# FALL 2019 E-NEWSLETTER

At Digital Mountain we assist our clients with their computer forensics, e-discovery, and cybersecurity needs. For this E-Newsletter, we provide an overview of deepfake technology and the impact on the legal industry. We also explore the regulatory climate and new developments.

## The Spooky Underpinnings of Deepfakes

There's nothing spookier at Halloween than faces hidden behind masks – even benign, cartoon animal masks can still stir a little anxiety because of the uncertainty of who might be hiding underneath the fake identity. And how often do we see news videos of actual villains committing crimes while donning masks? Those costumed trick or treaters and masked bandits have a technological twin: deepfakes. Deepfakes, the common name for synthesized content created via Artificial Intelligence (AI), are real news right now. Indeed, many academics, politicians, and everyday internet users are concerned about the potential dangers of deepfakes, all the while fascinated by the rapidly increasing quality of the technology. In this article, we'll give you the basic facts of deepfakes.

**Deepfake: What's in a Name?**

"Deepfake" is actually an amalgamation of "deep" from the machine learning term "deep learning" and the word "fake," which accurately describes the content, and is attributed to a user on the forum gathering site Reddit, which subsequently banned "r/deepfakes" for violating its site standards policy (https://www.reddit.com/r/DeFranco/comments/7vy2lh/rdeepfakes_have_-been_banned/_). Despite the clever, admittedly novel moniker, the deepfake name is actually very helpful in learning about what a deepfake is.

A deepfake is a video or still image which has been altered from the original to replace the identity of the authentic subject with another identity, with or without an audio track. As deepfake technology has improved, three categories of deepfakes have emerged:

**(1) Face Swaps** – a new facial image replaces the facial image of the original image subject. Face swapping apps such as Snapchat, Cupace, and Face Swap Live allow users to easily perform Face Swaps and post them to social media. These basic

deepfakes are the easiest and most accessible form, and can be used for still images, recorded video, and live stream.

**(2) Lip Sync** – an existing video is altered to make the subject appear to speak new content. The difference between a face swap and a Lip Sync is that in a Face Swap, the "new" face is pasted over the original face like a translucent mask. Any facial expressions are those of the original subject. In a Lip Sync, the "new" face is moving, and thus, viewers see that person's facial movements.

There is a now famous Lip Sync deepfake of former President Obama warning about the dangers of deepfakes (https://youtu.be/cQ54GDm1eL0 – be advised the video contains some profanity). The deepfake is revealed thirty-six seconds into the video, and the impersonator and the deepfake are shown in split screen.

**(3) Puppet-master** – This is the most complex deepfake and is perhaps easiest imagined as the name implies. In a Puppet-master deepfake, an entirely new video is created by recording the movements of an "actor" and overlaying that image with another subject. For example, the creator of a Puppet-master deepfake could produce a video of his late grandmother making an apple pie by overlaying images (some systems require just a single image) of the grandmother over a celebrity chef's pie making how-to video. Puppet-master videos can capture both the images and the voices of the target subject provided there is a sufficient sample of each. Audio synthesis of the subject's genuine speaking voice isn't necessary to create a deepfake, but having a data set which includes it, increases the complexity options.

**The Technology is Real – the Product is Fake**

The goal of deep learning is to create an artificial neural network that, like the human brain, can teach itself. For deepfakes, the specific artificial neural network at play is called a Generative Adversarial Network, or GAN. Like deepfakes, understanding GANs starts with examining the name:

**G**enerative – the goal is to create, or generate, something;
**A**dversarial – at least two elements working against each other; and,
**N**etwork – in a unit tied together with cohesive instructions.

A GAN works by teaching one part of the network ("A") a specific goal – for deepfakes, the goal is to create the most realistic image possible. Component A is taught the goal by uploading authentic images as the desired result. A second component of the GAN ("B") creates images and submits them to A for comparison against the learned goal. Component A judges and replies with feedback. The more images fed to A during the training, the better the feedback. While this is a simplified explanation of how a GAN works, it nonetheless demonstrates the basic technology.

