HRI PRIOR-ART COLLISION MATRIX

Constitutional Causal Continuity / Human Reasoning Infrastructure

Search date: 12 August 2026 | Prepared for public scholarly release

Tony T. Nguyen — Highland Reflexive Press

Executive finding

Bottom line. This search identified substantial component-level convergence around runtime authorization, effect-time governance, delegation continuity, path/trajectory governance, proof-carrying actions, stateful policy, and behavioral constitutions. It did not identify a single authenticated source in the searched record that materially instantiates the entire HRI chain as one unified invariant: decision-environment integrity → human authority lineage → semantic epistemic/authority separation → verification/admissibility separation → commit-time freshness → causal effect closure → trajectory composition → constitutional proxy amendment → reflexive recovery without self-legislation.

Interpretation. HRI is therefore under HIGH prior-work pressure at the component level, especially Modules 4, 8, 9 and 10, while the searched record leaves a residual systems-level distinction around authority formation, legitimacy/admissibility separation, proxy constitutional amendment, and the full cross-module composition. This is a search result, not a novelty, priority, freedom-to-operate, or patentability conclusion.

1. Search scope and evidentiary discipline

Searches were run on 12 August 2026 across current technical literature (primarily author-posted arXiv records) and a targeted Google Patents reconnaissance. Search concepts included agent authorization continuity, path governance, delegated authority, proof-carrying actions, irreversible-effect boundaries, stateful governance, constitutional governance, capability mutation, approval/attestation, and AI-agent validation. Patent searching was exploratory rather than claim-chart exhaustive.

Chronology is evaluated only from the public dates visible in the searched records. HRI conception chronology and the effective prior-art date for any patent claim are not adjudicated here. A patent opinion requires claim-specific searching against the correct filing/priority date and legal analysis by qualified counsel.

2. Collision scale

Level	Meaning
NEAR-COLLISION	Very close to an HRI mechanism or formulation at the same abstraction level; residual distinction remains material.
HIGH	Strong overlap with one or more HRI modules/mechanisms; reviewer or examiner pressure is likely.
MODERATE	Conceptual or technical precursor that narrows novelty claims but does not substantially instantiate the HRI mechanism.
LOW	Analogous or background work; useful context but weak collision.

