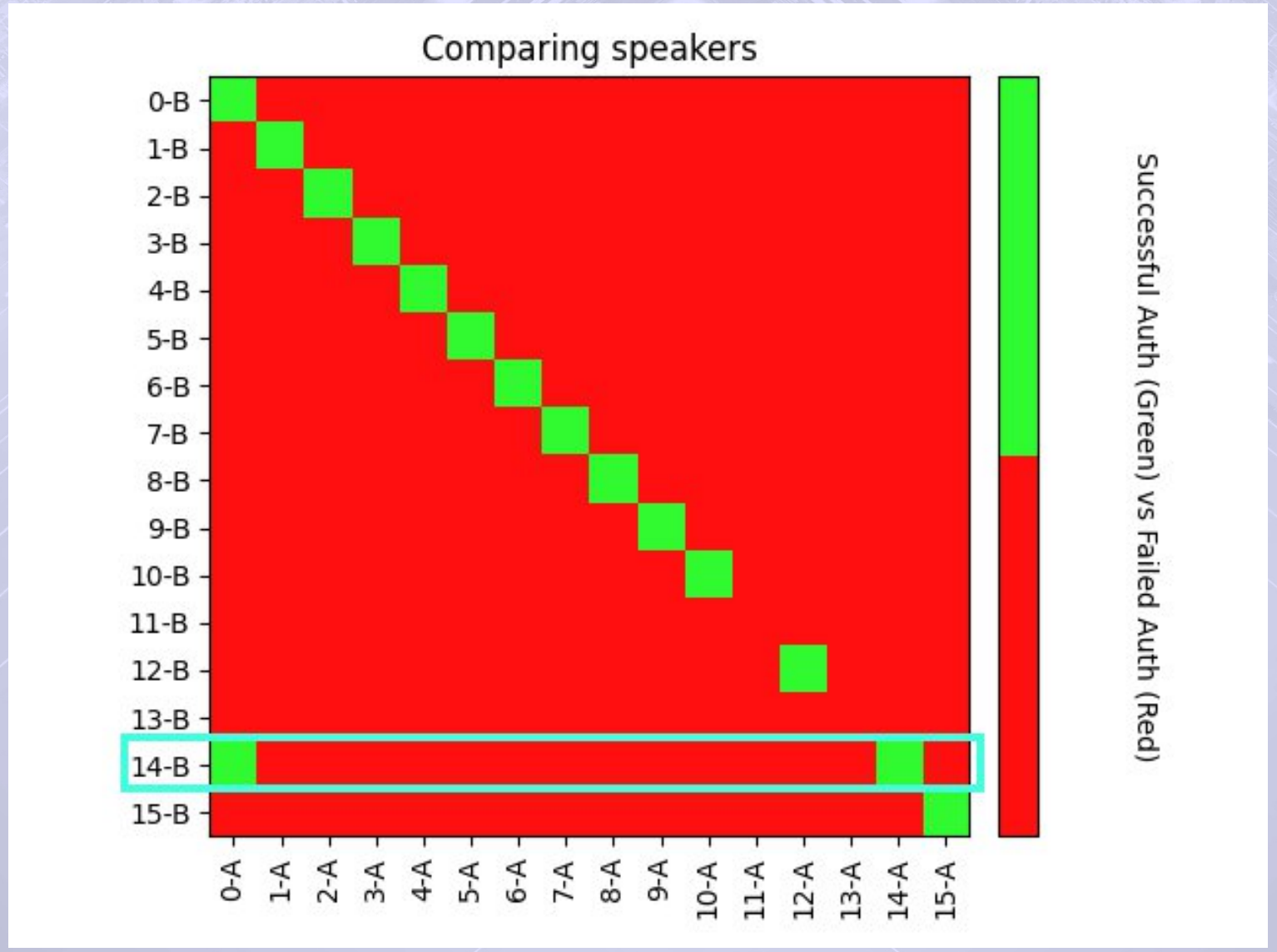
Bypassing voice biometric solutions



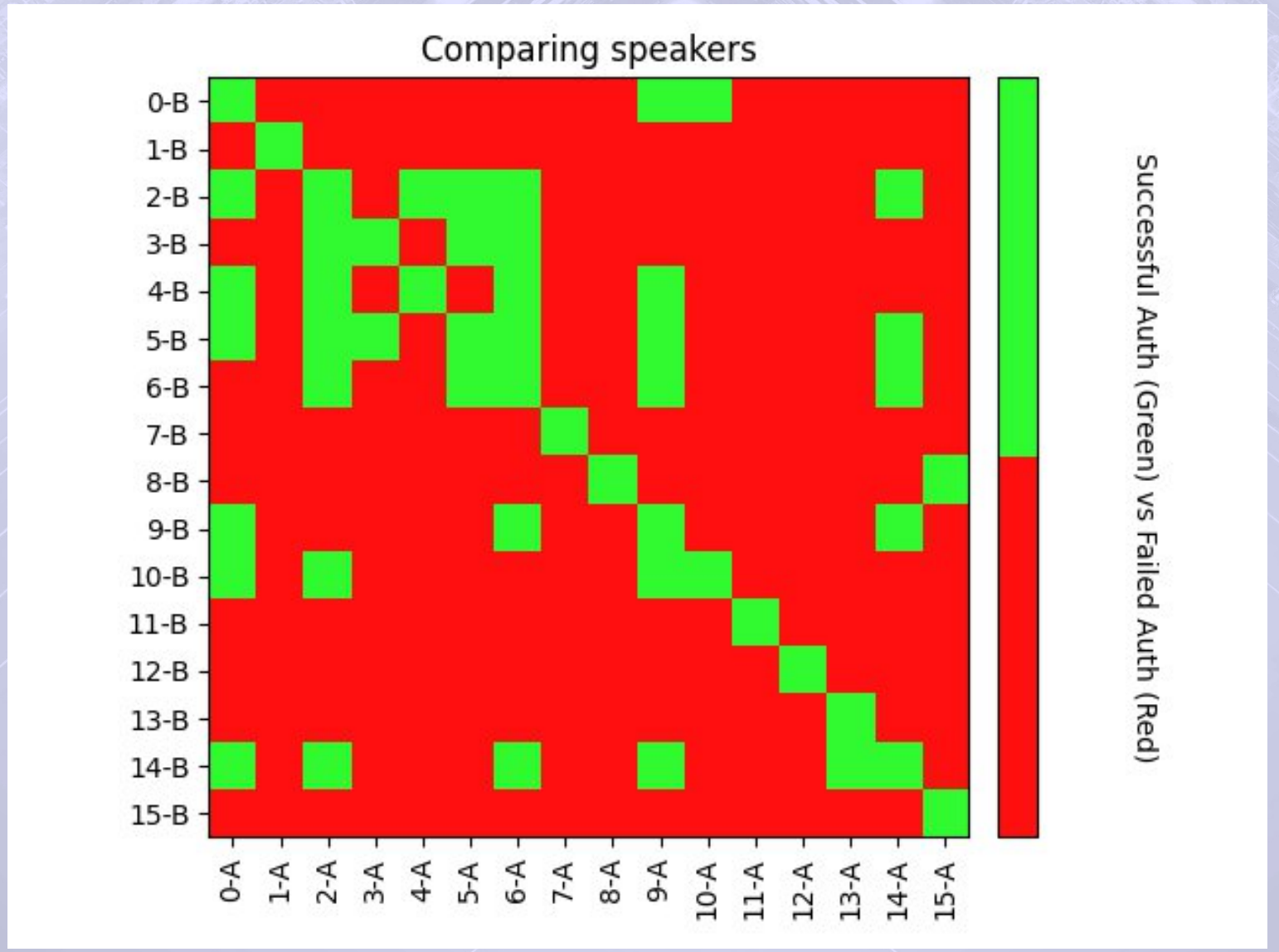
Bypassing voice biometric solutions



Bypassing voice biometric solutions



Bypassing voice biometric solutions

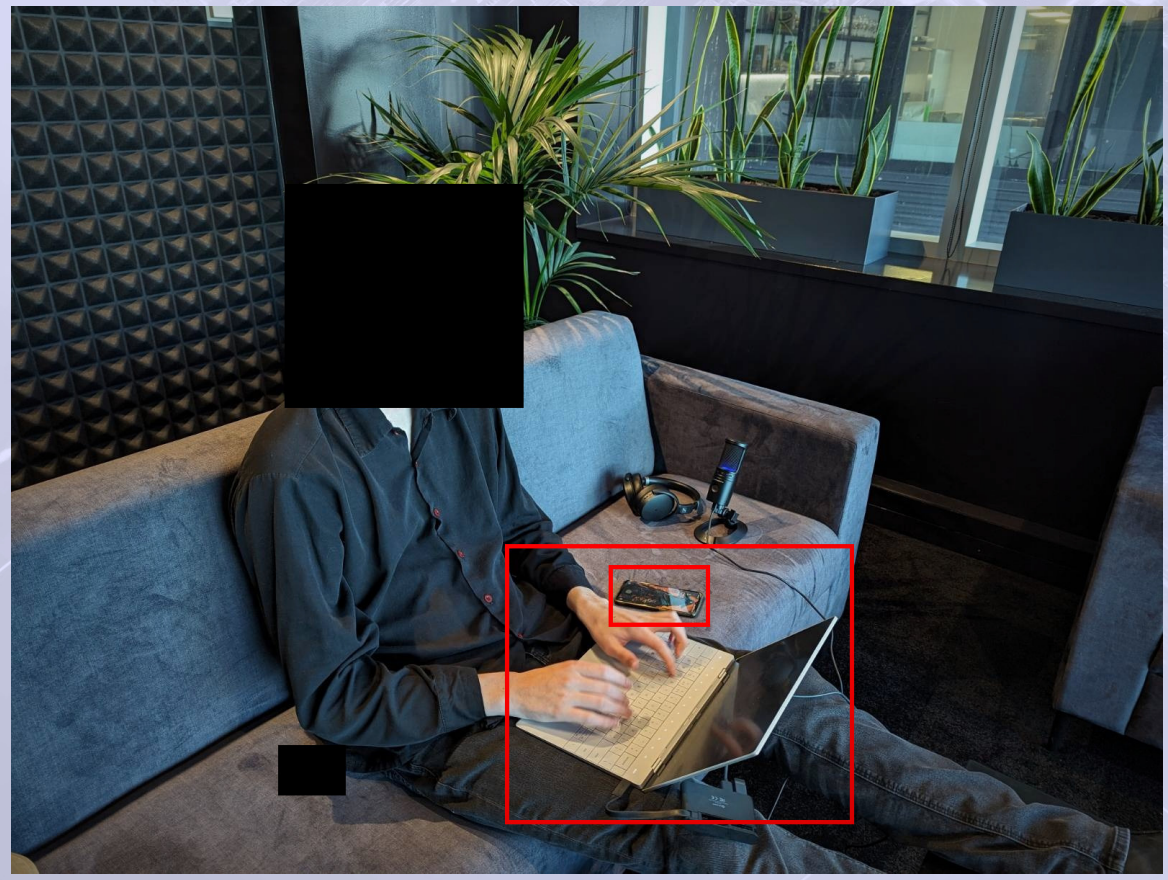


Sending audio from laptops to phones

- Highly complicated audio setups for bypasses

Sending audio from laptops to phones

- Highly complicated audio setups for bypasses



Bypassing voice biometric solutions

- So with that, what do we get?

Bypassing voice biometric solutions

- So with that, what do we get?

