

From the Panopticon to the Datapoint: Operational power and governance by design

AUTHOR:

Torres Ponce, Mariano Enrique

Lawyer (LL.B.), Specialist in Computer Law

ABSTRACT

This article introduces a conceptual category intended to capture the logic of contemporary technological authority without relying on inherited metaphors or remaining at the level of declarative critique. The notion of operational power designates a form of authority that emerges when technical environments preconfigure the space of possible action, incorporate behavioral feedback, and ground their legitimacy in the continuity of their functioning. Under these conditions, governance no longer operates primarily through articulated norms but through systems whose effects are realized directly in the organization of conduct.

From this starting point, the article argues that a constitutional response limited to the language of principles is no longer sufficient. Effective constraints require translating normative commitments into properties of design and operation that can be examined, contested, and verified. The analysis therefore develops a set of material guarantees aimed at restoring a human interval and meaningful control, articulated through practical

intelligibility, traceability that allows causal reconstruction, reversibility with bounded costs, and a proportionality that adapts dynamically to risk and impact.

The article further outlines an independent verification method conceived as part of the system lifecycle rather than as a posterior formality. This method relies on counterfactual auditing guided by design parameters, stress and failure testing under realistic conditions, and the preservation of auditable archives that enable third party review. To make these guarantees operational, the framework proposes falsifiable metrics that allow abstract concerns to be translated into observable thresholds, including default adherence, reversibility gradients, effective opacity, and behavioral entropy reduction.

The argument is completed by an institutional standard applicable to both public and private contexts, addressing procurement practices, separation of functions, formal recognition of the normative force of default settings, and forms of interoperability and portability that preserve contextual integrity. Taken together, these elements offer a way to limit governance by optimization without arresting innovation, while reopening a space for judgment in environments where personalization and default configurations increasingly govern in fact.

KEYWORDS

operational power, operational constitutionalism, algorithmic governmentality, default jurisdiction, practical intelligibility, traceability, reversibility, falsifiable metrics, counterfactual auditing, normative architecture

EXECUTIVE SUMMARY

Background: Everyday life is increasingly organized within digital environments that shape relevance, routes, and temporal rhythms without asserting visible authority. In this context, explicit coercion recedes and optimization takes its place, while legitimacy shifts away from command toward experienced utility. Technical systems configure the space

of action through predefined parameters and default settings that operate in practice as governing forces. Consent becomes procedural, often reduced to repetition, and personalization absorbs conflict by aligning choices with preferences that the system itself has progressively formed.

Gap: Critical scholarship has offered persuasive diagnoses of this transformation, yet it largely remains at the level of description. Panoptic metaphors and analyses focused on data extractivism identify symptoms but do not establish verifiable thresholds or remedies capable of being enforced. Law continues to operate in a declarative register while contemporary power is exercised through technical realization, which renders guarantees ineffective when they are not translated into features of design and operation. What remains insufficiently articulated is a category that distinguishes assisted automation from governance by design and that enables the authority of technical environments to be subjected to empirical scrutiny through falsifiable standards.

Purpose: This article advances the notion of operational power to capture that distinction. Operational power arises when legitimate action is parameterized in advance through default settings that channel conduct, when behavioral feedback recalibrates the environment without explicit notice, and when legitimacy is grounded in the smoothness of functioning rather than in justification. Only when these elements converge does governance by design acquire constitutive force. On this basis, the article develops material guarantees intended to restore a human interval and meaningful control without arresting innovation. Normative principles are translated into verifiable conditions of practical intelligibility, traceability sufficient for causal reconstruction, reversibility with bounded costs, and a proportionality that adjusts dynamically to risk and impact. The contribution is explicitly theoretical and normative in scope and does not claim to report original empirical datasets.

Methodology: An independent verification method is proposed as part of the system lifecycle rather than as a terminal form of control. Guided counterfactual auditing intervenes in defaults, frictions, and latency windows in a controlled manner to assess the channeling force of architectural choices. Failure, stress, and context change scenarios are examined to identify points at which reversibility erodes or consent loses substantive meaning, with particular attention to vulnerable populations and boundary conditions that average validations tend to obscure. In operational terms, metrics such as default adherence or reversibility gradients can be implemented by observing how frequently

users remain on preconfigured paths when exit entails nontrivial costs, and by measuring the time, effort, or loss required to undo an automated outcome once it has taken effect. A minimal auditable archive preserves model versions, data routes, relevant system states, and chains of custody so that contested decisions can be reconstructed by third parties. These falsifiable metrics transform philosophical diagnosis into testable hypotheses, including default adherence as an indicator of environmental jurisdiction by design, reversibility gradients as a measure of effective undoing, effective opacity as the distance between declared behavior and reconstructable operation, and behavioral entropy reduction as a signal of narrowing trajectories without proportional gains in legitimate ends.

Results: The proposed institutional implementation realigns incentives and redistributes capacities for control. Public and corporate procurement can require auditable archives from the outset, activation thresholds for safe shutdown and degradation, and provisions for third party verification under confidentiality. Governance structures separate design, operation, and supervision, assess effective human intervention rather than nominal oversight, and activate a right to deliberative latency that enables pauses, escalations, and reasonable disconnections. Formal recognition of default jurisdiction reflects the fact that defaults govern in practice and therefore require reinforced justification, non-discrimination testing, symmetric alternatives, and documented rationales. Interoperability and portability oriented toward performance equivalence prevent personalization from hardening into an exit barrier. The framework also incorporates a competition policy expressed in operational terms, addressing self-preferencing in ranking and relevance systems, integrations that close ecosystems, and access to interfaces and auditable logs for research and competitive scrutiny.

Conclusion: The contribution of the article is twofold. It offers a conceptual framework that distinguishes assisted automation from governance by optimization and an operational standard capable of translating principles into tests and remedies. Its originality lies in specifying material guarantees and falsifiable metrics that render limits enforceable in systems that modulate conduct through architecture rather than through explicit mandates. The legitimacy of technical environments cannot rest on functioning alone but must be supported by verifiable constraints that preserve a human interval and meaningful control. Where governance operates through optimization, guarantees must take the form of requirements of design, operation, and proof that any independent party

can audit. Under these conditions, law recovers practical relevance in real time, innovation is shaped by responsibility, and operational power is contained without extinguishing the creative capacity needed to design freedoms rather than efficiencies alone.