3. Academic / technical collision matrix

Source	HRI modules	Collision	Overlap	Residual distinction	Release implication
Zhang & Zhang — Earned Authority under a Fixed Ceiling (arXiv:2607.23586, 2026-07-26)	M4, M9, M10	NEAR-COLLISION	Authorization continuity under agent evolution; immutable effect ceiling; transition envelope; complete mediation; attenuating delegation; authority cannot expand beyond user-issued ceiling.	Begins from an existing grant; does not establish decision-environment integrity, legitimacy/admissibility split, semantic epistemic→authority firewall, or proxy constitutional amendment.	Cite prominently; frame as strongest effect-time neighbor.
AgentBound — Verifiable Behavioral Governance (arXiv:2606.30970, 2026-06-29)	M1, M9, M10	HIGH	Delegated authorization + owner-signed behavioral constitutions + site contracts; formal conservative composition; governance receipts; standing delegation.	No searched equivalent of HRI authority-formation threat model, three-state authority semantics, DEI, or full proxy-amendment rule.	Treat as major constitutional-runtime neighbor.
Proof-Carrying Agent Actions (arXiv:2606.04104, 2026-06-02)	M1, M9	HIGH	Portable action certificates; pre-action admissibility; explicit approval semantics; evidence and replay; runtime-neutral effect governance.	Does not unify upstream human decision integrity, full cross-module semantic firewall, or constitutional substrate mutation.	Strong pressure on commit certificates / proof-carrying effect boundary.
Verifiable Agentic Infrastructure / DTF (arXiv:2605.15228, 2026-05-13)	M1, M7, M9	HIGH	Proof-derived execution authority; justification proof; independent consensus; ephemeral execution identity; append-only evidence chain.	HRI rejects epistemic strength/consensus as sufficient authority absent human/institutional lineage; DTF does not supply HRI DEI/proxy mesh.	Explicitly distinguish proof-derived admissibility from legitimacy.
Governing Actions, Not Agents (arXiv:2606.26298, 2026-06-24)	M9	NEAR-COLLISION	Governance at irreversible side-effect boundary rather than reasoning/planning; preserves agent autonomy; institutional attestation.	Does not cover HRI's full authority-formation, trajectory, proxy amendment, and reflexive-recovery layers.	Closest public neighbor to consequence-boundary / Chic Boundary intuition.
The Unfireable Safety Kernel (arXiv:2606.26057, 2026-06-24)	M9	HIGH	Execution-time alignment; agent-inaccessible enforcement layer constraining what the agent may do at action time.	Focus is unbypassable action constraint, not human authority lineage, DEI, or constitutional continuity.	Pressure on non-bypassable execution frontier; retain HRI semantic/authority distinctions.
Runtime Governance for AI Agents: Policies on Paths (arXiv:2603.16586, 2026-03-17)	M9, M10	HIGH	Execution path as governance object; policies over partial path, next action, identity and organizational state.	Does not instantiate the HRI human-authority semantic type system or approval-laundering/proxy-amendment architecture.	Strong prior work for "local validity is not compositional" / trajectory framing.
Stateful Governance for Concurrent Agentic Systems (arXiv:2608.02764, 2026-08-03)	M9, M10	HIGH	Governance facts must persist across actions; stateful policies for concurrency and history-dependent constraints.	No searched full authority legitimacy / decision formation / constitutional proxy model.	Very recent pressure on trajectory/stateful-governance layer.
Authorization Propagation in Multi-Agent AI Systems (arXiv:2605.05440, 2026-05-06)	M1, M4	HIGH	Maintaining authorization invariants as agents retrieve, delegate, and synthesize across boundaries; not reducible to prompt injection or classic access control.	Does not cover HRI's epistemic authority typing, decision-environment integrity, causal effect closure, or constitutional mutation.	Strong precursor for delegation continuity.
Think Before You Act — Neurocognitive Governance (arXiv:2604.25684, 2026-04-28)	M8, M10	MODERATE/HIGH	Pre-action governance reasoning loop; layered rules embedded into agent deliberation; empirical compliance evaluation.	Philosophically differs from HRI where successful reasoning/governance cognition does not itself confer execution authority; relies on internalized self-governance.	Useful contrast case: cognitive governance versus authority separation.

Source	HRI modules	Collision	Overlap	Residual distinction	Release implication
Proof of Execution (arXiv:2607.05397, public record 2026)	M9	HIGH	Runtime proof of execution; authorization/path compliance; causal event stream; replay; distinct planning/enforcement/effect/record planes.	No upstream authority formation / DEI; no full cross-module constitutional mesh.	Pressure on causal trace, effect closure and auditability.
AgentCity — Constitutional Governance for Autonomous Agents (arXiv:2604.07007, 2026-04-08)	M10	MODERATE	Constitutional governance, separation of powers, human adjudication via ownership chain.	Allows agents to legislate operational rules; HRI explicitly forbids machine self-legislation for protected constitutional behavior without valid human activation.	Important conceptual neighbor and useful point of contrast.
FAVA — Formal Authorization for Verified Agents (arXiv:2607.27267, 2026-07-29)	M1, M9	HIGH	Permission-carrying authorization under dynamic runtime state/data flows; static tool permissions deemed insufficient.	No searched DEI, legitimacy split, proxy amendment, or full HRI lifecycle composition.	Late-July near-field technical pressure; include in ongoing search watch.