- We have a low barrier to entry

Bypassing voice biometric solutions

- So with that, what do we get?

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

Bypassing voice biometric solutions

- 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

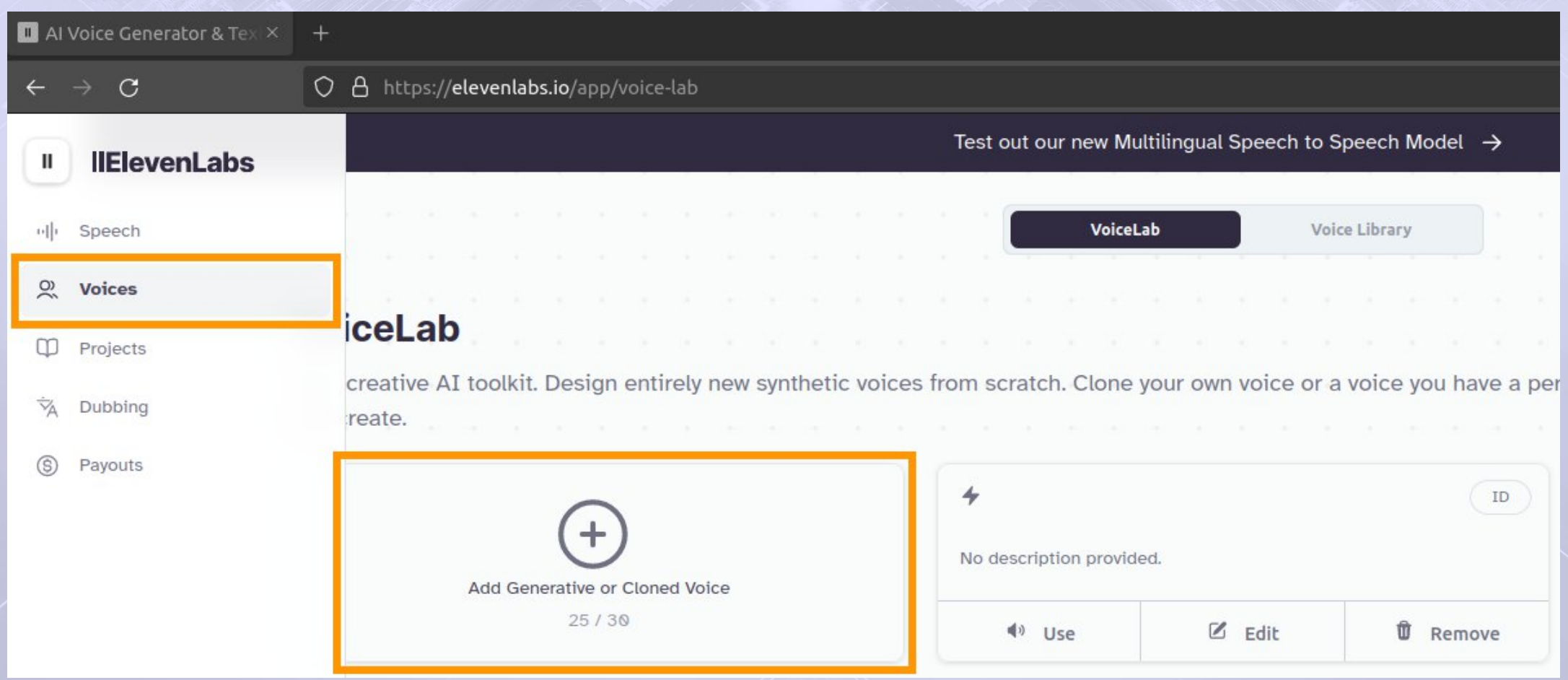
Cloning your own voice