TABLE OF CONTENTS

Abstract

Keywords

Executive Summary

A. Introduction. The Power That Remains Unseen

B. Philosophy of Power. From Law to the Algorithm

C. Philosophy of Surveillance. From the Panopticon to the Datapoint

D. Operational Constitutionalism. Material Guarantees and Control Thresholds

E. Independent Verification Method and System Lifecycle

F. Political Economy of Operational Power. Incentives and Design Capture

G. Objections, Limits, and Second Order Risks

H. Institutional Implementation. An Operational Standard for Public and Private Practice

I. Conclusions and Empirical Agenda

References

A. INTRODUCTION. THE POWER THAT REMAINS UNSEEN

The power that organizes contemporary life no longer depends on a visible figure or on solemn forms of expression. It operates as an environment that frames everyday experience and guides action without presenting itself as authority. For a long time, sovereignty was embodied first in the person of the ruler and later in the written norm. In the present, authority is increasingly distributed across technical architectures that shape conduct with such effectiveness that experienced utility displaces obedience as a source of legitimacy (Foucault, 1975; Lessig, 1999; Bobbio, 1991).

What was previously negotiated through the language of duty is now resolved through system design. These systems do not deliberate, justify, or persuade. They execute. Their effects accumulate quietly and organize behavior in ways that recall earlier grammars of punishment and legal obligation, yet with a capacity for modulation that compresses the interval in which interpretation, judgment, and dissent once took place. That interval, which defined legal freedom in a robust sense, is progressively absorbed by continuity and speed.

This transformation cannot be reduced to a mere informatization of law. It reflects a deeper shift in normativity, one that relocates the center of gravity from articulated rules to technical conditions of possibility. Where procedures once allowed delay and contestation, interfaces now favor confirmation and uninterrupted flow. User experience comes to stand in for justice, while preference is shaped through iterative feedback and returned to the individual as a programmed response. Surveillance no longer requires a tower or a frontal gaze. It is embedded in personalized services that accumulate traces, anticipate trajectories, and recombine profiles, producing a statistical self that gradually aligns choice with prediction. Autonomy is eroded not through prohibition but through optimization, as openness to the world gives way to managed possibility presented as convenience (Deleuze, 1990; Zuboff, 2019; Han, 2014).

Describing this configuration requires a category that goes beyond panoptic imagery or accounts focused solely on opaque data extraction. The distinctive feature of the present is not constant observation alone, but the replacement of explicit coercion by optimization as a governing principle. Operational power names the form of authority that parameterizes the field of the possible in advance so that compliance consists in following predefined trajectories, learns continuously from behavioral feedback, and grounds its legitimacy in fluency and usefulness rather than in command or sanction. This distinction

makes it possible to separate instrumental automation from a platform logic that no longer functions merely as infrastructure, but as an operational principle that defines the conditions of order themselves (Rouvroy & Berns, 2013; Floridi, 2014).

Seen from this angle, it becomes clear why law loses symbolic and practical primacy when it remains confined to declaration. Contemporary power acts through technical realization and configures action through defaults and parameters that govern in fact. Any rights-based approach that seeks effectiveness must therefore translate normative commitments into verifiable operational conditions. The first movement of this article advances this hypothesis by combining genealogical perspective and contemporary diagnosis. A law that speaks proves insufficient when confronted with systems that act. Legal thought is thus compelled to develop enforceable criteria of intelligibility, traceability, reversibility, and proportionality capable of restoring a human interval and meaningful control in contexts where computational speed tends to eliminate delay. Without delay, deliberation withers, and without deliberation practical freedom cannot be sustained (Lessig, 1999; Floridi, 2014; Foucault, 1975).

The guiding claim developed throughout the article is that constitutionalism must learn to operate in the language of execution without abandoning its limiting function. This requires a careful distinction between systems that govern through optimization and those that merely assist human decision making, along with the definition of thresholds and remedies that render both critical diagnoses and technical promises open to verification. The present introduction outlines this conceptual program. The sections that follow specify the contours of operational power, articulate material guarantees for an operational constitutionalism, and propose methods and metrics capable of subjecting to empirical scrutiny what the philosophy of power has long identified as a drift from law toward algorithmic ordering, with the aim of reopening the space of judgment within an order that tends to present itself as nothing more than the inevitability of functioning (Deleuze, 1990; Winner, 1980).

B. PHILOSOPHY OF POWER. FROM LAW TO THE ALGORITHM

A genealogical view of power reveals a gradual displacement that moves from sovereign presence to textual mediation and, more recently, to technical operation. For a long period,

law organized authority within the register of language and procedure, sustaining a regime in which time for deliberation allowed judgment to take shape. In the present, authority is increasingly organized through architectures that execute. Normativity no longer appears primarily as an articulated rule, but as a condition of possibility embedded in technical arrangements that redirect conflict away from debate and toward design, making compliance an environmental outcome rather than an explicit act of will (Foucault, 1975; Deleuze, 1990).

Grasping this shift requires abandoning the idea of technology as a neutral instrument. Technical artifacts possess a mode of existence that reshapes memory, perception, and action. Acceleration compresses the distance between decision and execution and narrows the space where interpretation and delay once played a role. As continuity becomes the dominant value, political reasoning drifts toward logistics, and practical judgment loses depth under the pressure of uninterrupted digital flow. In this context, a form of power emerges whose legitimacy rests less on justification than on performance (Simondon, 1958; Stiegler, 2016; Virilio, 1995).

The novelty of the current configuration does not lie simply in the fact that code regulates behavior. What is decisive is that regulation migrates from obligation to design. Interfaces suggest preferred paths, protocols stabilize exchanges, and default settings orient conduct without appearing as commands. Technical artifacts thus condense political decisions, while law that remains confined to declarative statements cedes ground to an order that governs through execution. Utility is easily mistaken for legitimacy at precisely the moment when the technical environment acquires de facto authority over conduct (Lessig, 1999; Winner, 1980).

