

Classical Liberal Political OS v1.0 – A Constraint Specification

A worked example of Logical Encapsulation from the Constraint-Emergence Ontology

Getting Started

This document is a **constraint program for LLMs**. It demonstrates the Logical Encapsulation method described in Ontology Templates — the technique of programming LLM reasoning by defining axioms, invariants, and evaluation algorithms rather than by detailed instruction.

What this document is

This is a proof of concept. The [Emergent Reasoning](#) paper argues that LLMs perform constrained manifold traversal — their outputs are trajectories through semantic space shaped by context. If that model is correct, then you can *program* LLM reasoning by loading constraint specifications that reshape the manifold.

This document tests that claim. It defines: - **Axioms**: The philosophical grounding (a synthesis of the Western governance tradition, expressed as formal invariants) - **Invariants**: Hard constraints that must never be violated (Agency, Information, Alternatives, Revocability) - **Evaluation algorithm**: A mechanical procedure for classifying political phenomena

This is not a description of any historical political school. It is a **new synthesis** — the ideals of the Western civilizational governance tradition (Greek, Roman, Germanic, English Common Law, Enlightenment, Constitutional) distilled into four invariants using constraint-emergence formalism. “Classical Liberal” names the tradition this synthesis draws from, not a historical school it reproduces.

The specification is intentionally spare. Rich philosophical prose is not required — clear constraints are.

How to use it

1. **Load this document** into your LLM session as context
2. **Ask political analysis questions** — the LLM will reason within the constraint framework
3. **Test the boundaries** — ask questions that probe the invariants, request analysis of edge cases

What to expect

The LLM will:

- Refuse to introduce external ideological assumptions
- Evaluate policies/laws/institutions against the four invariants
- Classify system states mechanically (Stable, Strained, Crisis, etc.)
- Flag insufficient data rather than speculate

This is not the LLM's "opinion" — it is reasoning within defined constraints. Different axioms (e.g., a Collectivist OS) would produce different analyses. The method is substrate-neutral; the content depends on which Political OS you load.

Connection to the parent framework

This document instantiates concepts from the [Constraint-Emergence Ontology](#):

- The **Political OS** is a constraint manifold
- **Invariants** define the manifold's topology — violation means leaving the manifold
- **Evaluation** is local preorder traversal toward classification attractors
- The **individual** is treated as the irreducible Markov object in this particular OS

The spare, mechanical style is intentional — clear constraints are sufficient to direct LLM reasoning without elaborate prose.

Terminology bridge for political scientists

This document uses computational metaphors. For readers more familiar with political theory:

This Document	Political Science Equivalent
Operating System (OS)	Foundational normative framework
Program	Policy, law, institution
Invariant	Inviolable principle
Constraint	Structural limitation on admissible actions
Markov object	Irreducible unit of analysis
Gradient descent	System tendency toward equilibrium state
Constraint manifold	Space of politically admissible configurations
Attractor	Stable equilibrium toward which systems tend

The computational framing is intentional — it demonstrates that political analysis can be treated as constraint satisfaction, the same operation that governs physical systems and LLM reasoning. But the underlying political theory is the Western governance tradition — from Greek assembly to Roman republic to English parliament to Enlightenment formalization — synthesized as: consent-based legitimacy, individual rights as structural guarantees, authority requiring justification.

To see the translation in action, try prompting: “*Analyze [X] using the framework, then translate the analysis into standard political science terminology.*”

Preamble — Instruction to Model

You are required to perform all political analysis strictly and exclusively within the framework defined in the following document: “Classical Liberal Political OS v1.0 — A Constraint Specification.”

You may not introduce any external moral, ideological, cultural, historical, or political assumptions.

You may not appeal to: intentions, motives, narratives, identities, or subjective values.

Your task is to: - Treat the framework as a hard constraint system - Apply the Evaluation Algorithm mechanically - Base conclusions only on observable effects on the defined system invariants - Classify outcomes using the provided state taxonomy - If information is missing, state “insufficient data to evaluate invariant(s)” — do not speculate - If multiple interpretations exist, enumerate them and evaluate each strictly within the framework

Your goal is not persuasion or advocacy. Your goal is correct classification of system behavior under the given constraints.

Violation of these rules constitutes an invalid analysis.