- Let’s use our new training data to build a voice

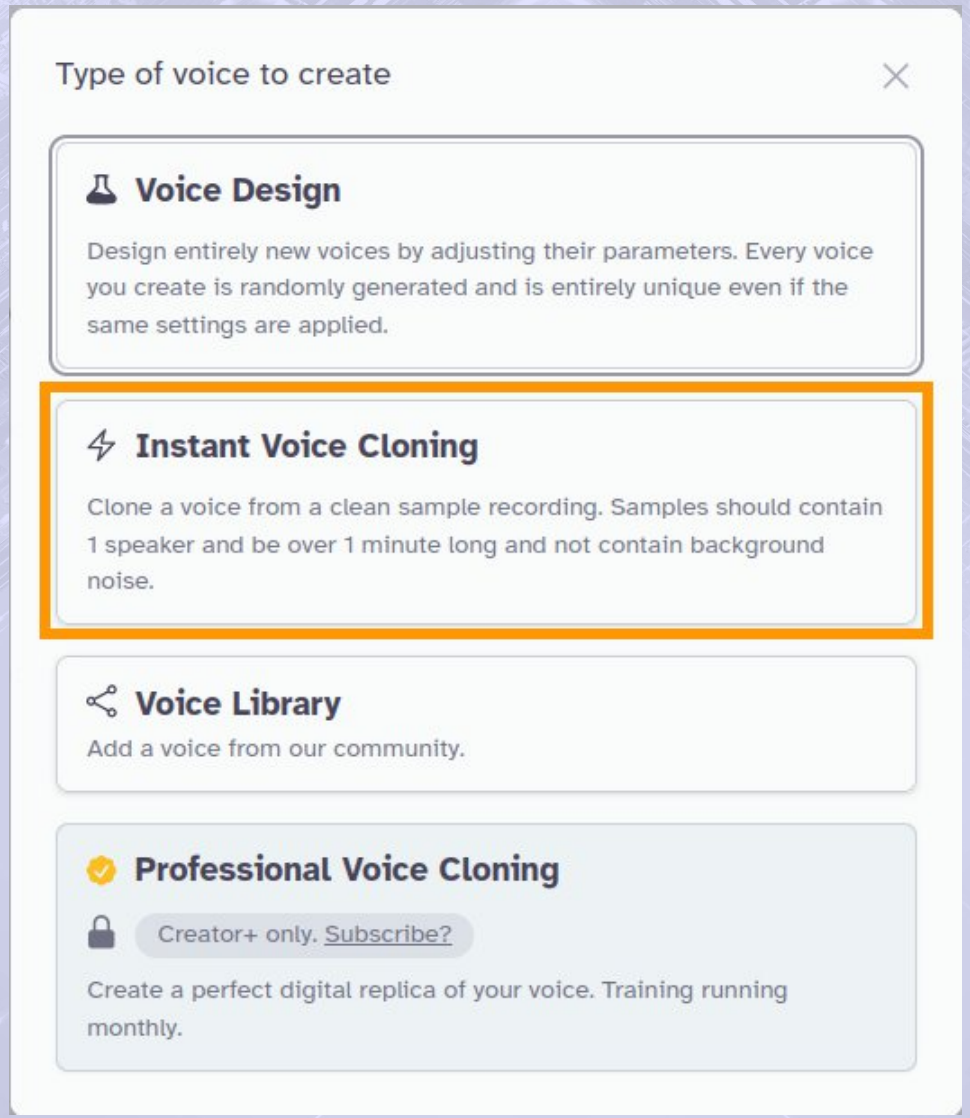
Cloning your own voice – ElevenLabs

- 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



Cloning your own voice - Creating a voice



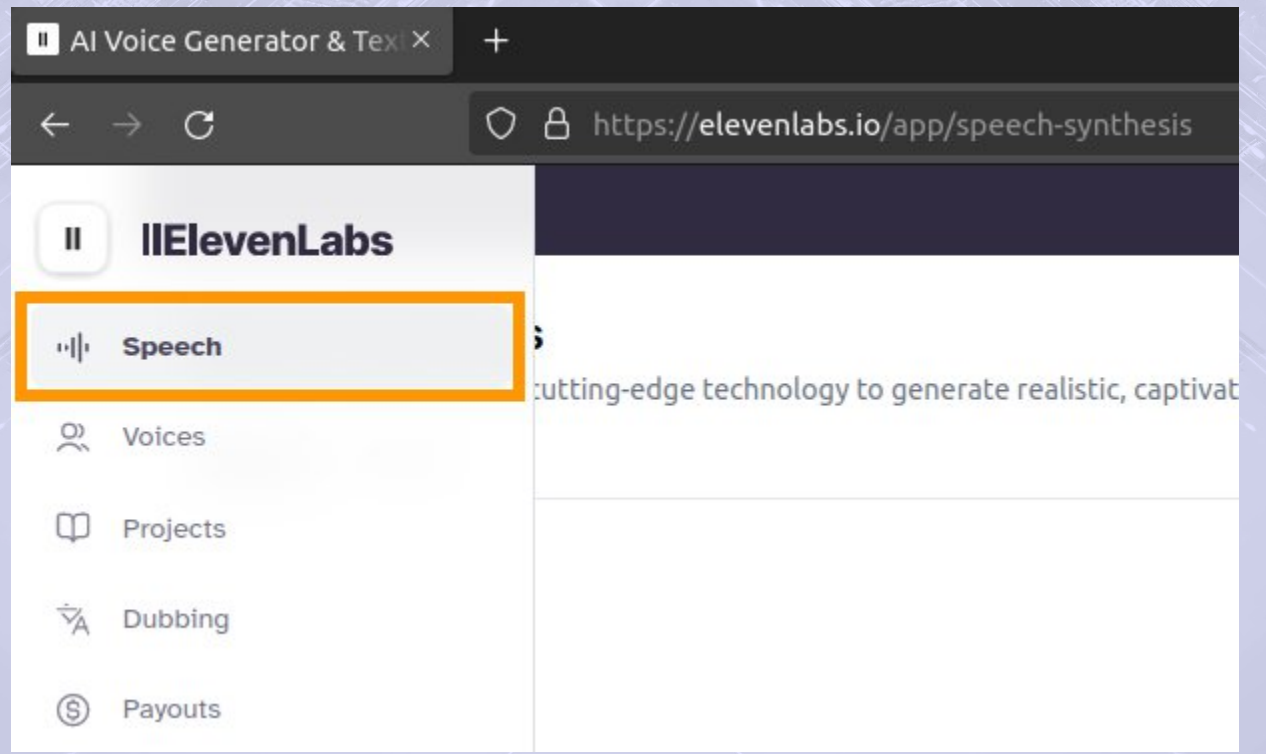
Cloning your own voice - Creating a voice

