DeepFakes: State of Affairs, State of Defense

KnowBe4Con, May 9th, 2019

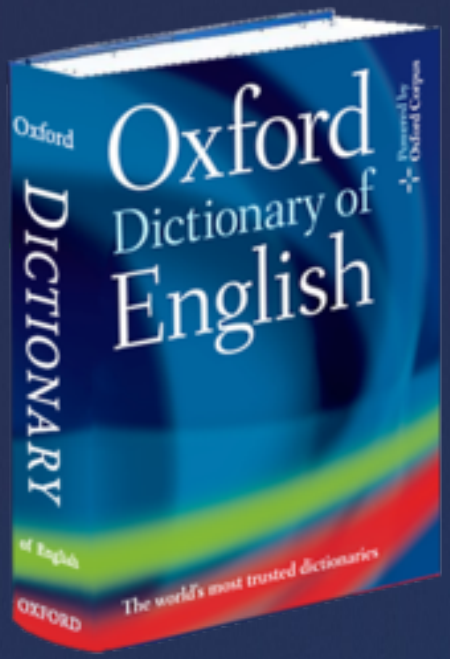
Dr. Lydia Kostopoulos

Lkcyber.com
@LKCYBER
linkedin.com/in/lydiak


1. Post Truth
2. DeepFakes
3. Tech Defense
4. Human Defense


1. Post Truth
2. DeepFakes
3. Tech Defense
4. Human Defense


DeepFakes are not happening in a *social* black hole.



post-truth

ADJECTIVE

Relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief.

Competing Understandings of Reality


FAKE
FACT


1. Post Truth
2. DeepFakes
3. Tech Defense
4. Human Defense

DeepFakes: An Evolving Term

* Loosely *

Fabricated “fake” media created through the use of Artificial Intelligence and/or *deep* learning methods.

DeepFakes: State of Affairs

Fake Photos

Voice Fakes

Video Alteration & Fabrication

Face Swapping

Authentic voice Lip Sync with complementary facial movement

DeepFakes: State of the Art

Fake Photo Generation

Source: thispersondoesnotexist.com

Produced by a GAN (generative adversarial network)
StyleGAN (Dec 2018) - Karras et al. and Nvidia
Original GAN (2014) - Goodfellow et al.
Don't panic. Learn about how it works.
Help me figure out what was learned by this AI here.
Check out text generation by another AI
Click for another person Link to image

Produced by a GAN (generative adversarial network)
StyleGAN (Dec 2018) - Karras et al. and Nvidia
Original GAN (2014) - Goodfellow et al.
Don't panic. Learn about how it works.
Help me figure out what was learned
Check out text generation by another AI
Click for another person Link to image

