Methods Addendum

GOPEL Infrastructure Specification and HAIA-RECCLIN Operational Model
for AI Provider Plurality Implementation

By Basil C. Puglisi, MPA
February 2026, revised November 2026
me@basilpuglisi.com | basilpuglisi.com | github.com/basilpuglisi/HAIA

1. Purpose

This addendum accompanies AI Provider Plurality: An Infrastructure Mandate for Democratic AI Systems. It provides the technical specification for GOPEL, a non-cognitive governance infrastructure designed to make multi-AI operations auditable, accountable, and interoperable at national scale.

GOPEL is general infrastructure. It is not limited to any single governance methodology. Any organization operating multiple AI platforms benefits from deterministic dispatch, cryptographic audit trails, checkpoint gates, and tamper-evident records. HAIA-RECCLIN is one governance implementation that runs on GOPEL infrastructure and demonstrates its operational feasibility.

The relationship between GOPEL and HAIA-RECCLIN is analogous to the relationship between the highway system and a trucking company. The highway system is public infrastructure that serves all vehicles. The trucking company is one operator that uses the highways and demonstrates they work. Congress is not being asked to fund a trucking company. Congress is being asked to fund the highway system.

1.1 Evidence Discipline

This document applies a three-tier evidence structure to distinguish between what is proven by others, what we have built as working concepts, and what we are asking Congress to fund.

Tier	Contents	Language Rule
Tier 1: PROVEN (by industry, academia, observable reality)	Single AI systems produce flawed outputs: hallucinations, bias, confabulation, alignment failures. This is established by peer-reviewed research and acknowledged by the companies themselves. Different AI systems produce different outputs on the same inputs. This is the nature of independently trained systems. Bias exists. Corporate concentration exists	Stated as established fact. Cited to industry and academic sources. We do not need to prove any of this
Tier 2: WORKING CONCEPTS (theories showing promise in development)	GOPEL: governance infrastructure treating multi-AI disagreement as signal. CBG: checkpoint process telling humans when to verify. HAIA-RECCLIN: implementation demonstrating feasibility. Documented instances where multi-AI comparison caught errors individual platforms missed. Observable behaviors in active development	"Working concept showing promise." "Theories in development." "Observable operational behaviors." Never: "proof," "proven," "validated," "evidence," "benchmark"
Tier 3: THE ASK (what Congress should fund)	Build infrastructure that protects the American public from flawed AI. Invest in diversification of AI systems to counteract bias and corporate control. Fund development of GOPEL as national AI infrastructure. Mandate API accessibility. Invest in small AI platforms to guarantee competitive plurality. Mandate verifiable inference attestation for AI platforms processing regulated data (Document 5)	Frame as infrastructure obligation. "The public needs protection." "AI is critical infrastructure." The precedent is highways, FAA, FCC, SEC. The ask is proportionate to the obligation

1.2 Methodology Origin

The governance methodology emerged from fifteen years of systematic practice, not recent adoption. BasilPuglisi.com began in 2009 with rigorous editorial standards and evolved through multiple stages: human-led writing, search-assisted sourcing, then the Factics methodology (pairing facts with tactics and measurable outcomes, first stated publicly in October 2012 and first in print in November 2012). This foundation was established before AI adoption in 2022.

When AI systems became available in 2022, they entered a workflow already governed by editorial rigor, constitutional thinking, and measurable accountability. The governance architecture described in this document represents the formalization of principles developed across more than nine hundred articles and refined through operational use across ten independent AI platforms.

2. GOPEL: The Infrastructure

GOPEL (Governance Orchestrator Policy Enforcement Layer) is a working concept for a non-cognitive agent that provides governance infrastructure for any multi-AI workflow. The specification exists, and a reference implementation was published on February 23, 2026. The operational experience from HAIA-RECCLIN demonstrates feasibility. The reference implementation is not deployed, and the national infrastructure has not been built.

2.1 Non-Cognitive Design

The agent performs zero cognitive work. This is a security architecture decision, not a limitation. For the purposes of this specification, cognitive operations are defined as evaluation, ranking, weighting, prioritization, summarization, semantic transformation, and filtering. The agent performs none of these.

