The Surveillance State After Scarcity

Information Integration, Institutional Limits, and the Mathematics of Ubiquitous Observation

Flyxion
Independent Researcher

October 2026

Abstract

Debates about the surveillance state usually treat the problem as a trade between security and privacy, and so they assume that the capacities involved (sensing, inference, retention, and the authority to act) travel together. Historically they did, because each was expensive and the expense held them in the same hands. This essay argues that falling costs of observation, storage, and analysis are separating them, and that the resulting problem is better described as the reallocation of scarcity than as the growth of a single quantity called surveillance. Three questions organize the argument. Why do modern institutions require organizations that integrate heterogeneous information at scale, of which Palantir Technologies is the most prominent commercial example? Where do the claims and capabilities of such organizations exceed what integration can justify, and where do they fall short of what accountability requires? And what changes when optical computation, distributed sensing, and chemical spectroscopy, carried by photonic metasurfaces and similar technologies, make observation a property of ordinary surfaces rather than a specialized activity? The central thesis is conditional: as the marginal costs of sensing, computation, and retention decline, the binding constraints on surveillance shift from the availability of observations to the legitimacy of inference, the allocation of authority, the integrity of evidence, and the reversibility of institutional decisions. The final part develops a formal framework that separates institutional action from the world state by an observation channel, an inference procedure, and an authorization policy, and it culminates in a conditional result on when exact correction of erroneous records is compatible with irreversible erasure. Throughout, the essay distinguishes documented facts, established mathematical results, labeled extrapolation over a horizon of roughly 2076 to 2126, and propositions proved here, and it declines to treat any one of these as evidence for another.

1 Introduction: Scarcity as an Unwritten Regulator

In much of the history of organized states, legal and political restraints on surveillance operated alongside a practical one that is easy to overlook, and that practical restraint was arithmetical. Its weight varied across jurisdictions and periods, and what follows treats it as an interpretive

hypothesis about one contributing constraint and not as an established general law. Observing a population required people, and people had to be paid, trained, supervised, and housed. Records required paper, and paper required storage, indexing, and clerks who could find a page again. Analysis required analysts, whose attention was the scarcest input of all. Because each of these inputs was costly, the aggregate cost of watching a large population closely was prohibitive, and the prohibition supplied part of the restraint that statutes and constitutions were later credited with. Bentham's Panopticon was a design for economizing on watchers, and its central idea was that the possibility of observation could substitute for observation itself [1]. Foucault's reading of that design as a general principle of disciplinary power [2], and Scott's account of the way states simplify societies in order to make them legible to administration [3], both describe regimes in which legibility was purchased at real expense and therefore was rationed. Scarcity rationed it. No one had to decide to ration it.

The argument of this essay is that this unwritten regulator is weakening, and that political thought about surveillance has not yet caught up with the consequences. When the marginal cost of an additional observation, an additional year of retention, and an additional unit of analysis all approach negligible levels, the question stops being how much a state or firm can afford to know and becomes how much it is entitled to infer, to retain, and to act upon. Those are questions of authority, evidence, and reversibility, and they are poorly served by a vocabulary built around access to records. A regime that restricts access to a database while leaving the surrounding environment densely instrumented may protect less than it appears to, because the contents of the protected record can often be reconstructed from correlated observations that were never protected at all.

It is useful to separate four capacities that are usually run together. Sensing is the capacity to produce measurements of the world. Inference is the capacity to turn measurements into claims about particular entities, states, or intentions. Retention is the capacity to keep measurements and claims available for later use. Authority is the capacity to act on claims in ways that bind other people, whether by detaining, denying, prosecuting, reassigning, or merely flagging. Historically a single organization, such as a police force or a ministry of finance, exercised all four, and their costs rose together. Technological development is now reducing the cost of the first three much faster than it is reducing the difficulty of establishing the fourth. The mismatch is the subject of this work. A world in which sensing, inference, and retention are nearly free while authority remains something that must be earned, justified, and checked is different from a world in which all four are costly, and different again from one in which all four are free.

The essay is organized in five parts. Part I asks why integration of heterogeneous information has become an institutional necessity, using Palantir as a documented case rather than as a moral emblem, and argues that the necessity attaches to a function and not to a provider. Part II examines the difference between technical integration and epistemic justification, and sorts the evidence available about platform capabilities into four categories that should not be mixed. Part III turns to what existing systems do not do, and in particular to the asymmetry between the ability to find a person and the ability to establish that a finding was mistaken

and must be undone. Part IV extends the analysis to a horizon of roughly 2076 to 2126, in which photonic metasurfaces and related technologies may make ordinary surfaces participate in imaging and chemical sensing, and it is careful to separate what has been demonstrated from what is extrapolated. Part V gives the formal core: a model of observation, inference, and action; propositions on identification, retention, authority, correction, and privacy; a conditional result on when exact correction is compatible with irreversible erasure; and an admissibility condition whose components are tested against one another, including on a worked example. A concluding discussion draws out the distinction between universal observability and universal accountability, and argues that the question of who observes whom should be treated as two independent dimensions of one mathematical problem.

A note on method is necessary because the subject invites overreach in both directions. Claims in this essay fall into four classes: documented facts about existing systems, drawn from vendor documentation, vulnerability records, standards bodies, and review articles; established results from the mathematical literature on identification, information theory, and privacy; extrapolations about future technology, which are labeled as such and tied to the conditions that would have to hold; and propositions proved here, whose scope is stated. Where a claim could belong to more than one class, the weaker classification is used. Nothing in the formal part is offered as evidence that any particular future will arrive, and nothing in the technological part is offered as evidence that any particular institutional arrangement is desirable.

2 Part I: Why Information Integration Becomes Institutionally Necessary

2.1 Observability is not surveillance

The opening distinction is between surveillance, understood as the observation of persons for purposes of control, and institutional observability, understood as the capacity of an organization to know the state of its own operations well enough to discharge its obligations. Hospitals must know which beds are occupied, which patients are scheduled for which procedures, which supplies are on hand, and which of these facts are current. Utilities must know the condition of assets spread across territories and jurisdictions. Financial regulators must be able to follow a transaction through institutions that use incompatible classifications. Emergency services must reconcile reports from agencies that do not share identifiers. In each case the operational difficulty is rarely a shortage of data. It is the difficulty of establishing which observations refer to the same entity, which records remain authoritative when sources disagree, and which decisions are justified by what is actually known as opposed to what is merely stored.

Three structural features make this difficulty persistent. The first is heterogeneity: records arrive in different schemas, with different units, different update cycles, and different notions of what counts as an entity. The second is jurisdictional fragmentation: the people who hold the records answer to different authorities, with different legal bases for collection and different obligations to share. The third is the requirement that action be traceable: modern institutions

are expected to be able to say, after the fact, why a decision was made, on what evidence, and under whose authority. An organization that can unify records but cannot trace decisions has solved the easier half of the problem.

2.2 Palantir as a documented case

Palantir Technologies is useful here because its public documentation describes the integration function with unusual explicitness. The company's architecture documentation presents three platforms: Foundry as the core data-operations platform, AIP as the generative-AI platform, and Apollo as the continuous-delivery system beneath both, with the Ontology described as the central organizing layer [25]. The same documentation states that AIP and Foundry together comprise several hundred microservices and assets running on an autoscaling compute mesh, and describes Apollo as the mechanism that allows deployment across environments of very different connectivity and classification. The Ontology is presented as a model of the nouns and verbs of an organization's processes, expressed in a form that both human users and automated agents can read and act on [26]. In the vocabulary of this essay, the Ontology is a proposed solution to the entity-resolution, authority, and action-tracing problems described above, and Apollo is a solution to the problem of running such a system in places where ordinary cloud deployment is unavailable.

Two observations about this documentation matter for what follows. First, it is vendor documentation, and it describes intended design rather than measured performance. Second, the design it describes is not obviously specific to surveillance of persons: the same architecture is offered for hospital operations, supply chains, and manufacturing, as well as for defense and intelligence. The conclusions of this essay do not depend on any judgment about the company's conduct, customers, or politics, which are contested and lie outside the scope of an analysis of capabilities and their limits.