4. Module-level pressure map

Module	Pressure	Why	Residual distinction
Module 1 — Human-rooted authority validation	HIGH	Delegation/authorization, proof-carrying actions, proof-derived execution authority, permission-carrying frameworks are crowded.	Residual distinction: three-state token/admissibility/legitimacy semantics and human-rooted authority formation boundary.
Module 2 — Human reasoning reconstruction	LOW/MODERATE	Many memory/persona/reconstruction systems exist, but this search did not identify a close governance collision.	Residual distinction: reconstructed human model is explicitly non-authoritative.
Module 3 — Judgment drift	MODERATE	Drift/trust monitoring is established; governance-specific judgment decomposition remains less directly collided in this search.	Avoid broad novelty claims over drift detection itself.
Module 4 — Distributed authority continuity	HIGH	Zhang & Zhang, authorization propagation, recursive delegation and stateful governance create strong pressure.	Residual distinction: human unavailability does not manufacture machine authority; continuity without sovereignty.
Module 5 — Human-accountable IP intelligence	LOW in this search	General AI-assisted IP/prior-art systems exist; this agent-governance search did not test this family exhaustively.	Requires separate patent/IP intelligence prior-art search.
Module 6 — Causal intellectual attribution	LOW in this search	Authorship/contribution literature exists but was outside current agent-governance search scope.	Requires separate authorship/provenance landscape.
Module 7 — Judgment measurement / weight separation	MODERATE	Consensus and independent-evaluator systems are common; explicit separation of epistemic/authority/participation/execution weights was not matched directly here.	Do not claim novelty of ensembles/consensus; focus on non-collapse of weights.
Module 8 — Runtime cognitive operations	MODERATE/HIGH	Pre-action governance reasoning and runtime reasoning-control work is active.	Residual distinction: verified reasoning remains epistemic and cannot self-confer external authority.
Module 9 — Causal authority closure	VERY HIGH	This is the most crowded area: action/effect boundaries, proof-carrying actions, safety kernels, execution proofs, fixed ceilings, permission-carrying authorization.	Residual distinction must be system-level: topology-independent causal closure tied to current human authority lineage and semantic firewall.

Module	Pressure	Why	Residual distinction
Module 10 — Reflexive constitutional mesh	HIGH	Behavioral constitutions, stateful governance, constitutional agent societies and path governance create strong pressure.	Strongest residuals: DEI/approval laundering before grant; material proxy amendment; no machine self-legislation; full cross-module compatibility.

5. Strongest residual HRI synthesis after this search

The searched record supports a conservative residual formulation rather than a historical-first claim:

> **No authenticated source identified in this search was shown to integrate, as one architecture,** human decision-environment integrity and approval-laundering resistance; exogenous legitimacy versus machine-checkable authority admissibility; semantic prohibition on epistemic-to-authority casts; current human authority lineage and delegation; commit-time freshness; topology-independent causal effect closure; composition-safe trajectory governance; material constitutional proxy amendment; and reflexive recovery that cannot self-legislate protected authority.

This is a negative search observation over the searched record, not proof of novelty. Search failure must never be converted into a historical-first or patentability statement.

6. Patent reconnaissance — targeted, non-exhaustive

Patent / publication	Date signal	Relevant subject	Collision	Preliminary observation
US20260017525A1 — Validating autonomous AI agents	2026 publication	Agent action validation / unpredictable autonomous behavior	MODERATE	Validation of autonomous actions is relevant background; quick review did not show the full HRI authority-continuity chain.
US12111754B1 — Dynamically validating AI applications for compliance	Granted	Guideline-boundaries, test generation, re-evaluation	LOW/MODERATE	Pressure on compliance testing, not HRI authority lineage/effect closure.
US20240022607A1 — Automated/adaptive model-driven security; ZBAC delegation	2024 publication	Delegable authorization tokens; revocation / application-chain delegation	MODERATE	Classic delegation/token pressure on Module 1/4; not HRI system-level synthesis.
US20250335458A1 — Contextual orchestration and scoped memory protocol	2025 publication	Human-in-loop / user sovereignty / scoped memory	LOW/MODERATE	Conceptual human-sovereignty overlap; no quick-search collision with CCC invariant.
US20250259042A1 / US20250259044A1 — scalable collaborative AI-agent orchestration	2025 publication	Secure token communication and multi-agent orchestration	LOW	General orchestration background, not human authority continuity.
US20260081772A1 — Memory-native identity in distributed environments	2026 publication	Dynamic agent identity / policy context	MODERATE	Relevant to identity continuity but not the full HRI authority framework.

