SWCP Portal
September 2021

# Fakes in the 21st Century

Lies and deception have been going on as long as humans have roamed this planet, and since our closest living primate relatives seem to have a gift for it, too, maybe much longer. While lying is a universal strategy to gain advantages amid the complicated mazes of social living, even **chimpanzees** seem to realize that deception is wrong as it hurts others.

That hasn't stopped it from happening throughout human history. Traces exist in every age and place, and deception is enshrined in myths and legends. People still talk of the **Trojan Horse**, which fatally fooled the ancient Trojans into a false sense of security. Remains of the **great wooden horse** used by the wily Greeks may have been recently found – unless it was faked long ago to scam ancient tourists.

Deceptive documents came early, as those who knew enough to create real ones could also hoax them. So up to half of all **monastic charters** in the Middle Ages were faked to give them big tax breaks.

Some scams have surprisingly long pedigrees. Appeals for help from "**Nigerian princes**" began as pleas from "Spanish prisoners" in the French Revolution. Fake news can have long lasting evil effects, too, as in the infamous **Protocols of the Elders of Zion.**

Originally written as a mean **satire** by unknown Russian anti-semites around 1900, it was cleaned up and presented as the leaked secret plans of a vast international Jewish conspiracy to take over the world. Used by the Nazis as a justification for the Holocaust, its poison is still being spread online to this day.

Voices are also relatively easy to fake, especially when flaws in recordings can help disguise them. From the start, however, photography embodied the ideal of objective reality – that a snapshot showed just what was actually there at that moment.

## From photos to photoshopping

However, even that is a bit of an oversimplification. Photos only ever show just one small bit of the world from a particular viewpoint at a single instant in time, so choices must be made. Plus early technology required long exposures, so pictures were usually carefully composed and set up in advance.

The first photographic fakes made with darkroom tricks were those of **ghosts and apparitions**, made popular by spiritualism after the Civil War. Later, movies with special effects showed that anything was possible on film, though it took over a century to go from **Georges Méliès**' charming fantasies to Tom Hanks appearing with JFK in ***Forrest Gump***.

However, as any filmgoer who has sat through the closing credits of a sci-fi epic knows, concocting convincing wonders takes a great deal of time, effort, and money. So most fraudulent film efforts are quick and cheap knock-offs using common editing techniques. Voices can be slightly slowed to sound drunk and actions sped up to look more aggressive, color changed to make people seem sick or enraged, etc.

Early visual propaganda was often very heavy-handed. During France's surrender in 1940, looped newsreel footage of the German dictator excitedly stamping his foot made it look like **Hitler danced** a jig. Under Stalin, Russian censors crudely airbrushed commissars fallen out of favor from official group photos. Their practices inspired the "**memory holes**" of Orwell's prescient dystopian novel ***1984***, down which papers holding inconvenient history would be sent for incineration, still a potent metaphor today.

The reason why is that the digital revolution and the internet made hoaxing so much easier. Once paper pages, audio recordings, and filmed images were replaced by digital media, it all just became data, pixels that could be manipulated even utterly transformed without any permanent basis to reference left like a film negative or a handwritten manuscript.

Old practices became new concerns when given a boost with **Adobe's Photoshop**, the leading graphics program. The name became a verb back when debates raged about the ethics of enhancing models to perfection or darkening O.J.'s mugshot. But the same tools that are used to make fakes can often aid in detecting the subtle differences between pixels from the original image and later alterations.

**One way** to tell if a picture was photoshopped is to look for what can't be easily changed. Shadows, light direction, and perspective should be consistent across the whole image. In small images, distortions and pixelation are more obvious, but highly-intent artists often overlook glaringly obvious mistakes, such as missing parts or leaving too many limbs.

Users can also try looking for other versions by reverse searching images on Google. Recently, more

help has been offered online with free photo analyzing websites such as **FotoForensics**, to look at images directly uploaded or from their web addresses.

