The Unfalsifiable Delusion

How mass surveillance infrastructure creates disparate impact on people with paranoia-spectrum conditions. A case study under ADA Title II.

Travis Gilly · Real Safety AI Foundation

Gilly, T. (2026). The Unfalsifiable Delusion: How Mass Surveillance Infrastructure Creates Disparate Impact on People with Paranoia-Spectrum Conditions. Working paper (v2), Real Safety AI Foundation.

"They are watching me everywhere I go" is no longer delusional. It is accurate.

CBT for persecutory delusions depends on falsifiability: the clinician helps the patient find evidence the belief is unfounded. Mass surveillance has eliminated that precondition for the general population.

SOURCES Freeman et al., The Lancet Psychiatry (2015); Garety et al. (2008); Goldstein-Street, Washington State Standard (2026).

PART I

Clinical Foundation

Reality testing, the falsifiability requirement, and a diagnostic paradox.

I

The intervention does not fail because the patient resists treatment. It fails because the environment changed.

The classification presupposes disconfirming evidence exists to find.

DSM-5 defines delusions as fixed beliefs "not amenable to change in light of conflicting evidence." When the world changes to match the belief's content, this is not clinical improvement. It is one symptom rendered diagnostically ambiguous by an environmental shift.

2020
"No cameras are specifically tracking you." The belief was a product of illness.

2026
"Yes. Flock cameras log your vehicle. DHS geofencing captures your phone." The belief is confirmed.

SOURCES American Psychiatric Association, DSM-5 (2013); Freeman (2016); Freeman (2007); Freeman et al. (2015); Garety et al. (2001).

2M+

Americans with active persecutory ideation as a current clinical feature. A conservative estimate.

Schizophrenia, schizoaffective disorder, delusional disorder, PTSD, and bipolar I with psychotic features.

This excludes prodromal phases and subclinical vulnerability. And it is not static: environmental exposure can expand the population by inducing new-onset conditions, addressed in Part III.

SOURCES World Health Organization (2025); Freeman (2007); Vassos et al. (2012); March et al. (2008); Selten et al. (2013).

A 119-page human-rights analysis of AI surveillance. Zero mentions of "disability."

SURVEILLANCE POLICY

Treats the surveilled population as undifferentiated. No examination of disability-specific harm.

CLINICAL PSYCHIATRY

Never examines whether real-world surveillance changes affect reality-testing efficacy.

DISABILITY ADVOCACY

Should be the connective tissue. Has not identified mass surveillance as a disability issue.

SOURCES Ünver, European Parliament, PE 754.450 (2024); Pot-Kolder et al., The Lancet Psychiatry (2018); Freeman et al. (2016); Burns et al., The British Journal of Psychiatry (2025).

PART II

The Surveillance Landscape

Flock Safety, DHS geofencing, and the three mechanisms of harm that follow.

II

This surveillance is not secret. It is publicized, debated, and marketed.

Weather apps and games feed a pipeline that ends at a federal warrant-free geofence.

Flock Safety ALPR

Every vehicle logged, stored, searchable without a warrant. Individuals logged hundreds of times a month.

DHS Webloc / Penlink

Consumer app data sold through brokers, drawing a warrantless digital fence around any area.

No exclusion mechanism

Children's phones, psychiatric patients' phones, every phone in the boundary is captured.

SOURCES Electronic Frontier Foundation (2025); Cox, 404 Media (2025, 2026); U.S. House Subcommittee on Cybersecurity, IT, and Government Innovation (2026); Zuboff (2019).

Credibility contamination: the clinician who was wrong once is no longer trusted about anything.

The treatment does not fail because the patient is treatment-resistant. It fails because the environment removed the tool the treatment depended on. And paranoia does not compartmentalize: confirmed macro-level surveillance becomes the evidentiary foundation on which every other persecutory belief becomes harder to challenge.