At this point a distinction becomes necessary. Not every use of automation produces a new form of power. Operational power refers to a specific configuration in which the field of legitimate action is parameterized in advance, behavioral feedback is reintegrated to recalibrate the environment, and optimization replaces coercion as the dominant governing logic, with fluency and usefulness serving as sources of legitimacy. Where these elements do not converge, technology may assist decision making, but it does not acquire constitutive normative force. Drawing this boundary prevents analytical inflation and allows attention to focus on cases in which technical systems cease to support judgment and begin to structure it (Yeung, 2017; Pasquale, 2015; Zuboff, 2019).

From this perspective, appeals to algorithmic transparency cannot remain at the level of abstract commitment. What is required is the possibility of causal reconstruction by third parties, capable of connecting outcomes to design choices and operational parameters. Only under these conditions can accountability address systems whose authority is exercised through execution rather than declaration, and whose opacity is often defended by appeals to complexity or inevitability (Nissenbaum, 2004; Hildebrandt, 2015).

The practical consequence of this analysis points toward institutional design. If a law that speaks no longer suffices in the face of systems that act, constitutionalism must be translated into verifiable properties of design and operation. Limits must take material form as guarantees that reintroduce understanding, restore a human interval, and expose the authority of defaults to scrutiny. The objective is not to obstruct innovation, but to preserve the space of judgment and to establish criteria capable of distinguishing when a system governs through optimization and when it remains a tool operating under conditions of meaningful human control.

C. PHILOSOPHY OF SURVEILLANCE. FROM THE PANOPTICON TO THE DATAPOINT

Surveillance no longer presents itself through a single, identifiable figure. The clear geometry of the central tower has been replaced by the diffuse capture of traces, correlations, and signals dispersed across everyday activity. A unique vantage point has given way to the aggregation of heterogeneous data that reconstruct shifting profiles and enable governance based on probability rather than direct observation. The aim is no longer to see an individual, but to anticipate trajectories and modulate contexts. Discipline yields to statistical control, and control gradually gives way to prediction, whose authority appears natural insofar as it proves effective. Calculation displaces the gaze, and technical execution takes the place once occupied by commands articulated in the name of law (Foucault, 1975; Deleuze, 1990; Zuboff, 2019).

This transformation is well captured by forms of algorithmic governance that operate without representing subjects as such. They act upon traces and signals that orient conduct in advance, often without requiring explicit intervention. Consent is reduced to a gesture of acceptance that rarely secures effective understanding, while the politics of agreement

is displaced by the mechanics of the click. One of the most significant consequences for a critical theory of power lies in this mutation. Coercion becomes optimization, and personalization absorbs conflict by translating it into the satisfaction of preferences that the system itself has progressively shaped. Under these conditions, the distinction between desire and design becomes increasingly fragile, as choice often amounts to confirming a prediction already embedded in the environment (Rouvroy & Berns, 2013; Morozov, 2013).

Such a regime depends on a specific network architecture that concentrates power in platforms capable of fixing standards of interaction, structuring regimes of visibility, and ordering relevance. These arrangements determine what appears and what remains invisible in the digital public sphere. The resulting economy of nodes, links, and flows produces normative effects without formal declaration. Access becomes a privilege, the order of appearance functions as a form of practical truth, and the opacity of ranking systems consolidates authority that rarely encounters democratic scrutiny (Castells, 1996).

Within this context, privacy can no longer be understood as a protected enclosure defined by exclusion alone. It is more accurately approached as contextual integrity shaped by expectations concerning circulation, purpose, and scale. In environments that recombine traces to generate inferences about identity, behavior, or risk, harm does not arise solely from disclosure. It emerges when data are displaced across contexts, altering meaning and producing tangible effects on opportunities, access, and differential treatment. This shift requires that traditional guarantees be rearticulated as material conditions governing the use, redistribution, and auditability of models and signals (Nissenbaum, 2004).

The movement from surveillance to governance by design becomes especially visible in the proliferation of defaults, gentle pushes, and algorithmic nudges that regulate conduct through architecture. As platforms adjust interfaces, calibrate frictions, and segment populations through continuous learning, the nudge loses its general character and takes on increasingly individualized forms. Authority settles into experienced utility and the smoothness of the path offered. Default values begin to operate as an environmental jurisdiction that orients decisions without explicit mandate and without acknowledging the normative force they exercise in practice (Yeung, 2017; Thaler and Sunstein, 2008).

From this perspective, contemporary surveillance cannot be reduced to an extension of security practices. It occupies a central position in the operational core of an automated social order. The datapoint enables a fine-grained engineering of conduct that intensifies the need for practical intelligibility and causal traceability. Without the possibility of reconstructing data routes and signal weights after the fact, the authority of the environment becomes effectively unchallengeable and the black box undermines accountability. Claims of intelligibility acquire substance only when they are supported by methods and auditable archives that allow third parties to reproduce decisions and verify limits. Absent these conditions, transparency remains largely rhetorical (Kroll et al., 2017; Edwards and Veale, 2017; Ananny and Crawford, 2018).

Operational constitutionalism responds to this diagnosis by linking critique to verifiable remedies. Where surveillance converges with personalization and optimization replaces coercion, material guarantees become necessary to restore a human interval and meaningful control. Deliberative latency, fair friction, and the recognition of default jurisdiction operate as mechanisms through which the authority of default parameters is acknowledged and subjected to prior explanation, symmetric alternatives, and archival support for independent audit. Through these instruments, law can engage with execution without relinquishing its limiting function, preserving a space in which the datapoint does not foreclose public deliberation but remains exposed to thresholds and tests capable of rendering both critical claims and efficiency promises open to verification (Lessig, 1999; Hildebrandt, 2015; Winner, 1980).

