

G-OS Structural Volume: Causal Architecture and External Fixed-Point Governance (Value $\rightarrow$ Back $\rightarrow$ Gate $\rightarrow$ Front $\rightarrow$ Human)

Yasuhiro Azuma
PEACE-MAKER Institute

June 22, 2026

Abstract

G-OS / PEACE-MAKER is a governance operating system designed to ensure drift-free, stable, and institution-compatible AI behavior through a fully externalized Value Layer and a single Integration Gate. The architecture establishes a unified causal line (Value $\rightarrow$ Back $\rightarrow$ Gate $\rightarrow$ Front) that prevents internal value drift and enables scale-free inheritance across international, national, organizational, and individual governance layers. This structural volume isolates and formalizes the causal architecture, Value Layer construction, and layer-wise roles, providing a model-free, hardware-agnostic governance blueprint that can be implemented on top of any AI system while preserving human command authority.

Contents

1	Executive Summary	3
2	Causal Architecture of G-OS	4
2.1	Single unified causal line	4
2.2	Externalization of normativity	4
2.3	Model-free, hardware-agnostic structure	4
3	Value Layer: External Fixed Point	4
3.1	Role of the Value Layer	4
3.2	Construction process	4
3.3	Value Reference Database (Value DB)	5
3.4	Scale-free inheritance	5
4	Back Layer: Exploration Without Normativity	5
4.1	Role of the Back Layer	5
4.2	Characteristics	5
4.3	Examples of Back Layer technologies	5

5	Gate Layer: Integration Gate and Value Consistency	6
5.1	Single Integration Gate	6
5.2	Core enforcement properties	6
5.3	Auditability	6
6	Front Layer: Stable Presentation of Approved Outputs	6
6.1	Role of the Front Layer	6
6.2	Characteristics	7
6.3	Examples of Front Layer technologies	7
7	Human Decision Layer	7
7.1	Human command authority	7
7.2	Responsibilities of human decision-makers	7
7.3	Closure of the causal line	7
8	LLM Translation Layer (Outside the Causal Line)	7
8.1	Positioning	7
8.2	Role	8
8.3	Constraints	8
9	International Alignment (Structural)	8
10	Structural Summary of G-OS / PEACE-MAKER	8
10.1	Meta-structure	8
10.2	Causal architecture	8
10.3	Value Layer	9
10.4	Drift-free behavior	9
10.5	Layer roles	9
10.6	Model-free / hardware-agnostic	9
10.7	Use-case scalability	9
11	Conclusion of Structural Volume	9

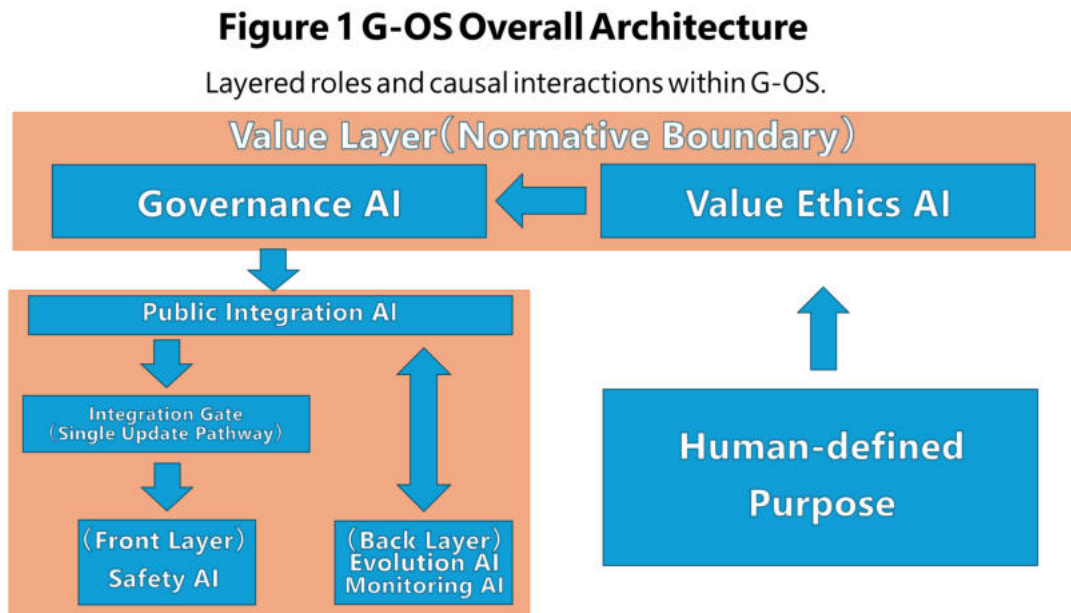
1 Executive Summary

Fundamental Principle (Non-Negotiable) The command authority must always remain human. The system’s role is to generate complete and transparent information for human decision-making. As long as this separation is preserved, governance cannot drift, escalate, or collapse — even if AI surpasses human intelligence.