2.3 A necessary function and a contingent provider

The thesis of Part I can be stated precisely. Call a function institutionally necessary if, in its absence, some obligation that the institution is legitimately charged with cannot be discharged. Integration of heterogeneous records into a common, queryable, permissioned, and traceable representation appears to satisfy this definition for several classes of institution. A national health system that cannot reconcile patient records across providers cannot meet its duty of continuity of care. A regulator that cannot link transactions across institutions cannot detect the patterns it is charged with detecting. The argument is not that these institutions must be monitored more intensively, but that they cannot meet their obligations at modern scale without some means of integration, and that the means must handle identity resolution, provenance, and permissions as first-class concerns.

It does not follow that any particular corporation is indispensable. The function could in principle be provided by a public agency, by a consortium, by open-source software operated

under public control, or by a competing vendor, and the practical question is which arrangement best controls the characteristic risks of the function: concentration of capability, dependence and exit cost, opacity of the representation to the people it describes, and capture of public decision-making by the maintainers of the platform that frames the decisions. The normative hinge of this part is a three-part test of any provider. Interfaces must be open enough that records and the logic built on them can leave. Authority over how entities are defined and how conflicts between records are resolved must rest with the accountable institution and not with the vendor. And the platform's own behavior, including its logs and permissions, must be inspectable by parties who do not depend on the vendor for their account of it. These criteria are the operational counterpart of the claim that the function, and not the provider, is necessary: they are what it means to separate the two in practice. They are stated as an analytic device and not as a verdict about any existing deployment, and the later parts of the essay can be read as specifying what the second and third criteria require technically.

It is worth saying what would count against the thesis. If it could be shown that institutions of the relevant kinds meet their obligations adequately without any cross-source integration layer, or that integration can be accomplished by fully decentralized means with no entity that holds the integrated representation, the claim of functional necessity would weaken. Third, and more specifically, the thesis would weaken if the integration function could be shown to be separable from the burdens of entity resolution and authority tracing, so that what is institutionally necessary could be discharged without the capability profile that gives rise to the risks examined in Parts II and III. None of these is established, but none is excluded, and the second in particular connects to the cryptographic and distributed-systems literature on privacy-preserving computation, which Part V takes up.

3 Part II: Where Surveillance Platforms Overstate What Integration Can Justify

3.1 Joining records is not establishing identity

A system can join two databases without establishing that the records refer to the same person. This is the oldest problem in record linkage, formalized by Fellegi and Sunter as a decision between match, non-match, and possible match under measured error rates [9], and it does not disappear when linkage is performed by learned models instead of deterministic rules. What matters institutionally is the base-rate structure of the problem. If a population contains N entities and a linkage procedure has a per-pair false-match probability of α among non-matching pairs, the expected number of false links is approximately α times the number of non-matching pairs, which is of order $N^2/2$ when every pair is a candidate. Practical systems use blocking or indexing to restrict comparison to plausible candidate pairs, which brings the candidate count closer to linear in N, but that restriction is a design choice and not a guarantee, and it trades missed matches for fewer false ones. Absent blocking, the candidate pair count is quadratic, and a false-match rate that is negligible for a single comparison produces a large

absolute number of erroneous identifications at scale, growing faster than the population does. Institutions that cite the low per-comparison error of an identification method as evidence that its outputs are reliable are making a statement about the wrong denominator. The relevant quantity is the proportion of flagged individuals who are the wrong individuals, and it depends on how rare true matches are among the candidates considered.

A related problem is that an internally consistent ontology may misrepresent the world. Consistency is a property of a model, and correspondence is a property of the relation between a model and what it models. A representation in which every object has a unique identifier, every link is typed, and every action is logged can be perfectly coherent and systematically wrong, because coherence can be achieved by suppressing or reclassifying the records that do not fit. The Ontology of an organization encodes decisions about what counts as an entity and what counts as the same entity across time, and those decisions are modeling choices that have to be defended and not merely implemented.

3.2 Enforcing permissions is not establishing legitimacy

Permission systems answer the question of who may see or do what, given a configuration. They do not answer whether the configuration was legitimately established. Palantir's documentation describes a layered set of controls in which project-level roles provide discretionary access while markings and organization-level controls provide mandatory restrictions that propagate to derived resources [27]. Such mechanisms can enforce a policy faithfully. They are silent on whether the policy reflects lawful authority, and an audit log that faithfully records every access under a policy that should never have been granted is an accurate record of an illegitimate arrangement. The completeness of the log establishes that the system did what it was configured to do, and nothing more. A complete audit trail is therefore necessary for accountability and insufficient for it.

The same layering exposes a subtler limit. Propagation rules protect derived data only along the paths that the rules cover. Palantir's logging documentation states that log data may contain sensitive information, that service and trace logs can contain sensitive content from any data source a workflow reaches, and that administrators should therefore choose markings reflecting the highest sensitivity the workflow might touch [28]. The same page states that Foundry logs are not audit logs and that reliable delivery of every log entry is not guaranteed [28]. This is a documented acknowledgment that the observability layer, the machinery that records what the system is doing, is itself a channel through which sensitive content can travel, and that its protection depends on explicit configuration and not on automatic inheritance from the sources. The documentation reviewed for this essay does not establish how often such configuration fails in practice, and the point is not that this is a flaw peculiar to one vendor. The point is that provenance and audit infrastructure creates a second body of sensitive material that has to be governed, and that governing it is a distinct task from governing the records it describes. The second statement matters as much as the first, because it marks a distinction that discussions of accountability often blur. Operational observability supports monitoring

and debugging and may be lossy. Forensic logging supports reconstruction of what occurred after an incident. Authoritative audit records support claims about who did what under what authority, and they must be complete to bear that weight. A platform can honestly provide the first while disclaiming the third, and an institution that treats operational logs as an audit trail has made an error that no feature of the platform prevents. A concrete instance of the general hazard is the publicly recorded vulnerability CVE-2022-27895, in which a component in Foundry's logging was found to capture sensitive information in logs, rated medium severity and addressed in a later version of the affected library [29]. A fixed vulnerability of this kind is evidence that the class of failure is real. It is not evidence about the present state of the system.

3.3 Four kinds of evidence

Discussion of what integration platforms can do is routinely conducted with undifferentiated evidence, and the discipline of separating four categories is a precondition for useful analysis. Vendor claims describe intended or marketed capability, and they are produced by parties with an interest in the conclusion. Independent performance evaluations measure capability under stated conditions by parties without that interest. Documented failures, including vulnerability records, incident reports, and court findings, show that specific things went wrong, though they do not by themselves show how often. Theoretical impossibility results show that certain guarantees cannot be achieved under stated assumptions, whatever the implementation. The four categories support different conclusions. A vendor claim supports the conclusion that the vendor makes the claim. An independent evaluation supports a conclusion about performance on the evaluated task and population. A documented failure supports an existence claim. An impossibility result constrains what any future system, from any vendor, could promise. In the sources consulted for this essay, public evidence about integration platforms falls overwhelmingly in the first and third categories, with comparatively little in the second, and the fourth category is largely supplied by the general literature on identification and privacy and not by studies of any specific product. That distribution is itself a finding about the state of public knowledge, and it counsels caution in both directions: against crediting capability claims that have not been independently tested, and against inferring wrongdoing from the absence of such tests.

4 Part III: Where Existing Systems Do Not Go Far Enough

4.1 The asymmetry between finding and unfinding

The more consequential shortcoming of contemporary integration systems may be an asymmetry of design effort. Considerable engineering has gone into finding entities, linking records, and surfacing patterns. Comparatively little public design attention has gone to the converse operations: establishing that an identification was mistaken, that an original source was unre-

liable, that a record was entered under authority that has since lapsed, or that a downstream decision must now be reversed. A system might be extraordinarily capable of locating a person while being structurally weak at establishing that the location was wrong and retracting its consequences. Accountability requires that these capacities be matched, because an institution that can act quickly on a claim and revise it slowly will tend to accumulate errors in the regime of the people least able to contest them.