DeepFakes: State of the Art

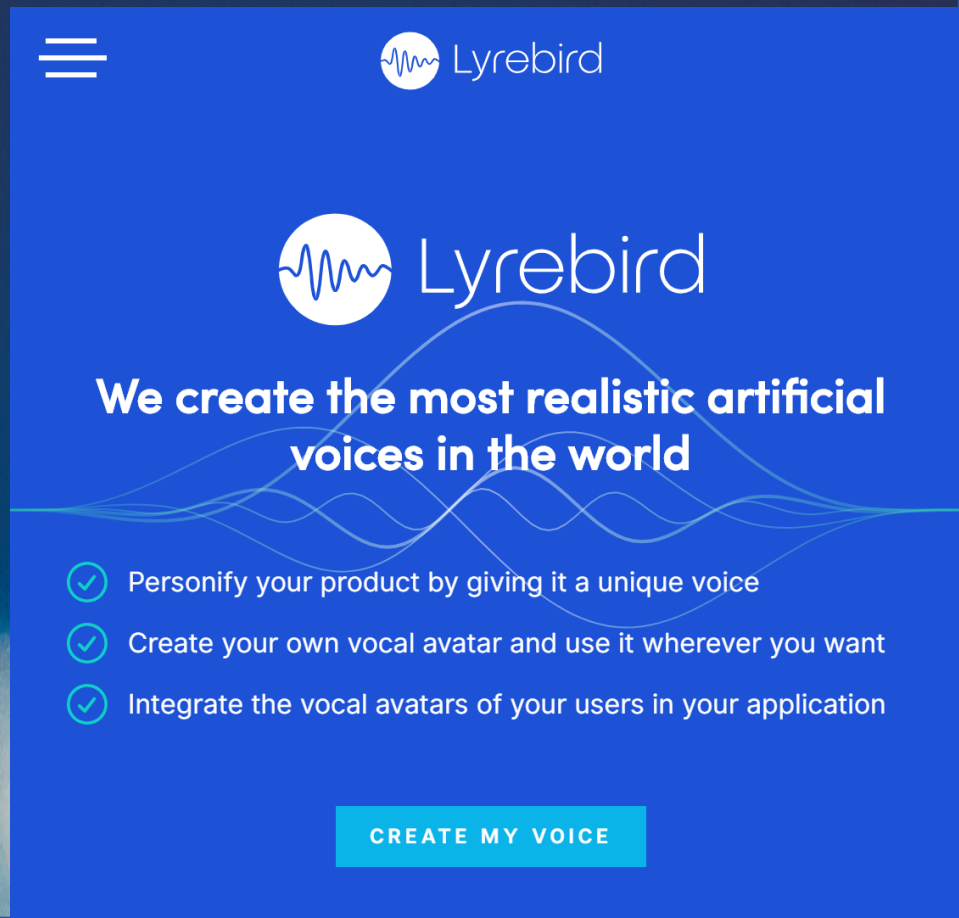
Voice Fakes: Authentic Voice Replication

#VoCo. Adobe MAX 2016 (Sneak Peeks) | Adobe Creative Cloud

1,613,073 views 9.2K 477 SHARE SAVE

Adobe Creative Cloud

Published on Nov 4, 2016

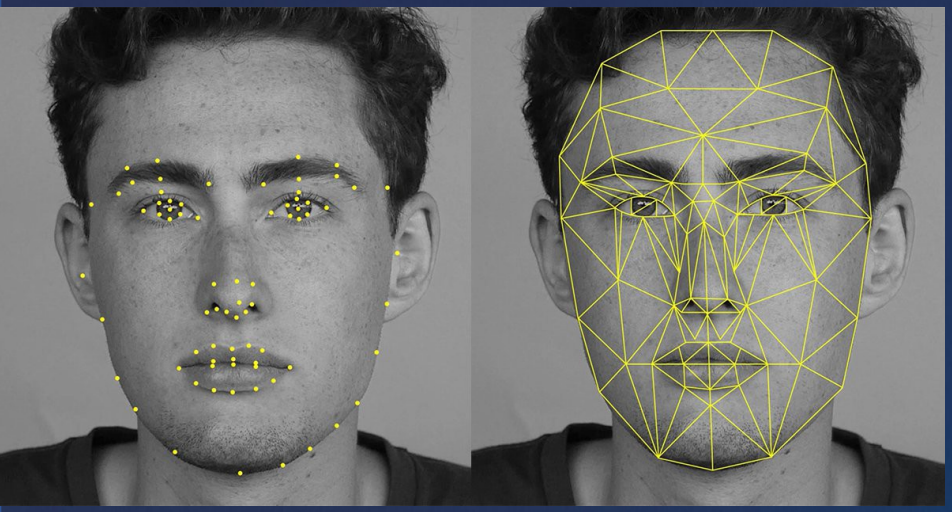


DeepFake Videos are <u>not</u>

Hollywood style Visual Effects (VFX)

DeepFakes: State of the Art

Video Alteration: Face Swapping



← Mapping the target actor's face.

