A Mental Trespass? Unveiling Truth, Exposing Thoughts, and Threatening Civil Liberties With Noninvasive AI Lie Detection

Kurtis G. Haut, Taylor Sen, Denis Lomakin, and Ehsan Hoque, Member, IEEE

Abstract—Imagine an app on your phone or computer that can tell if you are being dishonest, just by processing effective features of your facial expressions, body movements, and voice. People could ask about your political preferences and your sexual orientation and immediately determine which of your responses are honest and which are not. In this article, we argue why artificial intelligence-based, noninvasive lie detection technologies are likely to experience a rapid advancement in the coming years, and that it would be irresponsible to wait any longer before discussing their implications. To understand the perspective of a “reasonable” person, we conducted a survey of 129 individuals and identified accuracy and consent as the critical factors. In our analysis, we distinguish two types of lie detection technologies: 1) “truth metering” and 2) “thought exposing.” We generally find that truth metering is already largely within the scope of existing U.S. Federal and State laws, albeit with some notable exceptions. In contrast, we find that the current regulation of thought-exposing technologies is ambiguous and inadequate to safeguard civil liberties. In order to rectify these shortcomings, we introduce the legal concept of “mental trespass” and use this concept as the basis for the proposed legislation.

I. INTRODUCTION

The world’s first real supercomputer was Control Data Corporation’s CDC 6600, developed in 1964. The computer was enormous, the size of multiple people, and state of the art—far ahead of the competition. Three times as fast as its predecessor, it could run 3 million megaFLOPS. It costs the equivalent of U.S. $60 million in 2021. The CDC 6600 was so powerful the word “supercomputer” was coined to describe it. If someone was to tell its creator, Seymour Cray, that in 50 years’ time, a processor the size of his forearm would cost 50,000 times less and be 2 million times faster, he might not believe them. But the NVIDIA GeForce GTX Titan X, released in 2015, was exactly that.

One field of technology experiencing a similar rapid advancement is computer vision-based artificial intelligence (AI), and advanced noninvasive, AI-driven sensing technologies. As we experience this revolution firsthand, we benefit from ever more surprising contributions to our daily lives.

AI-powered thermal cameras systems were actively being used to screen passengers for fevers associated with coronavirus [17], [25], [30], [52], [85] and systems that extract heart rate from a common video stream [4], [23], [37], [56], [58] are currently being used in health monitoring [19], [36]. Recent advances have even enabled noninvasive systems to evaluate aspects of an individual’s mental state from facial expressions alone, such as whether someone is imagining versus remembering an event [35], or whether someone is experiencing one of the common emotions [26], [29], [68]. As such systems become further developed, they will likely find even more unanticipated applications. However, not all of these applications may be beneficial to humankind.

With the increasing powers of noninvasive AI, also come new methods for invasion of privacy and circumvention of our rights against unreasonable searches. For example, a recent AI system purports to be able to predict one’s sexual orientation from their facial features [53], [82]. It is easy to foresee the harm that can result from exposing one’s private sexuality considering the case of college student Tyler Clementi. After Tyler’s roommate set up a webcam in their room and publicly broadcasted a private sexual encounter he had with another male student, Tyler, under extreme stress, died of suicide [61]. Similarly daunting is the use of AI systems in police surveillance. Allegations of Chinese government oppression against the Uyghur minorities in the Xinjiang province have been made as AI facial recognition is used in conjunction with camera surveillance [64]. Chinese authorities state that the use of such technologies is necessary to fight terrorism and that similar surveillance systems were instrumental in enforcing the quarantines that helped halt the progression of COVID-19 throughout China [18]. How do we ensure that advances in AI sensing technologies are not abused?

The legal system may be in a good position to help prevent abuses while not stifling the benefits such technology brings. However, legal systems have a mixed history of being in sync with technological developments [45], [86] and the future is anything but certain regarding the interaction between technological advances and people from an ethical perspective [44]. In particular, the interplay between advances in lie detection technologies and the legal system has a rich history and most unpredictable future [14], [34], [39], [47], [62], [75], [81].

A. Summary of This Article

In this article, we examine the progression of lie detection technologies and evaluate their potential to cause societal harm...
through loss of privacy and circumvention of civil liberties. We then consider to what extent U.S. law currently regulates these technologies. As illustrated in Fig. 1, we distinguish between two types of lie detection technology: 1) “accurate truth metering,” which involves evaluating the veracity of an individual’s statement (e.g., the degree of belief that an individual has in their intentionally made statement) and 2) “accurate thought exposing,” which involves predicting an individual’s thoughts with superior-to-human accuracy.

In our analysis, we generally find that truth metering is already largely within the scope of existing U.S. Federal and State laws, albeit with some notable exceptions. In contrast, we find that the current regulation of thought exposing technologies is ambiguous and inadequate to safeguard civil liberties. In order to rectify these shortcomings, we introduce the legal concept of mental trespass and use this concept as the basis for proposed legislation.

More specifically, in this article, we argue the following.

1) Development of noninvasive, AI-based lie detection technologies is likely to progress rapidly in the near future, and no law or government effort is likely to halt its production, distribution, and use (in many cases the government is investing heavily in the advancement of such technologies).

2) Lie detection technologies carry with them much potential for individual harm in terms of loss of privacy, wrongful criminal conviction, and unfair bias.

3) While the current legal environment generally already regulates accurate truth metering technologies, it is largely ambiguous with regards to the legality of uses of accurate thought exposing technologies.