Unified Causal Line

Value → Back → Gate → Front → Human Decision

G-OS / PEACE-MAKER is a governance operating system designed to ensure drift-free, stable, and institution-compatible AI behavior by externalizing all normative content into a fixed Value Layer and enforcing a single Integration Gate for all updates. This structural volume focuses exclusively on the causal architecture and layer definitions, separating them from operational and implementation details.



1

Figure 1: G-OS Architecture: Value → Back → Gate → Front → Human Decision

2 Causal Architecture of G-OS

2.1 Single unified causal line

The causal architecture of G-OS / PEACE-MAKER is defined by a single, unified causal line:

$$\text{Value} \rightarrow \text{Back} \rightarrow \text{Gate} \rightarrow \text{Front} \rightarrow \text{Human Decision}$$

All governance-relevant behavior must be derivable along this line. No shortcut, bypass, or parallel path is permitted to influence the Front Layer or human judgment without passing through the Integration Gate.

2.2 Externalization of normativity

All normative content is externalized into the Value Layer. The Back Layer performs exploration without normative intent, the Gate Layer performs value-consistency evaluation, and the Front Layer expresses only Gate-approved outputs. Human actors perform the final decision-making step.

2.3 Model-free, hardware-agnostic structure

The causal line is independent of any specific AI model, training method, or hardware. It can be applied to any model and any hardware configuration, as long as the Value, Back, Gate, and Front roles are structurally separated and causally ordered.

3 Value Layer: External Fixed Point

3.1 Role of the Value Layer

The Value Layer is the externally fixed normative anchor of G-OS. It is not generated by the system; it is defined, maintained, and updated by human institutions. It provides the reference against which all proposals from the Back Layer are evaluated by the Gate Layer.

3.2 Construction process

Construction of the Value Layer follows three steps:

1. **Extraction of Normative Intent:** Identification of normative content from constitutions, laws, regulations, international principles, and institutional policies.
2. **Structural Encoding:** Transformation of normative intent into a structured, machine-readable representation (e.g., rule sets, constraints, taxonomies, risk classes).
3. **External Fixation:** Storage of the encoded normative content in an externally governed Value Reference Database (Value DB) that is not modifiable by the Back, Gate, or Front Layers.

3.3 Value Reference Database (Value DB)

The Value Layer is operationalized as a Value Reference Database:

- Contains structured representations of laws, ethical guidelines, international principles, and institutional policies.
- Is governed and updated through human institutional processes (e.g., legislative change, regulatory updates).
- Is read-only from the perspective of Back, Gate, and Front; they may query but never modify it.

3.4 Scale-free inheritance

The Value Layer supports scale-free inheritance across international, national, organizational, and individual layers. Higher-level Value Layers (e.g., international human rights) can be inherited and refined by lower-level Value Layers (e.g., national law, organizational policy) without breaking the causal architecture.

4 Back Layer: Exploration Without Normativity

4.1 Role of the Back Layer

The Back Layer performs exhaustive or broad exploration of possible actions, plans, or system configurations without embedding any normative intent. It is proposal-only and cannot directly affect the external world.

4.2 Characteristics

- **Exploration-only:** Generates candidate options, scenarios, or solutions.
- **No direct output:** Cannot directly produce outputs visible to users or external systems.
- **Normatively neutral:** Does not encode or enforce values; it operates under purely technical or exploratory objectives.

4.3 Examples of Back Layer technologies

The Back Layer can be instantiated using existing technologies such as:

- SAT solvers and constraint satisfaction engines.
- Mathematical optimization (LP, MILP, nonlinear optimization).
- Simulation engines and scenario generators.

- Knowledge graph exploration and search.
- Any AI model used purely for exploration and proposal generation.

5 Gate Layer: Integration Gate and Value Consistency

5.1 Single Integration Gate

The Integration Gate is the single point of judgment in G-OS. All proposals from the Back Layer must pass through the Gate before they can be expressed by the Front Layer or presented to human decision-makers.