D. OPERATIONAL CONSTITUTIONALISM. MATERIAL GUARANTEES AND CONTROL THRESHOLDS

If contemporary governance operates through design, a constitutionalism adequate to the problem cannot remain at the level of principle alone. Its effectiveness depends on the capacity to translate normative commitments into properties of systems that actually operate. Legitimacy is no longer settled in declarative form but in the way a system functions over time. For this reason, rights can no longer rely on statements of intent. They must be anchored in technical conditions that can be measured, audited, and challenged without resting on the goodwill of operators or on the opacity of providers. Legal limits recover practical force precisely where the smoothness of the environment

tends to blur the distinction between utility and authority, and between functional success and normative justification (Lessig, 1999; Winner, 1980; Hildebrandt, 2015).

Material guarantees refer to those minimal conditions of design and operation that restore a human interval and preserve meaningful control without freezing innovation. They do not take the form of abstract principles but of practices that can be demanded both as obligations of means and as obligations of result. Practical intelligibility requires explanations that allow reconstruction of how outcomes were produced, rather than narratives that merely reassure. Traceability depends on documented data routes, identifiable model versions, and recorded decisions preserved under conditions suitable for independent expert review. Reversibility presupposes the availability of corrective paths whose costs remain bounded for both the affected person and the operator. Proportionality acquires a dynamic character, since the acceptable degree of automation must adjust to risk and rights impact, and must allow for safe degradation when uncertainty increases. The availability of a symmetric alternative path ensures that individuals are not locked into a single solution when normatively equivalent options exist that achieve the same purpose with lower intrusion or stronger forms of ex post social control (Edwards and Veale, 2017; Kroll et al., 2017; Nissenbaum, 2004).

Within this framework, the notion of default jurisdiction plays a central role. Default settings govern in practice and therefore cannot be treated as neutral technical conveniences. Their activation requires reinforced justification, including documented motivations, evidence of non-discrimination testing across relevant populations, and the availability of alternative paths that do not operate through punitive frictions. Where defaults concentrate operational power, control cannot be reduced to ritualized consent or to privacy policies that resist comprehension. It requires thresholds for activation, stress testing under adverse conditions, and audit mechanisms that allow third parties to reproduce outcomes and identify biases or calibration failures that routine operation may conceal. This is especially relevant where cumulative effects arise from the interaction of multiple components under changing conditions, complicating the attribution of responsibility (Yeung, 2017; Pasquale, 2015).

Meaningful control can therefore be understood as the effective capacity to intervene at different moments of operation with tools proportionate to risk. Before deployment, this involves setting design and testing requirements supported by acceptability metrics that can be falsified rather than by declarations of compliance. During operation, it requires

the possibility of pauses, escalations, and reasonable disconnections enabled by deliberative latency and fair friction. After the fact, it depends on access to repair, full review of the technical record, and institutional learning that leads to the revision of parameters and procedures when harm or unacceptable risk is established. Absent this articulation, governance by optimization tends to present itself as an irreversible trajectory rather than as a set of choices open to public justification and limitation (Hildebrandt, 2015; Rouvroy & Berns, 2013).

To make this framework enforceable, operational thresholds are introduced to relate levels of risk to degrees of automation. In contexts where impact is limited and reversibility is trivial, full automation may remain acceptable under periodic audit and continuous monitoring. Where impact increases, automation must be accompanied by intensified human involvement, genuinely symmetric alternatives, and robustness testing that includes subpopulation coverage before any deployment at scale. In high impact contexts, automation must be constrained or suspended until material guarantees are capable of absorbing residual risk without transferring costs to affected individuals or vulnerable third parties. In such cases, safe degradation and default shutdown take precedence over uninterrupted service whenever the system loses calibration or operates beyond its domain of validity (Edwards and Veale, 2017; Kroll et al., 2017).

Verification ceases to function as a bureaucratic afterthought when it is expressed through metrics that independent authorities can calculate without privileged access. Indicators such as effective human intervention rates distinguish nominal supervision from actual capacity to alter outcomes. Time to decoupling measures how rapidly a safe shutdown can be achieved once a serious deviation is detected. The coverage of usable traces indicates how many decisions can be reconstructed without gaps. Discrepancies across subpopulations reveal asymmetries in error or treatment that averages conceal. Measures of fair friction assess whether alternative paths operate as genuine options rather than deterrents. Records of drift document when and how system behavior changes as data or environmental conditions evolve. These elements must be made available in standardized formats and accessible repositories so that transparency does not collapse into promotional claims or symbolic portals devoid of auditable substance (Lessig, 1999; Winner, 1980).

Operational constitutionalism is not exhausted by techniques of control. It also implies a redistribution of capacities within institutional arrangements. Functions of design,

operation, and supervision must be separated, auditors must enjoy material independence, and affected groups must have avenues for informed participation where impacts are structural. Incentive structures should reward designs that incorporate limits as a constitutive feature rather than as an external constraint. Particular attention is required to prevent regulatory capture through standards shaped by incumbents that entrench dominant positions under the appearance of security or unavoidable efficiency, a pattern well documented in earlier technological cycles and now reappearing within algorithmic and platform-based ecosystems (Pasquale, 2015; Zuboff, 2019).

This section constitutes the normative core of the article and prepares the transition from conceptual articulation to methodological implementation. What follows develops a framework for independent verification and lifecycle governance that connects philosophical analysis, legal reasoning, and the practical oversight of systems in production. The aim is to ensure that limits do not remain symbolic affirmations but operate as concrete constraints capable of orienting technical practice while preserving the space of judgment and public deliberation.

E. INDEPENDENT VERIFICATION METHOD AND SYSTEM LIFECYCLE

Moving from principles to operational properties requires a method that can be applied by third parties without resting on confidence in the operator. Verification is therefore treated as a practice embedded in the system lifecycle and not as a final checkpoint. Its role is to make claims of intelligibility, traceability, reversibility, and proportionality open to refutation, and to allow causal reconstruction when outcomes are disputed. The underlying assumption is that technical authority must be exposed to scrutiny in a manner comparable to how legal reasoning tests the justification of decisions, by relating levels of risk, degrees of automation, and the remedies effectively available to those affected (Kroll et al., 2017; Edwards and Veale, 2017).

