IRMI

EMPLOYMENT PRACTICES LIABILITY CONSULTANT

ARTIFICIAL INTELLIGENCE APPLICATIONS IN INVESTIGATIONS: DETECTING DECEPTION

By Don Phin, Esq.

As an investigator of sexual harassment and other workplace allegations, I'd love to have a Truth Machine to help support the accuracy of my findings, especially as it relates to credibility assessments. Today, the use of artificial intelligence (AI) in investigations is emerging as a vital tool in assisting businesses when they investigate workplace misconduct complaints, including sexual harassment allegations. Such applications can also be used in the interviewing/hiring process, particularly when assessing candidates for senior management-level jobs where the financial consequences of a "mis-hire" are considerable (not to mention positions having national security implications).

Last year, I wrote about the advancements in AI and how they were affecting the human resources (HR) function, especially hiring. I discussed how potential bias by programmers could be inherent in and thus impair the accuracy of any AI system. Aware of this challenge, programmers are now developing algorithmic techniques for achieving even greater levels of accuracy and, ultimately, fairness. While most of these programs are sought by police, courts, and military agencies to ferret out criminals and potential terrorists, they also have numerous applications to various HR and employment-related functions.

Indeed, these new forms of AI may soon function as Truth Machines and significantly assist corporations and their HR departments in making what are now some of the toughest—yet critical—decisions that could end up making or breaking a business.

In this article, I will discuss the following.

- Recent breakthroughs in detecting deception
- Ways in which AI can affect the future of employment practices liability (EPL) investigation practices and the interviewing/hiring process

This article appeared in the Winter 2019 issue of *EPLiC*. Copyright 2019 by International Risk Management Institute, Inc., 12222 Merit Dr., Suite 1600, Dallas, Texas 75251–2266, 972–960–7693, https://www.IRMI.com. All rights reserved. ISSN: 1529–840X. This material may be quoted or reproduced only with written permission from the publisher. The opinions expressed in *EPLiC* are those of the individual authors. Nothing published in *EPLiC* is to be construed as legal, accounting, or professional advice. If such advice is required, the services of a competent professional should be sought.

- Practitioners' views and opinions about the use of these tools in determining a person's credibility

From the Truth Machine to AI

In 1996, futurist James Halperin wrote the science fiction novel *The Truth Machine* (Random House, 1996). The Truth Machine (known as the ACIP) was 100 percent accurate in determining whether a person was lying or telling the truth. It was so effective that it could help to eliminate virtually all crime and dishonesty. Halperin takes us on a journey into the potential impact, misuses, and dangers inherent in his invention.

> In 2028 the ACIP was still used mainly in the court system, although Congress had also approved it for customs, immigration, food and drug testing and inspection, and several other government functions. But the inevitability of its invasion into other areas was becoming obvious. The ACIP was already popular; few voters sided against widening its use. Even then, its effects were impressively positive, far beyond even the most optimistic predictions of just a few years earlier. Savings in court costs had exceeded projections by over 30 percent, but those savings were dwarfed by the benefit of forcing lawyers (who had previously tended to be among the most intelligent and least productive members of American society) to redirect their energies to more worthwhile areas….
>
> Now any person could be truth-tested by reserving time at one of thousands of testing stations offering ACIPs for civil litigation or mediation. A politician who refused to submit to ACIP was unelectable….
>
> And the virtual elimination of crime was by far the greatest godsend of all.... Most were guilty of minor offenses such as traffic violations, underage drinking, illegal drug usage, minor tax evasion, cheating on expense reports, that sort of thing.

The Polygraph: A Forerunner of the Truth Machine

For years, the closest thing to a Truth Machine was the polygraph. Polygraphs are accurate at measuring respiration rate, pulse rate, blood pressure, and skin conductivity (usually affected by perspiration). However, *interpreting* these signals is more of an art than a science. Therefore, it is the subjective analysis of test results—rather than the issuing of a clear "yes, he was telling the truth" or a "no, he was lying" signal—that caused the use of lie detector tests to be prohibited in most contexts.

Fortunately, given recent improvement in AI-based facial, voice, and eye detection, we are getting a lot closer to creating a more legally viable Truth Machine.

Newly Developed Tools To Detect Deception

AI algorithms are being programmed to detect deception in numerous ways. Factors such as voice inflection, facial expressions, eye movements, and brain activity can all be helpful in ferreting out dishonesty. Let's talk about each device in more detail.