SOURCES Freeman, The Lancet Psychiatry (2016); Garety et al. (2001); Burns et al. (2025).

A rational decision that severs every lifeline at once.

Learning that geofencing tracks phones, a person discards theirs. Not delusional; evidence-based. But the phone connects them to therapists, case managers, and crisis services.

INVISIBLE FROM OUTSIDE

No one calls 911 when someone stops answering their phone. Weeks or months may pass before anyone recognizes a pattern of withdrawal.

Intent is not required for disparate impact.

SOURCES Pew Research Center (2023).

The algorithm cannot distinguish a burglar casing a building from a person circling it out of fear.

Erratic movement, irregular sleep, pacing during acute anxiety, sudden routine changes, all produce the same geolocation signature the system was built to flag as suspicious. Additional scrutiny then confirms the persecutory belief system it was never designed to account for, creating a closed feedback loop.

SOURCES Richardson, Schultz & Crawford, N.Y.U. Law Review Online (2019); Brayne, American Sociological Review (2017); Selbst, Georgia Law Review (2017).

PART III

Legal Framework

ADA Title II, Section 504, and the self-sealing paradox.

III

Disparate impact claims do not require discriminatory intent.

"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."

28 C.F.R. § 35.130(b)(3)

No intent required

Facially neutral policy, disproportionate harm, is enough.

Liability on the procurer

The municipality that chose the technology, not the private vendor, under state action doctrine.

Section 504 extends it

Any federally funded program, catching DHS geofencing and funded law enforcement.

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

If the belief is no longer delusional, the person may lose qualifying-disability status, and with it, standing to challenge the surveillance. But the paradox contains its own rebuttal.

What the surveillance takes from the diagnostic category on one end, it replenishes on the other, by inducing new cases in people with no prior history.

No court has tested this theory. Its legal viability is, at present, theoretical.

Same structure as industrial disease: liability from dose-response, not pre-existing vulnerability.

LOW
A single article. Transient discomfort.

MODERATE
Sustained behavioral modification, documented by Pew.

HIGH
Persistent hypervigilance, meeting subclinical thresholds. Some tip into clinical territory.

The government cannot argue accurate beliefs are non-delusional while its own infrastructure creates the clinical populations it would disclaim.

SOURCES Selikoff, Churg & Hammond, JAMA (1968); Vassos et al. (2012); Varese et al. (2012); Selten et al. (2013); Paksarian et al. (2017); Pew Research Center (2023).

"Everyone is watched" plus "you can be identified" equals "you are being watched, specifically."

Individual identification (AR facial recognition) converging with networked tracking constructs the precise content of referential persecutory delusions. AI behavioral inference, predicting intent, maps onto thought broadcasting. This paper's category, AI-induced harm requiring no interaction at all, means the reasonable-modification analysis under *Alexander v. Choate*, 469 U.S. 287 (1985), must account for the trajectory of convergence, not just current deployment.

SOURCES Moore et al., arXiv:2603.16567 (2026). **AUTHORITY** Alexander v. Choate, 469 U.S. 287 (1985).

"Treatment resistance" must be reassessed as environment resistance.

CLINICAL

Differential assessment of proportionate versus disproportionate surveillance beliefs; revised reality-testing protocols.

LEGAL

ADA Title II and Section 504 complaints with the respective civil rights offices; GAO assessment requests.

POLICY

Mandatory disability impact assessments; geofence exclusion for treatment facilities; documented accommodation review.

No longitudinal data	No study has tracked psychosis outcomes before and after surveillance deployment.
Untested survey	The proposed general-population pilot has not been administered.
Genetic confounding	Complicates urbanicity claims and may apply here, though it does not undermine the treatment-efficacy argument.
Untested paradox	The self-sealing paradox has not been tested in any court. Its legal viability is theoretical.

SOURCES Paksarian et al. (2017).

IN CLOSING

Mass surveillance is a disability rights issue, not merely a privacy issue.