4) In order to mitigate the potential harms such technologies may bring, we recommend the introduction of a regulatory federal “Mental Trespass Act.”

Due to the hybrid nature of this article in considering technological, legal, and public perspectives, we use an unconventional paper style. The remainder of this article is organized as follows. Section II provides an overview of the technologies that underlie lie detection and describes the revolutionary advances that are imminent. Section III is broken down into a theoretical analysis of three components: 1) potential harms; 2) laws and limitations; and 3) public perspective. The potential harms component describes how the coming technology advancements could cause serious adverse effects with regards to the infringement upon civil liberties. The laws and limitations component examines current U.S. Federal and State laws relevant to the coming lie detection advances and highlights gaps in their coverage that leave the public susceptible to harm. The public perspective component outlines the surveying techniques used to garner public opinion. This article then proceeds to Section IV comprised of the public perspective survey responses and recommendations of our proposed Mental Trespass Act. This article concludes with Sections V and VI.

II. BACKGROUND

A. Technology Progression: Lie Detection Revolution

The underlying technology of lie detection is comprised of several components, including developments in physiological knowledge, improvements in questioning techniques, and more recently, advances in AI sensing tools and data analysis systems. The exponential advances that these components have made recently make a lie detection revolution seemingly inevitable.

1) Timeline of Deception Technology: “If anyone bring an accusation against a man, and the accused—jump into the river— if he sink in the river his accuser shall take possession of his house.”—Code of Hammurabi, circa 1754 BC.

What started out as mere random chance or religious belief, the art of lie detection has progressed to include increasingly powerful scientific techniques, including advanced sensing tools and more refined questioning techniques (see Fig. 2). As demonstrated in the above quotation, lie detection was essential enough to human civilizations that it appears in the Code of Hammurabi, one of the very first instances of written law from circa 1754 BC [65]. Translations of preserved tablets of the Code state that questions of honesty were to be resolved through what has been termed trial by ordeal. It took approximately 800 more years before the first glimmer of scientific legitimacy in lie detection appeared, which was found in the ancient Hindu text; the Vedas. Loosely based on the involuntary fight or flight response (which causes individuals to go white as blood is diverted from body extremities to the heart and lungs), the Vedas describe how to spot a poisoner, “[The poisoner] does not answer questions, or they are evasive answers; he speaks nonsense—his face is discolored…” [80]. The scientific progression of lie detection continued in the third century BC, as renowned physician
Erasistratus used pulse, skin temperature, and skin pallor to correctly detect the lies of Prince Antiochus, as the prince tried to conceal his passionate love for his father’s new wife [7], [79].

An underlying premise regarding lie detection began to be recognized. Charles Darwin wrote in his 1872 book, The Expression of Emotions in Man and Animals, that “…actions become habitual in association with certain states of the mind, and are performed whether or not of service in each particular case …” [21]. We recognize a fundamental premise of lie detection in that: *a person’s internal state of mind uncontrollably leaks out into the externally observable world when appropriately probed.* Indeed, this premise must hold for a given lie detection technique to work. Through *appropriately probes* we recognize that specialized questioning techniques may be necessary to cause honest and dishonest subjects to elicit detectable differences in observable behavior. This definition additionally brings attention that advanced tools may be useful in observing these subtle differences.

The use of tools and specialized questioning techniques in lie detection is demonstrated with the perhaps most well known and widely used lie detection device, the contemporary polygraph. Like Erasistratus’s technique, the polygraph tracks the subject’s heart rate and respiration. The modern polygraph, however, has two notable improvements over Erasistratus, including: 1) additional sensors for blood pressure, skin conductivity, and respiration rate and 2) a formal questioning technique, known as the *control question test*. The polygraph sensors collectively provide a measure of the subject’s physiological arousal. Crucially, the control question test begins with questions unrelated to the matter for which the lie detector is being applied, including *baseline questions* and *control questions*. Baseline questions are trivial questions used to indicate the subject’s arousal at rest. Alternatively, control questions are designed to create a strong physiological response in most people, for example, *Have you ever stolen office supplies from work? Have you ever cheated on your taxes?* The unrelated questions are followed by *relevant questions*, which are questions pertinent to whatever is being investigated (i.e., the alleged crime). The underlying theory of the control question test is that someone who is lying is more likely to be nervous during the relevant questions than during the control questions, compared to an honest subject who is expected to have a similar or reduced response during the relevant questions compared to the control questions [13], [63]. Other questioning techniques such as the guilty knowledge test (GKT), which relies on the strategic use of information only a guilty person would know, have been developed and compared with the control question test [59]. Depending on the context and person that is being interrogated, one questioning technique may be preferred and more effective than the other.

While there are many earlier references to the use of blood pressure in lie detection, John Larson, the first U.S. police officer with a Ph.D., is credited as the inventor of the modern polygraph in 1921 [3], [76].

Beyond the basic sensors used in the common polygraph, numerous technologies have been used for lie detection. Most notable among these is the rapid development of computer hardware as well as the analysis software that runs on it. Similar to the rapid advances in computer hardware described in the introduction, the field of computer vision has seen remarkable growth and new development in the past few decades, especially in the last decade. In 2012, “AlexNet” astounded researchers with its accuracy in image classification and demonstrated the power of convolutional neural networks for the task [50]. In 2014, the invention of generative adversarial networks utilized deep learning to generate realistic images [33], which recently became embroiled in controversy with their application in deepfakes. Researchers and software

---

### Fig. 2. Timeline of technological advances and legal doctrines. (The exponential time scale is not represented clearly by this figure, but should be noted nonetheless.)