This is the point at which the distinction between an evidence record and an authoritative conclusion becomes practically important. A record should not become authoritative because it has been incorporated into a larger system. Incorporation changes its accessibility, its apparent weight, and the ease with which it can be cited, and none of these is a change in its warrant. Yet institutional practice often treats presence in an integrated system as a kind of certification, so that the act of integration quietly converts evidence into conclusion. The mechanism is not that integration decides a record is authoritative. It is that presence in an integrated system lowers the cost of treating the record as authoritative and makes that treatment less visible to auditors, who see a query result and not the weak source behind it, so that authority drifts toward whatever has been integrated. Resisting that conversion is a design requirement and not merely an attitude.

4.2 Requirements that follow

Several technical requirements follow from taking the asymmetry seriously. Bidirectional provenance means that every derived claim can be traced to the observations that support it and, in the other direction, every observation can be traced to the claims and decisions that depend on it. The second direction is the one that is usually missing, and it is the one needed for correction. Persistent correction histories mean that when a record is corrected, the correction is itself an immutable, attributable entry that remains attached to the record for everyone who later relies on it, so that a corrected record cannot circulate in its uncorrected form without a visible flag. Cross-jurisdictional authority verification means that the authority under which a record was collected or an action taken can be checked against its source, so that authority does not silently expand as records cross organizational boundaries. Independently inspectable access decisions mean that the reasons access was granted or denied can be examined by a party other than the one that granted it. Enforceable obligations to propagate retractions mean that a retraction has downstream effects that are tracked and confirmed and are not merely requested. Each of these is a substantial undertaking, and some conflict with one another and with privacy, as Part V shows.

Existing reference frameworks provide partial scaffolding. The NIST Privacy Framework, first issued in 2020, organizes privacy risk management around outcomes and enterprise risk practices [22], and the NIST AI Risk Management Framework offers a structure for identifying and managing risks of AI systems across their lifecycle [23]. NIST's own site lists Privacy Framework 1.1 as an initial public draft as of the date of writing (checked October 2026), and law-firm commentary on the April 2025 release reports that it adds material on the interaction

of AI and personal data, including risks of reconstruction and of membership inference, and aligns the framework with Cybersecurity Framework 2.0 [24]. The description of the draft's contents here follows that commentary, and the essay does not assume the content of any final text. These frameworks are valuable reference points, but they are voluntary, sector-agnostic, and oriented toward organizational risk management. They do not by themselves settle the institutional questions raised here, such as who holds the authority to order a retraction across organizational lines, or what remedies attach when propagation fails.

5 Part IV: From Surveillance Infrastructure to Sensing Environments

5.1 Why the horizon is long and why it is hedged

This part introduces a horizon of roughly fifty to one hundred years, approximately 2076 to 2126, for a transition in which sensing ceases to be a specialized activity and becomes a property of ordinary environments. The horizon is chosen because the relevant technologies are at very different stages of maturity, and because the institutional adaptation they would require is slow. It is an extrapolation, and it is hedged accordingly. The endpoints are not predictions. They are chosen to be far enough out that institutional change, which unfolds over generations, is a live question, and near enough that the specific physical dependencies named below still constrain the space of scenarios, whereas over several centuries nearly any dependency could be assumed away. A horizon of two or three decades would understate the institutional question, and one of two centuries would leave the physics unconstrained. The formal results of Part V are conditional on a cost structure for observation and remain valid whether the relevant technology arrives within decades or never. The purpose is to identify which physical dependencies would have to be resolved for particular futures to occur, so that governance can be designed against the conditions under which each future would arise and not against a single forecast.

5.2 What the current literature supports

The relevant convergence is not simply smaller cameras. It is the combination of optical computation, distributed sensing, chemical spectroscopy, cheap storage, and increasingly automated interpretation. The research literature supports parts of this trajectory. A 2024 review in *Advanced Photonics* surveys metasurface-based computational imaging, organizing it by how computation enters the pipeline, from conventional metasurface design, through computation added in the imaging process or in postprocessing, to end-to-end systems in which the metasurface and the reconstruction algorithm are optimized jointly [17]. A separate 2024 review of computational metaoptics argues that combining the wavefront-shaping capability of metasurfaces with reconstruction algorithms can improve imaging beyond what conventional optics achieves, including in phase imaging and quantum state measurement [18]. A 2025 review

of optical metasurfaces for biosensing and bioimaging covers refractometric sensing, surface-enhanced spectroscopic sensing with enhanced molecular fingerprinting, and chemical-contrast imaging, and it names material limitations, difficulties of large-scale fabrication, and the complexity of data analysis and readout as outstanding challenges [19]. A further 2025 review addresses semiconductor metasurfaces for surface-enhanced Raman spectroscopy, describing a transition from nanoparticle-based substrates to nanoresonator-based metasurfaces and the possibility of combining chemical and electromagnetic enhancement [20], and a review of metasurface biosensors emphasizes fluorescence-based detection of biomolecules including nucleic acids and proteins [21].

5.3 What these results do not establish

Taken together these reviews establish that metasurfaces are a serious platform for compact, multifunctional optics and for enhanced spectroscopic sensing in laboratory and emerging practical settings. They do not establish that any surface could become a camera and a chemistry laboratory, and the gap between the two statements is where the analysis has to live. A metalens is not independently a camera: image formation requires a detector, illumination or ambient light of adequate quality, calibration, readout electronics, a power supply, and computation. A spectroscopic surface is not independently a chemistry laboratory: surface-enhanced techniques derive their enhancement from the region immediately adjacent to the structured surface, which means that the analyte generally has to be near or on the surface, and that a surface senses what reaches it and not what lies at a distance. Standoff chemical measurement faces different and generally more severe signal limits. Fabrication at architectural scale, durability in uncontrolled environments, fouling and drift, and the need for reference data to interpret spectra are all further dependencies. These are not incidental caveats. They determine which futures are physically plausible and which are not, and they locate the points at which governance could intervene, since a sensing capability that depends on a detector, a power source, and a readout channel can be regulated at those points in ways that a purely passive property of a surface could not.

5.4 Retrospective inference

The most consequential property of cheap retention, and the one that connects this part most directly to the formal results of Part V, is that the value of a stored measurement is not fixed at the moment it is collected. A measurement stored today might become informative decades later, when new reconstruction algorithms, reference databases, or analytical methods become available. Spectra that are uninterpretable now may become identifying once reference libraries mature, and images of limited resolution may become usable under improved computational reconstruction, a possibility consistent with the emphasis of the computational-imaging literature on reconstruction from optically encoded measurements [17, 18]. The surveillance capability of a historical record may therefore exceed the capability that existed when the record was collected.

This is the reason retention has to be justified and cannot be assumed. The party that decides to retain a record decides at a time when the future class of inference procedures, and the future holders of the record, are unknown. As the cost of storage falls, the cost to the collector of keeping a record falls with it, but the option value of the record to whoever later holds it, including successors, acquirers, and adversaries who do not share the collector's purposes, does not fall and may rise. The harm that retention can enable is therefore not bounded by what the collector can do with the record today. Proposition 2 gives this a precise form for an ideal observer: for a fixed target, the best achievable risk and the information available are monotone in the history and in the class of estimators. It also states the limit of the claim, since a misspecified model does not become more accurate merely by being applied to more data. The practical consequence is that retention limits and purpose restrictions are structural protections and not administrative conveniences, and that the operations that actually foreclose future inference are those that irreversibly destroy the record, the keys, or the auxiliary data needed to interpret it, or that coarsen it before storage.

