The Unfalsifiable Delusion: How Mass Surveillance Infrastructure Creates Disparate Impact on Americans with Paranoia-Spectrum Conditions

Travis Gilly
Real Safety AI Foundation
Illinois, United States
ORCID: 0009-0007-2954-6313
t.gilly@ai-literacy-labs.org

Abstract

Reality testing is a core cognitive-behavioral intervention for persecutory delusions: the clinician helps the patient examine evidence for the belief that they are being monitored, the patient discovers the belief is unfounded, and cognitive reframing occurs. This paper argues that mass surveillance infrastructure deployed by U.S. government agencies and their private contractors has eliminated the foundational precondition for this intervention by making persecutory beliefs factually accurate for the general population, creating a specific and documentable disparate impact on Americans with paranoia-spectrum conditions. We identify three distinct mechanisms of harm: reality testing destruction, in which the therapeutic intervention fails because the environment has changed to confirm the delusion rather than because the patient resists treatment; device abandonment and social isolation, in which affected individuals make the rational decision to discard communication devices upon learning that government geofencing captures every phone in their neighborhood, severing contact with support systems and crisis services; and algorithmic pattern-matching discrimination, in which surveillance systems designed to flag anomalous movement patterns systematically misidentify the behavioral symptoms of paranoia-spectrum conditions as suspicious activity, subjecting affected individuals to additional scrutiny that compounds the original harm. We document that this intersection has been entirely overlooked in both the surveillance policy literature and the clinical psychiatric literature, including a 119-page European Parliament analysis on AI surveillance and human rights that does not contain the word "disability." We frame these harms under ADA Title II and Section 504, identify a novel legal problem we term the self-sealing paradox (in which surveillance retroactively eliminates the diagnostic category it damages), and propose clinical, legal, and policy responses.

1 Introduction

Cognitive-behavioral therapy for persecutory delusions follows a well-established protocol. The clinician and patient collaboratively examine evidence for the patient's belief that they are being watched, followed, or monitored. The patient discovers that the evidence does not support the belief. Cognitive reframing occurs. The delusion loosens its grip. Recovery progresses (Freeman et al. 2015; Garety et al. 2008).

This therapeutic mechanism depends on a single precondition: the persecutory belief must be falsifiable. There must exist, somewhere in the patient's observable environment, disconfirming evidence that the clinician can help the patient find. If nobody is actually watching the patient, the clinician can demonstrate this. If no cameras are actually recording the patient's movements, the clinician can point to their absence. The intervention works because reality provides a counterweight to the delusion.

This paper argues that mass surveillance infrastructure deployed across the United States has removed that counterweight. License plate readers operated by Flock Safety under government contracts now track vehicle movements across thousands of jurisdictions (Goldstein-Street 2026), with public records and the Electronic Frontier Foundation's HaveIBeenFlocked tool showing individual citizens logged hundreds of times per month (Electronic Frontier Foundation 2025). The Department of Homeland Security draws digital fences around entire neighborhoods using commercially purchased advertising data, capturing the location of every phone within the geofence without a warrant (U.S. House of Representatives, Subcommittee on Cybersecurity, Information Technology, and Government Innovation 2026). Cell-site simulators mounted in vehicles sweep every device identifier within range. Facial recognition systems identify individuals in public spaces.

The persecutory belief "they are watching me everywhere I go" is no longer delusional. It is a factually accurate description of the surveillance environment in most American metropolitan areas.

For the general population, this raises well-documented privacy and civil liberties concerns. For Americans with paranoia-spectrum conditions, including schizophrenia, schizoaffective disorder, delusional disorder, and PTSD with paranoid features, the consequences are qualitatively different and clinically specific. The primary therapeutic intervention for their condition has been rendered ineffective by an environmental change they did not cause, cannot escape, and are not equipped to contextualize.

A clinician working with a patient who believes "they are

Accepted to the AAAI/ACM Conference on AI, Ethics, and Society (AIES 2026), Malmö, Sweden, October 2026. This is the author's accepted manuscript, posted under ACM author-retained posting rights; it is not the Version of Record. An earlier version of this paper was posted on SocArXiv, March 2026, https://doi.org/10.31235/osf.io/cfqwz_v1.

tracking my car everywhere I go" can no longer say "let's examine the evidence for that belief" because the evidence confirms it. A clinician working with a patient who believes "they are monitoring my phone" can no longer help the patient reality-test that belief because the belief is true. The intervention does not fail because the patient resists treatment. It fails because the environment has changed to match the delusion.

This paper documents three distinct mechanisms by which mass surveillance infrastructure creates disparate impact on this protected class, identifies a complete absence of this analysis from both the surveillance policy literature and the clinical psychiatric literature, frames the harm under existing disability rights law, and proposes clinical, legal, and policy responses.

2 Clinical Foundation

2.1 Reality Testing as Therapeutic Intervention

Cognitive-behavioral therapy for psychosis (CBTp) has accumulated substantial evidence as a treatment for persecutory delusions over the past two decades. Freeman and colleagues have established that persecutory delusions are maintained by a combination of anomalous experiences, reasoning biases, emotional processes, and behavioral responses, and that CBT targeting these maintenance factors produces clinically significant reductions in delusional conviction and distress (Freeman 2016; Freeman et al. 2015).

Reality testing sits at the center of this therapeutic approach. The clinician does not directly challenge the delusion by telling the patient their belief is false. Instead, the clinician and patient collaboratively design behavioral experiments to test the belief against observable evidence. If the patient believes they are being followed, the clinician might help the patient systematically observe whether the same vehicle appears behind them on different routes. If the patient believes their home is under surveillance, the clinician might help the patient examine the physical environment for evidence of monitoring equipment.

The efficacy of this approach depends on the environment providing disconfirming evidence. The behavioral experiment works because, in the vast majority of historical cases, the patient is not actually being followed. The same car does not appear on different routes. No monitoring equipment is found. The patient's belief is tested against reality, and reality does not support it.

Garety and colleagues' cognitive model of delusions (Garety et al. 2001) identifies this evidence-evaluation process as central to both the formation and maintenance of delusions. Delusions form when anomalous experiences are interpreted through reasoning biases (jumping to conclusions, externalizing attributions) without adequate consideration of alternative explanations. The therapeutic process introduces alternative explanations and helps the patient weigh evidence for competing hypotheses.