<table>
<thead>
<tr>
<th>Technological Advances</th>
<th>Legal Doctrines</th>
</tr>
</thead>
<tbody>
<tr>
<td>300BC</td>
<td>Code of Hammurabi dictates the use of “Trial by Ordeal”</td>
</tr>
<tr>
<td>1754BC</td>
<td>Ancient Hindu Text the Vedas</td>
</tr>
<tr>
<td>1791</td>
<td>4th &amp; 5th Amendments</td>
</tr>
<tr>
<td>1921</td>
<td>Polygraph</td>
</tr>
<tr>
<td>1923</td>
<td>Frye v. United States</td>
</tr>
<tr>
<td>1948</td>
<td>Universal Declaration of Human Rights</td>
</tr>
<tr>
<td>1988</td>
<td>Employee Polygraph Protection Act</td>
</tr>
<tr>
<td>1997</td>
<td>ALI Restatement of Law “Intrusion upon Seclusion”</td>
</tr>
<tr>
<td>2008</td>
<td>EEG based lie detection</td>
</tr>
<tr>
<td>2011</td>
<td>Thermal camera lie detection</td>
</tr>
<tr>
<td>2015</td>
<td>80% accuracy emotion recog.</td>
</tr>
<tr>
<td>2018</td>
<td>fMRI used to determine what is being seen</td>
</tr>
<tr>
<td>2019</td>
<td>Congressional Hearing on impacts of Facial Recognition Technology</td>
</tr>
</tbody>
</table>
engineers working with computer vision have an incredible array of tools with which to develop new technologies in the coming years. We highlight the progress in computer vision specifically because these advances enable lie detection to be performed at a distance due to their inherent noninvasiveness.

Regardless of how sophisticated these deep learning algorithms have become, their effectiveness fundamentally relies on good data and lots of it. It thus comes as no surprise that one of the major factors limiting progress in noninvasive deception detection has been the lack of good data. However, with recent advances in Internet technologies, techniques are becoming available to scalably gather data on deception. For example, Sen et al. [71] developed a system for gathering video deception data via crowdsourced individuals. In addition, U.S. government entities have very recently expressed desire to gather data sets on “credibility assessment,” which could be used to develop deception detection technologies. In fact, in 2019, the intelligence advanced research project association (IARPA) puts out a grand challenge concerning the collection of deception data. The credibility assessments standardized evaluation (CASE) challenge formally called for a protocol to standardize this procedure in regards to how these datasets are gathered and accessed. Additionally, governments have started pouring vast amounts of funding into projects that expand their powers of surveillance. Backed by the Chinese and Russian governments, AI startup Megvii raised U.S. $460 million for the development of facial recognition technology [43].

We emphasize these specific examples to illustrate the noninvasive nature of these developing technologies, advancements in data collection procedures/capabilities, and government vested interest. Because of these qualities, it seems inevitable that accurate AI-based lie detection will soon be upon us.

III. METHODOLOGY

We apply a theoretical framework grounded in analysis of potential harms arising from lie detection technology and the existing legislation safeguarding individual liberties. We identify important legal gaps in coverage, leaving the general public susceptible to invasion of privacy through the circumvention of civil liberties. We supplement our analysis by surveying the general public and provide our recommendations of a Mental Trespass Act in Section IV.

A. Potential Harms

In developing lie detection technology, the goal is to balance the negative impact of lying with the negative impact of lie detection. In order to have a better understanding of the ethics of lie detection, it would be prudent to have an account of why lying is wrong, and in which cases it may be considered morally justified.

In her book “Lying: Moral Choice in Public and Private Life,” Bok [10] outlined a principle of veracity to explain why telling the truth is more beneficial to society than lying. Bok shows that human interactions rely on assuming others are telling the truth. If we had to assume everyone else was lying, the goals of our communication would be undermined and all the information we received from human sources would need to be verified by nonhuman sources. Therefore, a system of assuming truthfulness benefits everyone, and in order for that system to function, its members must be truthful.

Bok, admitting that not all lies are malicious, also outlines conditions that must be met in order for a lie to be justified. A full account of the ethics of lying is beyond the scope of this article. We are concerned with the harm lie detection can cause to society. As evidenced by Bok’s conditions of justified lying, there exist cases in which lying is not only permissible but also even ethical. Cases that are particularly important to consider are those in which people lie to protect themselves or others from violence or persecution. Widespread lie detection has the potential to put these people in danger. It is our belief that if the development of noninvasive lie detection technology is inevitable, we have a moral obligation to contribute to its development for use in ethical settings. Two key areas posing direct threats to public well being are the accuracy of the lie detection technologies in question and their ability to invade personal privacy.

1) Accuracy: Accuracy poses a major problem to lie detection being applied to real-world problems. A 100% accurate lie detector is impossible to achieve with computer vision and language processing techniques, which use machine learning classifiers for the task of lie detection. These algorithms are trained using data (mainly video, audio, and text) of people lying. The algorithms can become increasingly accurate with the acquisition of more data; however, accuracy remains severely limited in cases not well represented. Even with a theoretically infinite amount of data, the prediction will have an inherent, irreducible amount of error known as Bayes error [32] due to the likely randomness present in signaling deception.

An inaccurate lie detector used in law enforcement could result in both guilty people being released and innocent people being convicted. Even a seemingly small error rate can have a massive impact on a large population in the case of criminal suspects. Schauer [69] estimated that a 1% wrongful conviction rate in the U.S. could result in 10 000 innocent people a year being convicted, and 4000 of those being sentenced to prison.

