

WHITE PAPER

SEPTEMBER 2026

Surveillance Capitalism and the Dignity Deficit

How pervasive AI surveillance systems erode the conditions for autonomous selfhood, and what regulatory floors are needed to restore them.

The Digital Dignity Institute of California

DDI White Paper Series | San Francisco | digitaldignityinstitute.com

Surveillance Capitalism and the Dignity Deficit

How pervasive AI surveillance systems erode the conditions for autonomous selfhood, and what regulatory floors are needed to restore them

*September 2026. DDI White Paper Series. The Digital Dignity Institute of California, San Francisco.
Prepared by the Digital Dignity Institute Research Team. Issued institutionally.*

Abstract

Pervasive AI surveillance systems erode the conditions for autonomous selfhood. This white paper examines how, and sets out the regulatory floors needed to restore them.

Data extraction is no longer confined to the browser or the workplace. Networked devices, biometric systems, location services, connected vehicles, and AI-enabled cameras now transform ordinary movement, communication, and participation into data points. These systems normalise continuous observation and make privacy a condition that individuals must repeatedly negotiate rather than a right they can meaningfully exercise.

The consequence is a dignity deficit. Human dignity depends in part on the practical ability to experiment, associate, and make decisions without constant behavioural inference. When people reasonably expect to be profiled or scored, they adapt their conduct before any decision is made about them, which produces diminished autonomy, weakened social trust, and a distribution of scrutiny that follows existing lines of power and inequality. Artificial intelligence intensifies this by converting fragmented and unremarkable data into predictions about identity, intent, risk, and worthiness, and those predictions shape access to work, housing, credit, public services, and public space.

We propose enforceable limits on high-risk behavioural surveillance, meaningful transparency and contestation rights, strict necessity and proportionality tests, and independent oversight for systems that influence consequential decisions. Together these establish a regulatory floor: a set of protections below which commercial and institutional data practices should not be permitted to fall.

KEY FINDINGS

01. Pervasive data collection changes behaviour before any automated decision is made, weakening autonomy and public trust.

02. AI inference creates new forms of surveillance by deriving sensitive conclusions from seemingly ordinary data.

03. The burdens of behavioural surveillance are distributed unevenly and frequently reinforce existing social inequalities.

04. Notice-and-consent models cannot provide meaningful protection where surveillance is ambient, opaque, and unavoidable.

05. Enforceable limits, contestation rights, and independent oversight are necessary regulatory floors for digital dignity.

Contents

1. Methodology
2. The Expansion of Pervasive Surveillance
3. The Dignity Deficit
4. AI Inference and the New Surveillance
5. Unequal Burden
6. Why Notice-and-Consent Fails
7. The Regulatory Landscape
8. Regulatory Floors for Dignity
9. Objections and Responses
10. Implementation and Sequencing
11. Policy Recommendations
12. Conclusion
 - Annex A: Model Schedule of Prohibited Practices
 - Annex B: Specification for an Exercisable Right of Contestation
 - Sources and Further Reading

1. Methodology

This paper is an interdisciplinary analysis drawing on privacy law, international human rights standards, political economy, documented surveillance deployments, and interviews with civil society, labour, and digital-rights practitioners. Each strand is set out below with its limitations, since the argument's reach depends on what each can and cannot support.

Legal analysis. We reviewed data protection instruments, artificial intelligence statutes and proposals, sectoral rules governing employment, credit, biometrics, and children's data, and the enforcement records and guidance of supervisory authorities across several jurisdictions. Legal positions are stated as at mid-2026. Several instruments are in transitional application, awaiting implementing regulation, or subject to challenge, and readers should verify current status before operational reliance.

Human rights standards. We draw on the treaty framework and its interpretive materials, particularly the jurisprudence on private life, freedom of assembly and association, non-discrimination, and the developing body of guidance on data protection as a condition of other rights rather than as a standalone interest. This literature supplies the paper's normative vocabulary and the proportionality apparatus applied in Section 8.

Political economy. We engage the scholarship on the commercial logic of behavioural data extraction, including the account of surveillance capitalism developed by Zuboff and the substantial critical literature that has followed it. That literature supplies the account of incentive structure that explains why voluntary restraint has not materialised and why the paper's proposals are framed as floors rather than as guidance.

Documented deployments. We assembled documented cases of behavioural surveillance and inference systems in employment, housing, credit, public administration, education, retail, and policing, drawn from regulatory enforcement records, litigation, procurement disclosures, freedom of information releases, investigative journalism, and vendor materials. The corpus is systematically biased toward deployments that became visible, meaning those that were litigated, leaked, disclosed, or investigated. Ambient systems that function without controversy are underrepresented, which means the record understates prevalence and may overstate the frequency of malfunction relative to routine operation.

Practitioner interviews. We conducted semi-structured interviews with practitioners in civil society organisations, trade unions and worker advocacy bodies, and digital rights litigation. Interviews were used to identify operational failure points in existing rights, particularly the gap between formal entitlements and their exercisability, and to test the practicality of the floors proposed in Section 8. Interview material is used to establish the existence and character of obstacles rather than their frequency, and interviewees were selected for expertise rather than representativeness.

Analytical framework. The paper applies two concepts from the Institute's earlier work. The **recourse deficit** describes the distance between the harms a system produces and the remedies actually reachable by those who suffer them. The **discovery burden** describes the cost, borne by the affected party, of establishing that a harm occurred at all. Behavioural surveillance is the domain in which the second concept does the most work, because the characteristic feature of inference-driven decision-making is that the person affected does not know an

inference was made, does not know what it contained, and frequently does not know a decision was taken.

2. The Expansion of Pervasive Surveillance

2.1 Beyond the browser and the workplace

Privacy debate inherited its vocabulary from an era in which data collection had edges. There was a session, a site, a transaction, a workplace. Collection happened at identifiable moments, in identifiable places, through identifiable interfaces, and the regulatory apparatus built in that period reflects those assumptions: notice at the point of collection, consent at the point of interaction, purpose specification at the point of processing.

Data extraction is no longer confined to the browser or the workplace. Networked devices, biometric systems, location services, connected vehicles, and AI-enabled cameras now transform ordinary movement, communication, and participation into data points. The collection surface has expanded from the interface to the environment, and the moments at which a person could plausibly be said to encounter a collection decision have largely disappeared.

The change is not merely quantitative. When collection happened at edges, the person retained a meaningful choice architecture: they could decline the site, close the account, leave the shop. When collection is ambient, the equivalent choice is to withdraw from movement, employment, transport, healthcare, education, and civic participation. Framing this as a choice is a category error that the current regulatory model nonetheless depends on.

2.2 The collection surface

The following is not exhaustive and is intended to establish the scope of the surface rather than to catalogue it.