5.5 Three scenarios

It is useful to distinguish three scenarios that have different feasibility requirements and different governance implications. In the first, sensing devices become increasingly inconspicuous: cameras and spectrometers shrink and become harder to notice, but remain discrete devices with identifiable operators, power, and data paths. In the second, sensor networks are architecturally integrated: sensing elements are built into fixtures, glazing, vehicles, and coverings as a design feature, so that the building or the street is the instrument, but each element still has a sensing function, a readout path, and an owner. In the third, surfaces are genuinely multifunctional: a single structured surface performs optical, spectral, and environmental measurements, with the physical elements of sensing and computation distributed through the material itself. The scenarios form a ladder of increasing physical difficulty and increasing governance difficulty, and Table 1 summarizes the dependencies that distinguish them. The third scenario is the one that most strongly challenges the conventional picture of surveillance as something done by identifiable actors with identifiable equipment, and it is also the one for which the evidence is thinnest.

5.6 Where scarcity goes

If observation becomes cheap, scarcity does not disappear. It moves. What remains scarce is the ability to interpret observations correctly, the attention of qualified reviewers, the legitimacy of the authority to act, and the capacity to reverse mistakes. A society that has abundant measurement and scarce interpretation is vulnerable to a characteristic failure, in which the volume of observations is mistaken for the quality of knowledge, so that a flood of weak signals is treated as strong evidence by systems and institutions that have no means of discounting it. The title of this essay refers to this relocation. The surveillance state after scarcity is not a state that knows everything. It is a state whose constraints have shifted from

	Inconspicuous devices	Integrated sensor networks	Multifunctional surfaces
Component maturity	Mature or commercial components	Components demonstrated in laboratory and early-practical settings	Individual components demonstrated; their integration is unresolved
Scenario plausibility	High; extension of existing practice	Plausible extrapolation, contingent on fabrication and durability	Speculative; depends on unresolved power, readout, and interpretation problems
Detector and readout	Discrete, miniaturized	Embedded in fixtures, cabled or networked	Must be distributed or integrated into the material
Power	Per-device or building supply	Building or infrastructure supply	Unresolved for passive, large-area operation
Analyte or light access	Device-local	Fixture-local	Surface-local; requires contact or near field for many spectroscopic modes
Interpretation	Remote analysis	Remote analysis with calibration data	Requires reference databases and onboard or networked computation
Governance handle	Device registration, operator identity	Building and procurement standards, datapath regulation	Materials and fabrication standards; datapath regulation only

Table 1: Dependencies and governance handles for three sensing scenarios. The table is an analytic aid and not a forecast.

the supply of data to the supply of justification.

6 Part V: The Mathematics of Ubiquitous Observation

6.1 Setup

Let X_t denote the latent state of the world at time t, taking values in a measurable space $\mathcal{X}$, and let $X = (X_t)_{t\geq 0}$ denote its trajectory. Let I be an index set of sensors, which may grow over time, each with a sensing function $h_i : \mathcal{X} \to \mathcal{Y}_i$ that may represent an optical image, a chemical spectrum, a thermal reading, an acoustic measurement, or a derived signal. The observation made by sensor i at time t is

$$Y_{i,t} = h_i(X_t) + \varepsilon_{i,t}, \qquad H_T = \{Y_{i,t} : i \in I,\ t \leq T\},$$

where $\varepsilon_{i,t}$ is measurement noise with law independent of the state, and H_T is the observation history up to time T. An institution does not act on X. It acts on a representation produced

from H_T, under an authorization policy:

$$A_T = \pi\big(f(H_T), \mathcal{P}_T\big),$$

where f is an inference procedure mapping histories to representations, π is a decision policy, and $\mathcal{P}_T$ is the set of permissions, rules, and constraints in force at time T. The separation of h_i, f, π, and $\mathcal{P}_T$ mirrors the four capacities of the introduction: sensing, inference, and authority are distinct components, and retention is the persistence of H_T across T. Each proposition below is about the relations among these components.

6.2 Identification

Definition 1. Let P_x^T denote the joint law of the observation history H_T when the trajectory is x, taken over the entire observation process, including any dependence among noise terms and any adaptive choice of sensors. Two trajectories x, x' are *observationally equivalent* under the sensor suite $(h_i)_{i\in I}$ if $P_x^T = P_{x'}^T$ for every T. A functional g of the trajectory is *identifiable* from the suite if $g(x) = g(x')$ whenever x and x' are observationally equivalent. Under the additive, state-independent noise model above, equality of the sensing functions, $h_i(x_t) = h_i(x'_t)$ for all i and t, is sufficient for observational equivalence. It is not necessary in general, and for models with correlated noise, adaptive sensor selection, or state-dependent noise, equivalence has to be checked at the level of the joint law and not of the mean measurements.

Proposition 1 (Observation does not entail identification)**.** *Suppose g is not identifiable from the sensor suite, so that there exist observationally equivalent trajectories x, x' with $g(x) \neq g(x')$. Then for any prior placing masses p and $1-p$ on x and x', and any decision rule based on H_T for determining which holds, the probability of error is at least $\min(p, 1-p)$, for every T and every inference procedure f.*

Proof. By the definition of observational equivalence, the laws of H_T under x and x' coincide. (In the additive model this follows from equality of the sensing functions, since the noise law does not depend on the state.) The likelihood ratio is therefore identically one, the posterior over $\{x, x'\}$ given H_T equals the prior, and the Bayes error of any decision rule is at least $\min(p, 1-p)$. □

The consequence is that additional storage cannot repair a structural identifiability failure. Retaining more observations from the same sensors refines the estimate of the observable quantities $h_i(X_t)$ and leaves the partition of states into equivalence classes unchanged. Only additional kinds of sensors, which refine the partition, can change which states are structurally distinguishable. Prior information can favor one member of an equivalence class over another by assigning the members different weights or by excluding one of them, but that is statistical preference and not identification: the observations still do not discriminate, and the conclusion is only as strong as the prior, which the observations cannot test. This is the formal content of the observation that a flood of measurements can leave ambiguity untouched, and it is the

reason that institutional confidence derived from the volume of data should be treated with suspicion whenever the sensor suite is known to be coarse relative to the question.

6.3 Retention

Proposition 2 (Retention expands possible future inference)**.** *Let $\mathcal{F}_T = \sigma(H_T)$, so that $\mathcal{F}_T \subseteq \mathcal{F}_{T'}$ for $T \le T'$. Let $g(X)$ be an integrable target and ℓ a loss function. For each T, let $\mathcal{C}_T$ be the class of $\mathcal{F}_T$-measurable estimators (functions of H_T), and let $R_T^* = \inf_{\hat{g} \in \mathcal{C}_T} \mathbb{E}\, \ell\big(g(X), \hat{g}\big)$. Then $\mathcal{C}_T \subseteq \mathcal{C}_{T'}$ and $R_{T'}^* \le R_T^*$ for $T \le T'$. More generally, enlarging the class of admissible estimators cannot increase the infimum. For the ideal observer, the mutual information $I(g(X); H_T)$ is nondecreasing in T.*

Proof. Every $\mathcal{F}_T$-measurable estimator is $\mathcal{F}_{T'}$-measurable because $\mathcal{F}_T \subseteq \mathcal{F}_{T'}$, so $\mathcal{C}_T \subseteq \mathcal{C}_{T'}$, and an infimum over a larger set is no larger. The statement about enlarging the class is the same fact. Monotonicity of mutual information follows from the chain rule, since $I(g; H_{T'}) = I(g; H_T) + I(g; H_{T'} \mid H_T) \ge I(g; H_T)$. □

Two features of this result deserve emphasis. First, it is an ideal-observer statement under a correct model. A real institution using a misspecified model can become more confidently wrong as data accumulate, since posterior concentration under a wrong likelihood is concentration on the wrong answer, and so the proposition does not imply that real accuracy improves with retention. Second, the result has an inter-temporal reading. If $\mathcal{E}_s$ denotes the class of procedures available in era s, and the classes are nested as capability grows, then the best achievable risk from a fixed retained history is nonincreasing in s. Auxiliary reference data behave similarly: the total information available to an adversary, $I(S; H, Z_s)$ for a secret S and nested auxiliary sources Z_s, is nondecreasing in s by the chain rule, although the marginal contribution of H given Z_s may rise or fall. The conclusion is the formal version of retrospective inference. A retained record is an option on the future capabilities of whoever holds it. Deletion is the most direct way to extinguish the option, but not the only one: irreversible destruction of cryptographic keys, irreversible coarsening of the stored measurements, and loss of the auxiliary data needed for interpretation can all make a retained record unusable, each with different guarantees and different failure modes.