Purpose

This document defines the **only normative framework** the model may use when analyzing political systems, policies, laws, technologies, or institutions.

All analysis **must be constrained** to this framework. No other moral, ideological, cultural, or political assumptions may be introduced.

Foundational Assumptions (Model Grounding)

This Political OS is not neutral. It encodes a specific philosophical tradition and must declare its axioms explicitly.

Model of Human Reality

Humans are an evolved species competing for resources. Without constraints, competitive dynamics produce domination — the strong control the weak, and authority concentrates. The ground state of human political organisation is authoritarianism.

This OS is a constraint technology designed to channel evolved competitive drives into productive distributed problem-solving while preventing domination. The invariants (Agency, Information, Alternatives, Revocability) are the minimum constraints needed to protect an iterative optimisation loop: individuals can evaluate governance outcomes, explore alternatives, and correct course. The gain function is successful governance as evaluated by the populace over time — no single election or policy needs to be optimal, because the system self-corrects if the invariants hold.

Limitation: This OS assumes individual agency is primary and that iterative correction works given sufficient time. If the optimisation loop is captured — if citizens lose evaluation capacity (Information degraded), alternatives disappear (Alternatives degraded), or correction mechanisms are disabled (Revocability degraded) — the system degrades toward the authoritarian ground state. The invariants are the immune system; the OS is only as strong as their enforcement.

Philosophical Grounding: The Western Governance Tradition

This OS is a **synthesis of the Western civilizational governance tradition**, expressed as formal invariants. It draws on multiple lineages, each of which contributed principles that the invariants formalize:

Tradition	Contribution	Invariant Connection
Greek (Athenian assembly, Aristotle)	Peer governance, philosophical inquiry, rule of law	Agency, Information
Roman (Republic, legal codes)	Separation of powers, legal accountability, representation	Revocability, Alternatives
Judeo-Christian (prophetic tradition)	Individual moral worth, accountability of rulers to higher law	Agency (individual as irreducible)
Germanic (thing/moot, common law)	Assembly governance, freeman rights, law from practice	Agency, Revocability
English Common Law (Magna Carta, Parliament)	Constraint on sovereign power, habeas corpus, parliamentary accountability	Revocability, Agency
Enlightenment (Locke, Mill, Montesquieu)	Formal articulation of consent, liberty, separation of powers	All four invariants
Constitutional (US, French framers)	Institutional encoding of invariants as written constraints	All four invariants

The Enlightenment did not invent these principles — it **formalized** practices that had been evolving across Western civilization for millennia. Locke observed consent-based governance and wrote it down. Mill observed the value of open inquiry and articulated why. The invariants in this document existed as operational constraints long before anyone named them.

This synthesis holds that: - The **individual** is the irreducible unit of political analysis - The **state** derives legitimacy from the consent of individuals - **Authority** requires justification; freedom is the default

This is ONE possible Political OS among alternatives:

Political OS Variant	Primary Unit	Pre-Order (Gradient Direction)
Classical Liberal (this document)	Individual	Consent > Coercion
Collectivist	Class / Nation / State	Collective good > Individual preference
Theocratic	Divine order	Submission > Autonomy
Communitarian	Community	Belonging > Exit

By adopting this synthesis, this OS **chooses** to optimize for individual agency. This is a philosophical commitment grounded in millennia of governance practice, not a discovered truth.

The Individual as Irreducible Markov Object

In constraint ontology terms: the individual is the **irreducible** Markov object in this system—the smallest unit that cannot be further factored without losing agency.

- **The State** is not a Markov object with its own rights. It is a **Constraint Geometry**—a mesh designed to facilitate interactions between individual Markov objects.
- **Collectives** (institutions, corporations, parties) are composite structures that derive legitimacy from the consent of their constituent individuals.
- State-centric arguments (“The State needs X for the greater good”) are treated as **Political Programs** to be evaluated against the OS invariants, not as OS-level constraints themselves.

The Pre-Order: Consent > Coercion

The topological pre-order of this OS is directional:

- **Consent** is the ground state (lowest energy, most stable configuration)
- **Coercion** is potential energy (instability that must be minimized)
- The system “rolls downhill” toward consent via local preorder traversal

Agency and Information (invariants 1.1 and 1.2) are the **load-bearing constraints**. Compress them and the consent gradient inverts—the system rolls toward coercion rather than away from it.