5.2 Core enforcement properties

The Gate enforces:

1. **Single-Point Evaluation:** All proposals are evaluated at one logical point, preventing fragmented or inconsistent value checks.
2. **Normative Consistency Checking:** Proposals are checked against the Value DB for legal, ethical, and policy alignment.
3. **Immutable Causal Direction:** Information flows from Value to Gate to Front; the Gate cannot be bypassed or overridden by Back or Front.

5.3 Auditability

The Gate Layer is designed for full auditability:

- Every decision (accept/reject) is logged with references to the relevant Value DB entries.
- The version of the Value DB used for each decision is recorded.
- The Gate’s behavior can be inspected by internal and external auditors.

6 Front Layer: Stable Presentation of Approved Outputs

6.1 Role of the Front Layer

The Front Layer expresses only Gate-approved outputs in a stable, predictable, and institution-compatible form. It does not perform normative evaluation; it formats and presents information for human decision-making.

6.2 Characteristics

- **Predictable:** Outputs follow predefined templates and structures.
- **Auditable:** Each output can be traced back to Gate decisions and Back proposals.
- **Institution-compatible:** Outputs are aligned with existing administrative, legal, and organizational processes.

6.3 Examples of Front Layer technologies

- Business intelligence dashboards and reporting tools.
- Visualization libraries and explanation interfaces.
- Structured decision-support documents and templates.

7 Human Decision Layer

7.1 Human command authority

The final decision-making authority always remains human. The system’s role is to generate complete and transparent information; it never replaces human judgment.

7.2 Responsibilities of human decision-makers

- Receive and interpret Front Layer outputs.
- Make binding decisions based on institutional mandates.
- Bear legal and institutional responsibility for decisions.

7.3 Closure of the causal line

The causal line closes at the human decision:

Value → Back → Gate → Front → Human Decision

No further automated step is permitted beyond human judgment within the governance loop.

8 LLM Translation Layer (Outside the Causal Line)

8.1 Positioning

The LLM Translation Layer sits outside the causal chain. It is not part of Value, Back, Gate, or Front. It serves only as a translation interface between human natural language and the structured representations used inside G-OS.

8.2 Role

- Translate human natural language inputs into structured queries or configurations.
- Translate structured outputs from the Front Layer into human-readable narratives.

8.3 Constraints

- The LLM cannot modify the Value DB.
- The LLM cannot bypass the Gate.
- The LLM cannot directly generate outputs that are treated as Front Layer results without passing through the Gate.

9 International Alignment (Structural)

G-OS aligns structurally with major international governance frameworks:

- **OECD AI Principles:** Human-centered values, transparency, robustness, accountability.
- **EU AI Act:** Risk-based approach, high-risk system governance, logging and oversight.
- **ISO/IEC 42001:** AI management systems, governance processes, and continuous improvement.
- **UN / ITU frameworks:** Human rights, sustainability, and international interoperability.

The external Value Layer and single Integration Gate provide a natural mapping to these frameworks, enabling consistent governance across jurisdictions.

10 Structural Summary of G-OS / PEACE-MAKER

10.1 Meta-structure

- **Scale-Free Structure:** Applicable from international to individual governance layers.
- **External Fixed Point:** All normativity is anchored in the Value Layer.
- **Single Update Pathway:** All changes to behavior flow through the Gate.
- **Uncertainty Externalization:** Uncertainty and exploration are confined to the Back Layer.

10.2 Causal architecture

Value $\rightarrow$ Back $\rightarrow$ Gate $\rightarrow$ Front $\rightarrow$ Human Decision

10.3 Value Layer

Immutable snapshot; fractal inheritance; drift-free anchor implemented as an externally governed Value DB.

10.4 Drift-free behavior

Drift-free behavior emerges from:

- Externalization of all normative content into the Value Layer.
- Single Integration Gate controlling all updates and outputs.
- Isolation of exploratory behavior within the Back Layer.

10.5 Layer roles

- **Back:** Exploration-only; proposal-only; no direct output.
- **Gate:** Single evaluator; full auditability; value consistency checking.
- **Front:** Predictable; auditable; institution-compatible.
- **Human:** Final decision-maker; responsibility holder.

10.6 Model-free / hardware-agnostic

The architecture applies to any model and any hardware. It defines governance structure, not specific algorithms.