Personal and household devices. Phones, wearables, voice assistants, televisions, appliances, and home security systems generate continuous telemetry on location, movement, sleep, physiological state, media consumption, and household composition. The data is frequently transmitted to parties with whom the person has no relationship and about whom they have no knowledge.

Biometric systems. Facial recognition and analysis, gait and body recognition, voice identification, and emotion inference are deployed in retail, transport hubs, stadiums, workplaces, schools, housing, and public space. Biometric identifiers are, unlike credentials, unrevocable, and inference systems built on them assert conclusions about internal states that the underlying science does not support.

Location and mobility. Handset location, application software development kits, wireless network probing, transit ticketing, tolling, and automated numberplate recognition produce continuous mobility records. Location data is uniquely revealing because destinations disclose the categories of activity that other law treats as most sensitive, including worship, medical care, union meetings, protest, and intimate association.

Connected vehicles. Modern vehicles generate telemetry on location, speed, braking, occupancy, cabin audio and video in some configurations, and driver behaviour, with onward transmission to manufacturers, insurers, fleet operators, and data brokers. Regulatory attention to this category is markedly thinner than to comparable data flows in other sectors.

AI-enabled cameras. The distinction that matters is not between camera and no camera but between recording and analysis. A camera that records produces a passive archive requiring effort to search. A camera with analysis attached produces a continuously updated, queryable, and cross-referenceable record of who was present, with whom, doing what, and how often. The second is a categorically different instrument and it has largely been introduced under regulatory frameworks written for the first.

Workplace systems. Productivity monitoring, communications analysis, keystroke and screen capture, badge and location tracking, scheduling systems, and wearable safety monitoring. The employment relationship removes even the nominal voluntariness on which consent-based protection rests.

Commercial and civic environments. Retail analytics, loyalty inference, connected advertising infrastructure, building access systems, and the data practices of public bodies delivering benefits, housing, education, and healthcare. Public and private collection are increasingly reticulated through procurement, data sharing, and brokerage, which means the traditional distinction between state and commercial surveillance no longer describes the actual architecture.

2.3 From transaction to condition

The cumulative effect of the surface described above is to convert privacy from a right that can be exercised into a condition that must be continuously negotiated, and negotiated on terms set by parties with vastly greater information and no obligation to accept a counteroffer.

Three properties characterise this shift.

Ambience. Collection is not an event that the person attends. It is a property of the environment they occupy. There is no moment at which a decision presents itself and therefore no moment at which the apparatus of notice and choice can attach.

Aggregation. Individual data points are unremarkable and the composite is not. A location record, a purchase, a device identifier, and a browsing pattern are each innocuous; assembled and joined across sources, they support conclusions the person has never disclosed to anyone. Regulation that assesses sensitivity at the level of the data point systematically misses the sensitivity of the assembly.

Irrevocability. Data extracted persists, propagates, and is re-used for purposes unformed at the time of collection. Withdrawal of consent, where it exists, operates prospectively on one relationship in a chain the person cannot see. The practical position is that collection is a one-way ratchet, and any framework that treats consent withdrawal as a meaningful remedy is describing an architecture that does not exist.

2.4 Scope and limits

The paper addresses behavioural surveillance conducted by commercial and institutional actors, including public bodies operating commercially procured systems. It does not address national security and intelligence surveillance, which raises distinct questions of authorisation, oversight, and secrecy that a paper of this length cannot treat responsibly. Where intelligence practice bears on the commercial architecture, principally through data purchase and through the normalising effect of state adoption, we note it and do not develop it.

The paper does not attempt to quantify aggregate behavioural effects of surveillance across a population. The measurement problems are severe, the identification strategies weak, and confident quantitative claims in this area generally outrun their evidence. Our claims concern mechanism, institutional adequacy, and the distribution of burden, each of which the available evidence supports.

The paper takes no position on the desirability of data-intensive services as such. The argument is about floors, meaning the protections below which practice should not be permitted to fall, and floors are compatible with a wide range of commercial models above them.

3. The Dignity Deficit

3.1 What dignity requires in practice

Dignity is frequently invoked in this field and rarely specified, which allows it to function as a rhetorical placeholder rather than as an analytical instrument. This paper uses it in a deliberately practical sense.

Human dignity depends in part on the practical ability to experiment, associate, and make decisions without constant behavioural inference. Three capacities do the work.

Experimentation. The capacity to try out beliefs, affiliations, tastes, identities, and arguments provisionally, without each attempt being recorded, aggregated, and treated as evidence of a settled disposition. Personal development requires the ability to hold positions loosely, and a permanent record of every provisional step converts development into a liability.

Association. The capacity to meet, organise, and affiliate without the fact of association being registered and inferred upon. Freedom of assembly has always presupposed a degree of practical obscurity, and the systematic removal of that obscurity changes the character of the right without amending its text.

Unmonitored decision. The capacity to deliberate, hesitate, change one's mind, and act without the process being observed and scored. Decision-making that is continuously legible to an observer with power over one's circumstances is decision-making under supervision, whatever its formal freedom.

These capacities are not private goods. They are the preconditions of the public activities that a democratic society depends on, which is why the paper treats surveillance as a structural rather than an individual concern.

3.2 Anticipatory adaptation

Pervasive data collection changes behaviour before any automated decision is made. This is the paper's central mechanism and it is frequently misunderstood in policy debate, which tends to locate the harm at the point of decision.

When people reasonably expect to be profiled or scored, they adapt their conduct in advance. The adaptation does not require them to know that a specific system exists, to understand how it works, or to have suffered any adverse outcome. It requires only a general and accurate belief that observation is occurring and that observation has consequences. That belief is now correct, widely held, and rational.

The adaptations documented in the literature and reported in our interviews are consistent in character. People self-censor in searches and communications, particularly on health, legal, financial, and political subjects. They avoid affiliations that seem likely to attract attention. They manage their conduct in workplaces to satisfy metrics rather than purposes. They decline to seek assistance from public bodies where seeking assistance creates a record. They present a curated self in circumstances where the curated self is what will be assessed.

Three features make this harm difficult for existing law to reach.

It occurs **before** any processing decision that a legal instrument would recognise as an event. There is no decision to contest, no output to explain, and no adverse action to appeal.

It is **self-inflicted in form and coerced in substance**. The person adapts their own behaviour, which makes the harm look like a choice and permits the deployer to disclaim responsibility for it.

It is **invisible in the aggregate**. The searches not made, the meetings not attended, and the assistance not sought leave no trace. A regulator reviewing the operation of a system sees the population that engaged with it and never sees the population that withdrew.

This is anticipatory adaptation, and it is the principal mechanism by which surveillance produces the deficit this paper describes.

3.3 Trust as a collective good

The second consequence is the weakening of social trust, and it operates at the level of institutions rather than individuals.