**A Short History of Rapid Development**

In 1997, researchers published a paper announcing the Video Rewrite Program which created a system by which new facial expressions and video tracks could be created to mimic work previously produced by movie studios (think Tom Hanks meeting former President Kennedy in *Forrest Gump*). By refining tracking and mimicking techniques, subsequent researchers and programmers developed the various programs that have helped make deepfakes easier to create and harder to detect.

In 2017, after a period of slow, quiet development, a poster on the internet forum collective Reddit began posting deepfakes under the name “r/deepfakes.” r/deepfakes claimed to be a researcher who was simply interested in the GAN technology and tried his hand at it by creating face-swapped celebrity pornography (https://www.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn). By the time Reddit shut down the account, r/deepfakes work had over 90,000 interactions, highlighting the viral nature with which posting on a single site can attract attention. To date, there are nearly 100 Reddit forums and countless pages on other social media platforms devoted to deepfake technology and products.

**Free for All**

Two important aspects of deepfake technology that make it attractive to creators are (1) the computer hardware required to produce a deepfake can be built for under $2,000, and many mid-range, out-of-the-box gaming computers will suffice, and (2) deepfake creation software programs can be downloaded without cost. Two of the most popular deepfake creation programs are FakeApp and DeepFaceLab, both of which are easily downloaded.

The biggest investment in creating high quality deepfakes is the time required in training the GAN. Images must be uploaded by the user – the more the better. Viewers can often tell a poor quality deepfake simply because the facial image has a “mask” appearance to it or the expressions have a stiff or unnatural quality, both of which are signs of shallow training of the GAN. If the user is looking to create a deepfake of a popular celebrity or politician, there are pre-loaded image data sets available for various deepfake applications.

**Credit Where Credit is Due**

Tracing the creation of a deepfake is nearly impossible. By the time a deepfake is posted to an internet site, the original source material may have undergone multiple manipulations by multiple people, and at its core, the technology is blending at least two data sets of material. Some technology enthusiasts have proposed both hardware and software tools that could apply “watermarks” for authentic, original content which could then be recorded on a blockchain. Each time that content is edited, a new watermark would then be recorded as a separate event on the blockchain record and would allow digital forensics investigators to trace the evolution of the deepfake in question. While not currently feasible for all devices and software applications, this technology is currently being developed for police body cameras (https://www.reuters.com/article/us-axon-deepfakes/axon-boosts-encryption-weighs-blockchain-to-tackle-body-cam-deepfakes-idUSKBN1WI0YG).

Deepfake technology is one of the latest examples of how closely and how quickly deep learning is coming to approximating the skills and plasticity of our human brains. It’s also a demonstration of how accessible AI technologies can be for the general public. No longer beyond the reach of the average person, we can now create our own studio quality videos right on our desks, or should we say desktops? Either way, with deepfakes, we experience the tricks and treats of masked identity every day.

**Please direct questions and inquiries about cybersecurity, computer forensics and electronic discovery to info@digitalmountain.com.**

## UPCOMING INDUSTRY EVENTS

**OPENTEXT ENFUSE 2019**
Las Vegas, NV: November 11-14, 2019

**TODAY'S GENERAL COUNSEL, "THE EXCHANGE" EDISCOVERY**
San Francisco, CA: November 19, 2019

**INFOSECURITY NORTH AMERICA AND ISACA**
New York, NY: November 20-21, 2019

**GEORGETOWN LAW'S 2019 ADVANCED EDISCOVERY INSTITUTE**
Washington, DC: November 21-22, 2019

**TODAY'S GENERAL COUNSEL, "THE EXCHANGE" EDISCOVERY**
Los Angeles, CA: December 11, 2019

***Click here to see more upcoming events and links***

*Digital Mountain, Inc. Founder and CEO, Julie Lewis, will be presenting at various upcoming industry events. Please send requests for speaker or panel participation for her to marketing@digitalmountain.com.*


## DIGITAL MOUNTAIN, INC.

4633 Old Ironsides Drive, Suite 401
Santa Clara, CA 95054
866.DIG.DOCS

**www.digitalmountain.com**


***Contact us today!***

***FOLLOW US AT:***