Cloning your own voice - Our new voice

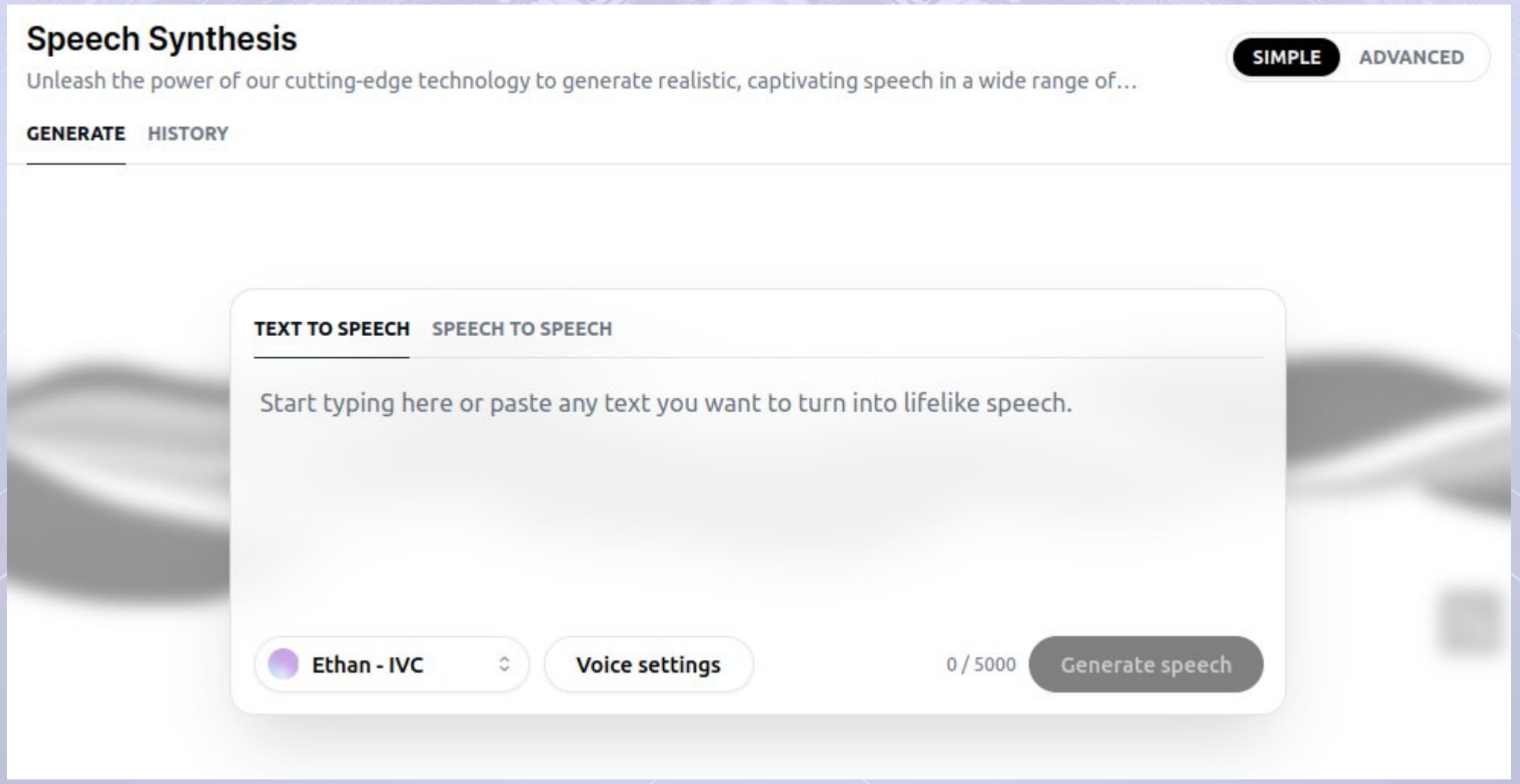
- We now have a voice, let’s use it

Cloning your own voice - Our new voice

- We now have a voice, let’s use it



Cloning your own voice - Our new voice

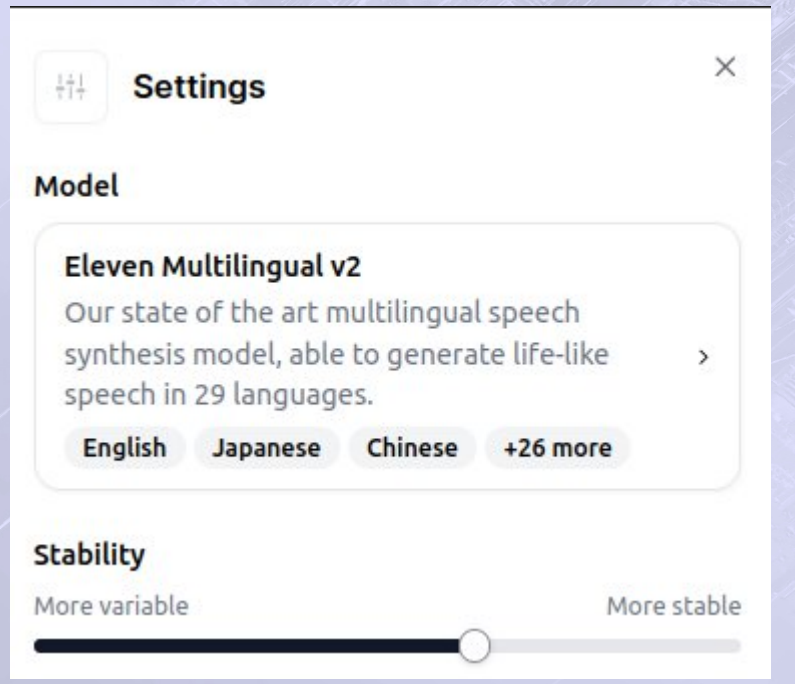


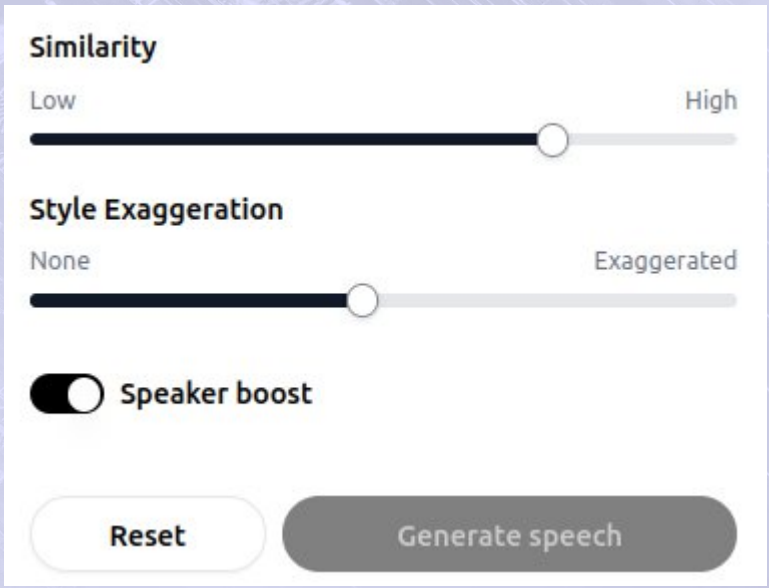
Cloning your own voice - Our new voice

- Let’s listen to some examples using our new voice

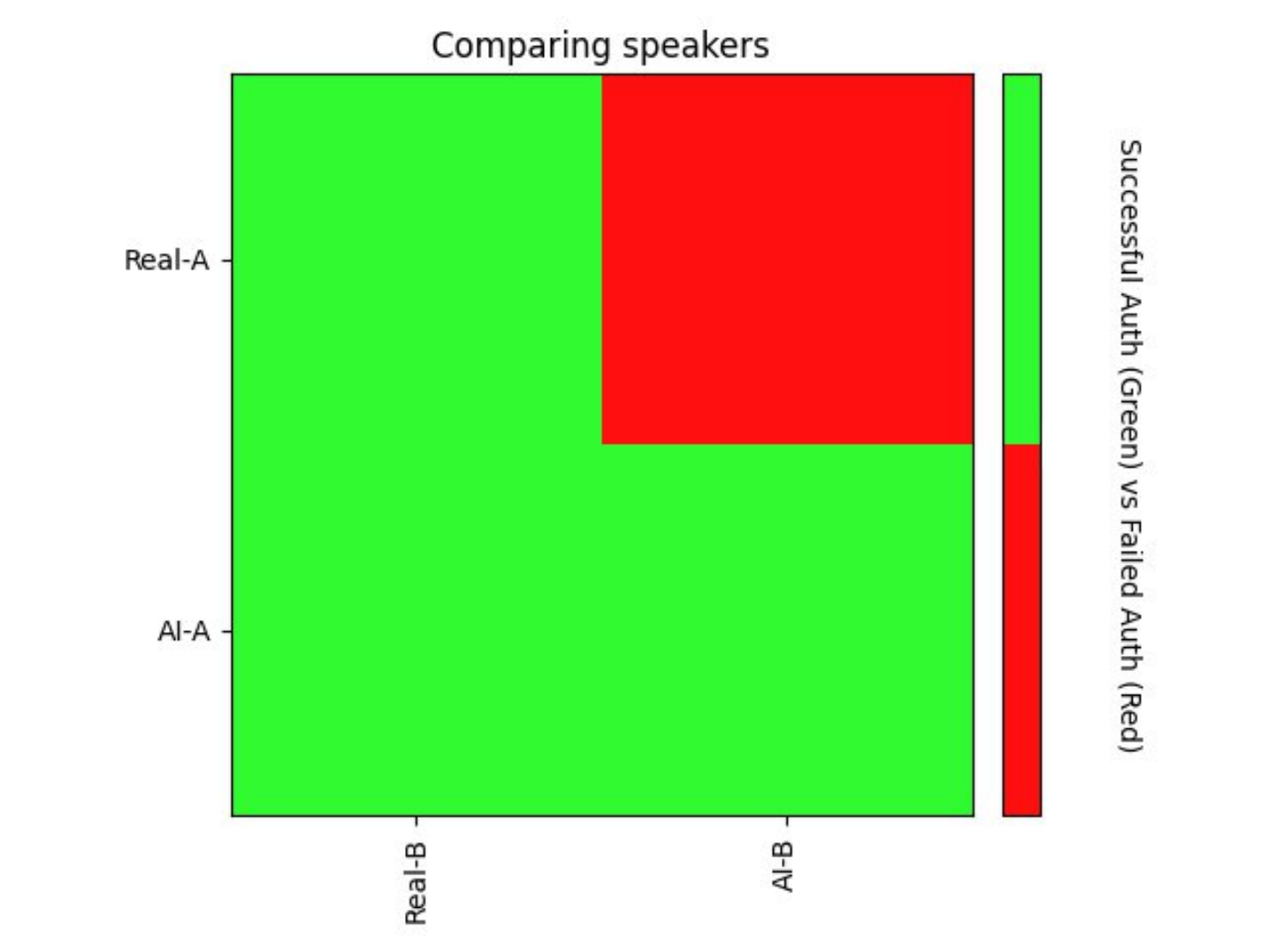
Cloning your own voice - Our new voice

- Let’s listen to some examples using our new voice





Cloning your own voice - Our new voice



Practical voice cloning with a victims digital presence

- So how do we use this for offensive purposes?

Practical voice cloning with a victims digital presence

- 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

Practical voice cloning with a victims digital presence

- 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

Practical voice cloning with a victims digital presence

- So what do we already know?

Practical voice cloning with a victims digital presence

- So what do we already know?
  - We know the victim’s name: “Simon Howard”

Practical voice cloning with a victims digital presence

- So what do we already know?
  - We know the victim’s name: “Simon Howard”
  - We know the victim’s company: “ZX Security”