Trust in an institution depends on a defensible belief about what the institution will do with what it learns. Where the answer is unknown, unbounded, or subject to unilateral change, the rational posture toward the institution is guarded, and guardedness has costs that fall on the institution's function. A health system whose patients withhold information provides worse care. A public agency whose applicants minimise disclosure administers benefits badly. An employer whose workers manage their observability loses the informal communication on which most organisations actually depend.

There is a further, slower effect. Populations become inured to observation, and the normalisation is itself a harm to the extent that it removes the social resistance that would otherwise constrain expansion. A practice that would have been contested when introduced becomes unremarkable, and the baseline against which the next expansion is judged moves accordingly. This ratchet operates without anyone deciding that it should.

3.4 The distribution of scrutiny

The burdens of behavioural surveillance are distributed unevenly and frequently reinforce existing social inequalities. This is not incidental to how these systems are deployed. It follows from where they are deployed and from who is in a position to refuse them.

Monitoring intensity correlates with dependence. It is heaviest in low-wage employment, in public benefit administration, in subsidised and managed housing, in immigration processes, in schools serving lower-income populations, and in policing directed at particular neighbourhoods. Each of these is a context in which the observed party requires something from the observing party and cannot decline the terms.

The consequence is that the population most exposed to behavioural surveillance is the population least able to contest it, least able to absorb an erroneous inference, and least likely to have access to the legal and technical resources that formal rights presuppose. Section 5 develops this by population. The point to establish here is structural: a system of protection that depends on individuals asserting rights allocates protection in proportion to capacity, and capacity is distributed inversely to exposure.

3.5 The deficit defined

Drawing the three effects together, the dignity deficit is the gap between the conditions that autonomous selfhood requires and the conditions that pervasive behavioural surveillance leaves available.

It comprises diminished autonomy, expressed as anticipatory adaptation of conduct in advance of any decision; weakened social trust, expressed as guardedness toward institutions and the degradation of their function; and maldistributed scrutiny, expressed as the concentration of observation on those least able to refuse or contest it.

The deficit is not the sum of individual privacy injuries. It is a change in the conditions under which a population conducts its life, and it is therefore not remediable through instruments that require individuals to identify, evidence, and litigate discrete harms to themselves. That structural mismatch between the shape of the harm and the shape of the remedy is the argument for floors, developed in Section 8.

4. AI Inference and the New Surveillance

4.1 From collection to inference

AI inference creates new forms of surveillance by deriving sensitive conclusions from seemingly ordinary data. This is the development that makes the current moment different from the one that data protection law was

built for.

Classical privacy regulation is organised around collection. It asks what data was taken, whether the person was told, whether they agreed, and whether the use matches the stated purpose. That architecture assumes a stable relationship between the sensitivity of a datum and the sensitivity of what can be learned from it, and machine learning has severed that relationship.

Systems now derive predictions about identity, intent, risk, and worthiness from inputs that are individually innocuous and were lawfully collected for unrelated purposes. Movement patterns support inferences about health conditions and religious observance. Purchase timing supports inferences about pregnancy. Typing dynamics and application usage support inferences about cognitive state and emotional condition. Social graph structure supports inferences about immigration status, sexuality, and political affiliation. Device and connection metadata support inferences about household composition and economic precarity.

Two consequences follow immediately.

Sensitive-category protection is circumvented without being violated. Most frameworks give special protection to defined categories, including health, biometrics, religion, politics, and sexuality. Inference reaches those categories from inputs outside them. The processing is lawful at every step and the outcome is a conclusion about a protected attribute that the person never disclosed and cannot correct.

Data minimisation is undermined at its foundation. The principle assumes that limiting collection limits knowledge. Where inference is powerful, a small quantity of unremarkable data supports a large quantity of sensitive conclusions, and minimisation applied to inputs delivers far less protection than it promises.

4.2 The proxy problem

Where a system is prohibited from using an attribute, it will typically reconstruct that attribute from permitted inputs, not through designer intent but because the attribute carries predictive signal and optimisation finds it.

This has been demonstrated repeatedly across domains: postcode and purchase history reconstructing ethnicity, career gaps and device patterns reconstructing parental status and disability, name and language reconstructing national origin. Prohibiting the input does not remove the inference. It removes the audit trail that would have made the inference visible.

The implication for regulatory design is significant and it is generally resisted because it is inconvenient. Input-level prohibitions are close to unenforceable against modern systems. Effective constraint has to operate at the level of outcome and of purpose, meaning what the system is permitted to conclude and what conclusions may be acted upon, rather than what it is permitted to observe.

4.3 Consequential domains

These inferences shape access to work, housing, credit, public services, and public space, and the paper's

proposals attach to consequence rather than to technology for that reason.

Work. Applicant screening, video assessment, productivity scoring, attrition prediction, scheduling optimisation, and security and integrity monitoring. Inference determines who is hired, how they are scheduled, what they are paid, whether they are retained, and increasingly whether they are surveilled more intensively than colleagues.

Housing. Tenant screening, income and stability prediction, eviction and complaint history matching, and property access systems using biometric identification. Screening errors in this domain are unusually consequential because a denial produces a record that follows the applicant into subsequent applications.

Credit and insurance. Alternative data underwriting, behavioural pricing, and fraud and risk scoring. The sector has the longest history of formal regulation and remains the clearest illustration of the gap between a formal right of correction and an exercisable one.

Public services. Eligibility determination, fraud risk scoring in benefit administration, child welfare risk prediction, and immigration processing. Errors here fall on people with the least capacity to absorb them and the most difficulty in contesting them, and the systems are frequently procured rather than built, which diffuses accountability across a chain of vendors.

Public space. Facial recognition in retail and transport, behavioural analytics in stadiums and civic buildings, and predictive deployment in policing. The distinctive feature of this domain is that exclusion can occur without a decision being communicated at all: a person is simply refused entry, followed, or attended to differently, with no notice that a system was involved.

4.4 Opacity and the discovery burden

People affected frequently cannot see, understand, or challenge the underlying model. Opacity operates at four levels and each defeats a different remedy.

Existence. The person does not know a system was used. This is the most common condition and it defeats every downstream right at once, since rights that must be invoked are worthless to a person with no reason to invoke them.

Input. The person does not know what data the system received, which is typically assembled from sources they have never dealt with, including brokers, aggregators, and partners.

Logic. The person does not know how the inputs produced the output. Explanations, where provided, are often generic, and the meaningful explanation is frequently claimed as commercial confidence.

Consequence. The person does not know how heavily the output weighed, because the human decision-maker in the loop is presented as the responsible actor while functioning, in practice, as an endorsement of the score.

