



DECEMBER 2024

## Police Departments Shouldn't Allow Officers to Use AI to Draft Police Reports

*Jay Stanley, Senior Policy Analyst with the ACLU's Speech, Privacy, and Technology Project*

A small but growing number of police departments are adopting software products that use AI to draft police reports for officers. Because police reports play such an important role in criminal investigations and prosecutions, introducing novel AI language-generating technology into the criminal justice system raises significant civil liberties and civil rights concerns. These concerns include the unreliability and biased nature of AI, evidentiary and memory issues when officers resort to this technology, and issues around transparency. In the end, we do not think police departments should use this technology.

Police reports play a crucial role in our justice system. As professor and police technology expert Andrew Ferguson explained in an in-depth law review [article](#) on AI-assisted police reports, “human observations and memorialized facts can change the tenor, severity, and trajectory of the criminal case,” and aggregate crime data drawn from police reports can “impact property values [and] municipal budget priorities, and stigmatize neighborhoods.” Police reports, which are often the only official account of what took place during a particular incident, are central to the criminal proceedings that determine people’s innocence, guilt, and punishment.

The most prominent of these AI products is “Draft One,” made by the police technology juggernaut Axon, which also dominates the police cloud storage and body camera markets. Draft One allows an officer to select a body camera video file, transcribes the audio from that file, and uses OpenAI’s GPT-4 “large language model” (LLM) to turn that transcript into a first-person narrative for the officer in the typical format of a police report. The officer can then edit the file before swearing to its veracity and submitting it.

As with any technology, the first question is “does it work?” So far the jury is still out. Analysts at the independent surveillance trade publication IPVIM [noted](#) customer reports that Axon’s product “performs worse with longer interviews, pursuits, and traffic accidents” and “in loud environments or when unrelated conversations interject (such as radio chatter).”

It’s also not clear whether it makes writing police reports more efficient, which is the technology’s major selling point. In the first study of the technology by criminologists, a

[randomized controlled trial](#) in one police department found that officers using the technology didn't write their reports any faster than those who wrote them the old-fashioned way.

That's just one study, of course — and sometimes institutions continue to use technology even when it's shown to be ineffective. Assuming this technology doesn't go away, we see four general problems with it.

### **1. AI problems**

As anyone who has experimented with AI knows, the technology is very quirky and unreliable. In particular, LLMs, while amazingly advanced at imitating human writing, are prone to unpredictable errors, called “hallucinations,” in which they get facts wrong or make them up entirely. Those AI errors may be compounded by transcription errors, including those resulting from garbled or otherwise unclear audio in a body camera video. The use of automatic speech recognition to transcribe body camera audio is separate from the use of LLMs to generate police reports, but is itself a product of AI and also prone to errors including simply [making things up](#).

In an ideal world, officers would carefully review an AI-generated first draft of a police report and correct any “hallucinations” or other errors. Unfortunately, we don't live in an ideal world, so one important question is how many such errors will creep into police reports and lead to injustices.

To counter the reliability problems of AI, Axon's product includes two purported safeguards. First, the AI always produces “insert statements” where it instructs the reporting officer to, essentially, fill in the blank with needed details. Second, a police department can set the product to randomly insert silly sentences into the AI-produced draft, such as mention of a flying squirrel entering the scene, which is intended to ensure that police officers are actually reading and editing the draft. On that safeguard, however, Axon's CEO [conceded](#) that “We're generally getting not-great feedback I would say on that — most agencies are saying no, they wouldn't use it.”

Not only is AI prone to making things up, it's also biased. Even if an AI program doesn't make explicit errors, it could still spin things in subtle, biased directions, perhaps in ways an officer doesn't even notice. Because LLMs are trained on something close to the entire Internet, they inevitably absorb the racism, sexism, and other biases that permeate our culture. While OpenAI and other companies that create LLMs put in place filters and other tools to try to keep those stark biases from showing up in the AI's output, they are not entirely effective. Many biases remain hidden below the surface and experimenters have [shown this](#) by using carefully constructed queries to bypass the filters.

Another potential source of bias could be data that is used to customize or “fine tune” a basic LLM for a specialized use such as police report generators. In other words, in addition to the basic training of an LLM on vast stores of human writing, the model can be given additional training on a large number of police reports so that it learns what those typically look like. But police reports vary widely, and it's possible that disproportionate training on certain types of reports may bias how the AI crafts a police report that is quite different from the average.

