DON'T BE DUPED

How to Spot Deepfakes

PROFESSOR V.S. SUBRAHMANIAN SHARES FIVE TIPS TO HELP AVOID GETTING TRICKED BY MODIFIED DIGITAL ARTIFACTS.

Deepfakes—digital artifacts including photos, videos, and audio that have been generated or modified using AI software—often look and sound real.

Deepfake content has been used to dupe viewers, spread fake news, sow disinformation, and perpetuate hoaxes across the internet. And though not all deepfakes are bad, the negative ones get the headlines and can have severe consequences.

V.S. Subrahmanian, Walter P. Murphy Professor of Computer Science at Northwestern Engineering, has five pieces of advice to avoid getting duped by dangerous deepfakes.

1. Automatically question what you see and hear

Anyone with internet access can create a fake, so anyone with internet access might become a target for deepfakes.

"Rather than try to detect whether something is a deepfake or not, basic questioning can help lead to the right conclusion," says Subrahmanian, founding director of the Northwestern Security and AI Lab and faculty fellow at Northwestern's Buffett Institute for Global Affairs.

2. Look for inconsistencies

For better or for worse, deepfake technology and AI both continue to evolve at a rapid pace. Ultimately, software programs will be able to detect deepfakes better than humans, Subrahmanian predicts.

For now, there are some shortcomings with deepfake technology that humans can detect. AI still struggles with the human body, sometimes adding an extra digit or contorting parts in unnatural ways. The physics of light can also cause AI generators to stumble.

3. Break free of biases

It's human nature to become so deeply rooted in preconceived notions that we take them as truth. In fact, people often seek out sources that confirm their own notions, and fraudsters create deepfakes that reinforce previously held beliefs to achieve their goals.

"Some people are more likely to consume social media information that confirms their biases. I suspect this filter-bubble phenomenon will be exacerbated unless people try to find more varied sources of information," Subrahmanian says.

"THERE ARE A LOT OF POSITIVE APPLICATIONS OF DEEPFAKES, EVEN THOUGH THOSE HAVE NOT GOTTEN AS MUCH PRESS AS THE NEGATIVE APPLICATIONS."

V.S. Subrahmanian
Walter P. Murphy Professor of Computer Science

4. Set up authentication measures

Audio deepfakes have been used to trick people into not voting by simulating a candidate's voice. This trick can get personal, as scammers have tried to steal people's money by recreating a relative's voice and calling and saying they need funds.

To avoid falling for this ruse, Subrahmanian suggests setting up authentication methods with loved ones. That means asking specific questions only that person would know, such as where they recently ate, or even a code word.

5. Understand that social media platforms can only do so much

Social media has changed the way people communicate with each other. They can share updates and keep in touch with just a few keystrokes, but their feeds can also be filled with phony videos and images.

Subrahmanian said some social media platforms have made admirable efforts to eliminate deepfakes. However, suppressing deepfakes could potentially suppress free speech. Subrahmanian recommends checking websites such as PolitiFact to gain further insight into whether a digital artifact is a deepfake or not.

BRIAN SANDALOW