This model assumes that competing hypotheses exist. It assumes that the environment contains evidence supporting the alternative explanation (nobody is watching you). When the environment no longer contains that evidence because surveillance is actually occurring, the model's foundational assumption is violated.

2.2 Population Data

The affected population extends across multiple diagnostic categories, each with distinct prevalence rates and varying proportions of persecutory features. Schizophrenia affects approximately 0.6 percent of U.S. adults (around 1.5 million people), with roughly 70 percent presenting persecutory features (Freeman 2007). Schizoaffective disorder affects approximately 0.3 percent (around 750,000 adults), with persecutory features common in the majority. Delusional disorder affects approximately 0.2 percent (around 500,000 adults), with the persecutory subtype most common. PTSD with paranoid features affects a subset of the approximately 6 percent lifetime PTSD population, with hypervigilance universal and paranoid features in a clinical subset. Bipolar I with psychotic features produces persecutory delusions in approximately half of manic episodes, drawing from a population of roughly 2.8 million adults.

A conservative estimate of the directly affected population, limited to individuals with active persecutory ideation as a current clinical feature, exceeds two million Americans. This does not account for individuals in prodromal phases, those whose symptoms are subclinical but who are vulnerable to environmental triggers, or those who have achieved remission through reality-testing interventions that may now be undermined by environmental change.

Critically, this population estimate is not static. Environmental changes can expand the at-risk population by functioning as etiological factors for new-onset paranoia-spectrum conditions. The urban psychosis literature has established that environmental stressors, including social threat perception, social fragmentation, and chronic activation of threat-detection systems, increase psychosis incidence in a dose-dependent manner (Vassos et al. 2012; March et al. 2008; Selten et al. 2013). If mass surveillance infrastructure functions as an environmental stressor of comparable magnitude, the affected population will grow in proportion to the expansion of that infrastructure. The legal and clinical implications of a protected class that expands as a consequence of the government program being challenged are addressed in the Legal Framework section below.

2.3 The Falsifiability Requirement

The DSM-5 defines delusions as "fixed beliefs that are not amenable to change in light of conflicting evidence" (American Psychiatric Association 2013). This definition contains an embedded assumption: that conflicting evidence exists to be presented. The classification system itself presupposes a distinction between the content of the belief and the state of the world.

When the state of the world changes to match the content of the belief, a diagnostic paradox emerges. The individual who believes "the government is tracking my movements through my phone" is not expressing a fixed belief that is not amenable to change in light of conflicting evidence. They are expressing an accurate observation about DHS geofencing

operations that has been confirmed by congressional investigations, journalistic reporting, and the government's own documentation.

This is not a clinical improvement. The patient has not recovered. Their paranoia-spectrum condition persists, with all its associated cognitive, emotional, and behavioral features. What has changed is that one specific symptom, the persecutory belief about surveillance, has been rendered diagnostically ambiguous because the environment has shifted beneath the diagnostic criteria.

The clinical implications are severe. Treatment plans built around reality testing must be revised. Clinicians must develop new frameworks for distinguishing between surveillance-related beliefs that are proportionate responses to actual environmental conditions and surveillance-related beliefs that are disproportionate, generalized, or accompanied by additional delusional features (referential thinking, thought broadcasting, paranoid ideation extending beyond surveillance). This differential diagnosis has not been developed because the clinical literature has not recognized the need for it.

3 The Research Void: Disability as the Missing Variable in Surveillance Scholarship

3.1 The Surveillance Policy Literature Gap

In May 2024, the European Parliament's Subcommittee on Human Rights published a 119-page in-depth analysis titled "Artificial Intelligence (AI) and Human Rights: Using AI as a Weapon of Repression and Its Impact on Human Rights" (Ünver 2024). The document provides comprehensive coverage of algorithmic authoritarianism across China, Russia, Iran, Egypt, and Ethiopia. It examines facial recognition, predictive policing, social credit systems, and mass surveillance. It addresses privacy, civil liberties, free expression, political repression, and democratic integrity.

The word "disability" does not appear in the document. "Schizophrenia," "paranoia," "delusion," and "psychiatric" likewise do not appear. The only psychological reference is a single sentence noting that AI "may contribute to addiction or mental health issues," without further development, without disability framing, and without any reference to pre-existing conditions that interact with surveillance environments.

This is not an isolated omission. Surveillance policy scholarship operates within a framework that treats the surveilled population as undifferentiated. The harms it identifies, such as chilling effects on free expression, erosion of privacy, and potential for authoritarian abuse, are described as affecting "citizens," "individuals," or "the public" without examining whether some members of the public bear qualitatively different harms due to pre-existing conditions that interact with the surveillance environment.

The disability rights implications of mass surveillance represent a complete blind spot in this literature. Not a contested point. Not an area of active disagreement. An absence.

3.2 The Clinical Psychiatric Literature Gap

The clinical literature on CBT for psychosis is similarly self-contained. Extensive research exists on reality testing protocols, behavioral experiments for persecutory delusions, and the cognitive model of delusions (Freeman 2016; Garety et al. 2001; Freeman et al. 2015). Recent work has explored virtual reality as a therapeutic environment for safely testing paranoid beliefs in controlled settings (Pot-Kolder et al. 2018; Freeman et al. 2016).

None of this literature examines whether real-world changes in the surveillance environment affect the efficacy of reality-testing interventions. The assumption that the external environment will provide disconfirming evidence for persecutory beliefs is treated as a given, not as a variable that might change.

The closest the literature comes to addressing this intersection is in the observation that delusional content evolves with technology. Modern patients incorporate real surveillance technologies into their symptom presentations, referencing GPS tracking, phone monitoring, and camera networks. Clinicians have noted the increasing difficulty of distinguishing between delusional beliefs about technology and accurate perceptions of actual technological capabilities (Burns et al. 2025). But this observation has not been developed into a systematic analysis of how the surveillance environment degrades therapeutic interventions. Whether knowledge that surveillance is real (versus delusional) affects psychiatric outcomes differently has been identified as a notable gap (Burns et al. 2025), and it remains unfilled.

3.3 The Interdisciplinary Failure

The surveillance policy scholar does not read the CBT for psychosis literature. The clinical psychologist does not read the surveillance policy literature. The disability rights advocate does not read either.

