UC Berkeley Criminal Law & Justice Center

Recommendations for the Use of AI Police Report Technology and CVSA

Berkeley, Jul 9, 2025

Executive Summary

The Berkeley Criminal Law & Justice Center (CLJC) investigated police use of Computer Voice Stress Analyzers (CVSA) and generative artificial intelligence (AI) for purposes of police report writing in Berkeley. The Center makes the following recommendations:

1. Berkeley Police Department (BPD) should cease its use of CVSA and decline to renew any related contracts, because the technology has not been proven to work and creates civil liberties risks for civilians as well as labor and employment risks for officers.
2. If BPD proposes adopting Draft One or similar AI police report writing software, City Council should await further research on the risks and benefits of the technology or partner with a research institution to design a pilot program to evaluate the technology as part of a tiered rollout.
3. If further research shows that the benefits of the technology outweigh the risks, City Council should create a policy for the future use of AI police reporting technology under its surveillance technology ordinance, § 2.99, with the following best practices:
    a. All reports written using AI technology should include a statement to that effect.
    b. Police officers should review and all AI generated reports in which they are involved, then digitally sign to attest to their accuracy under penalty of perjury as with current report protocol.
    c. The council should create and measure metrics for the technology's success, and adjust the § 2.99 surveillance technology use policy or cease its use if those metrics are not met.

Introduction

The Office of the Director of Police Accountability (ODPA) presented a potential policy review to the Police Accountability Board (PAB). The PAB, by majority vote[1] accepted the proposal for a policy review to be conducted by the ODPA. The PAB did so under its Charter[2] authority. This present report was initiated at the request of Hansel Aguilar, Director of Police Accountability for the City of Berkeley. Berkeley voters established the role of Director of Police Accountability to "provide oversight of local law enforcement, access records, investigate complaints from the public and recommend disciplinary actions."[3] Within the scope of this role, Director Aguilar provides oversight of the use and adoption of police technology.

The Criminal Law and Justice Center (CLJC) is a research center based within Berkeley Law with institutional expertise in the areas of policing and technology in criminal law. In fall of 2023, the CLJC investigated Berkeley's Community Control of Police Surveillance (CCOPS) ordinance and Berkeley Police Department (BPD)'s use of automatic license plate readers, GPS trackers, body-worn cameras, unmanned aerial vehicles, and fixed-location surveillance cameras. On March 7th, 2025, the CLJC hosted a conference focused on the use of artificial intelligence and other emerging technologies in the administration of criminal law.

The CLJC supports data-driven use of policing technologies, including novel technologies, for public safety purposes with common-sense guardrails to protect civil liberties. In keeping with that approach, this report draws on research from police technology manufacturers and civil liberties groups, criminal case law, and conversations with experts throughout the field of law enforcement. The first section of this report gives background on the legal framework for police use of technology, significant policy considerations, and CLJC's previous related findings in Berkeley. The second section discusses CVSA technology, issues with its use, and potential alternatives. The third section discusses available research regarding AI police reporting, related risks and legal issues, and suggests guardrails for its implementation. Finally, the concluding section summarizes CLJC's recommendations for next steps for the City of Berkeley to optimize public safety while protecting civil liberties in policing.

[1] 2024-05-22 PAB Meeting Minutes: https://berkeleyca.gov/sites/default/files/legislative-body-meeting-minutes/2024-05-22 Regular Meeting Minutes - signed.pdf - Page=4
[2] Berkeley City Charter Municipal Code 125 (13)(d): https://berkeley.municipal.codes/Charter/125(13)(d)
[3] Berkeley, CA Measure II, Police Accountability Charter Amendment (November 2020).

I. Evaluation Framework

This section non-exhaustively outlines some of the important legal, constitutional, and policy issues relevant to police use of technology. Sections II and III then apply those rules to the technologies at issue.