The agent performs seven deterministic operations and no others:

#	Operation	What It Does	What It Does Not Do
1	**Dispatch**	Sends identical prompts to selected AI platforms via API	Does not modify, prioritize, or sequence prompts based on content
2	**Collect**	Receives all platform responses without modification	Does not filter, rank, or evaluate responses
3	**Route**	Delivers responses to Navigator for synthesis	Does not choose which responses to forward
4	**Log**	Writes structured audit records for every operation	Does not summarize or interpret logged content
5	**Pause**	Stops at checkpoint gates, delivers governance package to human	Does not decide whether to pause (gates are preconfigured)
6	**Hash**	Computes SHA-256 cryptographic hashes for tamper detection	Does not evaluate content being hashed
7	**Report**	Counts approval rates, reversal rates, threshold triggers	Does not interpret what the counts mean

The security rationale is straightforward: if the agent cannot evaluate content, it cannot be manipulated through adversarial inputs, prompt injection, or model poisoning. The attack surface is reduced to message transport

and logging, both of which are deterministic operations amenable to formal verification. This materially reduces the cognitive manipulation risk: a superintelligent platform cannot corrupt a system that has no beliefs, preferences, or judgment to target. Residual risks in transport, identity, and human factors are addressed through deterministic infrastructure controls (see Section 5.6).

This design addresses the escalation scenario that Geoffrey Hinton and others have warned about: what happens when AI capability advances to the point of influencing corporate decision-making or operating beyond the comprehension of its operators. If a cognitive governance layer existed between the AI platforms and the human, that layer itself becomes a target for manipulation by a sufficiently advanced system. GOPEL eliminates this vulnerability by design. There is no cognition to manipulate. There is no judgment to influence. The governance infrastructure is architecturally intended to remain mechanically reliable regardless of how capable the AI platforms it governs become, subject to pilot validation of the deterministic controls specified in Section 5.6. This is the structural answer to the concentration-of-capability risk: not a smarter governor, but a governor that does not need to be smart.

2.2 Infrastructure Analogy

The government does not build cars. It builds roads. It does not fly planes. It created the FAA. It does not own the broadcast spectrum. It created the FCC. GOPEL follows this pattern. The AI platforms are the vehicles. GOPEL is the road. It does not generate AI outputs. It makes AI outputs auditable.

Any vehicle operating on public infrastructure must meet compatibility standards. Any AI platform operating on public cognitive infrastructure must maintain audit accessibility. This is not regulation of what AI says. It is a requirement that AI remains auditable.

2.3 What GOPEL Is Not

GOPEL is not a competing AI. It generates no content. It is not a filter. It blocks nothing. It is not a regulator. It enforces no content standards. It is not a product. It is infrastructure. The government builds it, maintains it, and makes it available. Organizations and agencies use it. AI companies maintain API compatibility with it. Citizens benefit from the accountability it creates.

3. HAIA-RECCLIN: The Implementation That Demonstrates Feasibility

HAIA-RECCLIN (Human AI Assistant with Researcher, Editor, Coder, Calculator, Liaison, Ideator, Navigator) is a governance implementation that runs on GOPEL infrastructure and whose operational experience supports feasibility. It organizes both humans and AI systems into seven functional roles that mirror constitutional checks and balances. It is one implementation. Others are possible for different organizational contexts.

3.1 Seven Functional Roles

Each role is independently assignable to different AI platforms. The framework requires exactly seven roles. No additions, no removals, no combinations. This is a framework invariant.

Role	Function	Core Responsibility
Researcher	Research	Gather verified data from multiple AI sources and human records. Cross-reference claims across platforms
Editor	Quality Control	Preserve accuracy, brand integrity, consistency, and traceability. Adapt content for audience
Coder	Technical Build	Write, review, and debug code. Implement technical specifications
Calculator	Quantitative Analysis	Validate quantitative or logical components. Mathematical modeling. Data processing
Liaison	Communication	Translate outcomes for policy, enterprise, and public use. Stakeholder coordination
Ideator	Innovation	Generate creative options, brainstorm novel approaches. Stress-test assumptions
Navigator	Synthesis	Synthesize multi-platform outputs with mandatory dissent preservation. Present trade-offs without forced resolution. The Navigator role is competitively held, not permanently assigned. Claude (Anthropic) currently occupies this role based on documented operational performance, including memory architecture, storage capability, and sustained output quality across workflows. Gemini and ChatGPT have been identified as capable substitutes. Under Agent Model 1, three independent platforms review Navigator output before final delivery. Under all models, Checkpoint-Based Governance places the human as final authority over every Navigator synthesis. The Navigator holds the pen. The human holds the authority. The architecture holds the Navigator accountable.

3.2 Anchor-Plus-Rotation Protocol

Working concept. Operational experience supports feasibility. Formal protocol not yet validated independently.

For each functional role, the infrastructure selects three platforms: one anchor designated for that role plus two from the remaining rotation pool. The rotation schedule changes with each task to prevent two-platform echo chambers. The rotation pool includes all available commercial AI platforms. Platform additions or removals update the pool without affecting the architecture.