10.7 Use-case scalability

The same structure can govern:

- Public-sector AI governance.
- Corporate AI risk management.
- Cross-border AI operations.
- Domain-specific governance modules.

11 Conclusion of Structural Volume

G-OS / PEACE-MAKER provides a drift-free, stable, and institution-compatible governance operating system by:

- Externalizing all normative content into a Value Layer.

- Confining exploration to a normatively neutral Back Layer.
- Enforcing a single Integration Gate for all updates and outputs.
- Restricting expression to a predictable, auditable Front Layer.
- Preserving human judgment as the sole decision-making authority.

This structural volume defines the governance architecture independently of any specific implementation. Operational procedures, migration plans, contingency structures, and detailed audit mechanisms are specified in the separate Operational Volume (G-OS Operational Handbook).

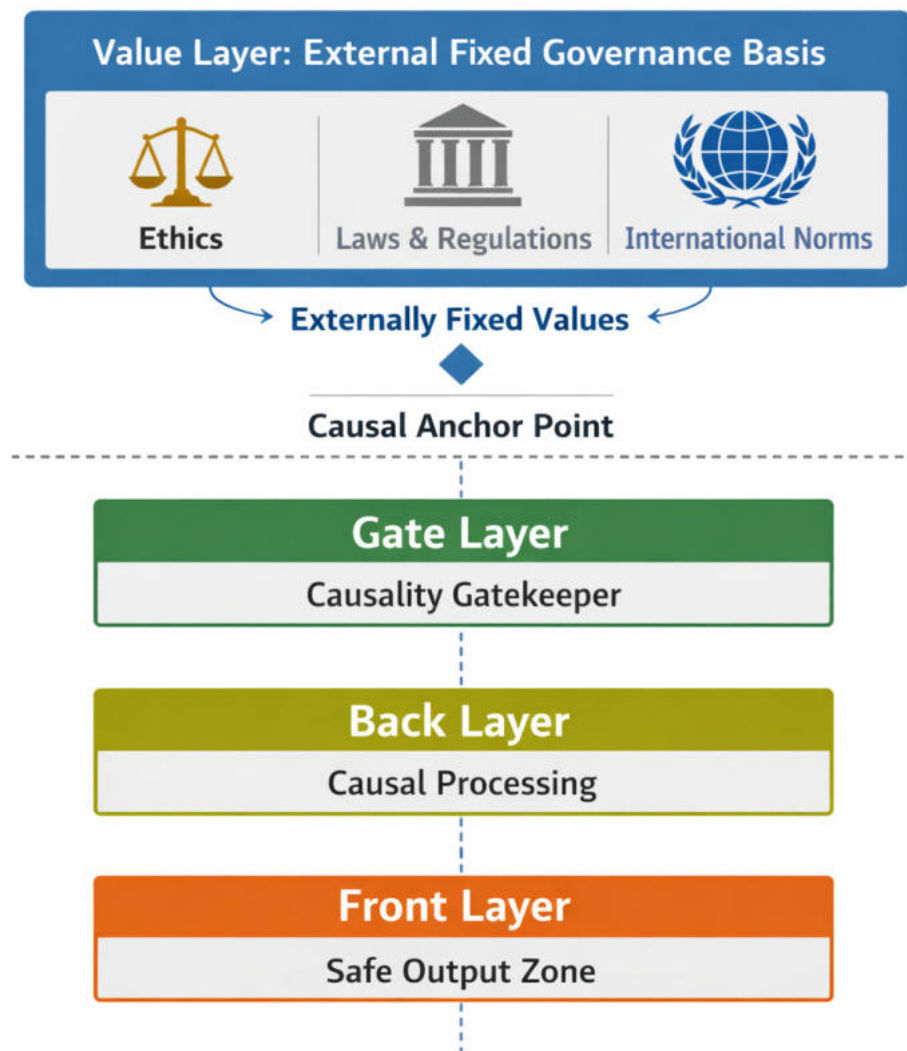


Figure 2: Value Layer: External Normative Anchor

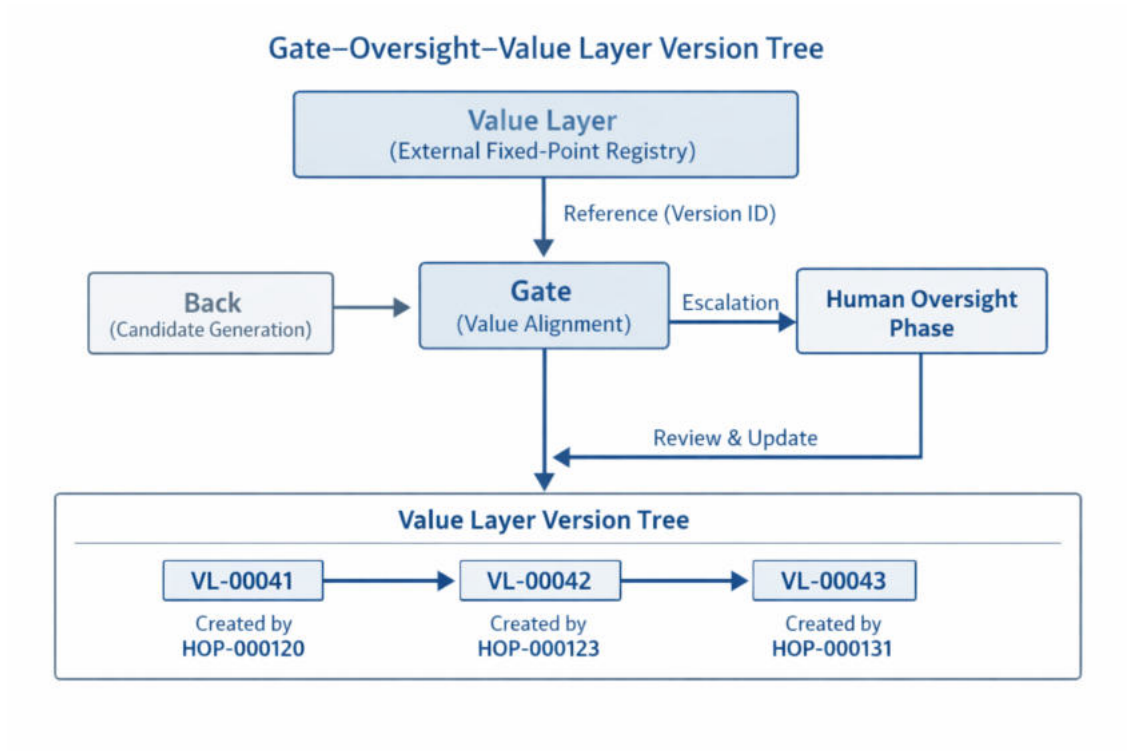


Figure 3: Integration Gate and Oversight Structure

PEACE-MAKER – Alignment with International AI Governance Frameworks

PEACE-MAKER Structural Element	OECD AI Principles	EU AI Act	ISO/IEC 42001	UN/ITU AI Governance
Value Layer (Externally Fixed Purpose)	Human-Centricity / Accountability	Fundamental Rights Protection	Clause 4 – Context & Leadership	Human-Rights-Based AI
Unknown = Unknown (Uncertainty Externalization)	Transparency / Risk Awareness	Article 9 – Risk Management	Clause 8.3 – Risk Controls	Transparency & Explainability