A. Legal Framework

Police use of emerging technologies must comply with federal, state, and local laws. The Supreme Court has held that police are subject to the constraints of the Fourth Amendment even when they use novel technologies.[4] When cities fail to constrain the use of novel technologies to comply with constitutional law, they risk costly civil rights and privacy lawsuits in unpredictable areas of law.

All police use of surveillance technology in Berkeley is subject to the city's CCOPS ordinance, section 2.99.[5] AI police report writing technology is a "surveillance technology" as defined by the ordinance because it is capable of capturing information associated with individuals or groups.[6] While AI police reporting technology does not literally have sensors to capture images or audio, its perceptive capacity relies on such technology because it reviews body worn camera footage and generates a novel form of information about groups and individuals.

Section 2.99 requires all surveillance technology used by BPD must go through an approval and assessment process.[7] City council must hear from various parties regarding the risks and benefits and proposed usage policies for the technology prior to its adoption.[8] City council approves technology only if it determines that the benefits outweigh the costs and concerns, appropriate safeguards are in place, civil liberties and rights are not unduly burdened, and the city does not have a reasonable alternative to the technology's use.[9] BPD is required to

[4] *See, e.g.*, *Kyllo v. United States*, 533 U.S. 27 (2001) (holding that police use of surveillance technology to gather information within a subjects home was a Fourth Amendment search.)
[5] Berkeley, Cal., Municipal Code § 2.99 (2018).
[6] *Id.* at § 2.99.020.
[7] *Id.*
[8] *Id.* at 2.99.030.
[9] *Id.* at 2.99.060.

submit detailed annual reports about its use of the technology, and city council may modify the usage policy as needed.[10]

B. Methodology

This report analyzes technologies through three primary lenses: validity and accuracy to achieve the stated objectives; ability to improve or enhance law enforcement work and responsibilities; and trade off with privacy and civil liberties. To assess validity and accuracy, this report relies on existing empirical research as well as information put forth by the creators of the technology, prominent civil rights groups, and the CLJC's own in-house expertise.

II. Berkeley Police Department should cease all use of Computer Voice Stress Analyzer technology, because it has not been proven to work.

Berkeley Police Department currently uses Computer Voice Stress Analyzer technology, which available research has proven to be ineffective. Berkeley should cease all use of this technology.

A. Effectiveness, risks, and benefits

Computer Voice Stress Analyzers, or CVSA, are hardware and software technologies that purport to verify whether a human speaker is lying or telling the truth by identifying "micro tremors" in the speaker's voice.[11] The manufacturer of CVSA technology, NITV (formerly known as the National Institute for Truth Verification, prior to a name change), claims that its most recently updated version, CVSA III, is "[t]he most accurate truth verification system in the world."

However, research shows that CVSA lie detection is about as accurate as random chance. A study funded by the National Institute of Justice found that CVSA detected only about 15% of lies about drug use in a field test.[12] The study also found that the test determined that 8.5% of

[10] Berkeley, Cal., Municipal Code § 2.99.070 (2018).
[11] *Computer Voice Stress Analyzer (CVSA)*, Hannibal Police Department, https://www.hannibalpd.com/about-us/computer-voice-stress-analyzer/#:~:text=Computer%20Voice%20Stress%20Analyzer%20(CVSA)&text=The%20CVSA%C2%AE%20uses%20only,saves%20each%20chart%20to%20file.
[12] Kelly R. Damphousse, Voice Stress Analysis: Only 15 Percent of Lies About Drug Use Detected in Field Test, National Institute of Justice Journal (March 16, 2008), https://nij.ojp.gov/topics/articles/voice-stress-analysis-only-15-percent-lies-about-drug-use-detected-field-test.

truthful participants were lying.[13] Another randomized study had half of participants steal money from a box, and the other half not steal, and then had half of all of the subjects lie.[14] CVSA accurately detected the truthfulness of the subjects' answers almost exactly 50% of the time: the same amount as if it had randomly guessed by chance alone.[15]