We don't know all the ways these kinds of biases could enter into the creation of a police report — and it's not clear that anybody else does either. But one example might be in transcription errors. [Studies](#) have found that automatic speech recognition systems perform much more poorly in accurately transcribing [African American Vernacular English](#) and other dialects of [Black English](#). Errors in interpreting Black English [can and have](#) hurt people [legally](#). Similar errors could come with those who speak with an accent. And these errors may be compounded when the transcriptions are subsequently run through an LLM to turn them into a draft report. Despite supposed officer review, this complex process may very well mean that many such errors make their way into police reports and are later used to prosecute innocent people.

Finally, AI has a privacy problem. Typically, people interact with LLMs by feeding them text or other documents along with questions or directions. When a user uploads a document or enters a prompt, both of those are transmitted to the LLM's operator, and what that operator does with that information is not subject to any legal privacy protections. Contractual arrangements or corporate policies may govern data handling, but if they exist at all they can vary widely, change over time, and lack enforcement. Corporations should not have access to internal police data that might, for example, give them an advantage when they are battling labor activists.

## 2. Evidentiary and memory problems

When an incident happens, it's critical that the record reflect the best and most diverse possible sources of evidence. One increasingly common source of evidence is any video that has been captured, including body camera video. But video evidence isn't objective — as every first-year film student knows, what it portrays can depend on camera angle, lighting, resolution, when it is turned on, the sensitivity and placement of the mic, and many other factors.

Another key potential source of evidence is the officer's memory of what took place. But human memory — unlike video recordings — is extremely malleable. Subsequent information about an event can literally change a person's memory of what took place.

This elasticity of human memory is why [we believe](#) it's vital that officers give their statement about what took place in an incident before they are allowed to see any video or other evidence, especially in a critical incident such as a police shooting. The filters of a human's focus will typically miss some things that are present in the bodycam recording — sometimes things that seem [entirely obvious](#) — just as a human will notice some things that a video misses. But if an officer reviews the recording before providing their statement, the recording may write over certain facts or details the officer might have otherwise surfaced through the process of memorializing the incident.

It's important to also capture the officer's subjective experience and memory of an incident — which may be pivotal to determining whether to file charges and later, in any prosecution — which will be based on all five of an officer's senses, as well as their perception of human nuances of the situation such as whether somebody is hostile or meek, frightened or bold. This subjective experience cannot be captured by a bodycam. While the body camera footage is not going anywhere, the officer's memory may be fleeting and unstable.

Another problem with allowing officers to review their bodycam footage before giving a statement is that it enables officers to lie, for example if they see that something illegal they did was not captured by a camera. That could also carry over to AI-drafted police reports based on body camera footage. For example, an officer who did not have probable cause to engage in a search or arrest might be able to justify an illegal act after the fact by the AI picking up on audio in the background that the officer never heard, and which could provide justification.

The point is to have as many independent sources of evidence for what took place as possible. The body camera video and the police officer's memory are two separate pieces of evidence. But if the police report is just an AI rehash of the body camera video, then you no longer have two separate pieces of evidence — you have one, plus a derivative summary of it that can reshape and contaminate the officer's memory or immunize the officer from accountability for illegal acts.

### **3. Technology transparency and discovery problems**

Given the novel, experimental nature of AI-generated police reports, transparency is crucial. It's important for the public to understand what's going on so that independent experts can evaluate the technology, and communities and their elected representatives can decide whether they want the police officers that serve them to use it. And it's vital that defendants in criminal cases be able to interrogate the evidence against them, including by understanding how it was generated, so that they can exercise their constitutional right under the Sixth Amendment to “confront their accuser.”

According to Axon, their system uses the same OpenAI model that powers ChatGPT. As an Axon exec [put it](#) in a webinar for prosecutors, “We turn off all the creativity, so that it sticks to the facts, and then we build basically a law enforcement layer on top.” That “law enforcement layer,” he explained, is where the LLM is instructed in such things as what a good police report looks like, what kinds of information to include and not to include, and when to give police officers prompts to add additional information.

That explanation leaves many unanswered questions. LLM prompts are the subject of much sophisticated discussion, experimentation, and analysis; exactly how a question and direction is phrased can have an enormous effect on the model's output. Clearly, Axon feeds the LLM the transcript of the audio portion of the body camera file that an officer has selected. But what is the prompt that accompanies that text? That's an example of the kind of element of an AI tool that ought to be public. If it's not, a police AI system could well contain an instruction such as, “Make sure that the narrative is told in a way that doesn't portray the officer as violating the Constitution.”

Any system that feeds information into our criminal justice system and has the potential to affect the fate of people accused of crimes must be open for experts, the media, and interested members of the public to inspect, study, and experiment with. Private companies like Axon are not subject to transparency laws like the federal Freedom of Information Act and the Privacy Act that can be used to pry information out of government agencies. But law enforcement agencies' reliance on Axon's technology means that its products must be open to further scrutiny, including through federal and state freedom of information laws.