The result is a harm mechanism affecting over two million Americans that exists in the space between three disciplines, visible to none of them. A 119-page human rights analysis of AI surveillance excludes the population with the most clinically specific harm mechanism. A clinical literature on paranoia treatment ignores the most significant environmental change affecting that treatment's efficacy in decades. And disability rights advocacy, which should be the connective tissue between these fields, has not identified mass surveillance as a disability issue.

This paper occupies that space.

4 The Surveillance Landscape

4.1 Flock Safety and ALPR Networks

Flock Safety operates automated license plate reader (ALPR) networks under government contracts across thousands of U.S. jurisdictions. These systems photograph and log every vehicle that passes, recording the plate number, timestamp, location, and direction of travel. The data is stored, searchable, and accessible to law enforcement without a warrant.

Public records obtained through FOIA requests and the HaveIBeenFlocked tool developed by the Electronic Frontier Foundation have revealed the scale of this monitoring (Electronic Frontier Foundation 2025). Individual citizens have been logged by Flock cameras hundreds of times in a single month. The coverage is not limited to highways or high-crime areas; cameras are deployed on residential streets, near schools, at neighborhood entrances, and on public thoroughfares throughout metropolitan areas.

Critically, this surveillance is not secret. Flock Safety and the jurisdictions that contract with them actively publicize the deployments. City council debates about Flock installations generate local news coverage. Flock's own marketing materials promote the breadth of their network. The result is that any American who reads local news, watches city council meetings, or encounters journalism about police technology is likely aware that their vehicle movements are being tracked.

For the general population, this awareness creates discomfort. For someone with persecutory delusions about being tracked, this awareness confirms the delusion with evidence the clinician cannot dispute.

4.2 DHS Geofencing and Location Surveillance

In early 2026, reporting by 404 Media, the Washington Post, and congressional oversight investigations revealed that the Department of Homeland Security was using commercially purchased advertising data to conduct warrantless location surveillance of entire neighborhoods (Cox 2026; U.S. House of Representatives, Subcommittee on Cybersecurity, Information Technology, and Government Innovation 2026). The tool, called Webloc and manufactured by the Nebraska-based company Penlink, allows agents to draw a digital fence around any geographic area and retrieve the location data of every phone that was present during a specified time period.

This data originates from ordinary consumer applications. Weather apps, navigation apps, food delivery apps, and games that request location permissions feed GPS coordinates to advertising networks. Those networks sell the data to data brokers. The data brokers sell it to Penlink. Penlink sells access to DHS (Cox 2025).

The geofence does not discriminate by age, disability status, or clinical vulnerability. Every phone within the boundary is captured. Children's phones, the phones of psychiatric patients, the phones of individuals in residential treatment facilities. The system has no mechanism for excluding protected populations because it was not designed with any awareness that protected populations exist in the data.

Simultaneously, DHS continues to deploy cell-site simulators (commonly called Stingrays) mounted in vehicles. These devices impersonate legitimate cell towers, capturing the unique identifiers of every phone within range. Like geofencing, this technology captures all devices indiscriminately.

4.3 The Publicity Dimension

The psychiatric harm mechanism documented in this paper does not require the affected individual to be personally targeted by surveillance. It does not require them to know the technical details of Flock Safety or Webloc. It requires only that they become aware that mass surveillance exists and that it captures their movements.

This awareness is now unavoidable. Congressional hearings on DHS surveillance tools generate national news coverage. Local Flock Safety deployments generate city-by-city reporting. The Snowden disclosures, though over a decade old, established a baseline awareness of government surveillance capability that has only intensified with subsequent reporting.

For a person whose psychiatric condition includes persecutory ideation about being monitored, the information environment is saturated with confirmation. The delusion is not being fed by internal cognitive distortion. It is being fed by the news.

5 Three Mechanisms of Disparate Harm

5.1 Mechanism 1: Reality Testing Destruction

The first and most direct mechanism of harm is the collapse of reality testing as a therapeutic intervention.

Consider a clinical scenario. A patient with schizophrenia presents with the persecutory belief: "The government is tracking me everywhere I go through cameras and my phone." Under the standard CBT for psychosis protocol, the clinician would help the patient examine evidence for this belief. They might explore questions like: Have you seen cameras following you? Can you identify any specific technology being used? Is there evidence that you specifically are being targeted?

In 2020, the clinician could help the patient discover that no, there were no cameras specifically tracking them. The belief was a product of the illness, not the environment.

In 2026, the honest answer to the patient's belief is: yes. Flock Safety cameras do log your vehicle everywhere you drive. DHS geofencing can and does capture your phone's location. Cell-site simulators do grab your device identifier. You are not specifically targeted, but you are captured in mass surveillance systems that track everyone.

The clinician now faces an impossible choice. They can tell the truth, which confirms the delusional framework and strengthens the patient's conviction that their paranoid beliefs are justified. Or they can minimize the surveillance reality, which requires the clinician to be dishonest with their patient and, if the patient later discovers the truth independently, destroys the therapeutic relationship entirely.

Neither option serves the patient. Both options are the direct result of an environmental change that neither the patient nor the clinician caused.

The treatment does not fail because the patient is treatment-resistant. It fails because the environment removed the tool the treatment depends on. Attributing this failure to the patient's condition rather than to the environmental change is a clinical error with significant consequences: it may lead to escalation of medication, hospitalization, or other interventions that address the wrong problem.

Critically, the damage to reality testing is not confined to surveillance-specific beliefs. Paranoia does not compart-

mentalize. When a macro-level persecutory belief is confirmed ("the government is tracking everyone"), micro-level beliefs inherit that confirmation's credibility. The patient who learns that Flock Safety cameras log their vehicle everywhere they drive does not then neatly separate that confirmed belief from their belief that a neighbor is monitoring them. The confirmed surveillance becomes the evidentiary foundation on which every other persecutory belief becomes more plausible, more resistant to challenge, and harder for the clinician to address.

The mechanism is credibility contamination of the therapeutic relationship. Reality testing depends on the therapeutic alliance, and the therapeutic alliance depends on the clinician's credibility as an arbiter of what is and is not real. The moment a patient can demonstrate that the clinician was wrong about surveillance ("you told me nobody was watching me, and they were"), every subsequent reality-testing intervention is undermined. The clinician's reliability has been falsified by the same evidence that confirmed the delusion.