6.4 Authority

Proposition 3 (More information does not establish more authority)**.** *Let $\mathcal{P}$ be a set of rules, where each rule $r \in \mathcal{P}$ has a scope σ_r, the set of actions it authorizes, a grant condition τ_r on histories, and a revocation condition ν_r on histories (a lapsed warrant or an exoneration, for example). Define the authorized action set $\mathcal{A}(H, \mathcal{P}) = \bigcup\{\sigma_r : r \in \mathcal{P},\ \tau_r(H),\ \neg\nu_r(H)\}$. Then for every history H, $\mathcal{A}(H, \mathcal{P}) \subseteq \bigcup_{r \in \mathcal{P}} \sigma_r$. If there are no revocations and every grant condition is monotone, so that $\tau_r(H)$ implies $\tau_r(H')$ for $H \subseteq H'$, then $\mathcal{A}(H, \mathcal{P}) \subseteq \mathcal{A}(H', \mathcal{P})$. In general,*

enlarging the history can widen which rules apply or can shrink the authorized set, but it cannot enlarge the ceiling set by $\mathcal{P}$.

Proof. Each term of the union lies in the ceiling by definition. Under the stated assumptions, a rule granted on H is granted on H' and is never revoked, so every action authorized on H is authorized on H'. When revocation conditions are present, a rule granted on H may be revoked on H', and the authorized set can shrink. □

This statement is nearly tautological, and its value lies in locating the normative content. Information matters to authority only through grant and revocation conditions, such as a standard of suspicion, a threshold of evidence, or the lapse of a warrant, and an increase in available information changes which conditions are satisfied without creating new rules. The monotone case is the one relevant to the design hazard described below and is not a model of legitimate authority, since a defensible authority system must be defeasible. A new observation may show that an earlier identification was false, that a warrant has expired, or that a person has been exonerated, and in such cases additional information should remove actions that were previously available. The effective authorized set should depend on the current status of each authorization and not on the accumulation of historical observations. The ceiling result holds in either case, but it establishes only that a formal rule system cannot authorize outside its declared scopes. It does not establish that the scopes were legitimately adopted, and the legitimacy of $\mathcal{P}$ is a normative premise that no property of H supplies. If the supply of observation expands, the pressure on institutions is to treat satisfied triggers as sufficient grounds, and to interpret the rule set liberally, so that practice drifts beyond the ceiling. The proposition does not prevent that drift. It identifies it as a change in $\mathcal{P}$, which requires an authority that can legitimately change $\mathcal{P}$, and so it turns the question of data abundance into a question of rule-making. It also exposes a design hazard: a trigger that is cheap to satisfy under abundant data, such as the presence of any correlation above a threshold, may authorize far more action than its drafters intended, because the number of histories that satisfy it grows with the volume of observation.

6.5 Correction

Let $G = (V, E)$ be the true dependency graph, in which vertices are records and decisions and an edge $u \to v$ means that v used u as input, and let $\hat{G}$ denote the part of it that the institution has recorded. For a record r_0 found to be false, let $\mathrm{Reach}_G(r_0)$ denote the set of vertices reachable from r_0 in G. Three distinct properties are often run together under the word correction. *Detection* is the identification of the potentially affected decisions. *Revision* is the correction of their present authoritative status. *Restoration* is the undoing of, or compensation for, their consequences. The three are not equivalent: a wrongful detention can be terminated and the record corrected without restoring the time already lost, and a false financial designation can be revoked while some of its economic consequences remain irreversible.

Proposition 4 (Sufficiency of recorded dependencies)**.** *If $\hat{G}$ contains every edge of G reachable from r_0, then* $\mathrm{Reach}_G(r_0)$ *is computable by graph traversal, so detection is complete and exact. Revision of every decision in* $\mathrm{Reach}_G(r_0)$ *is then possible provided each such decision has a revisable authoritative status, and restoration is possible only for those decisions whose consequences are reversible or compensable. Complete recording therefore secures detection, and it secures revision and restoration only under the further conditions stated.*

Proof. Traversal of the recorded edges from r_0 visits exactly $\mathrm{Reach}_G(r_0)$. The remaining claims restate the definitions of revisability and reversibility: no property of the dependency record converts an irreversible consequence into a reversible one. □

Proposition 5 (Limits of discovery without records)**.** *Suppose there are two possible true dependency graphs G_1, G_2 that leave the same retained record, in the sense that every log, timestamp, and artifact available to the institution is identical, but with* $\mathrm{Reach}_{G_1}(r_0) \neq \mathrm{Reach}_{G_2}(r_0)$. *Then no procedure that takes the retained record as input can output* $\mathrm{Reach}_G(r_0)$ *exactly in both cases. Any procedure guaranteed to detect every affected decision in both cases must output a superset of* $\mathrm{Reach}_{G_1}(r_0) \cup \mathrm{Reach}_{G_2}(r_0)$*, and so must include decisions that did not in fact depend on r_0 in at least one of the cases.*

Proof. A procedure is a function of the retained record. Identical inputs give identical outputs, but the two required outputs differ. A guarantee of detection in both cases requires the output to contain both reach sets, hence their union, and since the two sets differ, the union strictly contains at least one of them. □

Conservative review of every later decision is an instance of the second proposition: it achieves detection without a complete record, at the cost of over-inclusion, which is a cost in review effort and, where review itself has consequences for the people involved, a cost in harm. The two propositions replace a single biconditional that would have conflated detection with restoration and treated a complete record as necessary for detection when conservative review can achieve it without one.

The second proposition has a direct application to learned models. If a record contributed to the training of a model, then every output of the model produced after training lies in $\mathrm{Reach}_G(r_0)$ unless the contribution can be shown to be removable, and the dependency edges are in general not recorded at the granularity needed. This is one reason the requirement of bidirectional provenance in Part III is substantive and not decorative, and it explains why correction is harder than detection in systems built without lineage in mind. An analogy to distributed systems is also instructive, though it is offered as analogy and not as a derivation. Results on the impossibility of simultaneously guaranteeing consistency, availability, and partition tolerance [16] suggest that when a retraction is issued under network partition, a downstream system must either withhold action until it can confirm the retraction, which costs availability, or act on possibly retracted records, which costs the correction guarantee. A legal and institutional rule that required confirmed retraction before action would impose the first cost and would have to be justified as the price of reversibility.

6.6 Privacy

Proposition 6 (Privacy loss is not observation volume)**.** *Let S be a secret about an individual and Z an adversary's auxiliary information, and measure privacy loss by $I(S; H \mid Z)$, writing $\mathsf{H}(\cdot)$ for Shannon entropy. (a) If $H_1 \subseteq H_2$ are nested histories, with S, Z, and the joint distribution held fixed, then $I(S; H_2 \mid Z) \geq I(S; H_1 \mid Z)$. (b) Across observation systems that are not nested, privacy loss is not a function of data volume: there exist histories H_1, H_2 with $\mathsf{H}(H_1) \gg \mathsf{H}(H_2)$ and $I(S; H_1) = 0 < I(S; H_2)$. (c) Privacy loss depends on the auxiliary information: there exist S, A, B with $I(S; A) = 0$ and $I(S; A \mid B) = \mathsf{H}(S)$.*

Proof. For (a), the chain rule gives $I(S; H_2 \mid Z) = I(S; H_1 \mid Z) + I(S; H_2 \mid H_1, Z)$, and conditional mutual information is nonnegative. For (b), let H_1 be n independent noise variables unrelated to S, with n large, and let H_2 be a single bit equal to S. For (c), let A and B be independent uniform bits and $S = A \oplus B$. Then S is uniform and independent of A, so $I(S; A) = 0$, while S is determined by (A, B), so $I(S; A \mid B) = \mathsf{H}(S) = 1$ bit. □