The population most harmed is least equipped to advocate for itself. Not filing complaints. Not testifying before oversight committees. Throwing away their phones because the news confirmed the government is tracking them, then disappearing from every support system that was keeping them alive.

The window for intervention is now, before reasonable modification becomes practically impossible.

REFERENCES & AUTHORITIES

Gilly, T. (2026). *The Unfalsifiable Delusion: How Mass Surveillance Infrastructure Creates Disparate Impact on People with Paranoia-Spectrum Conditions.* Working paper (v2), Real Safety AI Foundation.

SCHOLARSHIP AND REPORTS

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). American Psychiatric Publishing.

Brayne, S. (2017). Big data surveillance: The case of policing. American Sociological Review, 82(5), 977–1008.

Burns, A. V., Nelson, K. A., & Wang, H. (2025). "The algorithm is hacked": Analysis of technology delusions in a modern-day cohort. The British Journal of Psychiatry, 1–5.

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., et al. (2016). Virtual reality in the treatment of persecutory delusions. The British Journal of Psychiatry, 209(1), 62–67.

Freeman, D., Dunn, G., Startup, H., et al., & Kingdon, D. (2015). Effects of cognitive behaviour therapy for worry on persecutory delusions in patients with psychosis (WIT). The Lancet Psychiatry, 2(4), 305–313.

Garety, P. A., Fowler, D. G., Freeman, D., et al. (2008). Cognitive-behavioural therapy and family intervention for relapse prevention and symptom reduction in psychosis. The British Journal of Psychiatry, 192(6), 412–423.

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A cognitive model of the positive symptoms of psychosis. Psychological Medicine, 31(2), 189–195.

March, D., Hatch, S. L., Morgan, C., et al. (2008). Psychosis and place. Epidemiologic Reviews, 30(1), 84–100.

Moore, J., Mehta, A., Agnew, W., et al. (2026). Characterizing delusional spirals through human-LLM chat logs. arXiv preprint arXiv:2603.16567.

Paksarian, D., Trabjerg, B. B., Merikangas, K. R., et al., & Pedersen, C. B. (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.

Pot-Kolder, R. M., Geraets, C. N., Veling, W., et al. (2018). Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders. The Lancet Psychiatry, 5(3), 217–226.

Richardson, R., Schultz, J. M., & 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, 15–55.

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., & 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. European Parliament, Directorate-General for External Policies. PE 754.450.

Varese, F., Smeets, F., Drukker, M., et al., & Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: A meta-analysis. Schizophrenia Bulletin, 38(4), 661–671.

Vassos, E., Pedersen, C. B., Murray, R. M., Collier, D. A., & Lewis, C. M. (2012). Meta-analysis of the association of urbanicity with schizophrenia. Schizophrenia Bulletin, 38(6), 1118–1123.

World Health Organization. (2025). Schizophrenia. WHO.

Zuboff, S. (2019). The age of surveillance capitalism. PublicAffairs.

GRAY LITERATURE AND POLICY

Cox, J. (2025, Sept. 30). ICE to buy tool that tracks locations of hundreds of millions of phones every day. 404 Media.

Cox, J. (2026, March 3). CBP tapped into the online advertising ecosystem to track peoples' movements. 404 Media.

Electronic Frontier Foundation. (2025). Automated license plate readers. EFF Street Level Surveillance.

Goldstein-Street, J. (2026, Feb. 4). WA Senate OKs guardrails for license plate readers. Washington State Standard.

Pew Research Center. (2023). How Americans view data privacy. Pew Research Center.

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.

CASES (full Bluebook citation on the slide where cited, not repeated here): Alexander v. Choate, 469 U.S. 287 (1985). **STATUTES:** ADA Title II, 42 U.S.C. § 12131 et seq.; 28 C.F.R. § 35.130(b)(3); Rehabilitation Act § 504, 29 U.S.C. § 794.