NITV argues that CVSA can be effective as an interrogation technique, but its use for interrogation creates a significant risk of false confessions. NITV's website highlights a study on the use of CVSA to elicit confessions in child sex crimes where offender admissions are "the only reliable way to prove guilt."[16] The authors used CVSA to identify vocal stress among adults suspected of child pornography possession, and asked further pressing questions if vocal stress was detected.[17] But many forms of truth detection generate false confessions by making the subject believe that their true statements are not believed and will lead to further consequences, especially among young or unsophisticated subjects.[18] Further, studies show that individuals who administer truth tests cannot differentiate between true and false confessions.[19]

B. Recommendations

Because CVSA produces results no better than chance and risks false confessions, the City of Berkeley should ban its use for any police purpose. The San Francisco Chronicle reported the same in a June, 2024 article highlighting the same issues with the technology that are outlined here.[20] For the above outlined reasons, CVSA is also not admissible in court, and the California Department of Corrections and Rehabilitation has agreed to cease its use.[21]

---

[13] *Id.*
[14] M. J. Janniro & V. L. Cestaro, *Effectiveness of Detection of Deception Examinations Using the Computer Voice Stress Analyzer*, 27 Polygraph 28 (1998).
[15] *Id.*
[16] Marigo J. Stathis & Maria M. Marinakis, *Shadows into Light: The Investigative Utility of Voice Analysis with Two Types of Online Child-Sex Predators*, Journal of Child Sexual Abuse, https://www.cvsa1.com/app/uploads/2022/08/Stathis-et-al-2020.pdf.
[17] *Id.* at 6.
[18] Charles R. Honts et al., *Polygraph Examiners Unable to Discriminate True and False Juvenile Confessions*, 1 Polygraph & Forensic Credibility Assessment 1 (2019).
[19] *Id.*
[20] Susie Neilson & Matthias Gafni, *It's Been Called Junk Science as Accurate as a Ouija Board. Why Are Police Agencies Still Using This Tool?*, San Francisco Chronicle (June 4, 2024), https://www.sfchronicle.com/california/article/law-enforcement-technology-investigations-18756947.php.
[21] Susie Neilson & Matthias Gafni, *California Prisons Will Bar Lie-Detector Test Compared to 'Ouija Board' After Chronicle Investigation*, San Francisco Chronicle (August 26, 2024), https://www.sfchronicle.com/bayarea/article/california-prisons-technology-19724054.php.

The risk of CVSA harming civil liberties is high if it is used on civilians or for interrogation purposes. Even if CVSA is used solely within BPD for purposes of investigating officer misconduct, it risks harming good police officers through false positive test results. The right to fair misconduct evaluation processes is critical for effective policing. Officers need to know that any allegations against them will be resolved through a fair process, not by junk science. The risk of good officers being fired or disciplined in error due to CVSA essentially flipping a coin as to their truthfulness is too high to merit its use.

At best, CVSA is a waste of the City's money. At worst, it could lead to false confessions and harm good police officers. The Criminal Law and Justice Center recommends that the City adopt a policy banning the use of this technology and cancel all contracts for its use.

II. AI Police Report Writing Technology

Berkeley Police Department does not currently use AI police report writing technology. As explained above, this technology is subject to Berkeley's surveillance technology ordinance. City Council should consider the following risks and benefits in its assessment and creation of a surveillance technology use policy, should it decide to move forward with deploying this emerging technology.

A. Effectiveness, risks, and benefits

The only AI police report writing technology currently on the market is Axon's Draft One. Axon markets Draft One as a "force multiplier" that uses a large language model (LLM) to create draft report narratives "in seconds."[22] Much is still unknown about how Draft One, or any AI police reporting, will work in practice. The most important unknown information, explained further below, are: (1) the impact of underlying bias in ChatGPT and biased training data; (2) the frequency and types of error as compared to human writers; (3) actual impact to police efficiency; (4) Axon's data use; and (5) courtroom evidentiary issues.