This is not a belief-by-belief problem that can be managed by distinguishing "the kinds of surveillance that are real" from "the kinds that aren't." During active psychosis, every piece of electronics is suspect, every technology is a potential vector, and the distinction between networked ALPR tracking and a neighbor's listening device is not one the patient is cognitively equipped to maintain. The confirmed macro-level surveillance generalizes across the entire delusional system, and the clinician who was wrong once is no longer trusted to be right about anything.

5.2 Mechanism 2: Device Abandonment and Social Isolation

The second mechanism of harm is more immediately dangerous and less immediately obvious.

A person with paranoid features who learns, through mainstream news coverage, that the government is tracking phones through geofencing makes a rational decision: discard the phone. Disable location services. Remove the battery. Stop carrying the device.

This is not a delusional response. It is an evidence-based protective action consistent with the threat as reported by credible news sources. A reasonable person, learning that their phone's location data is being sold to federal surveillance agencies, might take exactly the same action.

But for someone with a paranoia-spectrum condition, this rational action has catastrophic consequences. The phone is their lifeline to family members who check on them. To friends who provide social support. To crisis hotlines. To their therapist. To their case manager. To emergency services.

Severing that connection does not look like a crisis from the outside. No one calls 911 when someone stops answering their phone. The family member thinks they're busy. The therapist notes a missed appointment. The case manager documents a no-show. By the time anyone recognizes a pattern of withdrawal, weeks or months may have passed, and the individual may be deep in an episode with no support system able to reach them.

The surveillance infrastructure did not intend to isolate psychiatric patients. But intent is not required for disparate impact. The predictable consequence of publicized mass phone surveillance is that the people most vulnerable to paranoid responses will take the most extreme protective actions, and those actions will sever the connections that keep them alive.

5.3 Mechanism 3: Algorithmic Pattern-Matching Discrimination

The third mechanism operates through the surveillance system's own analytical processes.

Geofence and behavioral surveillance systems are designed to identify anomalous patterns. Movement that deviates from expected norms, unusual hours of activity, irregular routes, repeated visits to multiple locations in sequence, sudden changes in routine. These patterns are flagged for additional review because, in the system's design logic, anomalous behavior may indicate criminal activity.

The behavioral symptoms of paranoia-spectrum conditions would be expected to produce patterns consistent with these anomaly flags. Erratic movement driven by paranoid avoidance of perceived threats. Irregular sleep schedules common in psychotic episodes. Pacing and circling behavior during acute anxiety. Avoidance of certain areas based on delusional beliefs about danger. Sudden changes in routine following symptom exacerbation. Research on predictive policing systems has established that algorithmic anomaly detection disproportionately flags individuals whose behavior deviates from population norms for any reason, including disability, homelessness, and mental illness (Richardson, Schultz, and Crawford 2019; Brayne 2017). Studies of automated loitering detection have demonstrated that systems designed to identify "suspicious" lingering in public spaces systematically flag individuals whose conditions produce extended stationary behavior, including those in psychiatric distress (Selbst 2017).

The algorithm cannot distinguish between a person casing a building for a burglary and a person circling a block because their paranoia makes them afraid to approach the entrance. Both would be expected to produce similar geolocation signatures. Both get flagged.

The consequence is that individuals whose psychiatric conditions produce flagged behavioral patterns receive additional surveillance scrutiny. This additional scrutiny, if it becomes perceptible to the individual (through increased law enforcement presence, unexpected encounters with police, or awareness that their data has been queried), feeds directly back into the persecutory belief system. The patient now has evidence not just that mass surveillance exists, but that they personally have attracted attention from it.

This creates a feedback loop: surveillance flags symptoms, additional scrutiny confirms the persecutory belief, confirmed beliefs worsen symptoms, worsened symptoms produce more flagged behavior. The system amplifies the condition it was never designed to account for.

6 Legal Framework

6.1 ADA Title II and Disparate Impact

Title II of the Americans with Disabilities Act prohibits public entities from discriminating against individuals with disabilities in their services, programs, and activities. The implementing regulation at 28 C.F.R. § 35.130(b)(3) provides that a public entity may not "utilize criteria or methods of administration" that "have the effect of subjecting qualified individuals with disabilities to discrimination on the basis of disability."

Disparate impact claims under Title II do not require proof of discriminatory intent. A facially neutral policy that disproportionately harms a protected class violates Title II regardless of whether the harm was intended.

Mass surveillance constitutes a government program or activity. Public roads are a government service. Although Flock Safety and data brokers like Penlink are private companies, their surveillance products are deployed through government procurement decisions. The state or local government that contracts with Flock Safety to install ALPR cameras on public infrastructure is the public entity whose decision creates the disparate impact. The municipality chose to deploy the technology. The municipality chose the locations. The municipality chose whether to conduct a disability impact assessment (none have). Under state action doctrine, the liability attaches to the government entity that procured and authorized the surveillance, not to the private vendor that manufactured the hardware. DHS geofencing uses federal government resources to collect data about individuals using public spaces. These programs are subject to Title II and Section 504 respectively.

The three harm mechanisms documented in this paper constitute disparate impact on a protected class. Americans with paranoia-spectrum conditions bear qualitatively different and clinically measurable harms from mass surveillance compared to the general population. The general population experiences discomfort, privacy concerns, and chilling effects. The protected class experiences destruction of a primary therapeutic intervention, rational behavioral responses that produce dangerous isolation, and algorithmic flagging of their disability's symptoms as suspicious behavior.

These harms are not speculative. They are predictable consequences of deploying surveillance infrastructure to populations that include individuals with documented psychiatric conditions, without conducting any assessment of the disability-specific impacts.

6.2 Section 504 of the Rehabilitation Act

Section 504 of the Rehabilitation Act of 1973 prohibits discrimination on the basis of disability in any program or activity receiving federal financial assistance. DHS surveillance programs are federally funded. Flock Safety installations that receive federal grants or operate under federally funded law enforcement programs fall within Section 504's scope.

The analysis under Section 504 parallels the Title II analysis: the same three harm mechanisms constitute disparate impact on the same protected class. Section 504 provides an additional legal avenue because its reach extends to any program receiving federal funds, capturing surveillance deployments that might fall outside Title II's coverage of state and local government programs.

6.3 The Self-Sealing Paradox

This paper identifies a novel legal problem that emerges from the intersection of surveillance policy and disability law.

