Explainer: What are deepfakes?

Deepfakes are not just a threat to politicians - brands need to be on alert to the risks of fake marketing

Rosalyn Page (CMO) | **19 July, 2019 07:36**

Have you seen the video with US speaker, Nancy Pelosi, appearing to be drunk and slurring her words? Or perhaps you've seen the one with Facebook CEO Mark Zuckerberg joking about knowing the public's secrets? Or the videos of Rasputin singing Beyonce, Andy Warhol eating a Burger King burger or Salvadore Dali being brought back to life?

In 2018, former US president, Barack Obama, warned in a video about enemies making it look like anyone is saying anything at any time.

While seeing long-dead artists animated in 2019 can seem like a bit of fun, it hides a darker problem about authenticity on the Web. And in an era of heightened concerns about the fake news label being thrown around to undermine news that doesn't suit, deepfakes now look set to accelerate a crisis in trust in the Web.

Deepfakes are essentially videos, and in some cases audio, which purportedly show someone doing or saying something they haven't in real life. They may show real people such as politicians, historical figures or just anonymous individuals. And they may be entirely manufactured, such as the ones showing long-dead people like Andy Warhol or Salvadore Dali speaking or interacting with things they never could have during their lifetime.

The rise of deepfakes can be explained by the increasing sophistication of technology and the crisis in confidence around authenticity of news and information overload because "not only is it difficult to verify the authenticity or veracity of the information we encounter, but we have more information available to us than ever before," research fellow at the Oxford Internet Institute, University of Oxford, Brent Mittelstadt, told *CMO*.

Machine learning brings stills to life

Simon Smith, cyber forensic investigator and cybercrime expert witness, told *CMO* the term 'deepfake' comes from the deep learning tech used to manufacture the fake videos.

"The very best technology is used to map out every muscle movement of a person's face [if looking at the face only] and replicated into a learning algorithm and associated with a word, phrase, attitude or feeling," he explained.

"Once enough learning has been attained, it is possible to attain an almost life-like effect with the assistance of morphing graphical technology that takes into account the person's age, muscles that move when other muscles move, stretching and expressions to give a realistic approach."

The answer to why we are seeing and hearing about deepfakes now lies in a confluence of advances in technology and the work of the darker parts of the Web.

There's an exponential growth in computing power behind the surge of deepfakes, according to founder and co-CEO, Kablamo, a cloud-based enterprise software outfit, Allan Waddell. Adding fuel to the fire is the sophistication of artificial intelligence (AI) and pattern recognition on image datasets.

"It's taking a set of images, and it's been a large number of images, to create models to overlay on an existing people. Traditionally, the

more images you have, the more accurate it becomes," he said. "There's been breakthroughs in the number of images and datasets needed to create these [deepfakes]."

The rise of fake marketing?

Once upon a time, fake videos might have been created for a bit of humour, like politicians or celebrities with fake lipreading to have them say something which parodies themselves. Mostly they were harmless because they were easily identified as fake and exaggerated enough to defy believability.

However, such advances in machine learning technology have enabled the creation of realistic-looking videos. Combine that with the pervasiveness of social media, where fake news and videos can spread without proper scrutiny or verification, and you bring the issue of deepfakes to the fore.

"The technology has been used for many years to help animatronics by mapping out joints and movements in cartoons. This is one step above that and [in the wrong hands] could cause identity theft, false impersonation, setup for crimes a person did not commit and much more serious repercussions," Smith told *CMO*.

And it's not just politics and history that's at risk from deepfakes. Marketing, which relies on brand trust and making representations about a product, will need to protections in place to defend against the threat from deepfakes, too.

According to Smith, deepfake technology can simulate actors and be used to avoid the hiring of talent, leading to potentially misleading marketing. "It can likewise be used to mock a competitor's talent and make fun of [or put words in the mouth of] a competitor by a bad actor."

Oxford Internet Institute's Mittelstadt told *CMO* if deepfakes deceive an audience or fabricate events with public figures, they pose risks to brand image if there is audience backlash due to perceived or actual deception.

"Creating deepfakes for endorsements, however, feels problematic in a different way; the problem there is a lack of consent from the endorser. The deepfake would effectively be putting 'words into their mouths' without consent, which is clearly a problem from the perspective of freedom of expression and autonomy."

Brands that rely heavily on influencers using social media with images of themselves with products are easy potential targets for deepfakes, Waddell said.

"Brand ambassadors who are on Instagram promoting brands, their images are all over the Web. Those images are not coming from official sources. Their faces could be easily overlayed on other people and videos created," he warned.

And there's a hefty risk to brand reputation if a deepfake is created which falsely portrays the brand in a certain way, which may get out publicly and have lasting damage to the brand.

"All it could take is one blockbuster news statement that gets out and have huge brand reputation issues," Waddell said.

Of course, it's not all risks. There are potential opportunities for marketing using the technology that can create deepfakes for what Waddell calls "brand expansion".

"There is an opportunity to provide an authoritative source of company videos, say using images of me as CEO with an approved script, which means as a marketer you may not need to have your CEO on hand to create approved content," he said.

Countering deepfakes

Meanwhile, Smith see humans rather than technology as the best defence against harmful deepfakes and protecting potential risks to brand reputation. But that requires an understanding of the existence of the technology and being on guard against deepfakes.

"As in social engineering and cybersecurity, … I have to say as people are the weakest link in any computer system, people are the strongest validation point as well," he said. "All representations need to be validated by the alleged author in person. We live in a world now where we cannot believe what we read or what we see."

The problem of deepfakes are a data and trust issue to Waddell. "Trust in data, whatever the source, is something we've been dealing with for a long time. It's how you trust the source. Watermarking and preventative measures like getting videos from a trusted source is going to become increasingly important."

Others Mittelstadt such as believe that technology will need to play a role in developing more sophisticated detection tools if deepfakes become a widespread threat to brands and trust in the online world in general.

"As we've seen with the human cost and difficulty of content moderation in other spaces, human reviewers offer a limited solution, or at least if we begin to see (deceptive) deepfakes at scale. Automated tools to detect deepfakes will be needed in that case to have any realistic chance of detecting, labelling and removing them at scale."

Follow CMO on Twitter: @CMOAustralia, take part in the CMO conversation on LinkedIn: CMO ANZ, follow our regular updates via CMO Australia's Linkedin company page, or join us on Facebook: https://www.facebook.com/CMOAustralia.