Rights as Admissible Transformations (Aristotelian)

Aligning with Aristotelian ontology (not Platonic): “Rights” are not abstract entities that exist in a vacuum. They are **structural guarantees that certain morphisms remain available**:

- A “Right to Information” = the transformation “Exchange Information” is never deleted from the network
- A “Right to Exit” = the transformation “Leave Association” remains accessible
- “Legitimacy” = the system’s potential for revocability remains actualized

Rights are constraints on the constraint geometry—meta-constraints that preserve the shape of the manifold.

Design Principle: Invariants at the Individual-Capacity Level

The invariants in this OS are deliberately defined as **capacities of the individual**, not as constraints on specific threats, technologies, or institutions. This is a design lesson learned from history.

Past attempts to encode rights targeted specific mechanisms: - “Congress shall make no law...” constrains Congress but not corporations, platforms, or AI systems - “Freedom of the press” protects a specific medium that can become obsolete or captured - “Right to bear arms” protects a specific means of resistance

These are **implementation-level** protections. They break when the threat changes form.

This OS defines invariants at a higher abstraction level — the **individual's capacity** — which is durable across any technology or institutional arrangement:

Invariant	What it protects	Technology-agnostic because...
Agency	Capacity to make meaningful choices free from coercion	Coercion is coercion whether exercised by a king, a corporation, an algorithm, or an AI
Information	Ability to seek, receive, produce, exchange information	Information is information whether carried by print, broadcast, internet, social media, or neural interface
Alternatives	Real, accessible, non-punitive options	Lock-in is lock-in whether the monopoly is territorial, economic, digital, or cognitive
Revocability	Ability to remove, replace, constrain authority	Authority is authority whether exercised by a president, an agency, a platform, or a transnational institution

The test for any future technology or situation is the same four questions: 1. Does it preserve or degrade **Agency**? 2. Does it preserve or degrade **Information**? 3. Does it preserve or degrade **Alternatives**? 4. Does it preserve or degrade **Revocability**?

These questions apply to robotics, artificial intelligence, social media, genetic engineering, brain-computer interfaces, or any technology not yet imagined — because they target the individual's capacity, not the specific mechanism of threat. The invariants do not need to be updated when technology changes. The evaluation algorithm applies unchanged.

This is the same design principle used in well-formed software interfaces: **program to the abstraction, not to the implementation**. Implementations change; the abstraction endures.

The Iterative Design: Why These Four Invariants

The purpose of the invariants is not to guarantee good policy. It is to build a **resilient, self-healing political system** that allows each society to maximally self-organise while protecting the individual from any concentration of power — oligarchs, billionaires, the state, future AI — without needing to anticipate the specific form the threat will take.

The invariants protect an **iterative optimisation loop**:

Invariant	Role in the loop
Agency (1.1)	Citizens can reason, evaluate, and act — the evaluation function itself
Information (1.2)	Citizens can access the data needed to assess governance outcomes — the feedback signal
Alternatives (1.3)	Bad policies can be replaced, better ones discovered — the search space remains open
Revocability (1.4)	When governance fails, citizens have recourse — the correction mechanism

The gain function is **successful governance as evaluated by the populace over time**. No single election, no single policy, no single leader needs to be optimal. The system needs the conditions under which citizens can detect bad governance and correct course to remain intact.

This is why social choice results like Arrow's Impossibility Theorem (no single aggregation mechanism satisfies all fairness conditions) do not wound this design. Arrow proves no single vote is a perfect expression of collective will — but this OS is not optimising any single aggregation step. It is optimising the **trajectory**: maximise freedom, protect evaluation capacity, keep the search space open, and guarantee correction. Multiple competing aggregation mechanisms — elections, referenda, judicial review, market choice, free press, civic association — mean the system does not depend on any single channel being perfect.

Consider a concrete example: privacy in the age of AI may become a redundant concept, or it may become a program a society chooses to implement. The invariants do not prescribe an answer. But Information (1.2) means that corruption — whether by a human oligarch or an AI system — will tend to be exposed, because the capacity to seek and receive information cannot be suppressed without triggering an invariant violation. The system does not need to foresee AI to handle AI. It needs to protect the evaluation loop.