Operational experience documented that when a primary platform (Claude) was temporarily unavailable, tasks completed successfully through systematic role reassignment. This demonstrates the infrastructure's resilience to provider disruption, a critical property for national-scale deployment.

3.3 Operational Sequence

This sequence describes Agent Model 3 (manual) operations. Agent Models 1 and 2 automate these steps through the GOPEL agent.

1. Task Definition and Criteria Setting. The human defines objectives and evaluation metrics.
2. Independent Generation. Each AI produces outputs without cross-prompt exposure. Identical prompts dispatched to three or more platforms.
3. Comparative Review. Outputs are compared for consistency, evidence quality, and bias indicators.
4. Dissent Logging. All disagreements are documented. None are discarded. Disagreement is diagnostic signal, not noise.
5. Human Arbitration. The human determines factual accuracy, ethical alignment, and final content. The human is the final authority on every decision.

6. Factics Integration. Every verified fact is paired with a tactic (actionable step) and measurable outcome (KPI).
7. Audit Logging. Citations, rationales, conflicts, and decisions are recorded with cryptographic hash chaining for tamper detection.

This process converts disagreement into diagnostic signal. In documented operational experience, multi-platform comparison surfaced errors that single-platform workflows missed. In one documented instance, eight of nine platforms produced the same incorrect output. The governance process flagged the single dissenter, triggered human verification, and the dissenter was correct. The eight were overridden. This is a single instance, not proof of general superiority. It is an observable behavior that supports the feasibility of multi-AI governance.

4. Three Agent Models

The infrastructure supports three agent models that scale governance intensity to match organizational risk tolerance. These are not a quality hierarchy. Each serves a different operational context. Model selection is itself a governance decision, governed by the Checkpoint-Based Governance four-stage decision loop: AI contribution provides analytical support, checkpoint evaluation structures review, human arbitration retains final authority, and decision logging creates immutable accountability trails.

	Agent Model 1	Agent Model 2	Agent Model 3
Designation	Responsible AI	Hybrid, Through to AI Governance	Manual Human AI Governance
Checkpoint Density	Single endpoint checkpoint	Checkpoint per functional role	Full human orchestration at every step
Automation Level	Agent runs full pipeline. Human reviews final output	Agent pauses after each role. Human approves before proceeding	No agent. Human dispatches, collects, routes manually
Appropriate For	Low to moderate risk. Routine operations with established patterns	High-risk decisions. Employment, credit, healthcare, law enforcement	Highest-consequence decisions. Novel situations. Framework development and validation
Evidence Tier	Tier 2: Specified architecture. Alpha code built October 2026, not deployed	Tier 2: Specified architecture. Alpha code built October 2026, not deployed	Tier 2: Operational experience exists. Produced published book, case studies, audit documentation
Federal Pilot Path	Phase 4: After Agent Models 2 and 3 are validated in agency pilots	Phase 3: After Agent Model 3 establishes governance baseline	Phase 2: Immediate adoption. Agencies operate manually while infrastructure is built

Agent Model 3 produced the published book Governing AI: When Capability Exceeds Control (Puglisi, 2025, ISBN 9798349677687, 204 pages). That process documented 96% checkpoint utilization, 100% dissent documentation, 28 major checkpoint decisions, and 26 preserved dissents across five independent AI platforms over six weeks. These are documented process outcomes from a single project, not validated performance benchmarks. The book is the most concrete artifact: a 204-page publication produced through the governance process it describes.

Agent Models 1 and 2 are specified architecture. The specifications exist in the HAIA-RECCLIN Agent Architecture Specification v2.2 (EU Compliance Version), published on GitHub. Alpha code for both was built in October 2026 as reference code and has not been deployed. Federal investment builds them.

5. Audit Architecture

Working concept. Specified in the Agent Architecture document and carried in the GOPEL reference implementation of February 23, 2026, which is not deployed. This section summarizes the specification for Congressional review.

5.1 Tamper-Evident Audit Trail

The audit trail is a structured text file (JSON or Markdown), not a database. Any AI platform can ingest it. Any auditor can query it. Platform-independent design means audit evidence does not depend on the system that produced it.

Each audit entry receives a SHA-256 cryptographic hash that incorporates the previous entry's hash, creating a sequential chain. If any entry is modified or deleted after the fact, the chain breaks and the integrity status changes from verified to compromised. This is the same cryptographic principle used in financial transaction logging and blockchain systems.