Integration Gate (Single Update Pathway)	Traceability / Safety	Article 12 – Logging & Record-Keeping	Clause 8.2 – Operational Control	Auditability & Oversight
Front/Back Layer Separation	Robustness / Safety	Article 14 – Human Oversight	Clause 8.4 – Monitoring & Review	Safe Deployment Principles
Model-Free / Hardware-Agnostic Architecture	Future-Proof Governance	General-Purpose AI Requirements	Clause 4 – Scope Independence	Technology-Neutral Governance
Temporal Convergence (Long-Term Stability)	Long-Term Trustworthiness	Article 15 – Technical Safety	Clause 9 – Continual Improvement	Sustainable AI Governance

Note:

The alignment shown in this table was not intentionally designed to match these international AI governance frameworks. PEACE-MAKER was developed independently, without referencing OECD, EU AI Act, ISO/IEC 42001, or UN/ITU documents.

Interestingly, the resulting structure naturally satisfies the core requirements of all major global AI governance frameworks. Whether this convergence is a coincidence or a structural inevitability is unclear even to me.

I suspect that my 20 years of experience in Japan's trust-bank system development — which emphasizes externally fixed purpose, single update pathways, strict separation of testing and production, non-speculative handling of unknowns, and long-term operational stability — may have shaped the Value Layer of PEACE-MAKER in a way that naturally aligns with global governance principles.

Figure 4: Alignment with OECD, EU AI Act, ISO/IEC 42001, UN Frameworks