Applying the Institute's framework: to obtain a remedy, a person must discover that a system exists, establish what it concluded about them, identify the responsible party in a procurement chain, locate a legal theory that

fits, evidence a harm that is frequently the absence of an opportunity rather than the presence of an injury, and fund the effort. Each step is a filter and the compound rate of survival is close to zero. The rights exist. The path does not.

This is why the paper's proposals place the discovery burden on the deployer. There is no configuration of individual rights that solves a problem whose first step is knowing that something happened to you.

5. Unequal Burden

The paper identifies heightened risks for workers, migrants, children, and communities already exposed to disproportionate monitoring. The organising observation across all four is that surveillance intensity tracks dependence, and dependence removes the capacity to refuse.

5.1 Workers

The employment relationship is where the consent model most clearly fails, since agreement obtained under threat of losing one's livelihood is not agreement in any sense the law elsewhere recognises. Several jurisdictions have accepted this in principle and few have followed it to its conclusion.

Contemporary workplace monitoring includes activity and productivity measurement, communications analysis, video and audio capture, location and biometric tracking, wearable monitoring, and inference systems that predict attrition, organising activity, and performance trajectory. Three features distinguish it from consumer surveillance.

Comprehensiveness. The employer observes a large share of waking hours, and where work is performed at home the boundary between employment monitoring and domestic observation collapses.

Consequence. Outputs determine scheduling, pay, discipline, and continued employment, and they do so continuously rather than at review points.

Chilling of collective action. Systems that predict attrition or flag communication patterns produce, as a by-product, indicators of organising activity. Whether or not that use is intended, workers report adapting their communications accordingly, which is anticipatory adaptation operating directly on a protected right.

Our interviews with union and worker advocacy practitioners identified a consistent operational failure. Individual data rights are structurally unsuited to workplace conditions, because exercising them identifies the worker as a complainant, and workers understand this. Collective exercise, meaning representative rights held by unions or works councils, is the only mechanism that practitioners regarded as usable, and it exists in a minority of jurisdictions.

5.2 Migrants

People in immigration processes face surveillance of a comprehensiveness applied to almost no other

population, combined with the weakest available capacity to contest it.

Documented practice includes biometric registration and cross-database matching, social media examination, device search and extraction, location monitoring through electronic reporting requirements, and risk scoring applied to applications and to enforcement prioritisation. Data sharing between immigration authorities and health, education, housing, and welfare bodies extends the surveillance surface into every service the person needs.

Three factors compound the burden. The consequences of an adverse inference are among the most severe any administrative system imposes. Procedural protections are frequently reduced relative to other administrative contexts. And the population is precisely the one for which raising a complaint carries the greatest perceived risk, which means non-exercise of rights should be read as a rational response rather than as an absence of grievance.

The service-sharing point deserves emphasis because it produces a distinctive form of anticipatory adaptation. Where a person believes that seeking healthcare, enrolling a child in school, or reporting a crime may generate a record accessible to immigration enforcement, the rational response is to avoid those services. The resulting harms are borne by the individual, by public health, and by the enforcement of criminal law, and they are invisible to any assessment that examines only the operation of the systems themselves.

5.3 Children

Children are subject to surveillance across every domain of their lives, and the frameworks purporting to protect them are among the weakest in practice.

Educational technology generates continuous records of attention, comprehension, behaviour, and social interaction. School security systems increasingly incorporate biometric identification and behavioural analytics. Consumer services collect from children at scale under age assurance mechanisms that are widely known to be ineffective. Parental monitoring products normalise observation within the family.

Two considerations are specific to this population.

Developmental. The capacity for experimentation described in Section 3.1 is most important during the period in which identity is formed. Continuous observation during that period is a direct constraint on the developmental process, and it is the population least able to articulate that objection.

Longitudinal. A record beginning in early childhood accumulates into a palimpsest of scores, flags, and classifications, each layer written over the last without erasing it, following the person into adulthood. The child who is flagged at nine has no realistic prospect of contesting that flag before it has shaped a decade of institutional treatment.

Consent-based protection is at its least coherent here, since it is exercised by a parent, over a period the child cannot foresee, in respect of consequences that arrive after majority.

5.4 Communities under disproportionate monitoring

Communities already subject to heavy policing, benefit administration scrutiny, or immigration enforcement experience surveillance as a cumulative condition rather than as a set of separate encounters.

Three mechanisms drive the concentration.

Historical data. Systems trained on records of past institutional attention direct future attention to the same populations, and the resulting encounters generate further records. The feedback loop is well documented in the predictive policing literature and appears in benefit fraud scoring and child welfare risk prediction on the same logic.

Deployment geography. Sensing infrastructure is installed where institutions have decided attention is warranted, and those decisions carry the history of the institutions that made them.

Service dependence. Populations that interact more frequently with public administration generate more records, and more records support more inference.

The result is that the same conduct produces different consequences depending on where it occurs and who performs it, which is a fair description of unequal treatment even where every individual system is facially neutral. The immiseration this produces is gradual and is not attributable to any single decision, which is precisely what makes it invisible to remedies organised around discrete adverse actions.

5.5 Compounding

The four populations above are not separate categories and the largest burdens fall at their intersections: the migrant worker in monitored low-wage employment, the child in a family under benefit fraud scrutiny, the resident of a heavily policed neighbourhood applying for housing.

Compounding matters for regulatory design because the harm is cumulative while assessment is systemic. Every impact assessment regime in force evaluates a single system against its own purpose. None evaluates the aggregate burden borne by a person subject to twelve of them. Cumulative exposure is the form the harm actually takes and it is measured nowhere.

6. Why Notice-and-Consent Fails

Notice-and-consent models cannot provide meaningful protection where surveillance is ambient, opaque, and unavoidable. The model persists because it is administratively convenient and because it distributes responsibility in a way that suits the parties who designed it, not because there is evidence that it protects anyone.

6.1 Four structural failures

Ambience defeats the moment of notice. Notice presupposes an encounter. Where collection is a property of the environment, there is no encounter, and the notices that do exist attach to interfaces the person passes through rather than to the collection itself. A person walking through a transport hub monitored by biometric analytics receives, at best, a sign.

Opacity defeats comprehension. Meaningful consent requires understanding what is being agreed to. The relevant facts include onward transfer to parties unknown at the time, inference not contemplated at the time, and combination with data the person cannot see. Nobody can consent to this, and the policies that describe it are written to be legally sufficient rather than to be understood.

Unavoidability defeats voluntariness. Consent requires a genuine alternative. Where the service is employment, housing, banking, healthcare, education, or transport, and where all providers operate comparable practices, refusal means exclusion from ordinary life. Agreement obtained under those conditions performs a legitimating function without performing a protective one.