Adobe, perhaps worried about liability, has been **doing research** on image manipulation detection, too, specifically on the tools most often used to edit faces and expressions. Though Adobe claims their algorithm has a *99%* success rate compared to *53%* by humans, no public product has yet been offered.

## Fakery goes deep

Perhaps Adobe hasn't sold its detector because it relies on **Artificial Intelligence** (AI) that could be possibly used to make hoaxes even better. For it is AI that provides the next stage of evolution of hoaxing, which is **already worrying** many smart people.

These are called **deepfakes**, that is faked media of people, things, or events using a type of AI called "**deep learning**". This is a method for machines to train themselves by exposing large neural networks to huge numbers of samples, and letting the machines figure out the commonalities or differences.

To speed the process, deepfaking relies on an algorithm called an **autoencoder** which has two parts working together almost as in a game. One of them reduces the data to features shared by both source and target while the other part of the algorithm attempts to recover the originals from the reduced dataset. This allows an AI system to gradually learn the shared features of both, and then, by swapping the media source and target, create a deepfake.

The resulting media file could be used for all kinds of mischief, and naturally porn leads the way. So far, mixing faces of female celebrities with porn stars in action has been far and away the main use.

Most other deepfakes have been spoofs, replacing actors in movie scenes. But a good many have been made by concerned groups to show the power of the technology. **Obama** has mouthed words written by Jordan Peele, Facebook's **Zuckerberg** has bragged about controlling billions of peoples' data, and Richard Nixon even read a speech actually written in case the first **moon landing** ended in disaster.

This has many more uses than just putting words in politician's mouths – like having folks bust out **expert dance moves**, for instance. The technology is developing at a very fast pace, too, making it easier and cheaper to do all the time. People wanting to play with it can freely download the **DeepFaceLab** software for replacing faces or heads, making them look younger or manipulating their lips. There's even a free smartphone app called **Zao** that can place a person's face in a movie scene from just one selfie.

Deepfaked audio has already been used for crime. One UK-based energy company CEO was fooled by a **faked voice message** from the firm's German owners to send *$243,000* to a fraudster.

But deepfake technology can even generate synthetic characters indistinguishable from real people, as is vividly demonstrated at **This Person Does Not Exist**, a gallery of diverse high-resolution images. There are already a number of synthetic characters working in **social media**, acting not just as influencers but as propagandists of various persuasions.

Synthetic actors, unlike the living kind, take direction and are not temperamental, talking back or demanding special treatment. And they are already here: the aging Swedish supergroup ABBA just announced new concerts featuring their own younger **digital avatars**, and it is not the first such **show**.

But the studio most likely to develop artificial actors is **Disney**, since they already have. In their Star Wars universe, one of the many valuable franchises the mouse empire maintains, they've already used several **digital main characters** to replace actors too old or even dead. More will doubtless come soon.

The race is on between deepfakers and institutions fearful of what could happen. Corporations like **Facebook** claim great success in their efforts, but they give few details. Every time a new way of detection is **announced** – like the lack of blinking in fakes – the technology rapidly adapts. The adversarial game-like way deep learning works makes it easier.

Some fear that deepfakes might be used to blackmail important people. However, the opposite could occur. Celebrities may shrug off embarrassing photographic evidence as fakes, just as Prince Andrew has to **distance himself** from the Epstein affair.

Producing good fakes is still not easy. One reason that celebrities and politicians are most used is that it usually takes a huge quantity of samples to capture all the needed nuances. Though they are getting easier, faster and cheaper to do, the political hoaxes many have feared haven't happened yet but it's even **possible** that the net already may be full of AI fakery.

To be effective, a deepfake should inspire action without thinking, like **phishing** does. Therefore, the most dangerous would be those involving publicly-known people in shocking situations during a crisis. A fake government announcement during a national emergency, for instance, could spread confusion.

Healthy skepticism is good, but never forget the basic question that since ancient Rome remains the surest way to uncover frauds: *Who benefits from it?*