Practical voice cloning with a victims digital presence

- 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

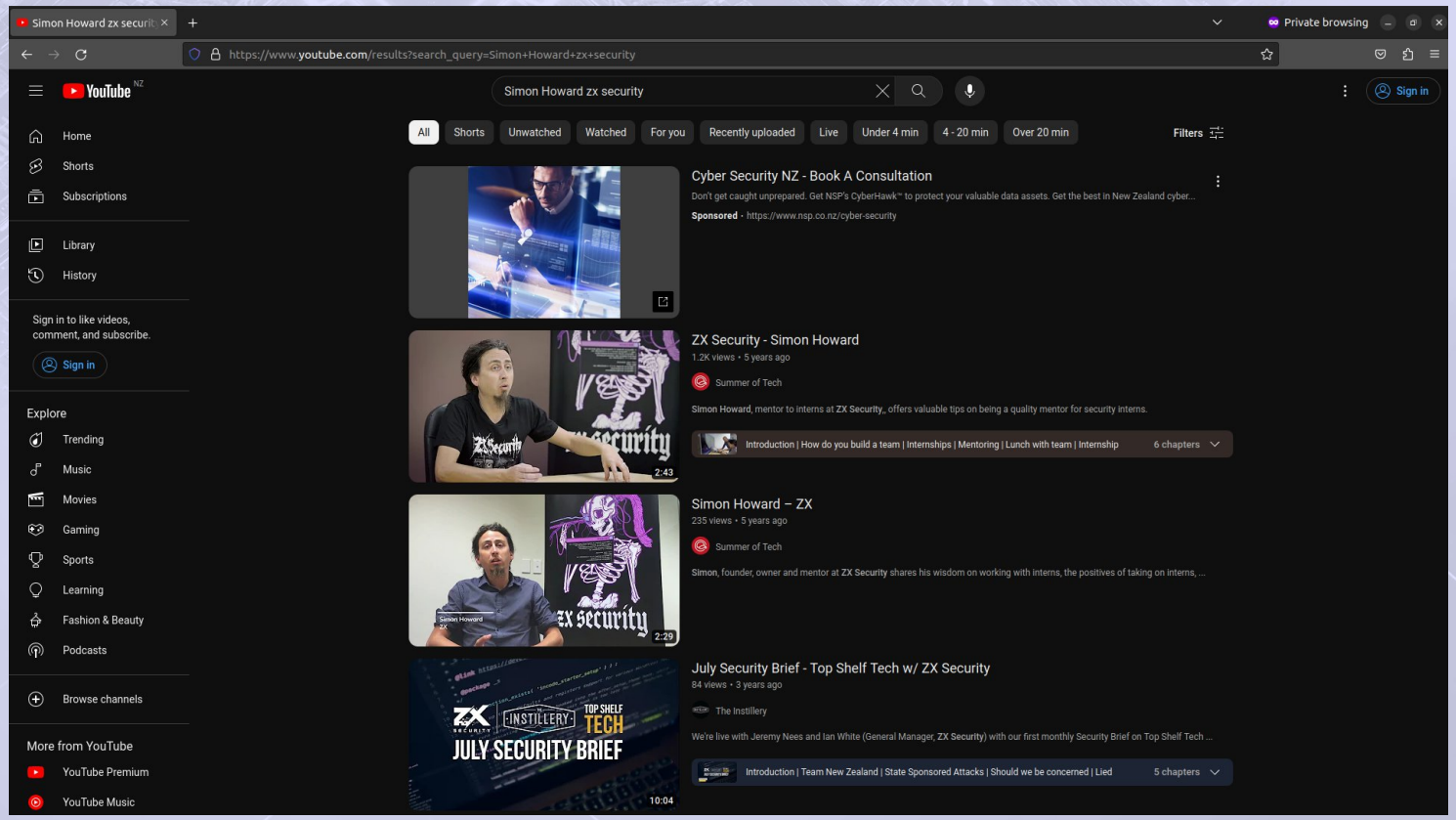
Practical voice cloning with a victims digital presence

- 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

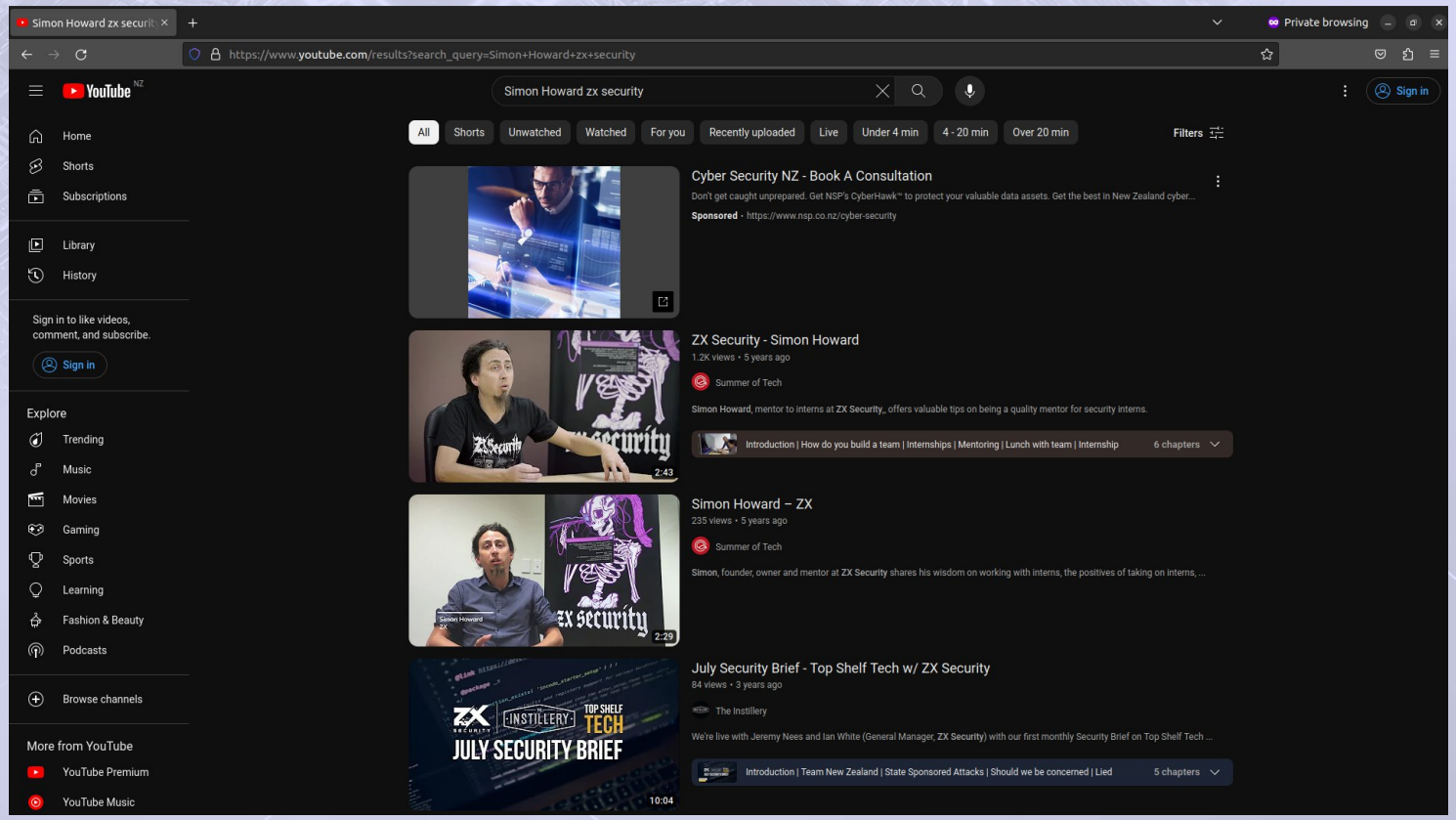
Practical voice cloning with a victims digital presence