The lack of transparency also threatens to become a significant legal issue when people are charged with crimes. In a 1963 decision, *Brady v. Maryland*, the Supreme Court ruled that prosecutors must turn over to defense attorneys any information in their possession that suggests defendants may not be guilty or is relevant to how they may be sentenced. The use of AI raises major questions about the scope of such sharing. For example, would such “Brady disclosures” need to include the original draft that OpenAI’s model presents to an officer before they edit it? Axon was asked in their webinar whether they preserve that original draft; an executive [replied](#) that “we don’t store the original draft, and that’s by design. . . . Because the last thing we want to do is create more disclosure headaches for our customers, and our attorney’s offices.” Police and prosecutors may view disclosure as a headache, but it’s vital for defendants seeking a fair trial — and can be constitutionally required. If a court rules that prosecutors’ sharing of such information is insufficient, it could lead to dismissal of charges against a defendant. It could also open up the possibility of challenges to prior convictions reliant on this technology.

And the questions about Brady disclosures go far beyond AI drafts to all the components of the machinery behind the AI-aided police report — such as details about Axon’s customization of and queries to an LLM and error rates in the generated output. It’s not hard to foresee computer scientists who work for OpenAI being pulled into a murder trial or the like due to questions raised by the defense about how the company’s LLM was trained. Adding even more complexity, these products continually change. This technology is so novel and complex that someone needs to be an expert just to know what questions should be asked or answered. We need to think hard before we allow this level of complexity and opacity to be injected into the criminal justice system.

Defendants, of course, are not this company’s customers. This basic reality reveals a systematic problem with the legal system’s use of for-profit companies, whose customer base largely represents one side in our adversarial criminal justice system. As a result, those companies are inherently incentivized to design those products in a manner to benefit that side and to disadvantage a criminal defendant. This problem is only exacerbated when it comes to novel, difficult-to-analyze AI products, which are shrouded in ever more layers of complexity and secrecy.

#### **4. A loss of the human disciplinary function that police reports play**

Finally, Andrew Ferguson [points out](#) that an important role of police reports as a disciplinary check may be lost with a transition to AI-drafted reports. In addition to the instrumental role that police reports play in the justice system’s machinery, forcing police to write down reasons for their exercise of discretionary power (like stops, frisks, searches, consent searches, etc.) reminds them of the legal limits of their authority. The act of writing reports functions as a form of internal mental discipline for police that continually reminds them of limits on their power. And what they write is often then reviewed by supervisors, who use what’s written to identify when an officer might not know or observe those limits. Those supervisors are often required to then take steps to address that issue. So police reports don’t just create external accountability, they also create internal mechanisms for enforcing limits on police power in real time.

A shift to AI-drafted police reports would sweep away these important internal roles that reports play within police departments and within the minds of officers. This is an additional reason to be skeptical of this technology.

### **Too many problems**

For these reasons, the ACLU does not believe police departments should allow officers to use AI to generate draft police reports. This technology removes important human elements from police procedures and is too new, too untested, too unreliable, too opaque, and too biased to be inserted into our criminal justice system.

The use of AI to draft police reports is likely to have different effects in minor and major allegations of wrongdoing. In more minor cases, as Ferguson points out, police reports are often the only account of an incident (besides the defendant's) that prosecutors, defense attorneys, and judges have access to as charging, plea bargain, and sentencing decisions are made. Changes in the generation of police reports, such as the use of AI with all its attendant problems, could increase the number of injustices in such cases.

In more serious cases, the police report is likely to be just one of many sources of evidence, including witness testimony and the body camera footage itself. But though the role of the police report may be smaller, its potential consequences could be much more significant. By default, Axon's software is set not to be used for any incidents involving felonies or arrests. But Axon's executives have publicly [stated](#) that "the DAs and the actual agencies that are doing these reports . . . are rapidly turning off the restrictions," and that "most of the agencies that are live right now are using it on all incidents." That means these AI-assisted police reports could already be affecting criminal outcomes in very serious situations.

If writing reports is a dull chore for police officers, there may be safer and more limited ways that AI could be used to help. For example, officers could make an audio recorded verbal narrative of what took place in a particular incident, and computers could transcribe those accounts and perhaps perform some light cleanup and formatting to create an editable first draft. With the audio of the officer's account preserved along with the written report, that would solve many of the issues highlighted above, while potentially making life easier for police officers — who, like most people, probably find it much faster and easier to speak than to write. AI has many potential functions, but there is no reason to use it to replace the creation of a record of the officer's subjective experience.

###