Voice Inflection

Voice-stress analysis systems, like the polygraph, recognize microtremor patterns when information is delivered under stress and, specifically, when those moments of stress are generated by an attempt to lie or deceive. A microtremor is a slight inaudible deviation in a person's voice that occurs when the subject is under stress. This stress can result from various sources, including non-deception-related stress (such as the act of being questioned, in general), surprise, grief, anger, fear, etc., and … lying.

An example of one such machine is the Computer Voice Stress Analyzer (CVSA). It claims to be the most accurate truth verification system in the world based on multiple technical and scientific studies. One of my former SEAL buddies told me about how CVSA was used to monitor phone calls in Afghanistan to identify when a caller was under stress.

What initially prompted my writing this article was listening to a podcast about how scientists at Adobe labs have created a program that will allow you to literally create a voice recording using a "cut and paste" technique, in effect manufacturing a deceptive recording.

To produce an accurate, deceptive recording (talk about an oxymoron!), the CVSA must first listen to somebody speak for as long as 40 minutes. Once it has done so, the CVSA will allow you to "cut and paste" content that sounds exactly like the speaker, even though the individual never used the words in the newly created combinations. The danger inherent in such a system can be envisioned in the following scenario. A disgruntled employee creates a fake voice track of their manager and places it on the Internet, and the manager's recombined words will then become viral in the same manner that fake news always does. Damage to the manager and likely his company's "brand" will be done before any repairs can be made. Frighteningly, someone investigating the manager's statements after they are made would be unable to detect, nor have any reason to think, that the recording was "manufactured" or even entirely fabricated.

The following are questions to consider.

- As with polygraphers, do we need to be certified or have some level of expertise in the art of interpreting the data produced by these AI tools?
- If someone can record the president saying one thing and then do a cut-and-paste job so it turned out to be just the opposite, how do we know that *any* voice recording evidence is accurate?

Given such technology, along with its potentially fraudulent uses, if I'm investigating a matter, I want to make sure there's a very clear chain of custody around any voice recordings. I want to make sure that, if somebody denies saying something, an analysis can be done to determine whether that particular recording was "manufactured."

Facial Expression

Researchers at the University of Maryland developed the Deception Analysis and Reasoning Engine (DARE), which uses AI to detect deception in videos. DARE looks at five microexpressions known to reveal lying: frowning, eyebrows raising, lip corners turning up, lips protruded, and head side turn.

While most people, including investigators, are able to detect lying over 60 percent of the time, DARE claims it pushes the accuracy up to the 80 and 90 percent range. DARE is now being used to detect deception in courtroom trial videos.

According to Raja Chatila, executive committee chair for the Global Initiative on Ethics of Autonomous and Intelligence Systems at the Institute of Electrical and Electronics Engineers, DARE should be used with caution. "If this is going to be used for deciding … the fate of humans, it should be considered within its limitations and in context, to help a human—the judge—to make a decision," pointing out that "high probability is not certainty" and not everyone behaves the same way.

Eye Movement

EyeDetect systems, by Converus, is a program that administers a 30-minute truthfulness test based on observations of eye movement. The company announced that analysis from its system will be accepted as evidence in a New Mexico court. It claims it can produce 86 percent accuracy in just 30 minutes. "It's a significant milestone to have EyeDetect test results admitted as evidence in court," said Converus President and CEO Todd Mickelsen in a statement. "Attorneys with strong cases can now use EyeDetect to exonerate their clients."

A sheriff's office in New Mexico also uses EyeDetect to screen job candidates, and the company has gotten some attention administering its tests to willing politicians.

Brain Scanning

Scientists have programmed machines to look at brain activity on magnetic resonance imaging scans (MRIs) or electroencephalograms (EEGs). An example of MRI use is No Lie MRI. According to their website, "No Lie MRI™ is a proprietary product that objectively measures intent, prior knowledge, and deception using algorithms to automatically analyze functional Magnetic Resonance Imaging (fMRI). The approach used by No Lie MRI™ has a verified accuracy that greatly surpasses all other truth verification/lie detection methods. Current accuracy is over 90% and is estimated to be 99% once product development is complete."