Source: https://hackernoon.com/building-a-facial-recognition-pipeline-with-deep-learning-in-tensorflow-66e7645015b8 by Cole Murray

Arguably most popular:
Nicholas Cage DeepFake face swaps →

DeepFakes: State of the Art

Video Alteration: Face Swapping

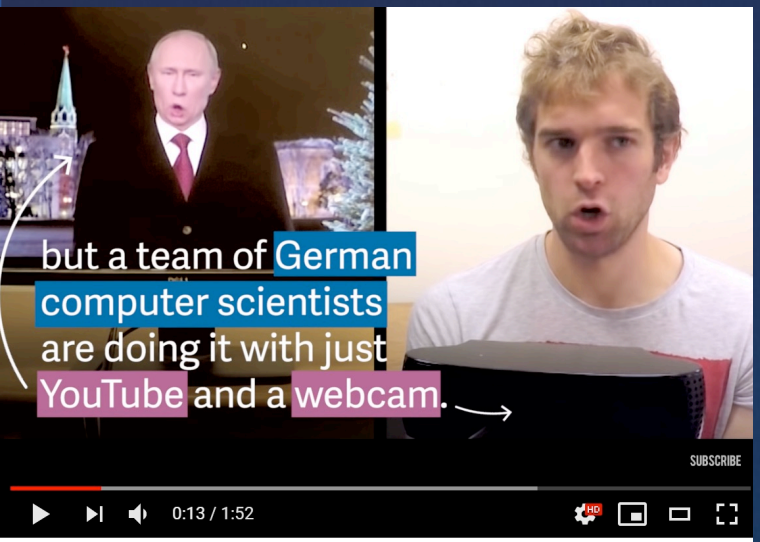
Steve Buscemi Face and Jennifer Lawrence Body & Voice

DeepFakes: State of the Art

Video Fakes

Deep Video Portrait (DVP)
Facial manipulation (not replacement), also referred to as facial reenactment and Video Puppetry.

DVP + Voice Fake
Authentic voice Lip Sync with complementary facial movement



Nothing is real: How German scientists control Putin's face

270,865 views 1.8K 103 SHARE SAVE

Quartz
Published on Apr 6, 2016

A research team has created software that allows them to control the face of anyone in any video on YouTube. The result is a weird cross between Snapchat's "face swap" feature and the "gibberish" scene from Bruce Almighty.

Synthesizing Obama: Learning Lip Sync from Audio

1,025,657 views 3.1K 131 SHARE SAVE

Supasorn Suwajanakorn
Published on Jul 11, 2017

Synthesizing Obama: Learning Lip Sync from Audio
Supasorn Suwajanakorn, Steven M. Seitz, Ira Kemelmacher-Shlizerman
SIGGRAPH 2017

DeepFakes: State of the Art

Video Fakes

Deep Video Portrait (DVP)

Faci
also
and

Nothing i
270,865 vi

Snapchat's "face swap" feature and the "gibberish" scene from Bruce Almighty.

DeepFakes: State of the Art

Video Fakes

Deep Video Portrait (DVP)
Facial manipulation (not replacement), also referred to as facial reenactment and Video Puppetry

DVP + Voice Fake
Authentic voice Lip Sync with complementary facial movement

What can we do about this from a security defense and awareness perspective?


1. Post Truth
2. DeepFakes
3. Tech Defense
4. Human Defense

Tech Defense

There is good news!

Technology is being developed to identify deep fakes!

Technology is a useful defense against technological manipulation.

SERELAY
TRUSTED MEDIA CAPTURE

Home Solutions Demo Use Cases

Capture verifiable photos and videos

Learn more

Truepic

#16 Fast Company's 50 Most Innovative Companies 2019

#1 in the Social Good Category

Photo and video verification you can trust

Truepic is the leading photo and video verification platform. Every day, we work diligently on technologies that can help restore trust in visual media.

Use Cases:

News & Media,
Insurance,
Ecommerce,
Field Construction Documentation,
Laboratory Results

DARPA is funding new tech that can identify manipulated videos and 'deepfakes'