5.2 Six Record Types

Record Type	What It Captures	Why It Matters
Request	Task assignment, role designation, model selection, human identity	Proves a named human authorized the task
Dispatch	Prompts sent, platforms selected, timestamps	Proves identical inputs went to independent platforms
Response	Complete, unedited platform outputs with receipt times	Proves what each platform actually said
Navigation	Navigator synthesis, convergence/conflict assessment, preserved dissent	Proves disagreement was surfaced, not suppressed
Arbitration	Human review decision, rationale, modifications, identity binding	Proves a named human made the final call
Decision	Final output, approval status, hash of complete transaction chain	Proves the entire chain from request to decision is intact and verifiable

5.3 Automation Bias Detection

Working concept. Threshold-based detection specified in architecture. The default thresholds are proposed and untested.

When a human operator approves AI outputs without substantive modification at rates exceeding a configurable threshold, the infrastructure triggers a mandatory review. Default thresholds (95% approval rate, less than 2% decision reversal frequency over three consecutive cycles, with mandatory audit initiation within five business days) are provided as starting points. Implementing organizations configure these thresholds based on industry context, risk tolerance, and oversight requirements. This addresses a documented

phenomenon: humans systematically defer to AI recommendations under volume pressure (EDPS TechDispatch #2/2025; Goddard, Roudsari, and Wyatt, 2011; Banovic et al., 2023).

Single-provider systems structurally cannot detect rubber-stamping because there is no comparison point. Multi-provider infrastructure with threshold monitoring provides the structural mechanism that single-provider systems lack.

5.4 Ten Default Reporting Fields

Working concept. Used operationally in Agent Model 3 manual workflows. Standardizes governance output format.

Field	Content
Role	RECCLIN functional role assigned to this task
Task	Understanding of the request in the agent's (or human's) own terms
Output	The substantive response content
Sources	Cited evidence with links. APA format preferred. A source is verified by its URL and matching content on the page, and a source that fails that test goes to the human arbiter as a question before the output is produced
Conflict	Documented dissent between sources or platforms. "None identified" if no conflicts found
Confidence	0 to 100 percent, with justification based on evidence quality
Factics	Fact paired with tactic and measurable outcome as integrated statement
Expiry	Time-sensitivity of the information. "Stable" or specific validity window
Recommendation	Primary suggestion plus alternatives
Decision	Specific choice requiring human approval, with recommendation and alternatives

5.5 Identity Binding

Audit records asserting human arbitration decisions must be cryptographically bound to an authenticated human identity. Federal pilots bind identity using existing federal identity infrastructure: PIV (Personal Identity Verification) or CAC (Common Access Card) backed authentication for arbitration records, digital signatures on decision entries, and key management aligned to agency information security policy. Enterprise deployments integrate with existing identity management systems (SSO, directory services). The objective is non-repudiation: any audit entry asserting a human decision can be traced to a specific individual, and that individual cannot plausibly deny the decision. Implementation specifics vary by agency and deployment context. The architectural requirement is constant: every arbitration and decision record carries authenticated identity binding.

5.6 Residual Threat Model

The non-cognitive design materially reduces the attack surface available to AI adversaries by eliminating the cognitive operations that manipulation, prompt injection, and social engineering require. Residual risks addressed by deterministic controls include: transport integrity (message signing and TLS verification on all API calls), key compromise (key rotation schedules and hardware security module storage for signing keys), API tampering (hash verification of dispatched prompts against audit records), log deletion or modification (append-only storage with cryptographic hash chaining, external backup with independent hash verification), and insider misuse (separation of duties, immutable deployment, role-based access control on audit file

storage). Each residual risk maps to a deterministic control rather than a cognitive judgment, maintaining the non-cognitive security boundary.

6. Operational Observations

All observations in this section are from single-practitioner implementation using Agent Model 3 (manual). They document the development journey of a working concept, not validated benchmarks. Independent validation through federal pilots is the purpose of the legislative ask.

6.1 What the Operational Experience Shows

Multi-AI workflows across ten independent platforms (ChatGPT, Claude, Gemini, Perplexity, Grok, Mistral, DeepSeek, Meta AI, Copilot, Kimi) produced the following observable behaviors:

- Different platforms produce meaningfully different outputs on identical inputs. This is Tier 1 (proven), not unique to HAIA-RECCLIN.
- When platforms disagree, the disagreement frequently identifies errors, biases, or gaps that no single platform surfaced alone. This is Tier 2 (working concept observation).
- A checkpoint process that flags disagreement and directs human attention to it produces documented instances of error correction. This is Tier 2.
- In one documented case study, eight of nine platforms agreed on an incorrect output. The governance process surfaced the single dissenter, triggered human verification, and the dissenter was correct. This is a single instance, not general proof.
- Platform loss (primary provider temporarily unavailable) was absorbed through role reassignment without task failure. This supports infrastructure resilience claims.