Unknown 1: Will bias in ChatGPT 4 impact Draft One's transcription of audio and video?

[22] *Draft One*, Axon, https://www.axon.com/products/draft-one.

Draft One uses ChatGPT 4 to write reports, an LLM developed by OpenAI.[23] LLMs are created by inputting massive amounts of existing language-based data into a computer program, which the program uses to create future predictions about novel data.[24] This means that characteristics of the training data shape the LLM's perception of future data. Researchers have repeatedly shown that LLMs are subject to the same social biases as their training data, and can make representations, use stereotypes, and even use derogatory language.[25]

ChatGPT 4, which forms the basis of Draft One, interprets input with racial bias that is difficult to detect and not yet possible to correct for in the underlying technology itself. For example, a recent study found that ChatGPT 4 exhibits significant bias against African American Vernacular English.[26] There is insufficient research on whether this bias impacts ChatGPT 4's ability to accurately transcribe audio, as required for DraftOne. The task of transcribing audio may not be subject to bias because ChatGPT will not be invited to draw conclusions, but only state events in the audio.

Axon writes that Draft One is trained with "diverse data that represents actual use cases in the field."[27] This could lead to less bias than exists in ChatGPT's baseline model, because the model in Draft One will be specialized to the task at hand. However, bias exists in human-written police reports too,[28] which could imbue another layer of issues into the language model.

---

[23] Ian T. Adams et al., *No Man's Hand: Artificial Intelligence Does Not Improve Police Report Writing Speed*, 14 Journal of Experimental Criminology at footnote 1 (2024), https://doi.org/10.1007/s11292-024-09644-7.

[24] *What is a large language model (LLM)?*, Cloudflare Learning, https://www.cloudflare.com/learning/ai/what-is-large-language-model/.

[25] Isabel O. Gallegos et al., *Bias and Fairness in Large Language Models: A Survey*, 50 Computational Linguistics 1097 (2024).

[26] Valentin Hofmann et al., *AI Generates Covertly Racist Decisions About People Based on their Dialect*, 633 Nature 147 (2024) (finding that ChatGPT 4 showed a level of bias "similar to archaic human stereotypes about African Americans that existed before the civil rights movement." The authors applied strategies that reduce overt racism in ChatGPT results, and found that those strategies did not affect covert racism as found in their study.)

[27] *Responsible AI for Modern Policing*, Axon, https://s3.amazonaws.com/a.storyblok.com/f/198504/x/70c48646ef/axon_responsible-ai-for-modern-policing-whitepaper-06a.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIU627ENUQT4RW23A%2F20250328%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20250328T222958Z&X-Amz-Expires=604800&X-Amz-SignedHeaders=host&X-Amz-Signature=c3f1a00a70b9597701f79f5ce379c5c97b9216e7ddb162eefecda49dfa13f348.

[28] C. Dominik Guss et al., *Problems with Police Reports as Data Sources: A Researcher's Perspective*, 11 Frontiers in Psychology (2020).

Unknown 2: How often do officers catch and correct AI errors during Draft One's officer review process?

Large language models make two major types of factual errors: hallucination errors and attribution errors.[29] Hallucination errors occur when an LLM creates an output that is entirely false or not supported by accurate data. For example, in transcribing audio from body camera footage, an LLM could imagine and report speech that never occurred. Attribution errors occur when an LLM connects two facts that should not be connected. For example, the LLM could attribute an officer's verbal command to the person being commanded. Both types of errors should be caught by human review, but the effectiveness of Draft One's review process is still unknown. Critical safeguards at the police department level, like training, supervisor review, and signing reports under penalty of perjury, can help mitigate these risks.

Axon advertises that any AI mistakes will be detected by police officer review of AI generated reports.[30] However, past studies of technology in policing shows that officers trust information generated by technology even when told not to.[31] Further research is needed to show how officers actually interact with the technology in practice.

Unknown 3: Does Draft One actually improve police efficiency?