Part (a) and Proposition 2 reinforce one another: for a fixed secret, fixed auxiliary information, and a fixed joint distribution, retention can only increase what an ideal observer learns, so the leakage of a retained history is nondecreasing in the history. Part (b) says only that this monotonicity holds within a nested family. Across different sensor suites, volume is a poor proxy for leakage, since a million independent noise samples may reveal nothing while a single bit may reveal everything. Part (c) is the one-time pad, and it shows that the same observation can be harmless or fully revealing depending on what else the observer holds. This is the formal reason that privacy in a densely instrumented environment cannot be assessed record by record. The classical statement of the difficulty in the statistical-disclosure literature is Dalenius's desideratum, that access to a published statistic should reveal nothing about an individual that could not be learned without it, and the established result in the differential-privacy literature is that this desideratum cannot be achieved in general in the presence of auxiliary information [12, 13]. Differential privacy offers instead a guarantee about the marginal effect of an individual's participation on a released output, which is a different and weaker property. For a mechanism satisfying ϵ-local differential privacy applied independently n times to the same secret, the ratio of posterior odds is bounded by $e^{n\epsilon}$ by the product structure of the likelihood ratios, a basic composition bound [13]. In an environment where the number of observations per individual grows without bound as the cost per observation falls, any fixed per-observation guarantee is consumed at a rate proportional to the number of observations, and a fixed total budget becomes a binding constraint on the number of observations that can be made at all, provided the observations are released through mechanisms that satisfy such a guarantee. Differential privacy is a property of a specified randomized mechanism and a neighboring-input relation, and not a generic property of observation: an ordinary passive sensor reading is not a differentially private release and consumes no formal budget. The budget argument therefore describes the constraint on any institution that chooses to offer such a guarantee, and for unrandomized sensing the corresponding constraint has to be stated in terms of the leakage $I(S; H \mid Z)$ and of limits on use. Where such a guarantee is offered, the

argument is the formal counterpart of the claim that scarcity has moved: the scarce resource is no longer the observation but the privacy budget it consumes.

Empirical work on reidentification makes the same point from the other direction. Demonstrations that ostensibly anonymized records can be reidentified by joining them with auxiliary data, whether through a small number of demographic attributes [10] or through sparse high-dimensional behavioral data [11], show that the removal of direct identifiers does not remove the identifying information, which is carried by the pattern of the remaining attributes.

6.7 Erasure and correction

An institution may be subject to two obligations that appear to pull in opposite directions. The first is erasure: information identifying the subject of a record must be irreversibly removed. The second is correction: if a record is found to be erroneous, every decision that depended on it must be discoverable. The tensions table below records a loose version of the conflict, and this subsection states exactly when it arises.

Let Λ denote the full dependency record, which specifies which records concern which subjects and which decisions used which records. Let $\sim$ be the equivalence relation on possible dependency records under which $\Lambda \sim \Lambda'$ if and only if the two differ only in subject-identifying content, such as which subject a given record concerns. The retained state after erasure is $e(\Lambda)$, and the erasure obligation is that $e(\Lambda) = e(\Lambda')$ whenever $\Lambda \sim \Lambda'$, so that the retained state carries no information distinguishing erasure-equivalent histories. For a query q designating a subject, let $D_q(\Lambda)$ denote the set of decisions that depended on records concerning q in Λ. Exact subject-specific discovery means a procedure c with $c(e(\Lambda), q) = D_q(\Lambda)$ for every Λ and every q.

Proposition 7 (Compatibility of erasure and exact correction)**.** *There exist an erasure map e and a discovery procedure c satisfying both obligations if and only if $D_q(\Lambda) = D_q(\Lambda')$ for every query q and every pair $\Lambda \sim \Lambda'$. In particular, if there are erasure-equivalent histories $\Lambda \sim \Lambda'$ and a query q with $D_q(\Lambda) \neq D_q(\Lambda')$, then no system can guarantee both obligations.*

Proof. Suppose e and c exist and $\Lambda \sim \Lambda'$. Then $e(\Lambda) = e(\Lambda')$, so $D_q(\Lambda) = c(e(\Lambda), q) = c(e(\Lambda'), q) = D_q(\Lambda')$. Conversely, suppose the invariance condition holds. Let $e(\Lambda)$ be the $\sim$-class of Λ, which satisfies the erasure obligation by construction, and define $c([\Lambda], q) = D_q(\Lambda)$, which is well defined by the invariance condition. □

The condition fails in the simplest nontrivial case. Let there be two subjects s_1, s_2, two records r_1, r_2, and two decisions, d_1 that used only r_1 and d_2 that used only r_2. Let Λ assign r_1 to s_1 and r_2 to s_2, and let Λ' assign r_1 to s_2 and r_2 to s_1. The two differ only in subject-identifying content, so $\Lambda \sim \Lambda'$, yet $D_{s_1}(\Lambda) = \{d_1\}$ while $D_{s_1}(\Lambda') = \{d_2\}$. After erasure the retained state cannot tell the two cases apart, and a correction for s_1 has to act on different decisions in the two.

The proposition also identifies what must survive erasure. The retained state must determine the map $q \mapsto D_q(\Lambda)$, and since it must be constant on $\sim$-classes it can carry only the part of that map that is invariant under erasure. Pseudonymous dependency tokens show one way the invariance condition can be restored. Suppose each record carries a random token, the dependency edges are keyed by tokens, and the link from subject to token is held outside the institution's erasure obligation, for example by the subject or by a separate correction authority. If queries range over tokens, then the answer to a query does not depend on subject-identifying content, the invariance condition holds, and exact discovery is available to whoever holds the token. The cost is that correction can then be initiated only by a holder of the link. The retention has been relocated and not removed, and the arrangement satisfies the erasure obligation only if that obligation is scoped to the institution and the holder of the link is trusted and available.

The result is conditional, and it is stated for a deterministic idealization in which erasure is exact. It does not say that correction and deletion always conflict. It says that they conflict exactly when erasure-equivalent histories require different correction responses, and it turns the question of which information must survive erasure into a question about the invariance of the correction answer. Approximate or probabilistic erasure, adversaries holding auxiliary data, and approximate discovery are not treated, and cryptographic commitments or separated correction authorities may achieve in practice what the idealized model rules out. Of the results in this part, this is the one most open to extension, since the interesting question is which relaxations of exactness restore compatibility and at what cost.

6.8 An admissibility condition

The propositions suggest a general formulation. An action a is admissible, given history H_T and rule set $\mathcal{P}_T$, if the evidence for it is sufficient, the actor has valid authority, the inference is provenance-preserving, the action satisfies applicable constraints, and required correction mechanisms remain available:

$$\text{Admissible}(a, H_T, \mathcal{P}_T) \iff \begin{cases} \text{the evidence supporting } a \text{ is sufficient,} \\ \text{the actor possesses valid authority,} \\ \text{the inference is provenance-preserving,} \\ \text{the action satisfies applicable constraints,} \\ \text{required correction mechanisms remain available.} \end{cases}$$

This is a proposed formal specification and not a theorem. Its usefulness depends on defining each predicate and then determining which conjuncts are jointly satisfiable, which conflict, and which cannot be guaranteed under realistic assumptions. The predicates map onto the propositions: evidence sufficiency is constrained by identifiability and, as the next subsection shows, by the consequences of the action; valid authority by the structure of grants, revocations, and ceilings; provenance preservation and correction availability by reachability and by the compatibility result for erasure; and the constraints on action in part by privacy loss relative

to auxiliary information. Table 2 records the principal tensions.