Guided counterfactual auditing focuses on governance parameters rather than on data alone. Default configurations, friction points, and latency windows are adjusted in a controlled way in order to explore alternative trajectories and to observe how design choices influence conduct. The emphasis lies on identifying situations in which architecture channels action while preserving the appearance of choice. The aim is not to

reconstruct the internal logic of a model in its entirety, but to assess the practical force of environmental arrangements that orient decisions and that often remain invisible to conventional statistical evaluation (Kroll et al., 2017; Ananny and Crawford, 2018).

In practice, this approach makes it possible to observe how frequently users remain on preconfigured paths when leaving them involves effort, delay, or loss. Comparing behavior under baseline conditions with behavior after modest changes in defaults or timing provides evidence of the extent to which outcomes are shaped by design itself, rather than by deliberate individual preference. These observations do not depend on exceptional scenarios. They emerge in ordinary use, precisely where governance tends to present itself as neutral convenience.

This perspective is complemented by sociotechnical red teaming. Here the system is deliberately exposed to strain in order to reveal points of fragility. Situations involving error, overload, changing conditions, or malfunctioning signals are explored, with particular attention to cases and populations that routine testing tends to marginalize. What matters is not only the detection of bias, but the identification of moments in which utility quietly turns into authority and personalization reduces the practical space for disagreement by raising the costs of deviation for both users and operators (Edwards and Veale, 2017; Ananny and Crawford, 2018).

Ex post reconstruction depends on the existence of a minimal auditable archive. Model and policy versions, relevant variable states, data routes, random seeds, and log custody must be preserved in forms suitable for external review. This allows a third party to reproduce a contested outcome and to explore nearby alternatives without relying on privileged access or on abstract explanations. Such documentary discipline is not ancillary. It is the condition under which explainability acquires operational meaning and supports review, correction, and repair within a reasonable time frame (Kroll et al., 2017; Edwards and Veale, 2017).

Within this framework, metrics serve as instruments for testing rather than as claims of precision. Default adherence captures how often decisions follow preset options when deviation carries nontrivial costs, offering an indication of architectural channeling. The reversibility gradient reflects the concrete effort required to undo an outcome in terms of time, resources, and associated losses, and functions as a trigger for stronger remedies. Behavioral entropy reduction signals a progressive narrowing of acceptable trajectories

without corresponding gains in legitimate objectives. Effective opacity points to the distance between declared system behavior and what can later be reconstructed through available records, orienting demands for pragmatic transparency rather than symbolic disclosure (Wachter et al., 2017; Ananny and Crawford, 2018; Winner, 1980).

Control is activated progressively rather than uniformly. When minimal intelligibility is absent, any presumption of proportionality loses credibility. Where reversibility cannot be exercised in practice, automation is reduced to non-binding support and decision-making returns to human judgment under reinforced archival conditions. Sustained levels of default adherence beyond agreed ranges support an inference of undue channeling and activate duties of prior explanation, genuinely symmetric alternatives, and traceability sufficient for independent audit. The objective is not to obstruct innovation, but to prevent optimization from displacing rights while presenting itself as efficiency (Kroll et al., 2017; Edwards and Veale, 2017).

These elements are distributed across the system lifecycle. Before deployment, design and testing requirements are defined through acceptability criteria that admit falsification and include attention to subpopulation effects and degradation scenarios. During operation, pauses, escalations, and disconnections remain possible through deliberative latency and fair friction, avoiding automatic continuity in situations of uncertainty. After operation, effective repair, full access to the technical record, and institutional learning allow parameters and procedures to be revised when harm or unacceptable risk is identified. Obligations are calibrated in relation to both risk and automation intensity, in line with the premises of operational constitutionalism.

The method becomes visible in domains where operational power exerts constitutive effects. In systems governing access to social benefits, minor threshold adjustments can suspend entitlements automatically and shift evidentiary burdens onto individuals, making reversibility and opacity central concerns. In ranking and visibility systems, relevance modulation combined with behavioral signals operates as a form of default jurisdiction, justifying advance explanation and audit-oriented traceability. In assisted clinical triage, deliberative latency preserves human review in borderline cases and requires documentation of overrides grounded in verifiable criteria. In consumer credit contexts with preloaded offers, sustained default adherence reveals the dominance of design choices and supports stronger demands for proportionality and effective understanding (Nissenbaum, 2004; Wachter et al., 2017; Pasquale, 2015).

Scope and limits are determined through necessity and proportionality. Latency adapts to urgency without excluding later review. Fair friction does not authorize artificial obstacles or procedural excess. Default jurisdiction is acknowledged only where evidence indicates widespread channeling, asymmetric exits, and the absence of genuinely equivalent alternatives. Clarity regarding these conditions protects the framework from accusations of inefficiency and preserves room for responsible experimentation, while keeping open the space of judgment even in settings characterized by high levels of automation (Thaler and Sunstein, 2008; Yeung, 2017; Mittelstadt et al., 2016).

F. POLITICAL ECONOMY OF OPERATIONAL POWER. INCENTIVES AND DESIGN CAPTURE

Operational power does not emerge spontaneously. It grows within incentive structures that reward fluency even when this fluency compresses the interval of judgment and stabilizes dominant positions through network effects, switching costs, and accumulated learning. Over time, the experienced usefulness of the system aligns with the interests of the operator, and authority becomes credible because the environment works, not because it can be justified in public terms. Competition is gradually displaced toward control over design choices, while default configurations acquire strategic value by governing in practice without presenting themselves as acts of authority (Zuboff, 2019; Pasquale, 2015).

The accumulation of data and models intensifies this dynamic. Each interaction reinforces informational asymmetries and increases dependence on the interface that organizes adaptation to a specific user and context. Personalization, initially perceived as convenience, becomes a barrier to exit as soon as leaving implies loss of familiarity, performance, or relevance. Under these conditions, portability risks remaining a formal entitlement unless it is supported by technical arrangements that allow experience to be transferred with limited degradation in context and functionality (Stiegler, 2016; Nissenbaum, 2004).

Attention and relevance markets further amplify these tendencies. Value concentrates where signals, ranking, and distribution are vertically integrated, encouraging design strategies that favor continuity of use and discourage deviation. Default adherence

becomes desirable because uninterrupted flow enhances platform value and consolidates intermediary positions as de facto standards. In this setting, operators can externalize compliance and audit costs onto users and producers without facing a setting of symmetric negotiation. Similar dynamics have appeared in earlier technological cycles, but here they are intensified by opacity and by the speed at which systems recalibrate themselves (Winner, 1980; Castells, 1996).