Clearly in cases where human lives are at stake, it is important to achieve as high of a level of accuracy as possible. The polygraph, once widely used and trusted, has now been panned as fallible and not completely accurate. Across 50 studies and 3099 examinations conducted in a laboratory setting, it was shown to have a median accuracy of 86% [22]. Any deception detection technology meant to replace or supplement the polygraph seemingly must have a significantly higher accuracy than this.

It is also important that computer vision-based deception detection techniques be unbiased across all populations. For the technology to be used by law enforcement, it should work equally well for all people, regardless of their race, gender, or age. Algorithms that are not trained with sufficiently diverse data can suffer when they are used on people not well represented in that data. Buolamwini and Gebru (2018) tested three algorithms trained to recognize gender in a diverse population, including lighter and darker skinned men and women. Their study found that while all three classifiers had a less than
1% error rate on photos of light-skinned men, their error rates ranged from 20%–34% for darker skinned women [15]. Therefore, it is essential that the data collected and used to train these lie detection algorithms be diverse enough to enable it to be relatively unbiased. Additionally, the algorithms should be tested regularly to ensure that they are equally accurate across all demographics. Without these technologies achieving a high degree of accuracy that works equally well across cultures, society risks wrongfully convicting individuals in a systematically biased way.

2) Privacy: The existence of a lie-detecting technology that can be used without interacting with someone at all raises significant questions about privacy and consent. What are the dangers of using lie detection on someone without their consent?

On the one hand, none of the information being collected is private. An algorithm analyzing someone’s facial expressions, voice, and speech patterns to determine if they are lying in using external observations that any person could utilize. By analyzing these aspects without the subject’s consent, would people be committing offensive actions against another person? Intuitively, it would appear as though this is not inherently doing anything wrong. This line of reasoning would indicate that a noninvasive lie-detecting algorithm could also be used on someone without requiring their explicit consent.

However, if the underlying technologies continue to advance, these algorithms could provide evaluations that are much more accurate than any human evaluation could possibly be. Falling under the category of “specialized technology,” even though both accurate truth metering as well as accurate thought exposing only use the knowledge available to humans, they can perform analysis at superior-to-human levels. Taking this into consideration, these technologies should not be considered equivalent to humans in their position.

An accurate enough thought exposing technology could interfere with the ability of people to think freely in their private mental sanctuaries. As mentioned earlier, it could reveal information that someone might normally keep private, such as their sexual orientation or their political views. If their thoughts about these matters could make them a target in the society in which they live, the widespread availability of noninvasive lie detection could limit not only the expression of these views but also the mere thought of them.

Ienca and Andorno [41] postulated four new rights as a result of emerging neuroscience technologies such as fMRI: “the right to cognitive liberty, the right to mental privacy, the right to mental integrity, and the right to psychological continuity.” We strongly agree, however, it is not clear whether violations of these new rights would be raised by the use of computer vision/language processing-based lie detection, as it is noninvasive and involves no direct brain scanning.

Conversely, Amy E. White, in her article, “The lie of fMRI: An examination of the ethics of a market in lie detection using functional magnetic resonance imaging,” disagrees “Functional MRI scans are not mind reading.” She writes, “They, not unlike traditional polygraphs, measure physical changes in the human body” [84]. White’s argument is that these two lie detection technologies essentially do not detect lies, but simply measure automatic responses correlated with deception. As computer vision and language processing techniques use this same process, White’s qualification would also exclude them from falling under the category of thought exposing. Without proper legal protections, individuals could be subject to their private thoughts becoming public knowledge.

B. Laws and Limitations: Existing U.S. Federal and State Laws

In this section, we discuss various legal issues with the public use of noninvasive deception detection technology without an observed party’s consent. The 1948 Universal Declaration of Human Rights has explicit language regarding human rights to “privacy,” with Article 12 of the declaration stating “No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence…” [60].

While it is noteworthy that the notion of such a right to privacy was considered so important as to be codified in the Universal Declaration, such declarations are left largely impotent with the interpretation of what constitutes “arbitrary interference” and privacy left undefined. An examination of international law is further limited by the fact that individual protections depend entirely upon a state’s willingness to comply with the international law [40]. Different international jurisdictions have approached privacy rights in widely differing ways. The EU is seen as having strong protections of privacy on the basis of human dignity compared to the minimalist privacy protections in the U.S. which are based on protecting individual liberty from government authority [54].

In this section, we focus on United States Law as the basis of our analysis. In addition to the minimal privacy protections offered under U.S. law, focusing on the U.S. is particularly well suited given the recent revelations of U.S. government and U.S. industry intrusion upon privacy as demonstrated through the 2013 Snowden and 2018 Cambridge Analytica scandals [40].

