Deepfakes, AI, and digital evidence

Mark Lanterman

Bench + Bar of Minnesota (July 2023)

MINNESOTA STATE BAR ASSOCIATION

OCTOBER 2023

BENCH+BAR of Minnesota

READING THE FINE PRINT

The extensive changes to Minnesota landlord-tenant law

Deepfakes, AI, and digital evidence

BY MARK LANTERMAN ✉ *mlanterman@compforensics.com*

MARK LANTERMAN is CTO of Computer Forensic Services. A former member of the U.S. Secret Service Electronic Crimes Taskforce, Mark has 28 years of security/forensic experience and has testified in over 2,000 matters. He is a member of the MN Lawyers Professional Responsibility Board.

With the ever-expanding prevalence of artificial intelligence, I'm sure that most of us have seen at least a few types of "deepfakes." Elvis Presley singing the latest top hits. Albert Einstein answering viewers' questions about life. Living portraits of old photographs. Or some more problematic examples, such as a menacing speech by Mark Zuckerberg or a video of a politician created to spread disinformation. Some may have even seen a video appearing to depict their company's CEO requesting an immediate wire transfer, as cybercriminals continue to use AI to bolster social engineering campaigns. It seems that just about everybody now has the ability to alter digital media, with varying degrees of believability.

Deepfakes, or digitally altered media that convincingly make one individual appear as another, have also had an impact on the courtroom. According to the Department of Homeland Security's paper, "Increasing Threat of Deepfake Identities," "Deepfakes... utilize a form of artificial intelligence/machine learning (AI/ML) to create believable, realistic videos, pictures, audio, and text of events which never happened. Many applications of synthetic media represent innocent forms of entertainment, but others carry risk."[1] While there have been cases of litigants attempting to enter a deepfake into evidence, the problem has also been reversed–litigants claiming that real evidence has been manipulated or fabricated.

Digitally stored information has repeatedly proved itself to be a pivotal source of evidence, often serving as a critical, unbiased witness. Nearly every case today involves ESI to some extent. When presented with this kind of strong, perhaps damning, evidence, people now have the ability to throw a new defense at the wall and see if it sticks: "It's not real." While a judge may reject the attempt,[2] the "deepfake defense" will still have consequences. As an NPR report about the phenomenon noted, "If accusations that evidence is deepfaked become more common, juries may come to expect even more proof that evidence is real."[3] Though the technology is relatively new, courts already have processes in place to handle fake evidence and can apply these same procedures to managing deepfakes.[4] But courts are less prepared to deal with proving that real evidence is, in fact, real. Furthermore, the better the evidence, the more likely that juries will feel required to verify its legitimacy. With the rise of common applications of artificial intelligence, the pressure is on to verify digital evidence as efficiently as possible.

Deepfakes present a host of legal concerns. From actors losing the rights to their own identities to reputational damage to manufactured evidence affecting the outcomes of custody disputes, we are just beginning to learn how to grapple with deepfakes and artificial intelligence. In the courtroom, well-communicated guidelines, strong authentication standards, and extensive training can address some of the risks. Expectations for juries surrounding the requirements for evidence verification should be well-established, and court-appointed digital forensic experts can manage and analyze digital evidence for both sides, helping to create an even playing field and manage costs.

Emerging laws and regulations will hopefully begin to help the legal community navigate new problems posed by these technologies. But developing tried-and-true methods to identify deepfakes reliably will undoubtedly remain a work in progress. Given how difficult it can be to spot a deepfake, the New York Times wrote recently, "Initiatives from companies such as Microsoft and Adobe now try to authenticate media and train moderation technology to recognize the inconsistencies that mark synthetic content. But they are in a constant struggle to outpace deepfake creators who often discover new ways to fix defects, remove watermarks and alter metadata to cover their tracks."[5]

In the meantime, members of the legal community should be on high alert for the possibility of altered digital media, from opposing parties and their own clients. Attorneys should strive to be especially vigilant in abiding by digital-evidence best practices throughout the entirety of a case. In the event that third-party verification is ultimately required, organizing original source material and making it readily available is essential. ▲

NOTES

1 *https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf*

2 *https://www.npr.org/2023/05/08/1174132413/people-are-trying-to-claim-real-videos-are-deepfakes-the-courts-are-not-amused*

3 *Id.*

4 *Id.*

5 *https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html*