Axon markets Draft One as improving police efficiency. The CLJC's in-house criminal law experts report that police report writing is a timely task that takes away from officers' ability to do more public facing safety tasks. If Draft One is effective, it would improve police efficiency and save overtime costs for BPD.

Existing research on the impact of AI on efficient report writing is not conclusive. One 2024 study found that AI police reporting was no faster than police report writing without AI.[32] There, a police department agreed to randomly select a group of officers to use Draft One and measure the speed of report writing compared to a control group.[33] The study found that there

[29] Notes from Larson Justice Conference.
[30] *Draft One*, Axon, https://www.axon.com/products/draft-one.
[31] Nathan Freed Wessler, *Police Say a Simple Warning Will Prevent Face Recognition Wrongful Arrests. That's Just Not True.*, ACLU (April 30, 2024), https://www.aclu.org/news/privacy-technology/police-say-a-simple-warning-will-prevent-face-recognition-wrongful-arrests-thats-just-not-true.
[32] Ian T. Adams et al., *No Man's Hand: Artificial Intelligence Does Not Improve Police Report Writing Speed*, 14 Journal of Experimental Criminology (2024), https://doi.org/10.1007/s11292-024-09644-7.
[33] *Id.* at 8.

was no statistically significant difference in the amount of time officers spent writing reports with or without AI.[34] The authors explained that many police departments already use some boilerplate language, and Draft One may require officers to verify important details, making the amount of time involved in each similar.[35]

This study may be flawed by the increased time needed for officers to become accustomed to new technology. It did not assess the accuracy of AI police reports compared to boilerplate language. Further research is needed to determine how effective Draft One may be in practice.

Unknown 4: What will Axon do with Berkeley's data?

Axon gets access to customer data uploaded to Draft One, and can use it to train its model.[36] Axon's use of customer data is subject to its data agreement, which it has the right to change at any time.[37]

Berkeley has made many commitments to its residents that represent Berkeley values. Most notably, Berkeley is a sanctuary city[38] and has resolved to protect the right to abortion.[39] Berkeley police are subject to oversight by the Police Accountability Board and the Office of the Director of Police Accountability. Axon is subject to no such restrictions, and is subject to the subpoena power of federal, state, or local prosecutors. Axon could hand over sensitive information about undocumented individuals caught on body-worn cameras to ICE, or hand over statements made on body-worn cameras by women who come to Berkeley from other states to have abortions to federal prosecutors.

Police reports are generally confidential and protected from broader public access, which could mitigate this risk. However, giving complete access to reports and the underlying

[34] *Id.* at 11.
[35] *Id.*
[36] *Responsible AI for Modern Policing*, Axon, at 6, https://s3.amazonaws.com/a.storyblok.com/f/198504/x/70c48646ef/axon_responsible-ai-for-modern-policing-whitepaper-06a.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIU627ENUQT4RW23A%2F20250328%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20250328T222958Z&X-Amz-Expires=604800&X-Amz-SignedHeaders=host&X-Amz-Signature=c3f1a00a70b9597701f79f5ce379c5c97b9216e7ddb162eefecda49dfa13f348
[37] *See Axon Customer Experience Improvement Program*, Axon, https://www.axon.com/aceip.
[38] Resolution 67,763-N.S. Reaffirming Berkeley's Status as a City of Refuge.
[39] *See* Ordinance No. 7,224-N.S, Adopting the Operative Principles of the United Nations Convention on the Elimination of All Forms of Discrimination Against Women.

video files to a third party may waive or undermine confidentiality protections. To comply with existing Berkeley ordinances protecting privacy, including § 2.99, Berkeley would need a specific agreement with Axon regarding Axon's use of its data, and most importantly, its limitations.

Unknown 5: Will Draft One create courtroom evidentiary issues for police, prosecutors, or defense attorneys?