An example of using EEG is Farwell Brain Fingerprinting. What the tool assesses is the extent to which a person recognizes a given stimulus. Per their website,

> Farwell Brain Fingerprinting technology is a new scientific technology to detect whether specific information is stored in a person's brain. This technology can provide evidence to identify criminals and terrorists accurately and scientifically. Brain Fingerprinting testing measures brainwave responses to crime-relevant or terrorism-relevant words or pictures presented on a computer screen. To date, Brain Fingerprinting testing has not resulted in *any* incorrect determinations—there have been no false positives or false negatives. It has provided highly accurate results in over 200 tests, including tests on FBI agents and tests sponsored by the CIA and the US Navy. Brain Fingerprinting testing has been ruled admissible in court in a murder case.

Using brain scanning tools requires either going into an MRI or having a sensor cap placed on one's head to detect brain waves. I doubt we'll be seeing that in our conference rooms anytime soon.

Combined Tools

AVATAR was developed at San Diego State University (SDSU) and the University of Arizona. It is being used to ask interactive questions on a video terminal at border crossings. "The system can detect changes in the eyes, voice, gestures, and postures, to determine potential risk."

According to Aaron Elkins at SDSU, the system has between a 60 and 75 percent accuracy rate, with peaks up to 80 percent. Again, this is an improvement on most human evaluation methods (e.g., interviewing a job candidate), which seldom achieve even 60 percent accuracy, yet the results are still far from perfect. When people fail the AVATAR assessment, they are sent to secondary screening—that is, screening by human agents.

Bottom line: These tools can identify deception better than most humans and can be helpful in making credibility assessments. And a 100 percent accurate Truth Machine may be only a few years away.

Questions for the Investigators

Before I began writing this piece, I posed these questions on LISTSERV for the members of the Association of Workplace Investigators.

- If there is a deception detection tool (it may use voice, facial movements, or eye movements) that is over 80 percent accurate (they are pushing closer to 90 percent all the time), would you use it in an investigation? Assume the witness must consent to being recorded by video and audio.

- If you wouldn't use it, why not?
- What if the witness refuses consent?
- What if a witness seeks (or provides) an independent evaluation of their truthfulness done by one of these tools? Many years ago, I had an alleged harasser pull that stunt by offering results from a polygraph the alleged harasser had obtained independently. Would you consider those findings in your credibility determinations?

The Investigators Respond

Here are just a few responses I received. I tried to spread across private investigators, lawyers, and HR folks.

Timothy W. Armistead, D.Crim. CFS, Armistead Research and Investigative Services

(Timothy has written an extensive paper on this topic, "The detection of deception by linguistic means: Unresolved issues of validity, usefulness, and epistemology," *Policing: An International Journal of Police Strategies and Management,* Vol. 35 (2), 304–326; Vol. 34 (4), 588–605.)

None of these techniques assesses deception directly, despite some claims (and DARE's preview of its invention) to the contrary; they detect emotion, and there are significant questions about the accuracy and universality of their findings, the credentials of many of their alleged expert practitioners, the controls for various confounding variables in the studies at issue (in particular, selection bias), and the conflation of "lying" indicators with momentary confusion, minor evasion, indecision about how to answer a question which calls for vaguely understood nuances, offense taken by the personal nature of a question, and so on.... As the software is explained in its own remarks, the DARE researchers have produced an automated facial recognition tool, a la Ekman's thesis, along with a voice stress analyzer. This is quite an achievement, but it leaves open the most critical question: the ability of the alleged indicators to differentiate between lies and their close neighbors.... A final note re Phin's observation that human detection of deception is essentially at the random rate of 50–60 percent: That's true of untrained subjects, and it's a frequently replicated finding. The success rates are significantly higher for trained investigators, depending on training, method, experience, and which study one consults.

Rhoma Young, HR Consultant

I work with a lot of the leading-edge AI folks and those working with cognitive predictability. I think the analysis is far from where we would like it to be, to depend and rely on it. A parallel situation could be much like employers wanting the perfect test to give applicants on defining and predicting honesty or to be wonderful salespeople. They want a "simple and safe" way for a tool to make decisions for them in hiring. I do not think AI accuracy and reliability in reading people is anywhere near ready to "tell the truth," especially given the various cultural and learned behaviors we deal with all the time. It seems we still have to rely on our own skills and careful discernment to make our own assessments.

Kirsten Hume Scrimshaw, Barrister and Solicitor

Thanks for this interesting topic for discussion, Don. I'm going to give the classic lawyer's response—it depends. My main concern with AI right now is that it tends to incorporate cultural bias, which can be a key factor in assessment of credibility. I would need to be convinced that the "deception detection tool" itself has been shown to be equally accurate across different sets of subjects. If it was less reliable in assessing the credibility of those who are members of nonmainstream groups, or people who are traditionally subject to discrimination (e.g., POC [persons of color], LGBTQI [lesbian, gay, bisexual, transgender, queer, intersex] people, indigenous people, recent immigrants, people whose first language is not English, etc.), I would have serious concerns that it would simply perpetuate discrimination.