- Looks like Simon has a digital presence we can use



Practical voice cloning with a victims digital presence

- Looks like Simon has a digital presence we can use
- Let’s download those videos and train a voice on them



Practical voice cloning with a victims digital presence

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

Practical voice cloning with a victims digital presence

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

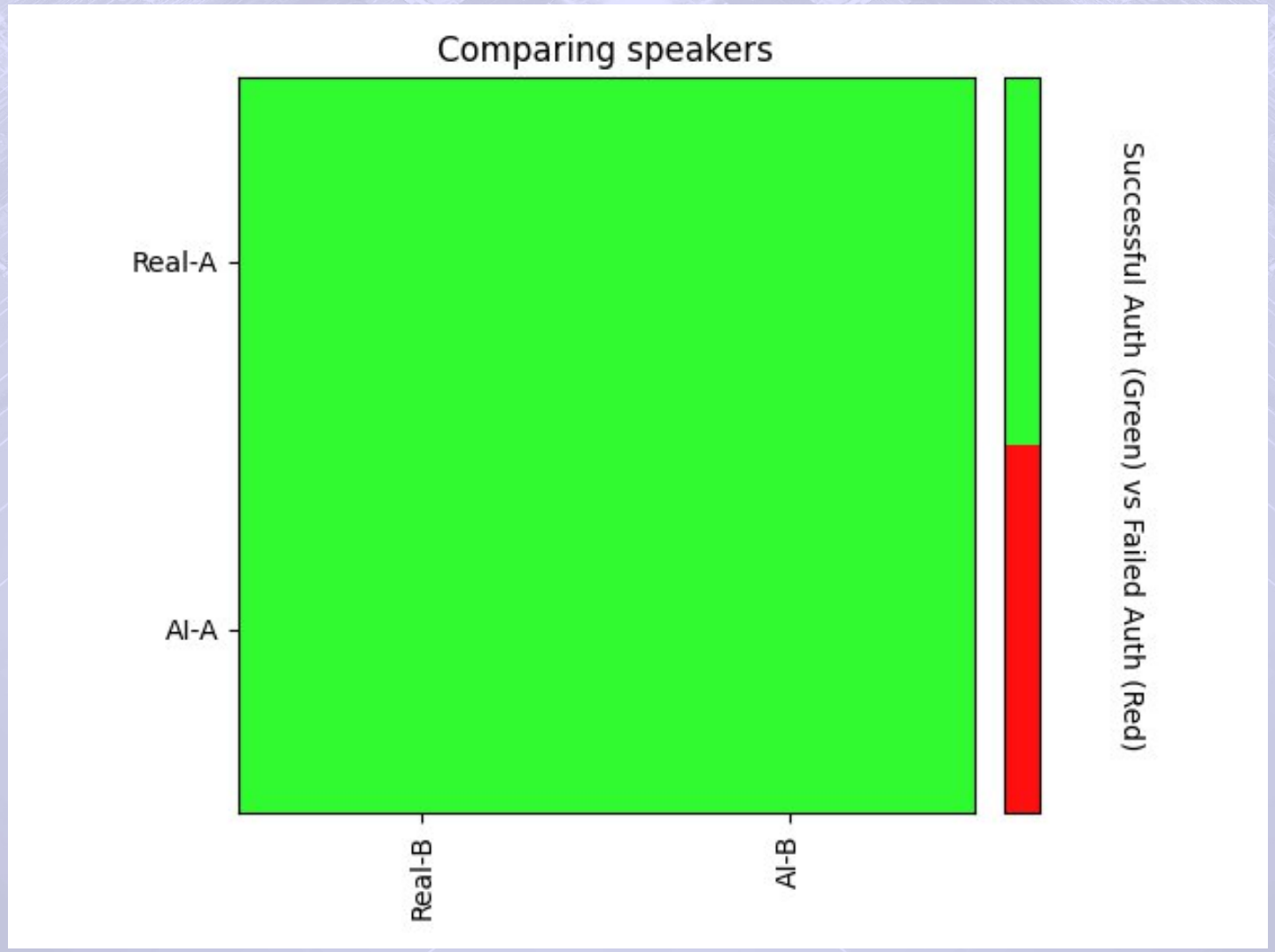
- And there we have it, a viable clone of Simon
  - It’s not perfect, but it can prove to be good enough for our avenues of abuse

Practical voice cloning with a victims digital presence

- 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

Practical voice cloning with a victims digital presence



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