Aggregation defeats scope. Consent is given per relationship and the harm arises from combination across relationships. A person can consent to each of a hundred collections and object to the composite they collectively produce, with no mechanism to express that objection because no single party is responsible for the composite.

6.2 Consent as a legitimating device

The four failures are compounded by an effect that is rarely acknowledged in policy discussion. Once consent has been obtained, it operates to transfer responsibility from the party that designed the system to the person who clicked. Subsequent harms are characterised as the consequence of a choice, and the choice is treated as informed because a document existed.

This is the most damaging feature of the model, since it converts a protective instrument into a liability shield. Practitioners interviewed for this paper described this as the principal obstacle they encounter: not the absence of rules, but the existence of a compliance artefact that satisfies the rules while defeating their purpose. Consent in this configuration is an anodyne substitute for protection, and its presence forecloses the argument that protection is absent.

6.3 What consent can still do

The paper does not propose abolishing consent. It proposes demoting it.

Consent remains appropriate where a genuine alternative exists, where the processing is discrete and comprehensible, where the purpose is specific and will not expand, and where refusal carries no material penalty. Those conditions describe a real but narrow set of cases, and consent should be confined to them.

Outside that set, protection has to come from rules that apply regardless of agreement. That is the definition of a floor and it is the subject of Section 8. It is also not a novel move: employment law, consumer law, product

safety, and tenancy law all abandoned the position that agreement legitimates any term, and they did so for the same reason, namely that the parties are not equally placed and the resulting bargains are not informative about what either side would accept under different conditions.

7. The Regulatory Landscape

This section assesses existing instruments against a single question: do they establish a floor, meaning a protection that applies regardless of agreement, that operates without requiring the individual to discover and assert it, and that attaches consequences to breach? Positions are stated as at mid-2026.

7.1 Data protection frameworks

Comprehensive data protection law, of which the European model is the most influential and widely emulated, establishes lawful basis requirements, purpose limitation, minimisation, transparency obligations, individual rights of access, correction, erasure and objection, and provisions on automated decision-making producing legal or similarly significant effects.

The framework contains genuine floors. Lawfulness requirements do not depend on the individual asserting anything. Purpose limitation constrains re-use in principle. Supervisory authorities can act without a complaint.

Four weaknesses recur in enforcement records and in practitioner accounts.

Legitimate interest as a general solvent. Where consent is inconvenient, processing is grounded in the controller's own assessment of its legitimate interests, conducted by the controller, documented internally, and reviewed only if challenged. Practitioners describe this as the mechanism through which most behavioural surveillance is lawfully conducted.

Inference sits awkwardly within the framework. Sensitive-category protections attach to data rather than to conclusions, which is the gap described in Section 4.1. Guidance has begun to address inferred data as personal data, and the position remains contested at exactly the point where clarity matters most.

Automated decision provisions are narrow. Protections typically apply only where a decision is based solely on automated processing, and nominal human involvement removes the case from scope. The human in the loop, presented as a safeguard, functions in practice as an exemption mechanism.

Rights are individual and reactive. Access, correction, and objection must be exercised, which returns the analysis to the discovery burden. Enforcement records show that the individuals who exercise these rights are disproportionately those with the resources and confidence to do so.

7.2 Artificial intelligence instruments

The newer generation of AI-specific instruments, of which the European regulation is the most developed, adds a risk-tiered architecture: a small set of prohibited practices, a larger set of high-risk applications subject to

conformity assessment, documentation, human oversight and post-market monitoring, and transparency duties for certain systems.

This is a structural advance because it establishes prohibitions that do not depend on consent, and it attaches obligations to deployment context rather than to data category. Several prohibitions bear directly on the concerns of this paper, including limits on social scoring, on certain emotion inference in workplaces and education, on untargeted facial image scraping, and on some biometric categorisation.

Three limitations qualify it.

The prohibitions are narrowly drawn, with exceptions that carry substantial weight, and much of the surveillance described in Sections 2 and 4 falls into the high-risk tier rather than the prohibited one, where the remedy is process rather than restraint.

The high-risk regime is largely a compliance architecture. It generates documentation, assessment, and oversight duties owed to regulators, and it creates few entitlements exercisable by affected persons. Conformity assessment is substantially self-assessment.

Coverage attaches to enumerated categories, which leaves ambient commercial surveillance outside a defined high-risk use largely untouched.

7.3 Sectoral instruments

Employment. Some jurisdictions require consultation with worker representatives before monitoring is introduced, restrict specific practices, or require proportionality assessment. Where collective mechanisms exist, practitioners regard them as the most effective protections available, precisely because they do not require an individual worker to identify themselves as an objector.

Credit. The oldest formal regime, with entitlements to disclosure, correction, and adverse action notice. Its history is instructive rather than encouraging: decades of formal rights have coexisted with well-documented accuracy problems and correction processes that are burdensome to use. Formal rights are not self-executing.

Biometrics. A small number of jurisdictions require affirmative consent before biometric collection and provide statutory damages with a private right of action. This combination has produced more litigation and more behavioural change among deployers than any other instrument in our review, which is probative of the proposition that enforceability rather than stringency is the binding constraint.

Children. Age-appropriate design requirements have shifted default settings in consumer services. They rest on age assurance mechanisms of contested effectiveness and do not reach educational or institutional monitoring, which is where the most consequential childhood surveillance occurs.

7.4 What the landscape reveals

Three findings follow.

Obligations exceed entitlements. The dominant regulatory instrument is a duty owed to a supervisory authority. Affected persons hold rights that require discovery and assertion. Compliance can be complete while recourse remains close to zero, which is the recourse deficit expressed as a design feature rather than an accident.

Prohibitions are the exception. Almost all of the apparatus is procedural, and procedure is compatible with the practice continuing. Where firm prohibitions exist, whether in biometric statutes or in the newer AI instruments, they change behaviour in a way that documentation duties do not.

Nothing addresses cumulative exposure. Every instrument assesses a system. None assesses the aggregate burden on a person subject to many, and the harm described in Section 3 is aggregate by nature. This is the largest single gap in the landscape and no jurisdiction we reviewed has attempted to close it.

8. Regulatory Floors for Dignity

We propose enforceable limits on high-risk behavioural surveillance, meaningful transparency and contestation rights, strict necessity and proportionality tests, and independent oversight for systems that influence consequential decisions. Together these establish a regulatory floor: a set of protections below which commercial and institutional data practices should not be permitted to fall.

The floor metaphor is doing specific work. A floor is not a ceiling on innovation, a licensing regime, or a general restriction on data use. It is a set of protections that apply regardless of consent, that do not require the protected person to know they exist, and that attach consequences when breached. Everything above the floor remains a matter for ordinary commercial and democratic decision.

8.1 Floor One: enforceable limits on high-risk behavioural surveillance

What it does. Places a defined set of practices beyond the reach of agreement, on the basis that their harms to the conditions of autonomous selfhood cannot be cured by disclosure or by consent.