6.2 Metrics from Working Concept Development

The following observations are from documented operational experience. They support feasibility, not proof.

Observation	Value	Evidence Tier and Qualification
Cross-platform consistency	0.96 ICC across 5 platforms, 4 dimensions (Case Study #001)	Tier 2. Consistency measurement from single-practitioner workflow. Demonstrates platforms can produce convergent governance outputs. Not a psychometric reliability coefficient
Composite collaboration score	91.8 across 9 platforms (EOY 2025 Audit)	Tier 2. Operational observation documenting collaboration quality across platforms. Single practitioner. Not a validated benchmark
Checkpoint utilization	96% (28 of 29 checkpoints used in book production)	Tier 2. Documented process outcome from a single six-week project. Shows governance processes can be sustained operationally
Dissent documentation	100% (26 dissents preserved, none discarded)	Tier 2. Demonstrates the architecture preserves rather than suppresses disagreement. Single project
Continuity under stress	100% task completion when primary provider unavailable	Tier 2. Demonstrates infrastructure resilience. Role reassignment absorbed provider loss

These observations indicate that multi-AI governance is operationally feasible, that disagreement between platforms produces useful diagnostic signal, and that governance processes can be sustained across extended projects. Federal pilot programs will determine whether these observations replicate across organizations, domains, and operational scales.

7. Ethical and Privacy Safeguards

Data Sovereignty. All inputs and outputs reside within the owning organization's controlled environment. The infrastructure does not create new data stores. It creates audit records of data that already flows between humans and AI platforms.

Transparency Without Surveillance. Oversight records process logic, not personal content. The audit trail documents what decisions were made, by whom, with what rationale, based on what platform outputs. It does not surveil the humans operating the system.

Provider Independence. No exclusive contracts or single-vendor dependencies. The infrastructure requires minimum three independently trained AI models with materially distinct training pipelines for every task. Independence is verified through provider attestation, divergence statistics from standardized test suites administered through GOPEL, conflict-of-interest disclosures, and documented corporate independence. This is not a preference. It is an architectural requirement that prevents corporate capture.

Open Audit Publication. Non-classified summaries are released for public review, preventing ethics-washing. For classified applications, audit methodologies follow existing federal security frameworks with appropriately cleared oversight personnel.

GDPR Erasure Reconciliation. The audit specification includes a Bridge Record protocol that reconciles data subject erasure rights (GDPR Article 17) with audit trail integrity. Personal data is anonymized. Governance metadata (who decided what, when, under what authority) is preserved. The hash chain is maintained through documented Bridge Records.

Retention and Minimization. Workflows classified as high-consequence or under active investigation retain full content (prompts, outputs, synthesis) in the audit trail with encryption at rest and role-based access control. Routine workflows retain metadata, cryptographic hashes, and structured references sufficient to reconstruct the decision chain without storing full content indefinitely. Retention schedules are configured by the implementing organization based on regulatory requirements, agency records management policy, and operational risk classification. This tiered approach ensures audit integrity while preventing the accumulation of sensitive operational content beyond its governance utility.

8. Federal Pilot Roadmap

This roadmap frames GOPEL development as federal infrastructure investment. Each phase produces independent validation data. Phase completion is a Checkpoint-Based Governance decision with documented rationale.

Phase	Name	Activity	Validation Output
0	**Immediate (No Agent)**	Adopt Agent Model 3 governance manually in pilot agencies. Collect platform histories. Build governance muscle before automation	Operational feasibility data from agency context. Baseline governance metrics for comparison with later phases
1	**Audit Infrastructure**	Design and validate audit file schema. Test cross-platform ingestibility. Verify hash chaining and tamper detection	Validated audit schema. Proof that multiple AI platforms can query the same governance records
2	**Agent Core**	Build the logging engine first. Verify immutability, completeness (six record types), and reconstruction (any transaction's full chain retrievable)	Functional logging engine. Reconstruction test results
3	**Dispatch and Synthesis**	Add API dispatch. Implement anchor-plus-rotation. Connect Navigator synthesis pipeline. Verify all transactions flow through logging	Agent Model 2 operational data. Comparison with Agent Model 3 baseline. Checkpoint effectiveness measurements
4	**Full Operations**	Implement per-role gates with pause/continue states. Test Agent Model 1 and Agent Model 2 configurations. Validate arbitration interface	Agent Model 1 operational data. Cross-model comparison. Automation bias detection performance. Scaling feasibility assessment
5	**Compliance Validation**	Internal review against regulatory coverage matrix. Produce remaining organizational documents. Prepare for conformity assessment	Compliance documentation package. Readiness assessment for high-risk classification deployment

Phase 0 can begin immediately in any federal agency with no infrastructure investment. The human operates manually, uses multiple AI platforms, and applies governance principles described in this addendum. The published book (Puglisi, 2025) documents what this process looks like in practice. Phase 0 is how agencies demonstrate interest and generate baseline data that informs infrastructure design in subsequent phases.