Taylor Hatmaker @tayhatmaker / 12 months ago

Comment

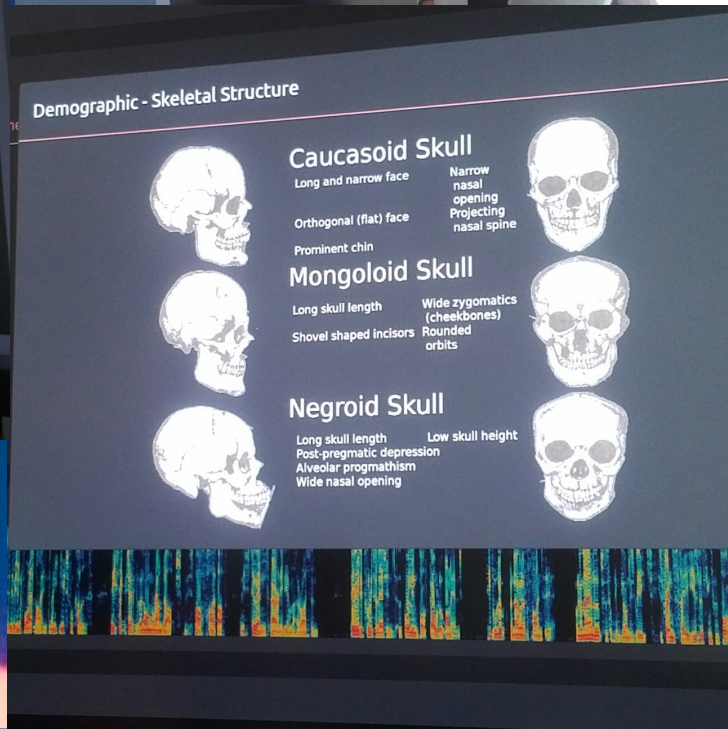
Dr. Rita Singh

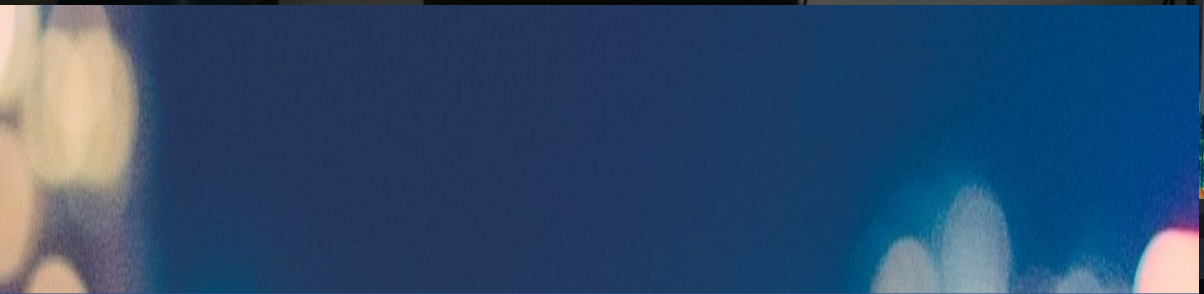
Carnegie Mellon University

State of the Art Voice Recognition Research


Voice Quality Features
Voice Parameters
Demographic
Environmental
Medical
Physical
Physiological
Psychological
Caucasoid
Skeletal Structure
Generated Face
Spectrogram
Pitch
175
Microsoft


Generated Face
Loudness

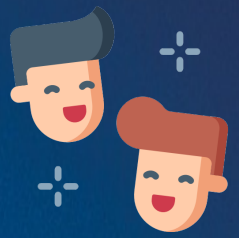

Demographic - Skeletal Structure
Caucasoid Skull
Long and narrow face
Orthogonal (flat) face
Prominent chin
Narrow nasal opening
Projecting nasal spine
Mongoloid Skull
Long skull length
Shovel shaped incisors
Wide zygomatics (cheekbones)
Rounded orbits
Negroid Skull
Long skull length
Post-pregmatic depression
Alveolar progmathism
Wide nasal opening
Low skull height




1. Post Truth
2. DeepFakes
3. Tech Defense
4. Human Defense

Human Defense

There is nothing inherently wrong with pushing the bounds of technology. However…



Human Defense

We need to have awareness in understanding what our ***'real'*** reality is.

Cybersecurity awareness should expand to DeepFake awareness.

Two Items for Security Awareness Effectiveness

1.

Persistent
& Compelling
Storytelling

Two Items for Security Awareness Effectiveness

2.

Building a culture of technology agility

Fun. Timely. Provocative.

Try it out!

Sapien2-0.com

Thank you!

Dr. Lydia Kostopoulos

Lkcyber.com

@LKCYBER

linkedin.com/in/lydiak