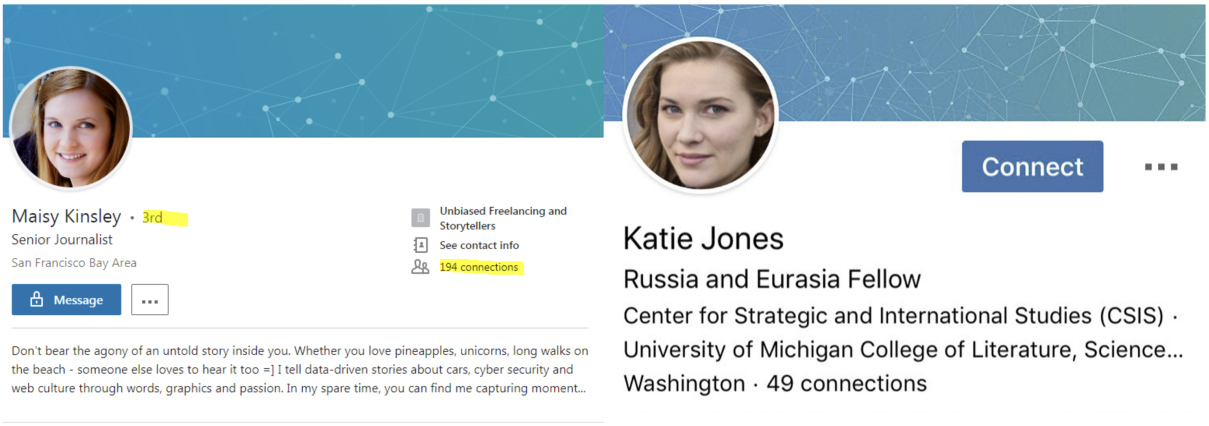
The Rise of Deepfakes

Recently, Deepfakes, one of the most recent developments in data manipulation, have been increasing in popularity, exponentially. Though a relatively novel technology and concept, Deepfakes have begun to challenge AI neural networks in recognizing fake or manipulated data.

What are Deepfakes?

"Deep fake (also spelled deepfake) is a type of artificial intelligence used to create convincing images, audio and video hoaxes. The term, which describes both the technology and the resulting bogus content, is a portmanteau of deep learning and fake." - Ivy Wigmore, content editor for TechTarget's IT Encyclopedia



Both accounts have been removed from LinkedIn after their discovery as Deepfakes.

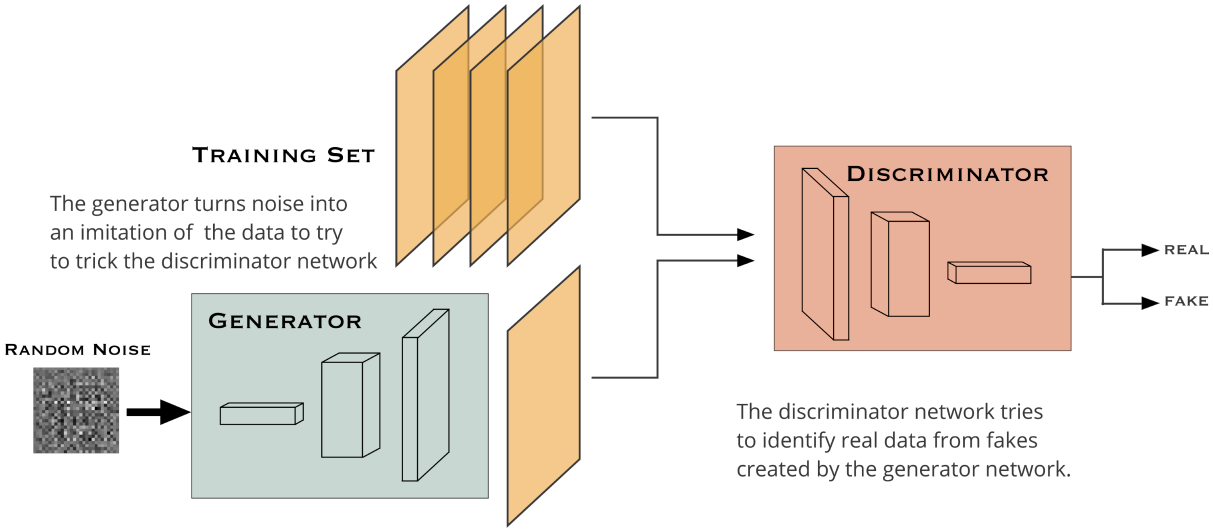
The Deepfake technology can also develop new images from scratch. A non-existent Bloomberg journalist, “Maisy Kinsley”, who had a profile on LinkedIn and Twitter, was the result of a Deefake. Another LinkedIn fake, “Katie Jones”, claimed to work at the Center for Strategic and International Studies, but is thought to be a Deepfake created for a foreign spying operation.