Mark Lujan Sr., Stockton PSPI

I would consider using such a tool once I knew more about how it works, and what its proven reliability is. Technology is all around us these days so why not take advantage of something that could help us in our investigations. As a past law enforcement officer who conducted numerous criminal investigation interviews with suspects, you gain a sense when someone is not being honest. I am sure you have heard the phrase "the sixth sense," and that is used a lot in law enforcement investigations. If you are good at speaking with various types of people, this sixth sense becomes quite acute over the years, but I am game for anything that might help improve that sixth sense reliability.

Claudia Viera, Esq., Mediator/Investigator

As part of your article, I would recommend that you also look at how accurate AI is across races—I have read that it is not as accurate for nonwhite faces (but I have not researched this topic so do not know if this is a valid critique). Obviously, I view accuracy across races/genders as necessary before such a tool should be widely used. Similarly, does the tool accurately depict honesty even when wearing eyeglasses, or when the person is an introvert who does not make good eye contact? Finally, I imagine that use of such tools would create heightened anxiety and might dissuade complainants from coming forward. However, in a high-profile he said/she said case, I can also imagine a lot of pressure to use such technology once it is available and has been proven accurate.

Conclusion

AI is an exciting new frontier. The main concerns surrounding use of the technologies discussed in this article include reliability, bias, and intimidation. The usefulness of AI often begins with a conversation about the data being used. Most of the data dealt with in a workplace investigation are testimony from witnesses. The fact is, data are almost always imperfect, and this is true for witnesses as well.

Not surprisingly, there is a backlash against the Orwellian nature of these tools. According to Jay Stanley, a senior policy analyst at the American Civil Liberties Union (ACLU), "[A]t the ACLU our opposition to lie detectors, which dates to the 1950s, has never been exclusively about their ineffectiveness or the specific technology of the polygraph. We have said since the 1970s that, even if the polygraph were to pass an acceptable threshold of reliability or a more accurate lie-detection technology were to come along, we would still oppose it because we view techniques for peering inside the human mind as a violation of the Fourth and Fifth Amendments, as well as a fundamental affront to human dignity."

My take on it is that I would be willing to use these types of technologies, especially in "he said/she said" sexual harassment-type scenarios. Why not use a tool that could "check my head"? Even if I can push up the results to 70 percent accuracy, I would hate to make a determination that is inaccurate, especially one that could also be potentially career ending.

I am less concerned about the intimidating nature of these tools. In fact, some degree of intimidation may actually help to prevent lying, which I believe is a good thing. Moreover, many investigations are already recorded by way of either audio or video. In my experience, after the first few minutes, the interviewee tends to ignore (or at least become more comfortable with) these devices, so that after a short time, the equipment has little or no impact on the results. Those audio recordings or videos can then be analyzed for their truthfulness.

Remember, we are not suggesting these tools be used to offer the extremely high standard of proof required in a criminal matter. Rather, my investigative findings are a "more likely than not" standard, as required in civil litigation. Accordingly, if someone accused of sexual harassment presented exculpatory results from the kinds of deception tests I have discussed, I would seriously consider such evidence.

Bottom line: at its current stage of development, I would not rely *solely* on the results of AI deception detection technology to arrive at a definitive conclusion in an investigation.

In fact, some of the shortcomings addressed in the survey responses above, particularly concerning the subtle differences in social behavior, communication style from person to person, plus gender and racial differences (and how these differences might skew AI test results), highlight the importance of the *human* skills that HR personnel and management must bring to the table in the event of an investigation and/or an EPL insurance claim. These skills and areas of expertise include knowing your people, listening to both sides of a story, and "managing" a conversation. However, I do consider such technology a valuable tool, which, combined with my knowledge and experience, can help me reach an accurate determination.

If you want to "geek out," and learn more about AI from my favorite resource, check out the Association for the Advancement of Artificial Intelligence.

Don Phin, Esq., *is coeditor of* EPLiC *and president of HRSherpas, Inc. He speaks frequently on HR risk management issues and has written numerous books on the workplace. To learn more, visit* www.donphin.com.