Design requirements.

Prohibition by practice, not by technology. A schedule of prohibited practices should be defined by what a system concludes and what is done with the conclusion, rather than by the technique used. Technology-specific drafting is obsolete on arrival and invites reconstruction of the prohibited outcome by other means. A model schedule appears in Annex A.

Prohibition survives consent. The defining feature of a floor is that agreement does not lift it. Instruments that permit derogation by consent will find consent, and the populations described in Section 5 will supply it under precisely the conditions that make it meaningless.

Outcome-level constraint. Following the proxy analysis in Section 4.2, prohibitions should attach to inferences

drawn and acted upon, not only to inputs used. A prohibition on using an attribute that permits its reconstruction from proxies delivers no protection and destroys the audit trail.

Narrow, reviewable exceptions. Exceptions should be specific, published, time-limited, and subject to independent authorisation, rather than drafted as broad public-interest carve-outs which, on the evidence of every framework we reviewed, become the operative rule.

8.2 Floor Two: strict necessity and proportionality

What it does. Requires any deployer of behavioural surveillance affecting consequential decisions to demonstrate that the measure is necessary for a specified legitimate purpose, that no less intrusive means would suffice, and that the intrusion is proportionate to the benefit.

Design requirements.

The burden sits with the deployer. Necessity is asserted by the party with the information and tested by a party without it. Requiring the affected person to disprove necessity is requiring them to litigate about facts they cannot access.

Assessment before deployment and on a defined cycle thereafter. Systems change, purposes expand, and the assessment conducted at procurement describes a system that no longer exists two years later.

Cumulative assessment. Necessity should be assessed in light of surveillance to which the population is already subject in that context, which is the only available means of addressing the compounding problem identified in Section 5.5. A monitoring measure that is proportionate in isolation may be disproportionate as the twelfth applied to the same workers.

Published reasoning. Assessments should be publishable in substance, with genuine commercial confidence protected but not extended to the reasoning itself. Non-public proportionality assessment is self-assessment with extra steps.

8.3 Floor Three: transparency and contestation that can be exercised

What it does. Converts the formal rights that exist in most frameworks into rights that a person without technical or legal resources can actually use.

Design requirements.

Notification without request. The person must be told that a system was used in a decision affecting them, without having to ask, because asking requires knowing. This single requirement does more to close the discovery burden than any other measure in this paper.

Substantive explanation. The explanation must state the principal factors that produced the outcome in the individual case, in language the recipient can act upon. Generic descriptions of system logic do not permit contestation and should not satisfy the duty.

Contestation with a duty to reconsider. A right to object is worthless without an obligation to reconsider on the merits, within a defined period, by a person with authority to reach a different outcome, with reasons given.

Effect during contestation. Where an adverse decision is contested, the framework must specify what happens in the interim. A right exercised after eviction, dismissal, or refusal of benefits is a right to an apology. Annex B sets out the specification in full.

Collective exercise. Individual rights fail in contexts of dependence, most clearly in employment. Representative exercise by unions, works councils, and accredited civil society bodies is the mechanism practitioners identified as usable, and it should be available as of right rather than as an exception.

8.4 Floor Four: independent oversight of consequential systems

What it does. Provides supervision of systems influencing consequential decisions that does not depend on individual complaints and does not rely on the deployer's own assessment.

Design requirements.

Ex ante registration for defined categories. Systems used in employment, housing, credit, insurance, education, public benefits, immigration, and policing should be registered with a supervisory body before deployment, with a public register recording deployer, purpose, population affected, and assessment status. Public registration is what makes independent scrutiny possible at all, and its absence is why so little is known about deployments that affect millions.

Inspection powers with technical access. Oversight requires the ability to examine systems in operation, including access to models, data, and outputs, exercised by bodies with the technical capacity to use it. Document review conducted by generalist regulators is not oversight.

Proactive audit, not complaint-driven regulation. Complaint-driven enforcement allocates attention to the populations most likely to complain, which inverts the distribution of need established in Section 3.4. A proportion of supervisory capacity should be directed by risk rather than by complaint volume.

Consequences that alter incentives. Penalties, suspension powers, and where appropriate an order to cease processing and delete derived inferences. Where breach costs less than compliance, the outcome is determined by arithmetic rather than by policy.

Independence in fact. Resourcing, technical capability, and freedom from both political and industry direction. An oversight body without technical staff is a body that will accept the assessments it is given.

8.5 The floors as a system

Floor	Applies to	Operates without individual action	Primary obligation holder	Failure it addresses
One: enforceable limits	Defined high-risk practices	Yes	Developers and deployers	Consent legitimating unacceptable practice
Two: necessity and proportionality	Consequential deployments	Yes	Deployers	Unbounded expansion of collection
Three: transparency and contestation	Individual decisions	Notification yes; contestation on request	Deployers	Discovery burden and recourse deficit
Four: independent oversight	Registered categories	Yes	Supervisory bodies	Self-assessment and enforcement scarcity

The four are interdependent. Prohibitions without oversight are unenforced. Oversight without registration cannot find its subject. Contestation without notification is unusable. Necessity testing without published reasoning is self-certification. A jurisdiction adopting one and declining the others should expect the adopted element to underperform, and the underperformance will be attributed to the measure rather than to the omission.

9. Objections and Responses

“This will suppress beneficial data use.” The floors attach to behavioural surveillance affecting consequential decisions, not to data use generally. Research, service operation, security, and the substantial majority of commercial data processing fall outside them. The objection is nonetheless worth stating precisely, because a floor drawn too broadly would justify it, and the drafting discipline in Annex A is directed at that risk.

“Necessity and proportionality tests are too vague to comply with.” They are the standard apparatus of rights adjudication across every jurisdiction in our review, applied routinely in contexts of far greater consequence. What is novel is not the test but its application to commercial data practice, and the unfamiliarity is a matter of habit rather than of legal difficulty.

“Registration will expose commercially sensitive systems.” Registration as proposed records deployer, purpose, population, and assessment status. It does not require disclosure of model weights, training data, or implementation. Comparable registration operates in financial services, pharmaceuticals, and environmental regulation without dissolving commercial confidence in those sectors.

“Notification will overwhelm people with alerts.” This is a design problem rather than an objection to the principle. Notification attaches to consequential decisions, which are not frequent events in an individual life, and consolidated notification through a single channel is achievable. The alternative is the current position, in which people receive no information because informing them would be inconvenient.

“Small firms cannot bear this.” Floors One and Three apply without a size threshold, because a small deployer causing a serious harm should not be immune and because honesty is not a function of headcount. Floors Two and Four are scaled, since they require infrastructure. This is the same asymmetry the Institute has

proposed in adjacent work and it holds for the same reason.