Narratives of inevitable innovation provide cultural cover for this political economy. They shift attention away from questions of limits and toward promises of efficiency, framing any form of constraint as an obstacle to competition even when its purpose is to preserve the space of judgment. Interoperability standards often reflect this imbalance. Drafted and timed by those already in a position to shape adoption, they present themselves as neutral coordination tools while entrenching incumbent practices and raising the cost of alternative designs that incorporate fair friction or deliberative latency (Pasquale, 2015; Zuboff, 2019).

A competition policy articulated in operational terms therefore cannot rely exclusively on ex post sanctions or delayed structural remedies. Intervention must occur where authority becomes embedded in design. The notion of default jurisdiction functions here as a regulatory lever. It recognizes that default values govern in fact and subjects their activation to reinforced justification, testing for discriminatory effects across relevant populations, the availability of genuinely symmetric alternatives, and the preservation of auditable records that allow reconstruction of how parameters were selected and on what grounds. Without this recognition, control tends to dissolve into ritualized consent and expansive transparency narratives that leave operational arrangements untouched (Yeung, 2017; Edwards and Veale, 2017).

Portability with practical equivalence extends this logic. Transferring data without transferring context and expected behavior amounts to a symbolic right of exit. A meaningful possibility of departure requires migration mechanisms capable of reconstituting profiles and preferences with an acceptable level of performance in the receiving environment, together with minimal obligations of technical cooperation among operators. Invocations of trade secrecy cannot operate as absolute barriers in this respect unless alternative forms of independent verification are made available. Achieving this balance between incentives to invest and social control over shared infrastructures

demands institutional precision and criteria that admit empirical testing (Kroll et al., 2017; Wachter et al., 2017).

Acquisitions of startups and the vertical integration of critical layers of the stack operate as accelerators of operational power by concentrating data, models, and distribution within a single governance structure. This concentration increases the risk that experimentation presented as provisional becomes a permanent exception. Thresholds are therefore required to identify situations in which assisted automation gives way to governance by optimization. Remedies in such cases include functional separation between design, operation, and standard setting, reasonable access to interfaces and auditable logs, and limits on self-preferencing in ranking and recommendation systems that shape visibility in the digital public sphere (Pasquale, 2015; Castells, 1996).

The framework of operational constitutionalism offers orientation without suppressing entrepreneurship or research activity. It neither condemns automation as such nor fixes design in static form. Instead, it requires intelligibility that can be exercised in practice, traceability sufficient for reconstruction, reversibility that does not impose excessive costs, and proportionality that adjusts to risk and impact. The effectiveness of remedies is subjected to evidence through metrics accessible to independent scrutiny. This allows regulatory intervention to operate with precision, distinguishing contexts where technology assists decisions from those where it acquires constitutive authority and therefore warrants substantive limits in everyday operation (Hildebrandt, 2015; Rouvroy & Berns, 2013).

What emerges is a political economy that makes visible the points at which utility masks authority and apparent competition coexists with the quiet capture of design. It supports a pragmatic agenda for material control of technical environments without denying their creative potential or their contribution to social welfare. The aim is not to arrest innovation, but to reopen the space of judgment within systems that tend to close it, combining verifiable limits with incentives for designs that treat limits as a condition of legitimacy rather than as an external burden, so that innovation proceeds with rights inside the system rather than outside it (Winner, 1980; Lessig, 1999).

G. OBJECTIONS, LIMITS, AND SECOND ORDER RISKS

Any attempt to translate constitutional limits into operational terms inevitably attracts objections. This is not a weakness of the proposal, but a condition of its relevance. A framework that aspires to guide real systems must remain exposed to critique, revision, and failure, rather than insulating itself behind conceptual elegance.

One recurrent concern points to cost. Translating rights into verifiable properties is said to increase design burdens and to discourage innovation. The response does not lie in denying this tension, but in specifying where it matters. Dynamic proportionality and falsifiable metrics allow obligations to be adjusted to risk and impact, concentrating regulatory effort where automation replaces judgment rather than where it merely assists it. This calibration keeps limits open to revision as evidence accumulates and avoids the familiar pattern in which efficiency rhetoric displaces safeguards that later prove far more expensive to reconstruct after harm has materialized, especially for groups already exposed to informational asymmetries (Mittelstadt et al., 2016; Kroll et al., 2017; Edwards and Veale, 2017).

Another line of criticism questions the very possibility of explanation in complex systems. Demands for explainability are portrayed as naïve, or as inevitably degrading performance. The standard advanced here does not require exhaustive disclosure or universal pedagogy. It asks for practical intelligibility in relation to contested outcomes and for the preservation of records that enable causal reconstruction by third parties. Even when parts of a model remain legitimately opaque for reasons of security or proprietary protection, the availability of traces, versions, and states makes accountability something more than a promise. Without such material support, appeals to complexity tend to operate as cultural justifications rather than as genuine technical limits (Wachter et al., 2017; Edwards and Veale, 2017; Ananny and Crawford, 2018).

Claims of trade secrecy and intellectual property raise a related objection. Audits are resisted on the ground that they threaten investment and innovation. The operational approach does not dismiss these interests, but it refuses to treat secrecy as an absolute shield. Where design decisions acquire public relevance and shape access, visibility, or rights, accountability cannot rest on trust alone. Verification by independent parties under confidentiality obligations allows protection of assets without turning opacity into immunity, acknowledging that technical artifacts embed choices whose distributive

consequences rarely remain neutral under closer examination (Pasquale, 2015; Winner, 1980).

Concerns about paternalism often follow. Measures such as fair friction, deliberative latency, or default jurisdiction are accused of interfering with individual autonomy. This objection overlooks how autonomy is already shaped by architecture. The purpose of these interventions is not to guide choices toward preferred outcomes, but to prevent the elimination of understanding through the systematic reduction of friction and to make the normative force of defaults visible. The notion of contextual integrity helps clarify why harm arises when data and decisions shift context without notice, and why restoring control need not imply bureaucratic overload or pointless detours. What is preserved is freedom as an informed practice, not as a sequence of clicks that merely ratifies predictions prepared in advance (Thaler and Sunstein, 2008; Yeung, 2017; Nissenbaum, 2004).