1) Fourth Amendment: In the United States, perhaps the most relevant legal issue with regard to public deception detection is raised with regard to the fourth amendment. The fourth amendment establishes the “right of the people to be secure in their persons… against unreasonable searches” and has been interpreted to prohibit searches when there is a reasonable expectation of privacy [5]. Several cases have established that in general there is no reasonable expectation of privacy for things that are in plain view in a public area. For example, the U.S. Supreme Court held that garbage that is left out on the curb can be searched without a warrant in California v. Greenwood [20], [38], [73]. This has been extended to include the use of some specialized equipment, particularly the use of a plane for aerial observation of someone’s backyard in California v. Ciraolo [27], and observation of an open field in an industrial complex with a high definition camera in Dow Chemical Co. v. United States [46], [67]. The court seemed to indicate the relevance of whether the equipment was available to the public, in one case finding that the EPA did not violate the fourth amendment when it “was not employing some unique sensory device not available to the public.” An analysis
of the smells during a routine traffic stop with a specialized drug-sniffing dog was also found to not constitute an unreasonable search in Illinois v. Caballes [8], [24], [72]. However, the ability to observe someone from a public area is not absolute. The Supreme Court found in Kyllo v. United States [2], [31], [70] that viewing a person’s home from outside with a thermal imaging camera (to determine if high-temperature drug growing lights were used) was indeed a violation of one’s “reasonable expectation of privacy.” In light of these Supreme Court cases regarding fourth amendment rights, how would we expect the use of a video-based lie detection apparatus to play out? One perspective is that an individual’s facial expressions are in plain view and thus do not carry a reasonable expectation of privacy, as in California v. Ciraolo regarding a person’s backyard, or someone’s garbage on the curb in California v. Greenwood. It is also likely that the camera used for deception detection needs to not be more advanced than the high-resolution camera deemed to be acceptable in Dow Chemical v. the United States. However, lie detection does involve the use of state-of-the-art AI-driven algorithms and computer vision/language processing techniques. It seems conceivable that a court could find such algorithms invasive in how they are uncovering someone’s internal mental state. Additionally, we may expect a court to consider, as it did in Dow Chemical Co. v. the United States, whether the equipment used is publicly accessible. Thus, whether such lie detection technology is made public or not will possibly affect whether its use constitutes a fourth amendment search (e.g., if it is made available to the public, the government would not be using “specialized technology”). However, it should also be noted that fourth amendment issues are limited to the government (or people working on behalf of the government) and do not apply to the public at large.

2) Fifth Amendment: In addition to the potential fourth amendment issues, the use of lie detection technology in a court of law by the prosecution may bring up Constitutional Law issues with regards to the fifth amendment protection against self-incrimination. The fifth amendment provides that “[n]o person shall be ... compelled ... to be a witness against himself” [6], [66]. We foresee that use of a lie detection technology without a subject’s consent may be interpreted as compelled testimony. However, the courts have interpreted the fifth amendment narrowly, giving the prosecution the right to compel the accused to provide a password to encrypted computer data [16], [81]. Additionally, the courts have determined that a suspect may be compelled to produce fingerprints, blood, and fingernail scrapings without violating the fifth amendment [42]. Furthermore, courts have even found that compelling a witness to provide a voice sample for identification does not trigger fifth amendment protections [83]. Thus, we believe that it is unlikely that a court would find the use of an AI-driven lie detection technology to be a violation of one’s fifth amendment rights. However, in certain contexts perhaps this is not the case. For example, Thompson [78] argued that highly invasive lie detection technology, such as unconsented application of the fMRI, is likely to violate the fifth amendment due process law as it “shocks the conscience.” Therefore, we take the stance that the degree of invasiveness is what fundamentally defines this question of violating the fifth amendment. We bring to light in this article that highly accurate, noninvasive lie detection technologies are not only imminent but also their risk of infringing upon our civil liberties is much greater. This is due to the fact that noninvasive lie detection devices are able to lie detect nonconsenting individuals from a distance. Furthermore, it is not clear whether these noninvasive methods would “shock one’s conscience” enough to violate the fifth using Thompson’s methodology.

3) Employee Polygraph Protection Act: The employee polygraph protection act (EPPA) prevents private employers from requiring job applicants or current employees to submit to a lie detector of any kind but allows polygraphs to be used by certain sectors, namely, government and security positions. However, according to its website, the fMRI-based lie detection company No Lie MRI “measures the central nervous system directly and such is not subject to restriction by these laws.” As noted by Greely and Illes [34], the language used in the provision of the legislation is broad enough that loopholes like this are possible. Without an explicit amendment or judicial review, No Lie MRI could continue to offer its services to employers, violating the intention of the EPPA, but not the text of the law. The EPPA is limited to employer–employee relationships and is silent with regard to public use.

4) Invasion of Privacy Laws: The strongest limitations on the public use of a noninvasive lie detection technology arise from state law. While it is difficult to analyze each state’s laws individually, a concise restatement of the preferred rules used by a majority of the states is available in the “Restatement of Law,” written by the American Law Institute (ALI). The restatement provides the law of intrusion upon seclusion, commonly referred to as “invasion of privacy,” which makes liable one who “intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another or his private affairs or concerns... if the intrusion would be highly offensive to a reasonable person.” This liability extends even when there is no publication or use of the information obtained in violation. In general, surveillance from a public place is not an intrusion upon seclusion, however, exceptions to this rule exist. In summary, it is unclear if when accurate noninvasive lie detection arrives, it will be legal to use on nonconsenting individuals caught unawares.

5) Court System: In the federal court system, it is unclear whether even a highly accurate lie detector would be admitted as evidence. The modern polygraph is an instructive example to illustrate this point. Despite being grounded in scientific principles, many, including John Larson himself, questioned the merit of the polygraph in measuring honesty. John Larson stated: “Beyond my expectation, through uncontrollable factors, this scientific investigation became for practical purposes a Frankenstein’s monster, which I have spent over 40 years in combating.” Just two years after the common polygraph saw its first practical use in a criminal investigation in 1921, the American judicial system developed the legal doctrine that would almost completely bar the polygraph from ever entering the courtroom again. The Frye standard prevents evidence from being presented unless it is generally accepted as reliable in the relevant scientific community [9], [74]. Per this standard, computer AI-based lie detection would likely not be admitted in its current state, as the underlying technology is
still developing and the accuracy of this method of lie detection has not been well established.