Finally, the legal admissibility of AI generated content is still unknown.[40] Some jurisdictions only use Draft One for minor incidents that do not lead to arrests, such as traffic collisions.[41] Axon's official website states: "Draft One serves the vast majority of reports, but excludes incidents where arrests took place and for felonies."[42]

Core constitutional protections likely make AI generated police reports inadmissible in criminal cases. The Sixth Amendment's confrontation clause guarantees criminal defendants the right to face and cross examine witnesses against them.[43] The officer who writes a police report is often a key witness against a criminal defendant at trial, where s/he is questioned about the statements he made in the report. Reports are often used to refresh recollection and as impeachment when the author testifies in court. The issue of applying the confrontation clause to AI products is yet unresolved by courts, but the right of meaningful impeachment of technology is questionable and complicated.[44] Should an Axon software engineer come into court and explain AI to the jury every time a Draft One police report is used? A defense attorney's cross-examination of such an expert is a poor substitute for the ability to ask questions of a human officer who wrote a report, and the risk of jury confusion is a threat to defendants' ability to exercise their trial rights.

Different risks come into play in cases where the right to counsel and confrontation of evidence do not apply. For example, there is no Sixth Amendment right to counsel in immigration court or DMV administrative hearings. Without the legal guardrails of criminal law, AI generated police reports are less likely to be thoroughly investigated for accuracy if reviewed by administrators.

---

[40] *See Draft One*, Axon, https://www.axon.com/products/draft-one, FAQs.
[41] *Using AI to Write Police Reports*, 18 Dispatch (Jan. 2025), https://cops.usdoj.gov/html/dispatch/01-2025/ai_reports.html.
[42] *Draft One*, Axon, https://www.axon.com/products/draft-one.
[43] U.S. Const. Amend. VI.
[44] *See* Andrea Roth, *Machine Testimony*, 126 Yale L. J. 1972, 2039 (May 2017).

B. Recommendations

The promise of Draft One for police efficiency, and the unknown factors regarding risks and effectiveness, create an opportunity for new study. If the City Council decides that this technology merits further exploration, it should partner with an established research institution to design a pilot study. A pilot study would help Berkeley determine whether and how to adopt the technology, and help other jurisdictions by contributing valuable data for their own decision-making. At the end of the pilot, the City should review the study's findings before approving or discontinuing use.

If Berkeley decides to adopt Draft One, Section 2.99 requires City Council to also adopt guidelines for its use. Some suggested guidelines are listed below.

Possible Guidelines

1. Disclaimers: Every AI generated police report should be marked as such, and carry the name of the officer who reviewed the report with a certification that he or she represents that the report is accurate.
2. Data segregation and retention: Berkeley should negotiate a contract with Axon that requires Axon to maintain the data separated to avoid commingling with other jurisdictions and delete all data within a set period of time.
3. Case type: AI generated police reports should not be used for cases involving arrests or possible felony charges.
4. External oversight: The PAB and Director of Police accountability should be granted access to data required to audit the use of AI technology.
5. Technical validation: Berkeley should engage with CLJC and technical experts from UC Berkeley to independently validate performance, accuracy, and bias mitigation capabilities before deployment, rather than relying solely on the manufacturer's claims.
6. Continuous monitoring: Beyond initial evaluation metrics, Berkeley should implement a continuous monitoring framework that regularly assesses the AI system's performance over time, identifies emerging patterns of errors or bias, and requires periodic revalidation as the system and its training data evolve.

Conclusion

Berkeley's commitment to evidence-based policing and civil rights protections make it a promising hub for exploration of novel police technologies. Berkeley's surveillance technology ordinance requires it to weigh the risks and benefits of AI police report writing technology before adopting it, and create usage policies for the technology if adopted. Draft One has the potential to improve police efficiency and save money by reducing overtime. However, at this point, many important factors are unknown about how the technology works in practice. If Berkeley chooses to adopt Draft One, the city should consider partnering with an established research institution to design a pilot study of the technology's effectiveness, benefiting both Berkeley and jurisdictions around the country.