9. Policy and Standards Alignment

The infrastructure is built to American standards for American public safety. Its architectural alignment with international frameworks, including the EU AI Act, NIST AI RMF, ISO 42001, and emerging global AI management standards, is by design, not by obligation. This positions the United States as the origin of AI governance infrastructure that other nations can adopt or align with, rather than as a follower importing foreign regulatory regimes. Agencies operating internationally or in transatlantic contexts can satisfy EU-level requirements through the same infrastructure without duplicative systems. Governance intensity remains selected by the operating organization based on risk context, not imposed uniformly by external mandate.

No formal conformity assessment has been completed. Alignment claims describe design intent, not certification status.

Authority	Requirement	GOPEL/HAIA-RECCLIN Mechanism
EO 14179 (Jan 23, 2025)	Removing Barriers to American Leadership in Artificial Intelligence. Revoked prior administration AI restrictions	GOPEL infrastructure removes barriers by providing governance through engineering rather than restriction. Non-cognitive design imposes zero content regulation
EO 14365 (Dec 11, 2025)	Ensuring a National Policy Framework for Artificial Intelligence. Minimally burdensome national standard to prevent 50-state patchwork. Builds on EO 14179	GOPEL provides the national infrastructure standard the order calls for. Engineering that makes less regulation safe. Compatible with state procurement carve-outs preserved in the order
OMB M-25-21	Governance, transparency, public trust	Role accountability, dissent logs, public audit summaries. Ten reporting fields standardize output
OMB M-25-22	Competitive acquisition, avoid vendor lock-in	Minimum three providers. Rotation protocol. Vendor independence audits. Interface interoperability
NIST AI RMF 1.0	Govern, Map, Measure, Manage risk	Govern (role accountability, dissent logs), Map (task and risk scoping per provider), Measure (disagreement rates, consistency), Manage (remediation, vendor independence)
EU AI Act Art. 14	Human oversight for high-risk systems	Three agent models with configurable checkpoint density. Agent Model 2 provides human review at every processing stage
EU AI Act Art. 12	Automatic logging	Six record types. Append-only. Cryptographic hash chaining. Platform-independent format
prEN 18286:2025	AI management systems quality standards	Checkpoint-Based Governance provides decision framework. Audit architecture provides evidence. Agent Models provide scalable implementation
ISO 42001	AI management system requirements	Documentation requirements, risk assessment, continuous improvement mapped to CBG four-stage loop
GPAI Code of Practice 2025	General purpose AI provider obligations	Cross-provider comparison as continuous evaluation. Dissent logging as transparency mechanism. API accessibility supports interoperability obligations
AISI (US/UK)	Red-teaming and evaluation practices	Multi-provider comparison functions as continuous cross-model probing. Human arbitration captures failure modes in audit trail
DMA	Contestability vs. gatekeepers	Multi-provider orchestration. Exit clauses. No exclusivity that blocks comparative evaluation

Interoperability Requirements. To prevent walled gardens, the infrastructure specification includes: provider-neutral prompt and IO schemas, exportable citation formats, latency budgets for cross-model calls, and contract terms forbidding exclusivity that blocks comparative evaluation. This maps to M-25-22 acquisition requirements and DMA contestability.

10. API Accessibility Mandate

Any AI company providing services to federal agencies, operating within federally regulated sectors, or participating in high-consequence decision pipelines (including employment, credit, healthcare, education, law enforcement, and national security) must maintain API accessibility compatible with the federally maintained governance infrastructure. Refusing access or deliberately degrading interoperability in these contexts constitutes a regulatory violation enforceable by the Federal Trade Commission. For commercial applications outside these scopes, voluntary adoption of GOPEL API compatibility is encouraged through procurement preference, consistent with OMB M-25-22 acquisition standards that already require open data formats, interoperability, and avoidance of vendor lock-in.

This mandate does not restrict AI capability or output. It requires audit compatibility, the same interoperability obligation that vehicles meet to operate on public roads. This is not regulation of what AI says. It is a requirement that AI remains auditable and contestable when used in consequential decision pipelines.