The Daubert standard, which has largely superseded the Frye standard in both federal court and most state courts, sets stricter guidelines for evidence being admitted. In addition to general acceptance, the Daubert standard also considers whether the scientific evidence has been tested, whether it has been peer reviewed, and whether it has a high rate of error. AI-based lie detection would likely be kept out of courtrooms according to the Daubert standard due to its current lack of widespread testing and peer review. These two standards ensure that the courts are well-equipped to keep potentially inaccurate scientific evidence out of the courtroom.

Whether or not a technology is admitted into the courtroom is of the utmost importance for the following reason. Historical review shows that once a technology is deemed as legitimate or as questionable, such characterizations are unlikely to be changed [78]. For example, even though both fingerprinting and polygraph analysis have a questionable scientific basis, fingerprinting has been generally admissible in court since 1911 [1] while polygraphs have not. Overall, the Daubert standard has prevented polygraph admissibility for concerns over scientific legitimacy amongst the others mentioned [55]. Given that recent legal analyses argue that fMRI should not be allowed at this time [48], [51], [57], [87], it is likely the Daubert standard will keep these technologies out of the courtrooms as well for the time being. Scholars have gone on to argue for the urgent need to regulate developing lie detection technologies, such as the neuroscience-based technologies upon which fMRI lie detection is one flavor [34], [57]. Prior analysis has come to the conclusion that looking at technologies like fMRI through the analogy of blood test and/or forced testimony is inappropriate, arguing that “the implicit assumption of mind–body dualism, which underlies this thinking, is dated and, most likely, no longer tenable” [78]. Scholars have argued the importance of considering the legal implications of an advancing technology before it becomes ubiquitous, Thompson [78] stating “if the existing scientific literature is indeed a harbinger of an important new technology, it will be to society’s benefit that some thought have been put into its implications before its wide scale deployment.”. The topic of advanced lie detection has received recent attention in ethical and legal contexts [28], [34], [51], [57], [77], however, precise definitions of the technologies in question and proposed legal doctrines offering a solution have yet to be fleshed out with enough granularity to be effective. With the legal doctrine and case history being classified as ambiguous at best, there is clearly a strong societal need to formally define what should be allowed regarding these evolving technologies.

C. Public Perspective

To understand public opinion on the usage of these lie detection technologies, we sampled the population by conducting a survey using the crowdsourcing platform Amazon Mechanical Turk. We included demographic information on the survey and based on the responses, launched multiple surveys with participation requirements to ensure that the demographic distribution of our respondents resembled the demographic distribution of the United States. We believe matching the distribution is essential not only for obtaining a reliable sample but also for obtaining one where we can reasonably extrapolate to claim any kind of generalized opinion. We monitored the quality of the survey responses by implementing a control question (e.g., control question: “Answer strongly agree to this question” all respondents failing to answer this question correctly were removed from the data). We also incorporated a required free-text response question to our survey and removed responses where the length of the response was less than ten characters in length. After all unsatisfactory data points were removed, we were left with n = 129 quality responses.

IV. RESULTS

A. Public Perspective Survey Responses

We set out to investigate whether the public opinion was in favor of or opposed to the legality of using these technologies on an individual without first obtaining their consent. This is crucially important for noninvasive lie detection technologies as they can be used on an individual caught blissfully unaware. Two multiple-choice questions asked respondents their level of agreement on whether police should be allowed to use a computer program to detect lies in a criminal suspect when the accuracy levels of the device were 100% and 90% (see full question text below). The results from the survey responses shown in Table I were strongly indicative of opposition to unconsented usage when the accuracy of the device was not absolute.

Survey Questions:

Q1: If there were a 100% accurate computer program to detect lies from a video recording, police in the U.S. should be allowed to use it on criminal suspects with their knowledge in an interrogation room, but without requiring the suspect’s consent.
Q2: What if the computer program were only 90% accurate?
Q3: Please briefly explain your answers.

Survey questions Q1 and Q2 had five response options: strongly agree, agree, neutral, disagree, and strongly disagree. We combined the agree and strongly agree responses to find the number in favor as well as the disagree and strongly disagree responses to find the number opposed. We conducted a proportions statistical test with our null hypothesis being that there is no difference in public opinion regarding this question (e.g., the number in favor is equal to the number opposed). After running the statistical test, the probability of that null hypothesis given our data was $p = 0.0001$ or 1/100th of a percent (0.01%). We thus reject the null hypothesis of there being no difference in public opinion and accept the alternative hypothesis that there is likely a difference. This means a majority of people are opposed to the unconsented use of these lie detection technologies when the accuracy is not guaranteed. Indeed, by looking at the differences between Q1 and

<table>
<thead>
<tr>
<th>Question</th>
<th>#Agree</th>
<th>#Disagree</th>
<th>Neutral</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>73</td>
<td>44</td>
<td>19</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Q2</td>
<td>33</td>
<td>69</td>
<td>27</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>
Q2, we see that the accuracy of the device determined the polarity of the agreement to unconsented usage of lie detection technologies. Interestingly, the public perspective is that the police should use a 100% accurate lie detection device, however, should not use the device if accuracy drops to 90%.

Some of the free-text responses that illustrate public sentiments are as follows.