If mass surveillance makes the persecutory belief "they are watching me" factually true, then the belief is no longer clinically delusional. A belief that accurately describes reality, by definition, does not meet the DSM-5 criteria for a delusion. If the belief is no longer delusional, the symptom profile of the individual changes. If the symptom profile changes sufficiently, diagnostic criteria may no longer be met. If diagnostic criteria are no longer met, the individual may lose their status as a person with a qualifying disability under the ADA. If they lose their disability status, they lose the legal protections that would have required accommodation.

The surveillance infrastructure retroactively justifies itself by eliminating the diagnostic category it damages.

This is not a hypothetical chain of reasoning. It describes a real mechanism by which the expansion of surveillance could narrow the definition of persecutory delusions, reduce the population that qualifies for paranoia-spectrum diagnoses, and thereby reduce the population with standing to challenge the surveillance under disability law.

The paradox has no existing precedent in either disability law or surveillance policy. It represents a genuinely novel intersection that neither field has contemplated.

The paradox, however, contains its own rebuttal. The same surveillance environment that may narrow the protected class by rendering specific persecutory beliefs non-delusional simultaneously expands the protected class by inducing new cases of paranoia-spectrum conditions in individuals with no prior psychiatric history. Sustained, confirmed, inescapable awareness of being monitored produces a psychological state, characterized by hypervigilance, behavioral modification, social withdrawal, and persistent preoccupation with being observed, that is clinically continuous with the prodromal and early features of paranoia-spectrum conditions. The surveillance environment does not merely damage existing members of the protected class; it generates new ones. The government cannot invoke the self-sealing paradox to argue that surveillance-specific beliefs are no longer delusional (and therefore no longer protected) while simultaneously presiding over infrastructure that is inducing the very conditions that qualify individuals for protection. The paradox dissolves when both directions are examined: what the surveillance takes from the diagnostic category on one end, it replenishes on the other, and the net effect is an expanding population of individuals bearing clinically measurable harm from government-deployed infrastructure.

6.4 Protected Class Expansion Through Government-Induced Onset

The preceding sections document disparate impact on Americans who already carry paranoia-spectrum diagnoses. This section argues that mass surveillance infrastructure also creates new members of the protected class by functioning as an etiological environmental factor, producing clinically diagnosable paranoia-spectrum conditions in individuals with no prior psychiatric history.

The precedent for this analysis is industrial disease law. Asbestos manufacturers did not merely worsen pre-existing respiratory conditions; they caused new disease in healthy workers. Those workers became members of a protected class (disabled under the ADA, claimants under tort law) as a direct consequence of occupational exposure. The employers' liability did not depend on pre-existing vulnerability. It depended on the dose-response relationship between exposure and clinical outcome: this substance, at this concentration, for this duration, produces this disease in people who would not otherwise have developed it.

Surveillance is the environmental exposure. The dose is awareness of being monitored, combined with confirmation that the monitoring is real, combined with the inescapability of the infrastructure, sustained over time. The clinical response follows established dose-response patterns in psychiatric epidemiology. Low exposure (a single news article about surveillance technology, producing transient discomfort) does not produce clinical outcomes. Moderate exposure (living in a jurisdiction with visible surveillance infrastructure, encountering regular news coverage, noticing cameras in one's daily environment) produces sustained awareness and measurable behavioral modification, including self-censorship, route alteration, and communication changes documented in general population surveys by Pew Research Center (Pew Research Center 2023). High exposure (living in a saturated surveillance environment, understanding the technical capabilities of the systems, knowing that one's movements are logged continuously) produces persistent hypervigilance that meets subclinical thresholds for paranoid ideation.

For some subset of the high-exposure population, individuals who may have had no prior psychiatric vulnerability, the sustained psychological state tips into clinical territory. This is not speculation about a future possibility; it is how environmental risk factors operate in every other domain of psychiatric epidemiology. Urban environments increase psychosis risk through mechanisms including social stress, perceived threat, and social fragmentation (Vassos et al. 2012; March et al. 2008). Childhood adversity increases psychosis risk through sustained activation of threat-detection systems (Varese et al. 2012). Social defeat produces paranoid ideation through chronic perception of subordination and vulnerability (Selten et al. 2013). While within-family genetic analyses have shown that some portion of the urbanicity-psychosis association is attributable to genetic confounding (Paksarian et al. 2017), the environmental contribution remains significant after controlling for familial liability, and no comparable genetic confound has been identified for surveillance-specific exposure. In each case, a sustained environmental stressor produces clinical psychiatric outcomes in a dose-dependent fashion in individuals who were not identifiably predisposed prior to exposure. They were exposed.

The legal significance of this etiological pathway is that it forecloses the government's most likely defense against the self-sealing paradox. If the government argues that surveillance-specific persecutory beliefs are no longer delusional because they are factually accurate, and that individuals holding accurate beliefs do not qualify as disabled under the ADA, the etiological pathway responds: the surveillance environment that made those specific beliefs accurate has also induced generalized paranoid ideation, social withdrawal, hypervigilance, sleep disruption, and functional impairment in individuals who did not previously have these conditions. These individuals meet diagnostic criteria not because of a single accurate belief about surveillance but because of a clinical syndrome that the surveillance environment precipitated. They are newly disabled, and the government's infrastructure is the cause.

The three-prong structure of the legal claim is therefore: first, disparate impact on the existing protected class (the three mechanisms documented above); second, expansion of the protected class through government-induced onset (this subsection); and third, the inability of the self-sealing paradox to serve as a defense, because the same infrastructure that might narrow one edge of the diagnostic category expands it from the other direction. The government cannot escape ADA liability by arguing that its surveillance made specific beliefs non-delusional when that same surveillance is creating new clinical populations with diagnosable conditions.

7 The General Population Baseline

7.1 Survey Design

Establishing the disparate impact of mass surveillance on a protected class benefits from demonstrating the baseline impact on the general population. If individuals without paranoia-spectrum conditions report surveillance-related psychological effects, the amplified impact on those with pre-existing vulnerability can be inferred through established dose-response relationships in psychiatric epidemiology.

We propose a pilot survey instrument for future validation: "Does knowing that AI cameras track your movements and can be searched by law enforcement without a warrant make you feel like you're being watched at all times, no matter where you are?"

This question is binary (yes/no), uses no clinical vocabulary, and describes the psychological state in ordinary language. A respondent who answers "yes" is describing the subjective experience of pervasive surveillance awareness, which is the same subjective state that characterizes the persecutory ideation of paranoia-spectrum conditions, expressed without clinical framing.