Without this mandate, AI companies can shut down API access to governance tools that make their outputs comparable and auditable. The fact that provider plurality can be killed by providers is itself the clearest demonstration that provider plurality requires legal protection.

10.1 Small AI Investment

Plurality only works if there are enough providers to sustain it. If the government mandates multi-provider governance but only four or five mega-platforms exist, that is an oligopoly with an audit trail, not checks and balances.

Federal investment in small AI platforms, modeled on SBIR, STTR, and DARPA funding for emerging technology companies, creates the supply-side complement to the demand-side infrastructure mandate. Any company receiving investment maintains GOPEL API compatibility. This is the same model the government uses for defense contractors, rural broadband, and regional aviation: fund competition so infrastructure serves everyone.

10.2 Verified Inference Attestation

AI platforms processing regulated data must return verifiable evidence of processing conditions alongside inference responses. The Verified AI Inference Standards Act (VAISA), Document 5 of this package, specifies the evidence standard that makes existing obligations enforceable at the AI inference boundary, phased from federal procurement through voluntary NIST standard to contingent legislative mandate. VAISA provides the specific technical content of the API accessibility mandate in this section.

11. Adaptive Governance and Continuous Review

AI technology evolves faster than traditional policy cycles. To prevent infrastructure obsolescence, GOPEL implementation requires built-in mechanisms for continuous assessment.

Agencies implementing GOPEL infrastructure should produce biennial reports to Congress documenting operational metrics, emerging risks, and recommended updates. These reports should include aggregate cross-platform disagreement patterns across agencies, documented instances where multi-provider comparison surfaced errors single-provider workflows missed, continuity and resilience outcomes when providers experience disruptions, cost-benefit comparison of multi-provider governance versus single-vendor approaches, and emerging technological capabilities requiring infrastructure adaptation.

This adaptive governance model ensures infrastructure remains aligned with technological evolution while maintaining democratic accountability. Quarterly internal reviews feed into biennial Congressional reporting, creating feedback loops between operational practice and strategic oversight.

Federal coordination mechanisms, such as those demonstrated by the White House Task Force on AI Education (2025), provide templates for interagency collaboration. A permanent interagency working group on AI governance should coordinate GOPEL implementation across OSTP, OMB, GSA, NIST, and FTC, ensuring consistent standards while allowing agency-specific adaptation.

Known implementation challenges include multi-provider coordination overhead, agency adoption learning curves, sustained API compatibility enforcement across commercial providers, and the need for investment continuity beyond initial pilot phases. These challenges are anticipated, not disqualifying. Phase 0 and Phase 1 baselines become the comparison benchmarks for all subsequent biennial Congressional reporting, creating a continuous measurement chain from first manual pilot to scaled automated operations. This ensures that effectiveness is not assessed against theoretical ideals but against documented, incrementally improving operational reality.

12. Conclusion

Single AI is proven flawed by industry research, academic study, and the companies' own safety evaluations. Different AI systems produce different outputs on the same inputs. This is observable reality, not theory. Bias exists. Corporate concentration exists. The American public interacts with these systems daily in finance, healthcare, education, and government services.

The logical infrastructure response is multi-AI governance with diversification investment. The government builds the road. The AI platforms are the vehicles. Citizens choose their vehicles. The government guarantees the road is safe.

GOPEL is that road. A non-cognitive governance layer that dispatches, collects, routes, logs, pauses, hashes, and reports. Seven deterministic operations. Zero cognitive work. The architecture cannot be co-opted because there is nothing to co-opt. Every decision cryptographically bound to a human identity and independently verifiable.

HAIA-RECCLIN is one implementation whose operational experience supports the feasibility of this infrastructure. A published book produced through the governance process it describes. Case studies documenting observable behaviors. Consistency measurements across ten independent platforms. Working concepts showing promise and a specification ready for development.

Executive Orders 14179 and 14365 call for a national standard that does not paralyze innovation. The specification provides a starting point for one. This is not a proposal for more regulation. This is the engineering that makes less regulation safe.

This proposal does not claim to be the complete answer. It is a pioneer path that combines established concerns and proven suggestions from multiple directions into one operational architecture with one goal: safe use of AI for the American public. Geoffrey Hinton warned about capability exceeding control. The American antitrust tradition established that concentrated power requires structural checks. Federal infrastructure precedent demonstrated that public safety and private innovation coexist when the government builds the road and the market builds the vehicles. Automation bias research proved that humans defer to machines under volume pressure. None of these observations originated here. What originated here is the combination of these

concerns into a single working architecture with a defined specification, documented operational experience, and a development path.

AI will never be absolute and without risk. The infrastructure is designed to manage risk, not eliminate it. What is needed is federal investment to build, pilot, validate, and improve the infrastructure that protects the American public. Not a finished product. A starting point. The country needs to start.