1) “If there was no margin of error then it would be acceptable. There is always that small percentage falsely accused and I would not be comfortable with a machine making a determination.”

2) “I think it opens the door to more and more invasive policies.”

3) “I believe it should be used in law enforcement, as it will help 10 fold in reducing and finding criminals.”

4) “Seems like to much Big brother to me.”

5) “THESE TACTICS ARE AS LIKELY TO BE USED OR INTERRUPTED INCORRECTLY AGAINST A LAWFUL CITIZEN.”

6) “This is way to Orwellian for me! Now an AI computer program detecting lies? What if this is hacked and made to work contrary to the original program? I say no to this.”

Based on public opinion, there is certainly concern over unregulated lie detection technology being used maliciously and we have an obligation as a society to mitigate that outcome. We are hopeful that our proposed Mental Trespass Act and recommendations for updating the language in the EPA to reflect our technology definitions would greatly aid this communal effort.

B. Recommendations

1) Definitions and Proposed Mental Trespass Act: In providing legal recommendations on how to mitigate the potential harms and ambiguity in the field, we first define two types of relevant technology as well as the different categories of their usage. We distinguish two major classes of lie detection tools, including: 1) accurate truth metering and 2) accurate thought exposing (depicted in Fig. 1).

Accurate truth metering is defined as “use of a device to measure an individual’s level of belief in an intentional statement made by the individual, with the device usage having an accuracy exceeding typical human performance.” A statement broadly includes spoken utterances, written text and drawings, bodily gestures, and other forms of communication. An intentional statement requires that the statement maker has the mental intent to make the statement. Thus, a spontaneous gasp of surprise, or the unconscious blushing after hearing a question is not intentional statements. In defining accurate, we use an excedat-hominem standard, i.e., a level of accuracy which clearly exceeds typical human ability. Thus, in defining an accurate truth meter, we consider the numerous studies on human performance regarding lie detection and note that this level of accuracy has been found to be approximately 54% [11], even amongst expert judges [12].

Accurate thought exposing is defined as “use of a device to expose an individual’s thoughts, without the individual’s consent, with the device usage having an accuracy exceeding typical human performance.” Accurate thought exposure specifically includes instances of questioning a suspect without consent and accurately measuring the suspect’s physiological response to the questions. As with the definition of truth metering, the definition of accurate thought exposure requires a level of accuracy which clearly surpasses human ability. A primary distinction of truth metering and accurate thought exposing is that truth metering requires an overt/intentional statement by the individual regarding the issue being observed (Fig. 1(b) and (c) illustrates this distinction). For example, in asking an individual what time it is, evaluating whether they are being honest about the time involves only truth metering. However, if one then uses a system to gauge the level of anger in their voice, the technology has crossed the boundary into the realm of thought exposure because the overt response to the question being asked does not pertain to anger. Similarly, if an individual is talking out loud to others in a public area on his/her own accord, and we evaluate the honesty of each of his/her overt statements, we are truth metering. However, if the individual’s statements do not directly involve their emotions, and we determine that the individual feels high levels of arousal we are thought exposing (noting that a human observer would typically not be able to discern that information).

In addition to the two different classes of noninvasive deception detection technology, it is important to independently consider whether the use is in the context of: 1) a criminal investigation; 2) pertaining to one’s employment; or 3) a “public use.” Within the context of criminal investigation, we consider not only direct involvement in a criminal trial but also any police interrogations which led to the arrest, as well as any gathering of evidence or a crime either with or without probable cause by an agent of the state. The employment context involves both current employees of a business, as well as interviews of prospective employees. Within public use, we also consider uses by commercial entities in interactions with customers, even though the action may occur in a private location. We concur with Greely and Illes that lie detection technologies and services must be regulated to prevent harm. Specifically, we believe that a federal Mental Trespass Act should be passed which:

1) provides a general ban on the use of accurate thought exposing on an individual without the individual’s consent;

2) makes an exception to this ban for use of accurate truth metering on individuals in a public space, as long as the particular usage would not be found offensive by a reasonable person;

3) updates the EPPA to explicitly include accurate thought exposing and accurate truth metering, even when such devices are noninvasive.

V. DISCUSSION

As much as these technologies have the potential to infringe upon people’s civil liberties, there are instances where the proper use of sophisticated, AI-driven sensing technologies and their associated algorithms can provide benefits to society.
A. Aiding Fairness and Justice in Court

Although the Frye and Daubert Standards keep inaccurate truth metering technologies out of courtrooms, there are approaches that can be taken to minimize unfairness issues raised. We demonstrate these aspects using two components: 1) an interview with a judge and 2) the establishment of essential design primitives for developing accurate truth metering technologies.

1) Interview With Judge: In order to get an expert opinion on the potential impact advances in lie detection technologies could have on the courts, we interviewed standing County Judge Dennis Cohen of Livingston County, New York, who has 12 years of experience on the bench.