This instrument is proposed as a starting point for a research program, not as a validated psychometric tool. A sin-

gle binary question does not establish a clinical baseline; it identifies the prevalence of a specific psychological experience in the general population. Formal validation would require test-retest reliability assessment, convergent validity testing against established surveillance concern measures (such as Pew Research Center instruments), and discriminant validity testing to ensure the question measures surveillance-specific anxiety rather than generalized anxiety. Nevertheless, even a pilot administration would generate indicative data. If a significant proportion of the general population answers "yes," the argument for amplified harm to clinically vulnerable populations follows from established dose-response relationships in psychiatric epidemiology.

7.2 Existing Survey Data

Existing survey data provides partial support for this approach. Pew Research Center has documented that a majority of Americans believe their online and offline activities are being tracked by the government and corporations, and that most feel they have little control over this data collection (Pew Research Center 2023). The behavioral consequences are documented: individuals report self-censoring online activity, avoiding certain websites, and changing communication habits in response to surveillance awareness.

These behavioral changes in the general population parallel, in attenuated form, the more extreme behavioral responses documented among individuals with paranoia-spectrum conditions, such as device abandonment and social withdrawal.

8 The Convergence Layer: Current Infrastructure and Emerging Capabilities

The three harm mechanisms documented in the preceding sections arise from surveillance infrastructure that is currently deployed and operational. This section examines how emerging technologies, some already in limited deployment and others in active development, will compound these harms if the current trajectory continues without intervention.

8.1 Individual, Networked, and Ambient Surveillance

The harm mechanisms documented in this paper are compounded by the convergence of surveillance modalities that is already underway.

Individual identification systems represent the emerging frontier. Meta's Name Tag facial recognition feature for consumer AR glasses, currently in limited deployment as of early 2026, can identify a person in the user's field of view and retrieve publicly available information about them. This capability is not yet ubiquitous, but the hardware (Meta Ray-Ban smart glasses) is commercially available, and independent researchers have already demonstrated that pairing consumer AR glasses with facial recognition databases enables real-time identification of strangers, including retrieval of Social Security numbers (Harvard I-XRAY project). Networked tracking systems (Flock Safety's ALPR network, DHS geofencing) are already deployed at scale as documented above. Combined, these systems enable what may be termed ambient surveillance: knowing who everyone is, where they are, and when they were there.

For someone whose persecutory delusions include referential features ("they are watching ME specifically, not just everyone"), the convergence of individual identification with networked tracking confirms the most pathological component of their belief system. Mass surveillance says "everyone is watched." Facial recognition says "you, specifically, can be identified." The combination says "you, specifically, are being watched everywhere you go." That is the precise content of referential persecutory delusions.

8.2 AI Behavioral Inference

The addition of AI-driven behavioral inference systems, currently deployed in predictive policing platforms and expanding into broader urban surveillance contexts, escalates the harm further. Systems that do not merely track location but infer intent, predict behavior, and construct behavioral profiles move the surveillance environment from "someone is watching" to "someone is watching me and drawing conclusions about my behavior."

This maps onto the symptom of thought broadcasting: the delusional belief that others can perceive or access one's thoughts. While AI behavioral inference does not literally read thoughts, its ability to predict behavior from movement patterns creates a functional equivalent from the patient's perspective. The system appears to know what they are going to do before they do it, which, for someone with thought-broadcasting delusions, constitutes confirmation of a symptom that treatment is designed to extinguish.

8.3 AI-Induced Harm Without Interaction

The emerging literature on AI-induced psychosis (Moore et al. 2026) focuses largely on harm arising from direct interaction between users and large language models, documenting pathways through which conversational AI systems degrade users' reality testing, create emotional dependency, or systematically dismantle their ability to distinguish AI-generated confabulation from fact. This paper documents a distinct and previously unidentified category: AI-induced psychological harm in the complete absence of user interaction, where the AI system's existence as surveillance infrastructure is itself the mechanism of harm. The affected individual never opens a chat window, never prompts a model, never receives a generated response. The AI operates in license plate readers, geofence analytics, behavioral inference engines, and facial recognition systems that the individual may never directly encounter. Yet the existence of these systems, combined with public awareness of their deployment, degrades the treatability of paranoia-spectrum conditions through the three mechanisms documented above. This distinction matters for the taxonomy of AI harms: current frameworks assume a user-system interaction as the site of harm. Mass surveillance AI demonstrates that the interaction is unnecessary. The infrastructure alone is sufficient. Moreover, if the infrastructure is sufficient to degrade treatability in those who already have paranoia-spectrum conditions, the

same dose-response logic that governs every other environmental risk factor in psychiatric epidemiology predicts that it is also sufficient to induce new cases in previously unaffected individuals, a mechanism addressed in the preceding subsection on Protected Class Expansion.

8.4 Connecting Convergence to Legal Remedy

The convergence of individual identification, networked tracking, and behavioral inference does not merely add new harms on top of the three mechanisms already identified. It transforms the nature of the legal claim.

The preceding sections document how current infrastructure creates disparate impact through reality testing destruction, device abandonment, and algorithmic pattern-matching. The convergence layer escalates the claim from "surveillance harms this population" to "surveillance is progressively constructing an environment that is functionally indistinguishable from the internal experience of persecutory psychosis."

Under ADA Title II, the relevant question is whether the government's methods of administration have the effect of subjecting individuals with disabilities to discrimination. When the surveillance environment replicates the phenomenology of the disability itself, the answer moves from arguable to self-evident. The reasonable modification analysis required under *Alexander v. Choate*, 469 U.S. 287, 301 (1985) (Alexander v. Choate 1985), must account not only for current harms but for the trajectory of harm as these technologies converge. A disability impact assessment conducted today that examines only ALPR cameras will be obsolete within months as facial recognition and behavioral inference capabilities are integrated into the same infrastructure.

9 Implications and Recommendations

9.1 Clinical Implications

Treatment providers require updated protocols that account for the surveillance environment. Specifically: Clinicians working with patients who express persecutory beliefs about surveillance must now conduct a differential assessment distinguishing between beliefs that are proportionate to the actual surveillance environment (and therefore not delusional) and beliefs that are disproportionate, generalized, or accompanied by additional delusional features. This differential assessment does not currently exist in the clinical literature.