Related Documents

This Methods Addendum is part of the AI Provider Plurality Congressional Package:

- Document 1: Summary Flyer (elevator pitch for infrastructure proposal)
- Document 2: Ethics for Oversight (constitutional and philosophical case)
- Document 3: AI Provider Plurality (legislative framework and policy mechanism). Document 3 includes funding mechanism options and phased appropriations placement. Phase 0 requires no new appropriation because it operates as a manual governance pilot using existing agency resources.
- Document 4: Methods Addendum (this technical specification and operational experience)
- Document 5: Verified AI Inference Standards Act (VAISA) (attestation API requirements and legislative framework for AI inference data protection)
- Document 6: Overwatch: The Cognitive Monitoring Shield for GOPEL (working paper to support AI Provider Plurality)

Funding and Appropriations Placement. This Methods Addendum is a technical specification and does not include appropriations figures by design. Detailed funding estimates, phasing, and appropriations language for Phases 1 through 5 are provided in Document 3. Congress retains full authority over funding mechanism selection, including but not limited to phased milestone-gated appropriation, competitive SBIR/STTR-style grants, user-fee sustainability models, and procurement-driven standards adoption. These options are presented in Document 3 as a menu, not a prescription.

Supporting technical documents available for Congressional staff review:

- HAIA-RECCLIN Agent Architecture Specification v2.2, EU Compliance Version (full GOPEL specification, GitHub)
- HAIA-RECCLIN Academic Working Paper, EU Regulatory Compliance Edition (literature positioning, limitations)
- Governing AI: When Capability Exceeds Control (Puglisi, 2025, ISBN 9798349677687, published book)
- Human-AI Collaboration Audit, End of Year 2025 (operational observations across 9 platforms)
- Checkpoint-Based Governance (four-stage decision loop, basilpuglisi.com/cbg)
- Agent Architecture CBG Case Study v1.1 (21 governance events, documented error detection)
- GOPEL: The Code Behind the Policy (reference implementation, February 23, 2026, basilpuglisi.com/gopel-the-code-behind-the-policy and GitHub)
- HAIA Agents: From the Definition Problem to Running Code (Alpha code for Agent Model 1 and Agent Model 2, basilpuglisi.com/haia-agents-running-code)

References

Banovic, N., et al. (2023). Being Trustworthy is Not Enough: How Untrustworthy Artificial Intelligence (AI) Can Deceive the End-Users. Proceedings of the ACM on Human-Computer Interaction.

European Data Protection Supervisor. (2025). TechDispatch #2/2025: Automation Bias. EDPS.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2011). Automation bias: A systematic review of frequency, effect mediators, and mitigators. Journal of the American Medical Informatics Association, 19(1), 121-127.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33(2-3), 61-83.

Kreps, S. E., Kriner, D. L., & Schneider, B. R. (2024). Automation bias in AI-assisted decision-making for national security. International Security.

Navarrete, R., et al. (2024). Automation bias in criminal justice risk assessment: compounding effects with racial bias. AI and Ethics.

Puglisi, B. (2025). Governing AI: When Capability Exceeds Control. ISBN 9798349677687.

Puglisi, B. (2026). HAIA-RECCLIN Agent Architecture Specification v2.2, EU Compliance Version. github.com/basilpuglisi/HAIA.

Puglisi, B. (2026). HAIA-RECCLIN Multi-AI Framework Updated for 2026. basilpuglisi.com.

Sabol, D. A. (2025). Federal policy for generative AI in U.S. education. White House Task Force on AI Education.

Canonical Definitions

HAIA = Human AI Assistant (governance layer for human-AI collaboration). Full expansion: Human AI Assistant.

RECCLIN = Seven functional roles: Researcher, Editor, Coder, Calculator, Liaison, Ideator, Navigator

GOPEL = Governance Orchestrator Policy Enforcement Layer (non-cognitive infrastructure agent)

CBG = Checkpoint-Based Governance (four-stage decision loop)

Factics = Facts + Tactics + KPIs (evidence-based decision methodology, first stated publicly October 2012, first in print November 2012)

#AIassisted using the HAIA Ecosystem | CC BY-NC-SA 4.0
Free for personal, educational, and noncommercial research use with attribution. Commercial exploitation, paid productization, and enterprise commercialization require separate permission and licensing.
Suggested citation: Puglisi, B. C. (2026). Methods Addendum: GOPEL Infrastructure Specification and HAIA-RECCLIN Operational Model for AI Provider Plurality Implementation. basilpuglisi.com.