How does a Deepfake work?

Deepfakes, often face-swaps, rely on artificial neural networks. According to an article written by James Chen in Investopedia.com, "A neural network is a series of

algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria."

Deepfakes use GANs (generative adversarial networks), where two machine learning systems process work hand in hand by passing, receiving and altering data to attain a near-perfect product. One machine learning model trains on a data set and then creates video forgeries, while the other attempts to detect the forgeries. The 'forger' machine learning model creates these fakes until the 'detector' machine learning model fails to detect any discrepancies in the fake produced by the 'forger.'



How to recognize a Deepfake?

Poor-quality deepfakes are easier to spot. The lip syncing of a video might be bad or the skin tone could be patchy in some places. Furthermore, there could be a constant blur or flickering around the edges of transposed faces. Finer details, such as hair, eye color, eyelashes and jewelry, are particularly hard for the 'forger' to render well. Badly rendered teeth and shape of facial features can also be a giveaway, as can strange lighting effects.

Above is an example of a poor, and frightening, rendering of a deepfake. The user has attempted to place a fraction of Donald Trump's face on that of Hilary Clinton. Though rather obvious, there are several major and minor rendering errors that make the image above a deepfake. The first tell is the overall posture and positioning of the face. Hilary Clinton's face is facing left, while Donald Trump's eyes look right. The second tell is the overall is the slight blur around the jawbones. The third is the difference in skin color and skin type. Both people have different skin color's and skin types and this machine learning model has not blended them together well. While these are just the general, easily visible features, there are several more that can be found with a closer inspection.

How can Deepfakes impact people?

Near-perfect deepfakes can target any popular or well-reputed person from our society by creating altered clips or swapped faces with unknown people.

An article in BecomingHuman.ai explains that this can play a big role in elections where a politician can be the victim of Deepfakes. Such fake videos or images circulating on social media and other online channels can influence the audience at the time of state or central government elections in any country.

Another serious consequences of Deepfake is widely used in creating sexually explicit images or videos misrepresenting the celebrities in spoofs. And celebrities like popular actors or stars from the music industry are mainly target while creating the Deepfake fake adult videos or images. The creators swap actresses' faces into videos by adult film stars enabling the transfer of sexual fantasies from imaginations of people to the Internet.

https://www.youtube.com/watch?v=cQ54GDm1eL0&ab_channel=BuzzFeedVideo

The BuzzFeed video above is a Deepfake of Barack Obama.

Social media channels, one of the most preffered online platforms where such contents are posted without any authentication, can heavily influence public opinion, which present real challenges for society. Social media companies also hire people or Deepfake service providing companies to spot such contents and help them remove it before it can influence a large body of users.

Deepfakes are powerful tools that can be used for exploitation and disinformation. With advances making them more difficult to detect, these technologies require a deeper look. It is important for people to realize that not everything on the internet is factual and that they must search for the source or authenticity of controversial content posted on the internet.