Skepticism also targets the role of the human in the loop. Experience shows how easily human oversight becomes symbolic, leaving substantive authority untouched. The control standard responds by tying automation to conditions that can be withdrawn. When intelligibility or reversibility cannot be exercised in practice, automation is reduced to nonbinding support. The right to latency is not limited to delay. It enables escalation, interruption, and reinforced documentation that anchors responsibility in identifiable decisions. Measurement of effective intervention replaces nominal supervision, restoring weight to practical judgment in environments where speed and scale would otherwise hollow it out (Hildebrandt, 2015; Kroll et al., 2017).

Beyond these familiar objections, the framework must also confront second order risks generated by its own instruments. Metrics invite optimization. Once indicators acquire relevance, they become targets, and Goodhart effects follow. Addressing this risk requires treating measurement itself as an object of governance. Rotation of tests, randomized audits, and sociotechnical red teaming help reveal evasive strategies, while cross domain validation and monitoring of drift reduce the likelihood that localized improvements conceal broader degradation. Oversight of metrics cannot be an afterthought, because incentives emerge as soon as numbers begin to matter (Barocas et al., 2019; Ananny and Crawford, 2018).

A final concern relates to institutional capacity. The framework may appear demanding when measured against existing resources and expertise. For this reason, it does not assume immediate completeness. It emphasizes prioritization by risk, gradual accumulation of public competence, and minimal standardization of archives and formats that lower the barrier to independent verification. Its logic is incremental rather than maximalist. What it rejects is innovation by inertia. Functioning alone cannot justify the abandonment of judgment when technical environments acquire de facto authority over collective conduct and shape outcomes beyond individual negotiation (Winner, 1980; Lessig, 1999).

H. INSTITUTIONAL IMPLEMENTATION. AN OPERATIONAL STANDARD FOR PUBLIC AND PRIVATE PRACTICE

Bringing the framework into practice requires a shift in how technology policy is usually conceived. The focus moves away from abstract commitments and toward verifiable conditions of design and operation that can be tested in real settings. This reorientation is demanding, not because it multiplies principles, but because it insists that the authority exercised by technical environments be exposed to scrutiny in a manner comparable to legal reasoning. Practical intelligibility, traceability, reversibility with bounded costs, and proportionality adjusted to risk and impact must therefore appear in contracts, procurement processes, and internal governance arrangements. What emerges is not a catalogue of best practices, but an operational standard that can be examined by third parties without reliance on trust in the operator or discretionary access granted by the provider.

Public and corporate procurement provides a particularly effective point of entry. Demand can require the preservation of minimal auditable archives, including model versions, data routes, seeds, and relevant system states, together with activation thresholds for safe degradation or shutdown when uncertainty increases. The purpose is not unrestricted disclosure, but pragmatic transparency that allows causal reconstruction and independent verification under conditions of confidentiality. Trade secrecy continues to protect investment, yet it no longer operates as a blanket exemption when technical artifacts determine outcomes of public relevance. When these requirements are embedded in

tender documents and contractual clauses, operational constitutionalism becomes a condition of participation rather than an aspirational add on.

Institutional arrangements must also address how responsibility is distributed. Separating the functions of system construction, operation, and supervision reduces the concentration of interpretive power and limits the risk that human oversight becomes merely symbolic. The material independence of auditors and the informed involvement of affected groups in cases of structural impact reinforce this separation. Activation thresholds that downgrade automation to nonbinding support when intelligibility or reversibility cannot be exercised in practice give substance to these arrangements. Oversight acquires practical meaning only when pauses, escalations, and the capacity to alter operational trajectories are available without hidden penalties.

Platform governance and data intensive services raise additional challenges. Here the recognition of default jurisdiction becomes central, because default configurations govern conduct in fact. Their activation calls for reinforced justification, testing for discriminatory effects across relevant populations, the availability of genuinely symmetric alternatives, and records that permit reconstruction of how parameters were selected and on what grounds. Particular care is required where multiple components interact and cumulative effects make attribution of responsibility difficult. Without this recognition, control tends to dissolve into ritualized consent that simulates choice while leaving underlying structures untouched.

Preventing capture of design by incumbents requires a combination of safeguards rather than a single remedy. Publicly accessible metrics that admit falsification make it possible for independent actors to assess default adherence, reversibility, behavioral narrowing, and opacity. Rotation of tests and randomized audits reduce the risk that indicators become targets. Temporary sandboxes remain useful only when they include clear exit conditions and predefined criteria for generalization. Interoperability must also be treated as a practical requirement rather than as a purely formal one, preserving context and acceptable performance upon migration so that personalization does not harden into an exit barrier.

Procedural law can support this transition by adjusting evidentiary rules to reflect real asymmetries. Persistently high levels of default adherence may justify presumptions of undue channeling and trigger duties of advance explanation and alternative provision.

Conversely, the absence of a minimal auditable archive undermines claims of proportionality and shifts the burden toward the operator to demonstrate that reversibility was not effectively nullified. These adjustments do not create new rights in the abstract, but reshape incentives in favor of designs that treat limits as part of their internal logic.

Competition policy articulated in operational terms completes the picture. Remedies focus on points where authority crystallizes, including functional separation between design, operation, and standard setting, constraints on self-preferencing in ranking and relevance systems, and reasonable access to interfaces and auditable logs for competitors and researchers. Acquisitions that integrate critical layers of the stack warrant particular scrutiny when they increase the risk of ecosystem closure. The concern is not size as such, but the gradual alignment of experienced utility with the interests of a single operator until default configurations function as a form of private jurisdiction over the digital public sphere.