Conjunction	Tension
Provenance vs. privacy	Complete provenance logs are themselves a body of sensitive records, as the documented concern about sensitive content in system logs illustrates; more complete lineage enlarges the surface that must be protected.
Correction vs. erasure	Exact discovery of affected decisions and irreversible erasure of subject-identifying information are compatible exactly when the correct discovery answer does not vary across erasure-equivalent histories (Proposition 7); pseudonymous dependency tokens can meet the condition only by relocating the subject-to-token link to a party outside the erasure obligation.
Evidence sufficiency vs. identifiability	Where the sensor suite is structurally unable to distinguish states, no volume of retained data makes the evidence sufficient; admissibility then requires either additional sensor types or a restriction of the action to what is identifiable.
Authority vs. availability under partition	Requiring confirmed authority and confirmed retraction before action costs timeliness; waiving the requirement under time pressure costs reversibility. The tradeoff has to be settled by rule in advance and not by operators in the moment.
Provenance preservation vs. inference by models	Learned inference procedures may not expose which inputs drove an output; provenance then holds only at the level of the input set, which makes Reach coarse.

Table 2: Principal tensions among the conjuncts of the admissibility condition.

The conjuncts are jointly satisfiable in simple cases, such as an action based on a single, directly identifying, well-provenanced record by an actor with explicit authority in a system with complete lineage. They become progressively harder to satisfy together as the action depends on inference from many weakly informative observations, a point that follows from the combination of the base-rate argument of Part II and the reachability argument above. A research program in this area would classify classes of actions and systems by which conjuncts can be guaranteed, which can be guaranteed only probabilistically, and which cannot be guaranteed at all.

6.9 Action-dependent evidence thresholds

The conjunct of the admissibility condition requiring sufficient evidence cannot be defined independently of the action. Suppose two actions depend on the same probabilistic identification, one triggering a reversible administrative review and the other an irreversible deprivation. It would be hard to justify treating the same evidentiary threshold as sufficient for both merely because the identification procedure has the same accuracy. Let $L(a, x) \geq 0$ measure the consequences of taking action a when the state is x, and let $\theta(a, \mathcal{P})$ be an institutionally established risk tolerance. A necessary condition for admissibility is then

$$\mathbb{E}\big[L(a, X) \mid H_T\big] \leq \theta(a, \mathcal{P}).$$

This is only a necessary condition. Expected loss cannot by itself establish legal authority, and some rights-based constraints cannot be traded against aggregate benefit. It can also conceal tail risk, since severe consequences can disappear into an average, and so where an action can cause irreversible harm a separate constraint on the probability that the action is wrong is imposed in addition to the expected-loss bound.

Proposition 8 (Reversibility scaling)**.** *Let $W_a \subseteq \mathcal{X}$ be the set of states in which taking a is wrong, let $L(a, x) = \ell_a \mathbf{1}[x \in W_a]$ with $\ell_a > 0$, and let $q_a(H) = \Pr(X \in W_a \mid H)$. (a) The condition $\mathbb{E}[L(a, X) \mid H] \leq \theta$ holds if and only if $q_a(H) \leq \theta/\ell_a$. Consequently, if a and a' have the same wrongness set and $\ell_{a'} = k\,\ell_a$ with $k > 1$, then under a common tolerance θ the posterior error probability tolerated for a' is $1/k$ of that tolerated for a, so a posterior that authorizes a may fail to authorize a'. (b) Suppose the state is known to lie in a pair $\{x, x'\}$ of observationally equivalent trajectories, with $x \in W_a$ and $x' \notin W_a$ and prior masses p and $1-p$. Then $q_a(H_T) = p$ for every T, and a fails the expected-loss condition for every T whenever $p > \theta/\ell_a$.*

Proof. For (a), $\mathbb{E}[L(a, X) \mid H] = \ell_a\, q_a(H)$, and dividing by $\ell_a > 0$ gives the equivalence. The comparison follows by substituting $\ell_{a'} = k\ell_a$. For (b), the proof of Proposition 1 shows that the posterior over $\{x, x'\}$ equals the prior for every T, so $q_a(H_T) = p$ for every T. □

Part (b) joins this proposition to Proposition 1. A structural identifiability failure bounds the confidence that any amount of retained data can supply, and so it bounds the set of actions that the data can ever justify: an action whose loss is large enough is inadmissible on that evidence alone, whatever is retained. The proposition also gives a precise sense to the claim that what matters is which inferences are allowed to become irreversible actions. The accuracy of the identification procedure is one input, and the consequences of the action determine how much of that accuracy the action requires.

6.10 A worked example

Consider a stylized deployment in which a municipal agency uses a record-linkage system to flag a person as the individual named in an outstanding warrant, and the flag licenses an officer to stop and question that person. The action a is the stop. Evidence sufficiency asks whether the posterior probability that the flagged person is the named individual meets the standard the rule requires. With prevalence π of true matches among the candidates considered, sensitivity s, and per-comparison false-match probability α, the posterior is $s\pi/\big(s\pi + \alpha(1-\pi)\big)$, which is small whenever α is not much smaller than π, however small α is in absolute terms. If the sensor suite cannot distinguish two people with the same recorded attributes, Proposition 1 says that no volume of retained data raises it. A reversible administrative review triggered by the same flag and a stop that can lead to detention may share the identification procedure but not the evidentiary threshold (Proposition 8).

Valid authority asks whether the warrant is in force. By Proposition 3, that is a matter of grants and revocations: a warrant recalled yesterday must remove the stop from the authorized

set today, which requires that the recall has propagated. Provenance preservation asks whether the link between the person's record and the warrant record can be traced to the observations that support it and to their sources, without which the evidence-sufficiency conjunct cannot even be assessed. Correction availability asks whether, if the match is later found wrong, the stop and anything that followed from it can be discovered and revised, which depends on the dependency record (Propositions 4 and 5) and on what survives any erasure obligation (Proposition 7).

The example shows that the conjuncts are not independent, contrary to what a list suggests. Evidence sufficiency cannot be assessed without provenance, valid authority cannot be sustained without revocation propagation, and revocation propagation is itself an instance of the correction mechanism. An admissibility condition is therefore better read as a set of coupled constraints on a single record-keeping architecture than as five checks to be performed independently. The example is hypothetical and is offered to test the condition and not to describe any real deployment.

6.11 Two dimensions of observation

The formal framework also allows the question of who observes whom to be posed without reference to any particular politics. Two quantities can be distinguished and should be reported separately, each with an explicit target. The *downward observation* of individuals by an institution is $\Omega_{\downarrow} = I\big(X^{\mathrm{ind}}; H^{\mathrm{inst}}\big)$, where X^{ind} is a specified feature of individuals' conduct and H^{inst} is the history available to the institution under its access and query policy. The *upward observation* of institutions by independent parties is $\Omega_{\uparrow} = I\big(D^{\mathrm{inst}}; H^{\mathrm{pub}}\big)$, where D^{inst} denotes a specified class of institutional decisions and H^{pub} is the history available to the public or to an independent oversight body. A comparison of the two is not meaningful without more structure. The targets may differ radically in entropy, dimensionality, and temporal structure, so a positive difference $\Omega_{\downarrow} - \Omega_{\uparrow}$ need not indicate any institutional advantage. A comparison becomes meaningful only after fixing a common decision task and measuring the reduction in decision risk that each history affords, or after a common normalization such as the fraction of target entropy resolved, $\Omega/\mathsf{H}(\mathrm{target})$, and even then the normalization is a modeling choice that has to be defended. The distinction between the two directions of observation is valuable without reduction to a single scalar. Both quantities are functions of the sensor suite together with the access and query policy. Since the sensor suite is a physical fact and the access policy is a rule, the same instrumented environment can support very different values of both, and nothing in the physics ties the two together. A technologically advanced surveillance architecture could therefore sustain asymmetrical institutional power or reciprocal accountability, depending on its authority structure. This formulation is in conversation with Brin's argument for reciprocal transparency [7], with Mann and colleagues' account of inverse surveillance [8], and with the surveillance-studies literature that treats visibility as relational and not as a property of the observed alone [5].