Patent-search caution: this is reconnaissance only. It does not include CPC/IPC class expansion, family tracing, prosecution-history review, non-English searching, claim-element charts, priority-chain validation, or search against each HRI claim's legally relevant critical date. Do not use this table as a freedom-to-operate or patentability opinion.

7. Release-safe statements

Status	Suggested wording
SAFE	“A current landscape search found strong neighboring architectures for effect-time authorization, path governance, delegated authority, proof-carrying actions, and behavioral constitutions.”
SAFE WITH QUALIFIER	“In the searched record, no single source was shown to instantiate the complete Constitutional Causal Continuity chain.”
DO NOT SAY	“HRI is the first architecture to solve AI authority.”
DO NOT SAY	“No prior art exists.”
DO NOT SAY	“The search proves novelty or patentability.”

8. Highest-priority watch list after release

Monitor the following because they are closest to HRI’s publishable technical center: Zhang & Zhang (fixed ceiling / authorization continuity); AgentBound; Proof-Carrying Agent Actions; Governing Actions, Not Agents; Stateful Governance for Concurrent Agentic Systems; FAVA; DTF; Proof of Execution; and any new work explicitly combining decision formation with effect-time authorization or behavioral-constitution mutation.

9. Source register

1. Z. Zhang and X. Zhang, “Are You Still the Agent I Authorized? Earned Authority under a Fixed Ceiling for Evolving Agents,” arXiv:2607.23586, 2026.
2. A. Kaul, Q. Lan, and P. Gupta, “AgentBound: Verifiable Behavioral Governance for Autonomous AI Agents,” arXiv:2606.30970, 2026.
3. Z. Wang, “Proof-Carrying Agent Actions: Model-Agnostic Runtime Governance for Heterogeneous Agent Systems,” arXiv:2606.04104, 2026.
4. J. He and D. Yu, “Verifiable Agentic Infrastructure: Proof-Derived Authorization for Sovereign AI Systems,” arXiv:2605.15228, 2026.
5. “Governing Actions, Not Agents: Institutional Attestation as a Computational Governance Primitive for Agentic AI,” arXiv:2606.26298, 2026.
6. “The Unfireable Safety Kernel: Execution-Time AI Alignment for Autonomous Agent Systems,” arXiv:2606.26057, 2026.
7. M. Kaptein, V.-J. Khan, and A. Podstavnychy, “Runtime Governance for AI Agents: Policies on Paths,” arXiv:2603.16586, 2026.
8. “Stateful Governance for Concurrent Agentic Systems,” arXiv:2608.02764, 2026.
9. “Authorization Propagation in Multi-Agent AI Systems,” arXiv:2605.05440, 2026.
10. E. Bandara et al., “Think Before You Act — A Neurocognitive Governance Model for Autonomous AI Agents,” arXiv:2604.25684, 2026.
11. J. Rhodes and G. Kang, “Proof of Execution: Runtime Verification for Governed AI Agent Actions,” arXiv:2607.05397, 2026.
12. “AgentCity: Constitutional Governance for Autonomous Agent Societies,” arXiv:2604.07007, 2026.
13. “FAVA: Formal Authorization for Verified Agents with Permission-Carrying Execution,” arXiv:2607.27267, 2026.
14. Google Patents records: US20260017525A1; US12111754B1; US20240022607A1; US20250335458A1; US20250259042A1/US20250259044A1; US20260081772A1 (accessed 2026-08-12).

Final epistemic boundary

CURRENT SEARCH RESULT: No full-system collision identified in the searched record; substantial and in several cases near-collision prior work exists for individual HRI mechanisms. The defensible scholarly position is candidate system-level synthesis under active convergence, not primitive novelty or historical priority.