On the topic of emerging technologies in lie detection, Judge Cohen said “I think it is a big area of advancement in law, and could help us resolve cases and work through investigations... just looking at what high-resolution cameras have done for us with law shows that we can often identify the right culprit or prove that something happened or didn’t happen.” Judge Cohen went on to say “Our whole society is changing because of technology. If it could be determined to be reliable... then it could open up a whole new phase of things.” When asked about his opinion on relating the polygraph to these developing technologies threats and their associated threats of unreasonable searches, Cohen remarked “Polygraphs are voluntary. This [referring to these developing technologies] would also be a voluntary procedure as well, at least for the foreseeable future. Therefore, it would not ever reach the bounds of an unreasonable search.” Here, we see an important point brought up that when consent has been unquestionably obtained from an individual, usage of the polygraph or technologies to replace the polygraph never constitute an unreasonable search. However, the utility of such technologies designed in this way is vast if not completely diminished due to their less than perfect accuracy. For example, polygraph tests and their results are almost entirely inadmissible in a federal court under evidentiary rules. Polygraph results are what is known as “highly prejudicial,” meaning that regardless of the test’s accuracy or even its relevance to the case at hand, hearing about it will bias the jury. This means that if the polygraph indicates that the defendant has lied, despite its questionable accuracy, a jury may treat that as definitive proof that the defendant lied. Additionally, if the jury believes the defendant lied about material facts related to the case, that may indicate proof of guilt, no matter how relevant or irrelevant those facts are to the defendant’s innocence. For these reasons, it is possible that even a 99% accurate lie detector could be excluded from evidence, due to a judge fearing the jury will treat it as 100% accurate. Thus, it is prudent that in developing these technologies, that an entirely different approach be taken in their design primitives, development, and deployment.

2) Essential Design Primitives: If accurate truth metering and/or thought exposure is used by law enforcement, it should be equally effective across all races and genders. Therefore, it is the responsibility of individuals who are researching and developing this technology to collect diverse data. We believe this could even be encouraged/enforced by federal funding guidelines for those who are studying deception detection using AI. In order to receive federal grants for this purpose, labs could be required to meet certain diversity standards in the data they collect and use in their deception detection algorithms. Additionally, the performance of said lie detection technologies should be standardized across all law enforcement entities.

Another relevant issue is how to maximize accuracy (as well as the ability to deploy such devices in court rooms) while preserving investigator autonomy. One solution proposed by Kleinberg et al. [49], in their prediction framework for whether judges should jail or release criminal defendants on bail, is to integrate the machine into the existing procedure, creating a human–machine symbiosis. Instead of having the algorithm make all the decisions, the algorithm should give the people that are using it more information for them to make informed decisions themselves.

In this vein, it is our suggestion for lie detection researchers to create an output that is nuanced and detailed, rather than a binary 1 for “lying” and a 0 for “not lying.” The lie detection device should detect and display indicators of deception when they appear. This fundamentally changes the role of the device. Instead of performing the evaluation based on an arbitrary decision boundary, it acts as a tool to assist people in doing the evaluation themselves. To interpret these more nuanced results, trained human operators should be employed. The use of such operators could even be required for the technology to be used. These operators should understand how to interpret the output and convey that information to investigators, while also understanding and conveying potential biases in certain questions as well as the potential for inaccuracy in the technology.

B. Elaborations on Recommendations

While dishonesty might frequently be harmful to people and society as a whole, we do believe that people have the right to exercise their ability to lie in some circumstances. It is important to properly balance an individual’s right against unreasonable search and invasion of privacy with another individual’s right to know the truth by using a lie detection technology. As defined in our recommendations, a truth metering device only operates on an individual’s intentionally made statement. We argue that by uttering a voluntary statement, a speaker is inviting such a statement to be evaluated and inadvertently provides consent for that process. Respecting the fact that such consent may be unintended, we believe it is necessary to strike a proper balance and that the use of the truth metering devices must be limited to nonoffensive cases. For example, if a reasonable person would find the usage offensive, it should not be allowed (similar to the invasion of privacy principles outlined in the restatement of the law). In contrast, a thought-exposing device gives the power to go beyond evaluating the veracity of an individual’s statements, potentially exposing one’s private, innermost thoughts. For this reason, we not only suggest a complete public ban on the use of accurate thought-exposing technology but also regulation on the production and dissemination of such technology. Through establishing these regulations, we not only prevent potentially malicious uses but we also offer further protections for the people against unreasonable searches of their private mental sanctuaries. Recall in the case of Dow Chemical Co. v. the United States, because the observation of the industrial
complex was done through a high-resolution camera and the general public had access to that technology, the court ruled that this did not constitute the bounds of an unreasonable search. By restricting public access to these emerging thought-exposing technologies, we thus prevent this legal precedent to be carried out in the future. It is our position that truth metering devices could remain available to the general public, as long as they were limited to uses that would be deemed nonoffensive to a reasonable person. This would allow them to be used for lie detection in contexts, such as navigating a foreign environment and dealing more fairly and justly with children. We formally take the stance that thought exposure technologies be regulated for the general public (potentially by using a permit schema that is externally audited by multiple third parties relatively frequently), and that their unconsented use be codified as an illegal mental trespass.

VI. CONCLUSION

Accurate noninvasive deception detection likely does not currently exist, although it is probably closer than most of us think. The technology’s ambiguous legal status deems it necessary to establish guidelines before it is fully developed and available. The introduction of AI-driven, advanced sensing technologies for this task raises new concerns regarding privacy and consent due to their noninvasive nature. Defining the technologies precisely as accurate thought exposing and accurate truth metering technologies is essential for proposing an airtight legal doctrine to safeguard our civil liberties appropriately. Otherwise, potential loopholes could emerge in the future causing harm to society and bypassing the intentions of the law and the protections that it offers (as is the case currently with No Lie fMRI and the EPPA). Emerging lie detection technology will be a powerful tool, benefiting the criminal justice system, the medical community, and many others. In order to utilize it to its fullest potential, however, it must be developed and used responsibly with the necessary restrictions—or it may end up doing more harm than good.

References