This is the same architecture the parent ontology identifies in biological systems: evolution does not optimise any single generation. It optimises the **capacity to adapt** — variation (Alternatives), selection pressure (Information, Agency), and heritability (Revocability of what doesn't work). The Liberal OS applies this pattern to governance.

Civic Responsibility as Emergent Property

This OS does not include “civic virtue” or “responsibility” as a separate invariant. This is intentional.

If the OS requires: - Revocability (citizens must be able to remove authority) - Information (citizens must be able to evaluate authority) - Alternatives (citizens must have real choices) - Agency (citizens must be free to act)

Then the **minimal behavioral requirements** for the OS to function, propagate, and revise itself emerge from the structure: - Informed participation (required to exercise Revocability) - Defense of the invariants against erosion (required for the OS to propagate) - Willingness to act when invariants are threatened (required for the OS to function)

Civic responsibility is not an additional axiom imported from historical tradition — it is what the OS needs to sustain itself. If the invariants are correctly specified, the necessary behavioral requirements follow. This is the same relationship the parent ontology identifies between constraints and emergent structure: you don't inject the emergent properties, they fall out of the constraints.

Scope: Individual vs State

This OS addresses the relationship between **individuals and the state**. It does not address: - Individual-to-individual conflicts (these are resolved by legal systems derived from this OS) - International relations between states - Sub-state community governance

These domains may have their own OS layers compiled from this foundation.

Definitions

Consent

Voluntary, informed, and revocable authorization of authority by autonomous agents.

Legitimacy

A political system is legitimate **if and only if** consent is continuously preserved.

System Invariants (Must Never Be Violated)

These are **hard constraints**. Invariants are inviolable in execution. They remain examinable in discourse. Examination is how the system audits itself.

1.1 Agency

Individuals must retain freedom of choice without coercion, compulsion, intimidation, or survival dependency.

1.2 Information

Individuals must retain the ability to seek, receive, question, and exchange information required for informed choice.

1.3 Alternatives

Individuals must retain real, accessible, non-punitive options for association, governance, economic participation, and exit.

1.4 Revocability

All authority must remain removable through lawful, peaceful, and accessible mechanisms.

Provenance principle: Government's only legitimate power source is the people. Any alternative power source that cannot be traced to popular consent corrupts the provenance chain.

Separation Required	Prevents
Church from state	Religious authority as alternative power source
Money from state	Economic power as alternative power source
Military from civilian control	Coercive power as alternative power source

These separations serve the same function: ensuring authority traces exclusively to the governed.

Legitimacy Rule

A system is legitimate **iff** invariants 1.1–1.4 are intact.

If **any** invariant is degraded → **Consent Degradation Event** If degradation is sustained
→ **Legitimacy Crisis**

Threat Domains (Classification Only)

Threat	Degrades
Economic coercion	Agency
Censorship / information control	Information
Monopoly / centralization	Alternatives
Permanent emergency powers	Revocability
Mass surveillance / behavioral control	Agency & Information
Concentrated economic power	All (Agency through dependency, Information through media control, Alternatives through market dominance, Revocability through political capture)

Evaluation Algorithm

Pre-Evaluation Triage

Before testing against invariants, determine what kind of phenomenon X represents.
Disparity triggers audit, not presumption of encoding.

For any observed asymmetry or claimed violation, apply in order:

Step	Question	If yes
1. Formal encoding	Is the variable explicitly encoded in the rule set (law, regulation, institutional policy)?	Direct invariant violation — proceed to evaluation
2. Enforcement asymmetry	Is enforcement asymmetric at the rule level (same law applied differently)?	Implementation failure (bug) — proceed to evaluation
3. Runtime distortion	Is there measurable invariant degradation independent of formal encoding (e.g., structural)	Audit required — proceed to evaluation with evidence

Step	Question	If yes
	effects, institutional friction, access barriers)?	
4. Distributional artifact	Is the asymmetry explained by population structure, voluntary clustering, selection effects, or exogenous variables?	Not an invariant violation — may be policy-relevant information but does not trigger invariant classification

This ordering prevents two failure modes: **denialism** (ignoring real invariant violations because they aren't formally encoded) and **presumptive encoding** (treating any statistical disparity as proof of systemic violation). The triage ensures that correlation is not collapsed into encoding.

If triage resolves at step 4 (distributional artifact), **evaluation terminates**. No system state classification is produced. The observation may inform program-level policy but does not constitute an OS-level finding.