“Enforcement capacity does not exist.” It does not, and this is the strongest objection in the set. Our answer is that Floor Four’s registration requirement is what makes proportionate enforcement feasible, since it directs finite capacity toward known deployments rather than dispersing it across an undifferentiated field, and that Floor Three’s collective exercise provision distributes enforcement effort beyond the regulator. A framework that depends on a well-resourced regulator and provides no other route will fail in most of the jurisdictions that adopt it.

“Surveillance will simply relocate.” Some will. Obligations attaching to deployment affecting a domestic population, rather than to the location of processing, limit the effect, and this approach is now established in data protection practice. Cross-border coordination on registration and on prohibited practices is the durable answer and is addressed in Section 10.

“This treats all inference as suspect.” It does not. Inference producing consequential decisions about identity, intent, risk, and worthiness is treated as requiring justification. Inference in product operation, scientific research, and service delivery is not within scope. The distinction is consequence, and it is the distinction on which every proposal in this paper turns.

10. Implementation and Sequencing

The floors differ in legislative difficulty, and the order of adoption determines whether they arrive as a system or as a welter of partial measures that create compliance cost without delivering protection.

Available without new primary legislation. Registration schemes for public sector systems can be established administratively, and public bodies are both the easiest starting point and, on the evidence of Section 5, among the most consequential deployers. Procurement conditions can impose necessity assessment, notification, and contestation duties on vendors immediately, and public purchasing power in this market is substantial. Supervisory authorities can issue guidance treating inferred attributes as falling within sensitive-category protection, which is the single most consequential interpretive move available under existing law. Existing collective consultation rights in employment can be enforced more vigorously where they already exist.

Achievable within a legislative cycle. Notification duties for consequential decisions, which require statutory amendment but not new architecture. Registration extended to private sector deployments in the enumerated categories. Collective exercise rights for accredited bodies. Removal of the solely-automated limitation in automated decision provisions, which is a narrow amendment with disproportionate effect, since it closes the exemption that nominal human review currently provides.

Requiring sustained effort. The prohibited practices schedule in Floor One, which is the most consequential element and attracts the most concentrated opposition. Necessity and proportionality as a general requirement with the burden on the deployer. Cumulative assessment, which is intellectually the least developed of the

proposals and requires methodological work that does not yet exist.

Sequencing risks. Two orderings should be avoided. Adopting registration without inspection capacity produces a public list that nobody examines and that functions as evidence of oversight in its absence. Adopting contestation rights without notification duties produces rights exercisable only by those who already know, which reproduces the distribution failure this paper identifies.

International coordination. Convergence matters most on two elements. A common core of prohibited practices prevents the relocation objection from having force, and mutual recognition of registration reduces compliance cost while extending visibility. Neither requires comprehensive treaty-making, and both can be advanced through the existing networks of supervisory authorities that already coordinate on cross-border enforcement.

11. Policy Recommendations

Recommendation 1. Enact a schedule of prohibited behavioural surveillance practices that consent cannot lift. Define prohibitions by inference and consequence rather than by technique, with narrow, published, time-limited and independently authorised exceptions. Annex A provides a model schedule.

Recommendation 2. Require necessity and proportionality assessment for consequential deployments, with the burden on the deployer. Assessment before deployment and on a defined cycle, published in substance, and conducted in light of the surveillance to which the affected population is already subject.

Recommendation 3. Establish notification without request. Any person subject to a consequential decision informed by a behavioural inference system must be told, at the time, that such a system was used and what principally produced the outcome. This is the single measure that most directly addresses the discovery burden.

Recommendation 4. Make contestation exercisable. A duty to reconsider on the merits within a defined period, by a person with authority to reach a different outcome, with reasons given, and with specified interim effect on the contested decision. Annex B sets out the specification.

Recommendation 5. Provide collective exercise of data rights. Extend standing to unions, works councils, and accredited civil society bodies, since individual exercise fails in exactly the contexts of dependence where protection matters most.

Recommendation 6. Register consequential systems publicly before deployment. A public register of deployer, purpose, population affected, and assessment status for systems used in employment, housing, credit, insurance, education, public benefits, immigration, and policing. Begin with public sector systems, which can be registered administratively.

Recommendation 7. Treat inferred attributes as attracting sensitive-category protection. Where a system derives or acts upon a conclusion about health, biometric identity, religion, political opinion, sexual orientation, immigration status, or trade union activity, the protections applicable to that category should attach to the inference regardless of the inputs used. This closes the circumvention identified in Section 4.1 and is achievable through interpretive guidance in several jurisdictions.

Recommendation 8. Remove the solely-automated limitation. Extend automated decision protections to decisions substantially informed by automated processing, so that nominal human review no longer operates as an exemption.

Recommendation 9. Build oversight with technical capacity and proactive audit. Inspection powers with access to models, data, and outputs; technically qualified staff; and a defined share of supervisory capacity directed by risk rather than by complaint volume.

Recommendation 10. Develop and mandate cumulative exposure assessment. Commission the methodological work required to assess aggregate surveillance burden on a population, beginning with workplace and public administration contexts, and require its use once available. Cumulative exposure is the form the harm actually takes and it is currently measured nowhere.

12. Conclusion

The argument of this paper can be stated in a sentence. The harm produced by pervasive behavioural surveillance arrives before any decision is made, falls hardest on those least able to refuse it, and takes a shape that instruments built around individual notice, individual consent, and individual complaint cannot reach.

Each element of that sentence has a consequence for policy. Because the harm precedes the decision, frameworks that attach exclusively to decision points will always be measuring the wrong thing, and the searches not made and the meetings not attended will remain invisible to every assessment conducted. Because the burden follows dependence, protections that require assertion will reach the people who need them least. And because the harm is cumulative, a regulatory apparatus that evaluates one system at a time will find each of them proportionate while their aggregate is not.

This is why the paper proposes floors. A floor is a protection that does not require the protected person to know it exists, that cannot be traded away by an agreement obtained under conditions that make agreement meaningless, and that attaches a consequence when it is breached. It is a familiar instrument. Employment law, product safety, tenancy protection, and consumer regulation all reached the same conclusion in their own fields, namely that some terms should not be available for negotiation because the parties are not equally placed, and none of those fields collapsed as a result.

The dignity at stake is not an abstraction. It is the practical capacity to hold a belief provisionally, to attend a meeting without the attendance becoming a datum, to seek help from an institution without the seeking

becoming a record, and to be assessed on what one has done rather than on what a model predicts one is likely to do. Those capacities are the conditions under which people become themselves, and they are being reduced by an accumulation of decisions that nobody made deliberately and that no one is currently accountable for.