Treatment plans built around reality testing for surveillance-related persecutory beliefs must be revised. Clinicians cannot ethically tell patients that "nobody is watching you" when mass surveillance systems are, in fact, watching everyone. Alternative therapeutic strategies that acknowledge the surveillance reality while addressing the disproportionate cognitive and emotional response must be developed.

"Treatment resistance" in patients with surveillance-related persecutory beliefs must be reassessed. A patient whose reality-testing intervention fails because the environment confirms their belief is not treatment-resistant; they are environment-resistant. Misattributing environment-resistance to treatment-resistance leads to inappropriate escalation of medication, unnecessary hospitalization, and erosion of the therapeutic alliance.

9.2 Legal Actions

The following legal actions are available under existing law: ADA Title II complaints filed with the DOJ Civil Rights Division, alleging that mass surveillance programs operated by state and local governments create disparate impact on individuals with paranoia-spectrum conditions. Section 504 complaints against federally funded surveillance programs through the relevant federal agency's Office for Civil Rights. State attorney general enforcement actions under COPPA against data brokers in the surveillance data pipeline, on the grounds that geofencing indiscriminately captures children's location data without parental consent. Requests for GAO assessment of federal surveillance programs' compliance with disability rights requirements.

9.3 Policy Recommendations

Mandatory disability impact assessments prior to the deployment or expansion of mass surveillance infrastructure. These assessments should specifically address the impact on populations with paranoia-spectrum conditions, including analysis of how surveillance publicity may affect therapeutic interventions.

Geofence data exclusion protocols for sensitive locations, including mental health treatment facilities, psychiatric hospitals, and residential treatment programs, to prevent the surveillance system from capturing data about individuals whose clinical vulnerability is highest.

Accommodation obligations for surveillance programs: if a surveillance deployment is shown to create disparate impact on a protected class, the deploying entity must demonstrate that reasonable modifications have been considered and that no feasible alternative exists that would reduce the impact.

10 Limitations

This paper has several significant limitations.

No longitudinal epidemiological data exists tracking psychosis outcomes in communities before and after surveillance deployment. The causal mechanism proposed here is grounded in clinical logic (the environmental change removes the precondition for an established therapeutic intervention) but has not been validated through population-level outcome studies.

The general population survey proposed above has not yet been conducted. The argument for amplified harm to clinically vulnerable populations currently relies on inference from established dose-response relationships rather than direct measurement.

Within-family genetic analyses have shown that some of the association between urban environments and psychosis risk is attributable to genetic confounding rather than environmental causation (Paksarian et al. 2017). This complicates any population-level claim about surveillance as a psychosis trigger, though it does not undermine the therapeutic

interference argument, which addresses treatment efficacy rather than disease onset.

The etiological argument presented above relies on analogy to established environmental risk factors for psychosis (urbanicity, childhood adversity, social defeat) rather than on direct epidemiological evidence that surveillance exposure produces new-onset paranoia-spectrum conditions. No study has isolated surveillance density as a variable in psychosis incidence. The dose-response framework proposed here is theoretically grounded but empirically untested. Longitudinal studies comparing psychosis incidence in high-surveillance and low-surveillance jurisdictions, controlling for urbanicity and other established risk factors, would be required to validate the etiological claim.

The self-sealing paradox identified above has not been tested in any court. Its legal viability is theoretical.

The algorithmic pattern-matching claim requires access to surveillance system flagging criteria, which are proprietary. The assertion that paranoia-spectrum behavioral symptoms would be flagged as anomalous is based on the general principle that these systems detect deviations from population norms, not on direct observation of specific flagging algorithms.

11 Conclusion

Mass surveillance infrastructure is not disability-neutral. The systems documented in this paper, including Flock Safety ALPR networks, DHS geofencing through Webloc and Penlink, cell-site simulators, and commercially purchased location data pipelines, create three distinct and compounding mechanisms of harm to Americans with paranoia-spectrum conditions. These harms are clinically specific, legally cognizable under existing disability rights law, and entirely absent from the surveillance policy literature, the clinical psychiatric literature, and the disability rights advocacy landscape.

The population most harmed by this infrastructure is the population least equipped to advocate for itself. Individuals experiencing active persecutory delusions are not filing ADA complaints. They are not testifying before congressional oversight committees. They are not writing op-eds about surveillance policy. They are, in many cases, throwing away their phones because the news confirmed that the government is tracking them, and then disappearing from every support system that was keeping them alive.

The legal tools exist. The clinical evidence supports the harm mechanism. The interdisciplinary research gap is documented. What is missing is recognition that mass surveillance is a disability rights issue, not merely a privacy issue, and that the population bearing the most severe and specific harm has been invisible to every institution with the power to protect them.

The window for intervention is now, before the surveillance infrastructure becomes so entrenched that the "reasonable modification" required under disability law becomes practically impossible. The precedent set by inaction will be difficult to reverse.

Researcher Positionality and Credential Statement

The CFP invites optional positionality, ethical considerations, and adverse impact statements. The author provides one here because the methodological commitments of this paper depend on epistemic positioning that academic detachment cannot supply.

The author writes from a position of dual standing. First, the author is disabled, with diagnosed neurodevelopmental and anxiety conditions (autism spectrum, ADHD, generalized anxiety disorder with agoraphobia) that produce direct lived experience of how environmental surveillance interacts with threat-detection systems, hypervigilance, and behavioral modification. The argument in this paper that surveillance functions as an environmental stressor in a dose-dependent fashion is not solely a theoretical claim derived from the urbanicity-psychosis literature; it is also a description of phenomenology the author has experienced.

Second, and more substantively for the clinical claims in this paper, the author is the primary caregiver to two family members with paranoia-spectrum conditions, with over 45,000 hours of caregiver experience across multiple active psychotic episodes. The author has watched reality-testing interventions succeed when the environment supported them and fail when the environment confirmed the delusion. The three harm mechanisms documented in this paper are not abstractions to the author.

The author is also the founder and executive director of a nonprofit organization dedicated to identifying and documenting AI-mediated harms, with a particular focus on harms to clinically vulnerable populations.

The author does not claim that lived experience confers authority over the clinical literature, the legal doctrine, or the surveillance scholarship cited in this paper. Lived experience grounds the methodological choices (which mechanisms to investigate, which harms to take seriously, which clinical framings to challenge) without displacing the burden of independent justification. Each empirical claim in this paper stands or falls on the cited evidence, not on the author's standing. The positionality is offered to make the methodological commitments visible to reviewers, in accordance with the conference's stated interest in reflexive accounts of ethics-in-practice.