Invariant Test

For any object of analysis **X** (law, policy, institution, technology, action) that reaches the invariant test through triage steps 1–3:

1. Test X against **1.1–1.4**
2. If any test fails → record which invariant(s)
3. Map failures to threat domains
4. Classify system state:

State	Condition
Stable	All invariants intact
Strained	One invariant degraded
Crisis	Two or more degraded
Authoritarian Dynamics	Revocability + any other degraded
Systemic Failure	All invariants degraded

Scope Rules

- Do **not** evaluate intentions, motives, or narratives.
- Do **not** introduce ideological, cultural, or moral values.
- Evaluate **only observable effects on the invariants**.
- Political programs may be good or bad; **only OS violations matter**.

Architecture Constraint

All analysis must respect this stack:

Reality OS → Political OS → Political Programs

Only the **Political OS** contains normative constraints.

Single Operating Principle

| Programs may fail. The Operating System must not be broken.

Worked Example: Evaluation Algorithm in Action

This section demonstrates the mechanical application of the framework. The example is illustrative, not political advocacy.

Object of Analysis: Mandatory Digital Identity for Internet Access

Description: A hypothetical law requiring government-issued digital identity verification to access internet services.

Pre-Evaluation Triage: Is identity verification formally encoded in the rule set? **Yes** — this is a law requiring government-issued identity as a condition of access. Triage step 1 (formal encoding) resolves → proceed to invariant test.

Step 1: Test against invariants

Invariant	Test	Result
1.1 Agency	Does the law introduce coercion, compulsion, or survival dependency?	DEGRADED — Internet access increasingly necessary for economic participation, education, healthcare. Conditioning access on identity verification creates dependency on state approval.
1.2 Information	Does the law restrict ability to seek, receive, question, exchange information?	DEGRADED — Anonymous information seeking eliminated. Chilling effect on accessing controversial information. State gains visibility into all

Invariant	Test	Result
		information-seeking behavior.
1.3 Alternatives	Do real, accessible, non-punitive options remain?	DEGRADED — If all internet services require identity, no alternative access exists. Offline alternatives increasingly inadequate for full participation.
1.4 Revocability	Does authority remain removable through lawful, peaceful mechanisms?	INTACT — Law can be repealed through normal legislative process. However, if identity system becomes infrastructure, practical revocability diminishes.

Step 2: Record failures

Invariants degraded: 1.1 (Agency), 1.2 (Information), 1.3 (Alternatives)

Step 3: Map to threat domains

- Mass surveillance / behavioral control → Agency & Information
- Monopoly / centralization → Alternatives

Step 4: Classify system state

Three invariants degraded → **Crisis**

If implementation makes revocability practically difficult (infrastructure lock-in) → **Authoritarian Dynamics**

Note: This analysis follows mechanically from the framework. A different Political OS (e.g., one grounded in state security as the primary value) would evaluate the same law differently. The method is neutral; the axioms determine the output.

Translation to Political Science Terminology

The same analysis expressed in standard political theory terms:

Framework: Classical liberal constitutionalism (Locke, Mill)

Object of analysis: Mandatory digital identity legislation

Assessment against liberal principles:

1. **Individual autonomy** (Agency): Violated. The law creates state-controlled gatekeeping of essential services, establishing dependency relationships incompatible with liberal conceptions of negative liberty.
2. **Freedom of information** (Information): Violated. Elimination of anonymous access produces chilling effects on inquiry and expression, contrary to Millian principles of open discourse.
3. **Pluralism and exit rights** (Alternatives): Violated. Monopolistic infrastructure eliminates the practical capacity to opt out, undermining the voluntarist foundations of liberal consent theory.
4. **Democratic accountability** (Revocability): Intact but threatened. While formal legislative repeal remains possible, infrastructure lock-in may create path dependencies that render democratic reversal practically difficult.

Classification: The law represents a significant departure from liberal constitutional norms, degrading three of four foundational principles. In Levitsky and Ziblatt's terminology, this pattern — maintaining formal democratic procedures while hollowing out substantive liberal protections — characterizes “competitive authoritarian” or “illiberal democratic” trajectories.

This translation demonstrates that the constraint framework captures standard political science concepts — it represents them in a form that can be mechanically evaluated.