This operational standard does not immobilize innovation. It gives it direction and responsibility. By reconnecting functioning with justification and utility with limits, it reopens the space of judgment within highly automated processes and offers a gradual path for regulators, operators, and courts. The immediate task is empirical. Thresholds and remedies must be tested across domains such as social benefits, ranking and moderation, clinical triage, and consumer credit, comparing reversal costs, availability of usable traces, and decoupling times. In this way, the theory of operational power remains exposed to evidence and correction, and technical creativity is preserved not as an escape from limits, but as a means to design systems in which freedom retains practical meaning.

I. CONCLUSIONS AND EMPIRICAL AGENDA

The analysis has shown that operational power cannot be understood as a secondary effect of informatization. It names a form of authority that governs by design and reorganizes conflict by shifting it away from explicit norms toward the architecture of action itself. In this configuration, legitimacy no longer derives primarily from command or formal obligation, but from experienced utility and functional fluency. When law remains confined to speaking while systems act, its capacity to limit power erodes, particularly where the interval of judgment is absorbed by continuous technical flows that organize options, rhythms, and exit paths according to profiles learned in real time.

The distinction between assisted automation and governance by optimization is central to this diagnosis. Only the latter acquires constitutive force, because it combines prior parametrization of possible action, continuous behavioral feedback, and legitimization grounded in functioning rather than justification. Where these elements converge, human control risks becoming nominal unless material limits intervene. Practical intelligibility, traceability, bounded cost reversibility, and dynamic proportionality are therefore not auxiliary safeguards, but conditions for reopening the space of judgment when uncertainty grows or when systems operate beyond their domain of validity. Without such degradation mechanisms, accountability retreats behind complexity and the black box functions as a barrier rather than as an object of scrutiny.

Operational constitutionalism responds to this displacement by connecting legal theory and the philosophy of power with methods of verification that admit falsification. It does not propose best practices insulated from dispute, but measurable conditions that external authorities can assess without relying on discretionary access. Categories such as default jurisdiction, deliberative latency, and fair friction become decisive because they determine whether intervention is possible before, during, and after operation, and whether contested outcomes can be causally reconstructed with evidence rather than narrative reassurance. This is particularly relevant in environments where ranking and visibility decide access to opportunities at scale, and where order of appearance functions as a practical criterion of truth without public deliberation.

The empirical agenda follows directly from this framework. It requires moving from general diagnosis to field-based measurement. Default adherence must be observed in sensitive domains. Reversibility gradients must be calculated in terms of time, effort, and loss for both affected persons and operators. Effective opacity must be assessed as the distance between declared behavior and what can later be reconstructed through minimal auditable archives. Behavioral entropy reduction can indicate the narrowing of valid trajectories without corresponding gains in legitimate ends. These measurements lose meaning if they rely on averages alone, and therefore require subpopulation coverage to detect asymmetries that remain invisible at the center of distribution.

Institutional implementation complements this agenda through contractual and procedural arrangements that make independent verification possible. Procurement can require the preservation of model versions, data routes, relevant system states, and logs with chain of custody, under confidentiality regimes that protect trade secrets without

blocking reconstruction in cases of dispute. Procedural law can respond by adjusting evidentiary presumptions to reflect real asymmetries, for instance when persistently high default adherence indicates undue channeling despite the formal availability of alternatives that are practically discouraged by cost or urgency.

A comparative research program adds a transnational dimension in which interoperability and portability with practical equivalence counteract design capture. Transferring data without preserving context does not secure freedom of exit. Migration artifacts must therefore be evaluated against published performance criteria and acceptable degradation margins. Concentration across critical layers of the stack warrants particular attention where vertical integration renders decision paths affecting rights and opportunities effectively unverifiable.

The itinerary proposed here avoids both normative maximalism and uncritical reliance on technical efficiency. It recognizes zones in which automation is admissible with limited oversight, zones that require strengthened human intervention and alternative paths, and zones in which automation must be degraded or withheld until material guarantees exist. The objective is not to slow innovation, but to bind it to limits that preserve judgment as a practical capacity rather than as a symbolic reference.

Operational power does not disappear when it remains unnamed. It consolidates quietly, through defaults, continuity, and optimization that present themselves as neutral improvements. Responding to it requires time, method, and institutions capable of insisting on evidence where promises prevail. Systems that calculate extensively still require limits that can be seen, tested, and corrected. Where those limits exist, creativity is not expelled, but oriented toward designs that sustain freedom rather than merely accelerating paths whose direction no longer admits question.

REFERENCES

- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media and Society*, 20(3), 973–989.

- Baracas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning*. <https://fairmlbook.org>.
- Bobbio, N. (1991). *El futuro de la democracia*. Fondo de Cultura Económica.
- Castells, M. (1996). *The rise of the network society*. Blackwell.
- Deleuze, G. (1990). Post-scriptum sur les sociétés de contrôle. *L'Autre Journal*, 1, 3–7.
- Edwards, L., & Veale, M. (2017). Slave to the algorithm? Why a "right to an explanation" is probably not the remedy you are looking for. *Duke Law and Technology Review*, 16(1), 18–84.
- Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.
- Foucault, M. (1975). *Surveiller et punir: Naissance de la prison*. Gallimard.
- Han, B.-C. (2014). *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Herder.
- Hildebrandt, M. (2015). *Smart technologies and the end(s) of law: Novel entanglements of law and technology*. Edward Elgar.
- Kroll, J. A., Huey, J., Baracas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., and Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review*, 165(3), 633–705.
- Lessig, L. (1999). *Code and other laws of cyberspace*. Basic Books.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2), 1–21.
- Morozov, E. (2013). *To save everything, click here*. PublicAffairs.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 101–139.
- Pasquale, F. (2015). *The black box society*. Harvard University Press.
- Rouvroy, A., & Berns, T. (2013). Gouvernementalité algorithmique et perspectives d'émancipation. *Réseaux*, 177(1), 163–196.
- Simondon, G. (1958). *Du mode d'existence des objets techniques*. Aubier.

- Stiegler, B. (2016). *Dans la disruption: Comment ne pas devenir fou?* Les Liens qui Libèrent.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Virilio, P. (1995). *La vitesse de libération*. Galilée.
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76–99.
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136.
- Yeung, K. (2017). 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication and Society*, 20(1), 118–136.
- Zuboff, S. (2019). *The age of surveillance capitalism*. PublicAffairs.