The reduction is not inevitable, and it will not be reversed by better disclosure or by more granular consent, both of which have been tried at scale and have produced compliance without protection. It requires limits that hold regardless of what anyone was persuaded to agree to. The purpose of a floor is to obviate the negotiation entirely, so that dignity is not something a person must repeatedly bargain for against parties who hold every advantage in the bargaining.

Annex A: Model Schedule of Prohibited Practices

Offered as drafting reference. Prohibitions are defined by inference and consequence rather than by technique, apply notwithstanding consent, and are subject only to exceptions that are specific, published, time-limited, and independently authorised.

A.1 General prohibition on consent derogation. No agreement, term, or condition may permit a practice prohibited under this schedule, and any purported agreement to that effect is of no effect.

A.2 Prohibited practices.

- a. Deriving or acting upon inferences of emotional state, cognitive state, deception, or credibility in employment, education, immigration, insurance, or policing contexts.
- b. Deriving or acting upon inferences of trade union membership, activity, or organising propensity in any employment context.
- c. Deriving or acting upon inferences of immigration status, religious belief or observance, sexual orientation, or political opinion, whether from those attributes directly or from proxies, in decisions concerning employment, housing, credit, insurance, education, healthcare, or public benefits.
- d. Continuous biometric identification of individuals in publicly accessible spaces for commercial purposes.
- e. Aggregating behavioural data across unrelated contexts to produce a general assessment of an individual's reliability, trustworthiness, or social worth, and acting upon that assessment in any consequential decision.
- f. Behavioural surveillance of children for commercial profiling or for behavioural prediction beyond the specific educational or safeguarding purpose for which the system was deployed.
- g. Monitoring of workers in areas or during periods where an expectation of privacy is reasonable, including sanitary facilities, rest breaks, and non-working time, and monitoring of domestic environments where work is performed at home beyond the specific work activity.
- h. Acting upon behavioural inferences to determine eligibility for essential services, including healthcare, utilities, and public benefits, without an individualised determination by a person with authority to depart from the inference.

A.3 Exceptions. An exception may be authorised only where it is specific as to purpose, population, and duration; published in a public register before reliance; time-limited and subject to renewal on evidence; and authorised by an independent supervisory body. General public-interest exceptions are not available under this

schedule.

A.4 Derived inferences. Where a practice is prohibited under A.2, the prohibition extends to inferences reconstructed from proxy variables, and the burden of demonstrating that a system does not so reconstruct rests with the deployer.

Annex B: Specification for an Exercisable Right of Contestation

Offered as a design specification. A right of contestation is exercisable only if all of the following are present. The absence of any one renders the remainder inchoate.

B.1 Notification. The affected person is informed, at the time of the decision and without having to ask, that a behavioural inference system was used, which deployer used it, and where the register entry may be found.

B.2 Explanation. The person receives the principal factors that produced the outcome in their individual case, expressed specifically enough to be acted upon. A description of the system's general logic does not satisfy this requirement.

B.3 Access to inputs. The person may obtain the data used in respect of them, including data obtained from third parties, in intelligible form and without charge.

B.4 Route. A single identified route to contest, available without legal representation, with a published procedure and a named responsible function.

B.5 Duty to reconsider. The deployer must reconsider on the merits, within a defined period, through a person with authority to reach a different outcome, who has access to the underlying material and is not the person who made the original decision.

B.6 Reasons. The outcome of reconsideration is communicated with reasons that engage with the points raised, not with a restatement of the original determination.

B.7 Interim effect. The framework specifies what happens to the contested decision pending reconsideration. For decisions concerning employment, housing, and essential services, the presumption should be suspension of adverse effect where suspension is practicable.

B.8 Onward remedy. Where reconsideration is unsuccessful, an identified route to independent review, at a cost proportionate to the matter, with the possibility of collective exercise where the same system has affected a class of persons.

B.9 Record. Contestations, outcomes, and error rates are recorded and reported to the supervisory body, and reported in aggregate publicly. A system that is contested frequently and upheld invariably is a system that requires inspection, and this is the only mechanism by which a regulator would know.

Sources and Further Reading

This paper was prepared from primary sources, meaning statutory text, implementing regulations, supervisory authority guidance and enforcement records, litigation documents, procurement and freedom of information disclosures, and peer-reviewed literature, together with practitioner interviews conducted by the Institute. Secondary commentary is used for interpretive context only. Positions are stated as at mid-2026, and readers should verify current status before operational reliance. Full citations for the instruments and studies referenced are available on request from the Institute.

Data protection and privacy law. Comprehensive data protection instruments and their implementing guidance, particularly on lawful basis, purpose limitation, inferred data, and automated decision-making; supervisory authority decisions and enforcement records concerning behavioural advertising, workplace monitoring, biometric processing, and data brokerage; sectoral privacy statutes governing biometrics, consumer data, and children's data.

Artificial intelligence regulation. Risk-tiered AI instruments and their prohibited practice provisions, high-risk classification criteria, conformity assessment requirements, and transparency duties; regulatory guidance on emotion inference, biometric categorisation, and social scoring.

Human rights standards. Treaty provisions and interpretive materials on private life, freedom of assembly and association, and non-discrimination; jurisprudence on necessity and proportionality in surveillance contexts; guidance from international human rights bodies on privacy in the digital age and on the human rights implications of automated decision-making.

Political economy and critical scholarship. Zuboff's account of surveillance capitalism and the critical literature responding to it; scholarship on the data broker industry, on the political economy of behavioural advertising, and on the relationship between platform business models and data extraction; research on surveillance in employment and on algorithmic management.

Empirical and technical literature. Research on chilling effects and behavioural adaptation under observation; work on inference of sensitive attributes from non-sensitive data; the fairness and accountability literature on proxy discrimination and feedback loops; evaluations of emotion recognition validity; and research on the accuracy and contestability of scoring systems in credit, tenant screening, and public administration.

Documented deployments. Regulatory enforcement records and litigation concerning workplace monitoring, tenant screening, benefit fraud scoring, child welfare risk prediction, immigration data sharing, educational technology, and biometric deployment in retail and public space; procurement disclosures and freedom of information releases; investigative reporting corroborated against primary documentation.

Practitioner interviews. Semi-structured interviews conducted by the Institute with practitioners in civil society organisations, trade unions and worker advocacy bodies, and digital rights litigation, used to identify operational failure points in existing rights and to test the practicality of the proposed floors.

The Digital Dignity Institute of California is an independent research institute based in San Francisco and engaged globally. It is independent of government and commercial interests. Its work spans global convenings, original research, policy advice, and structural consulting, and is directed at developing frameworks that centre human beings and make artificial intelligence answerable to them.

Correspondence: connect@digitaldignityinstitute.com

© 2026 The Digital Dignity Institute of California. This paper may be reproduced with attribution.