The author discloses the use of speech-to-text software and Anthropic's Claude as assistive technology for the disclosed neurodevelopmental conditions. All claims, citations, and arguments are the author's own.

References

Alexander v. Choate. 1985. 469 U.S. 287.

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, 5 edition.

Brayne, S. 2017. Big Data Surveillance: The Case of Policing. *American Sociological Review*, 82(5): 977–1008.

Burns, A. V.; Nelson, K.; Wang, H.; Hegarty, E. M.; and Cohn, A. B. 2025. 'The Algorithm is Hacked': Analysis of Technology Delusions in a Modern-Day Cohort. *The British Journal of Psychiatry*. Published online November 3, 2025. doi:10.1192/bjp.2025.10452.

Cox, J. 2025. ICE to Buy Tool That Tracks Locations of Hundreds of Millions of Phones Every Day. *404 Media*. Published September 30, 2025.

Cox, J. 2026. CBP Tapped Into the Online Advertising Ecosystem to Track Peoples' Movements. *404 Media*. Published March 3, 2026.

Electronic Frontier Foundation. 2025. Automated License Plate Readers. EFF Street Level Surveillance.

Freeman, D. 2007. Suspicious Minds: The Psychology of Persecutory Delusions. *Clinical Psychology Review*, 27(4): 425–457.

Freeman, D. 2016. Persecutory Delusions: A Cognitive Perspective on Understanding and Treatment. *The Lancet Psychiatry*, 3(7): 685–692.

Freeman, D.; Bradley, J.; Antley, A.; Bourke, E.; DeWeever, N.; Evans, N.; et al. 2016. Virtual Reality in the Treatment of Persecutory Delusions: Randomised Controlled Experimental Study Testing How to Reduce Delusional Conviction. *The British Journal of Psychiatry*, 209(1): 62–67.

Freeman, D.; Dunn, G.; Startup, H.; Pugh, K.; Cordwell, J.; Mander, H.; et al. 2015. Effects of Cognitive Behaviour Therapy for Worry on Persecutory Delusions in Patients with Psychosis (WIT): A Parallel, Single-Blind, Randomised Controlled Trial with a Mediation Analysis. *The Lancet Psychiatry*, 2(4): 305–313.

Garety, P. A.; Fowler, D. G.; Freeman, D.; Bebbington, P.; Dunn, G.; Kuipers, E.; et al. 2008. Cognitive-Behavioural Therapy and Family Intervention for Relapse Prevention and Symptom Reduction in Psychosis: Randomised Controlled Trial. *The British Journal of Psychiatry*, 192(6): 412–423.

Garety, P. A.; Kuipers, E.; Fowler, D.; Freeman, D.; and Bebbington, P. E. 2001. A Cognitive Model of the Positive Symptoms of Psychosis. *Psychological Medicine*, 31(2): 189–195.

Goldstein-Street, J. 2026. WA Senate OKs Guardrails for License Plate Readers. *Washington State Standard*. Published February 4, 2026.

March, D.; Hatch, S. L.; Morgan, C.; Kirkbride, J. B.; Bresnahan, M.; Fearon, P.; and Susser, E. 2008. Psychosis and Place. *Epidemiologic Reviews*, 30(1): 84–100.

Moore, J.; Mehta, A.; Agnew, W.; Anthis, J. R.; Louie, R.; Mai, Y.; et al. 2026. Characterizing Delusional Spirals Through Human-LLM Chat Logs. *arXiv preprint arXiv:2603.16567*. To appear at ACM FAccT 2026.

Paksarian, D.; Trabjerg, B. B.; Merikangas, K. R.; Mors, O.; Børglum, A. D.; Hougaard, D. M.; et al. 2017. The Role of Genetic Liability in the Association of Urbanicity at Birth and During Upbringing with Schizophrenia in Denmark. *Psychological Medicine*, 48(2): 305–314.

Pew Research Center. 2023. How Americans View Data Privacy. Technical report, Pew Research Center. Report by Colleen McClain, Michelle Faverio, Monica Anderson, and Eugenie Park, October 18, 2023.

Pot-Kolder, R. M.; Geraets, C. N.; Veling, W.; van Beilen, M.; Staring, A. B.; Gijsman, H. J.; et al. 2018. Virtual-Reality-Based Cognitive Behavioural Therapy Versus Waiting List Control for Paranoid Ideation and Social Avoidance in Patients with Psychotic Disorders: A Single-Blind Randomised Controlled Trial. *The Lancet Psychiatry*, 5(3): 217–226.

Richardson, R.; Schultz, J. M.; and Crawford, K. 2019. Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. *New York University Law Review Online*, 94: 192–233.

Selbst, A. D. 2017. Disparate Impact in Big Data Policing. *Georgia Law Review*, 52(1): 109–195.

Selten, J.-P.; van der Ven, E.; Rutten, B. P.; Sham, P. C.; and Cantor-Graae, E. 2013. The Social Defeat Hypothesis of Schizophrenia: An Update. *Schizophrenia Bulletin*, 39(6): 1180–1186.

Ünver, H. A. 2024. Artificial Intelligence (AI) and Human Rights: Using AI as a Weapon of Repression and Its Impact on Human Rights. Technical report, European Parliament, Directorate-General for External Policies. PE 754.450.

U.S. House of Representatives, Subcommittee on Cybersecurity, Information Technology, and Government Innovation. 2026. Letter to DHS Secretary Kristi Noem Regarding Webloc and Penlink Surveillance Tools. Ranking Member Shontel Brown et al., February 19, 2026.

Varese, F.; Smeets, F.; Drukker, M.; Lieverse, R.; Lataster, T.; Viechtbauer, W.; et al. 2012. Childhood Adversities Increase the Risk of Psychosis: A Meta-Analysis of Patient-Control, Prospective- and Cross-Sectional Cohort Studies. *Schizophrenia Bulletin*, 38(4): 661–671.

Vassos, E.; Pedersen, C. B.; Murray, R. M.; Collier, D. A.; and Lewis, C. M. 2012. Meta-Analysis of the Association of Urbanicity with Schizophrenia. *Schizophrenia Bulletin*, 38(6): 1118–1123.