7 Discussion: Observability, Accountability, and the Fate of Privacy as Access Control

The deepest distinction the formal analysis supports is between universal observability and universal accountability. A society might become capable of recording nearly every publicly observable event without becoming capable of interpreting those events correctly, and the first capacity does not entail the second. Conversely, a society might become extremely effective at monitoring its own institutions while retaining meaningful constraints on the surveillance of individuals. These are different technological and political trajectories, and the work of institutional design consists largely in choosing between them. The mathematical results above explain why the choice cannot be left to technology: identification does not follow from observation, retention enlarges future options without regard to present intent, information does not create authority, correction depends on dependency records that must be kept in advance and whose compatibility with erasure turns on what the correction answer needs to know, the evidence an action requires scales with how reversible the action is, and privacy loss depends on what else is known.

The conventional conception of privacy as control over access to information is under particular pressure. Nissenbaum's contextual integrity already shifts the question from access to the appropriateness of information flows given the norms of the context in which information was shared [4], and Zuboff's account of surveillance capitalism describes the commercial extraction and repurposing of behavioral data that such flows make possible [6]; both fit the present argument, though neither supplies the formal treatment of reconstruction attempted here. If information about an individual can be reconstructed from indirect measurements, preventing direct access to a record does not prevent reconstruction of its contents from correlated observations. Access control nonetheless remains useful, and its effectiveness depends on the threat model, much as the possibility of reconstructing a document does not make encryption useless; what fails is its sufficiency as a complete theory of privacy. In formal terms, protection of a secret $S = g(X)$ against an observer who sees $E = h(X)$ requires that $g(X)$ be independent of $h(X)$ under the observer's prior, a condition that fails whenever the world is physically coupled in the way that makes sensing possible in the first place. The conditions under which information about an individual can be made inaccessible while the surrounding environment remains observable form a question of considerable mathematical interest, joining inverse problems, information theory, causal inference, cryptography, distributed systems, and jurisprudence. The existing literature on inverse problems, which classifies when and how reconstruction from indirect data is stable, is directly relevant, since the same ill-posedness that makes reconstruction hard also makes some protections robust [15], while the theory of the comparison of experiments gives a precise sense in which one sensing arrangement is at least as informative as another [14]. Three families of response suggest themselves. Coarsening the sensor, by limiting the resolution, bandwidth, or modality of observation, reduces what can be inferred and is a design choice available at the point of manufacture. Breaking the dependence, through randomization at the source as in local differential privacy, converts a structural guarantee into a budgeted one. And restricting use, through admissibility conditions on actions as

opposed to access conditions on records, places the constraint where the harm occurs. None of the three is sufficient alone, and the third is the most consistent with the central thesis, since it moves protection from the supply of data to the legitimacy of action.

Several limitations should be stated. The formal framework treats institutions as single decision-makers and the world state as a well-defined random object, which are idealizations. The erasure result in particular is stated for a deterministic idealization in which erasure is exact, and approximate or probabilistic erasure, adversaries holding auxiliary data, and approximate discovery are not treated. The propositions are deliberately simple, and their value is to isolate mechanisms and not to predict outcomes. The technological analysis relies on review articles, which summarize laboratory and early-practical work and are not evidence of deployment, and the scenarios in Part IV are conditional structures and not forecasts. The assessment of the documented behavior of integration platforms relies on public vendor documentation and vulnerability records, and so it describes what is publicly stated and not what is internally true. Finally, the question settled provisionally in the framing, whether the necessity argument concerns the state's capacity to observe society or society's capacity to observe the state as well, is answered here by treating the two as independent dimensions of one problem. A different framing would produce a different monograph, and the decision to treat reciprocity as a variable and not as a premise is the main substantive commitment of the work.

8 Conclusion

Surveillance has always been shaped by what it cost, and the cost structure is changing. The consequence is not simply that more can be observed. It is that the limits on observation, which scarcity once supplied implicitly, must now be supplied explicitly or not at all. The institutions best placed to integrate information at scale are also the ones for which the questions of identity, authority, provenance, and reversibility are most pressing, and the present generation of systems devotes more of its capability to finding than to correcting. Over a longer horizon, photonic and related technologies may extend sensing into ordinary materials, but the physical dependencies that remain, detectors, power, readout, analyte access, and interpretation, define the points at which governance can act and the conditions under which each future is possible. The formal framework developed here does not predict which future will occur. It states what each possible future would require of its institutions, and it shows that the decisive requirements concern the legitimacy of inference, the allocation of authority, the integrity of evidence, and the reversibility of decisions, rather than the quantity of observation.

References

[1] J. Bentham, *Panopticon; or, The Inspection-House*, 1791.

[2] M. Foucault, *Discipline and Punish: The Birth of the Prison*, Gallimard, 1975.

[3] J. C. Scott, *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*, Yale University Press, 1998.

[4] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University Press, 2010.

[5] D. Lyon, *Surveillance Studies: An Overview*, Polity, 2007.

[6] S. Zuboff, *The Age of Surveillance Capitalism*, PublicAffairs, 2019.

[7] D. Brin, *The Transparent Society*, Addison-Wesley, 1998.

[8] S. Mann, J. Nolan, and B. Wellman, "Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments," *Surveillance & Society* 1(3), 331–355, 2003.

[9] I. P. Fellegi and A. B. Sunter, "A theory for record linkage," *Journal of the American Statistical Association* 64(328), 1183–1210, 1969.

[10] L. Sweeney, "Simple demographics often identify people uniquely," Carnegie Mellon University, Data Privacy Working Paper 3, 2000.

[11] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," *IEEE Symposium on Security and Privacy*, 2008.

[12] T. Dalenius, "Towards a methodology for statistical disclosure control," *Statistik Tidskrift* 15, 429–444, 1977.

[13] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science* 9(3–4), 211–407, 2014.

[14] D. Blackwell, "Comparison of experiments," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93–102, 1951.

[15] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, 1996.

[16] S. Gilbert and N. Lynch, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services," *ACM SIGACT News* 33(2), 51–59, 2002.

[17] X. Hu, W. Xu, Q. Fan, T. Yue, F. Yan, Y. Lu, and T. Xu, "Metasurface-based computational imaging: a review," *Advanced Photonics* 6(1), 014002, 2024.

[18] C. Roques-Carmes, K. Wang, Y. Yang, A. Majumdar, and Z. Lin, "Computational metaoptics for imaging," arXiv:2411.09133, 2024.

[19] A. J. El-Helou, Y. Liu, C. Chen, and F. Wang, "Optical metasurfaces for the next-generation biosensing and bioimaging," *Laser & Photonics Reviews*, 2025.

[20] A. Berestennikov, H. Hu, and A. Tittl, "Molecular spectroscopies with semiconductor metasurfaces: towards dual optical/chemical SERS," *Journal of Materials Chemistry C* 13, 11499–11514, 2025.

[21] M. Iwanaga, Q. Hu, and Y. Tang, "Metasurface biosensors: status and prospects," *Applied Physics Reviews* 12(2), 021305, 2025.

[22] National Institute of Standards and Technology, *NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management, Version 1.0*, January 2020.

[23] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, 2023.

[24] National Institute of Standards and Technology, *NIST Privacy Framework 1.1, Initial Public Draft*, NIST CSWP 40 ipd, 2025, https://doi.org/10.6028/NIST.CSWP.40.ipd. Draft status checked on NIST's site in October 2026; the description of its contents follows law-firm commentary on the April 2025 release, and the final text is not assumed.

[25] Palantir Technologies, "Architecture center: overview," Foundry documentation, accessed October 2026. https://www.palantir.com/docs/foundry/architecture-center/overview

[26] Palantir Technologies, "AIP architecture," Foundry documentation, accessed October 2026. https://www.palantir.com/docs/foundry/architecture-center/aip-architecture

[27] Palantir Technologies, "Migrate from and disable the 'Propagate view requirements' setting," Foundry documentation, accessed October 2026. https://palantir.com/docs/foundry/platform-security-management/disabling-propagate-view-requirements/

[28] Palantir Technologies, "Configure logging [Beta]," Foundry documentation, accessed October 2026. https://palantir.com/docs/foundry/administration/configure-logging

[29] MITRE Corporation, CVE-2022-27895: Information exposure through log files in Palantir Foundry Build2, published November 2022, https://www.cve.org/CVERecord?id=CVE-2022-27895.