Universal Library of Innovative Research and Studies Research Article

ISSN: 3065-0003 | Volume 3, Issue 3

Open Access | PP: 14-19

DOI: https://doi.org/10.70315/uloap.ulirs.2026.0303003

Mapping C3P Control Objectives to ISO/IEC 42001 Requirements for AI Lifecycle Governance

Paul Gresham

Founder, Paul Gresham Advisory LLC, Purchase, NY, USA.
ORCID: 0009-0005-2975-1187

Abstract

This article examines how the four control objectives of the Continuous Compliance Control Protocol align with the requirements of ISO/IEC 42001 for managing an artificial intelligence system across its lifecycle, and how they generalize to autonomous agents. Organizations adopt artificial intelligence in regulated settings faster than they build the mechanisms that continuously prove governance, and a management system standard states what must be governed, while leaving open how durable evidence is produced as systems run. A baseline mapping establishes that the four control points meet the requirements of the standard for a conventional, human-in-the-loop deployment. The central finding is that the same four points generalize without modification to AI coding agents and to general agentic workflows, where a defined purpose, an auditable trail, verification, and safe delivery produce tamper-evident evidence at every stage. That evidence assembles into a chain of custody for the agent that supplies exactly the documentation an ISO/IEC 42001 audit expects. The work serves compliance leaders, audit professionals, and engineering teams who govern autonomous artificial intelligence.

Keywords: *AI governance, ISO/IEC 42001, Agentic AI, AI Coding Agents, Continuous Compliance, Chain of Custody, Auditability, AI lifecycle.*

INTRODUCTION

Artificial intelligence has moved from experimental use into the production delivery pipelines of regulated industries, where models generate code, agents execute multistep tasks, and inference services make decisions that carry legal and financial weight. Governance standards have appeared to address this shift, with ISO/IEC 42001:2023 establishing the first international management system standard for artificial intelligence (International Organization for Standardization, 2023). The pace of standard publication has outrun the pace at which organizations build the concrete mechanisms that demonstrate conformance while a system operates, and this gap is where the present analysis is situated.

A management system standard specifies what an organization must establish, document, and review, and organizes these duties into clauses for context, leadership, planning, support, operations, performance evaluation, and improvement, supported by an annex of controls aimed at risks specific to artificial intelligence (Gueorguiev, 2025). What the standard does not provide is an operational control plane that produces verifiable evidence as a continuous output of the delivery process. The Continuous Compliance Control Protocol, hereafter C3P, was introduced as exactly such a control plane for regulated software delivery, defining what must remain true for governance, auditability, integrity, and traceability across a delivery lifecycle (Gresham, 2026). C3P was designed for human-led delivery, and its author identified the intersection with ISO/IEC 42001 and the extension to artificial intelligence as the next line of inquiry.

This article carries that line of inquiry forward through a structured mapping. The contribution is threefold and analytical in character. First, each of the four C3P control objectives is decomposed into verifiable properties and matched against specific clauses and annex controls of ISO/IEC 42001. Second, every correspondence is classified by the degree to which the existing control plane meets the requirement, separating direct coverage from cases that require extension. Third, a concept set of new managed artifacts is defined for artificial intelligence, since models, datasets, prompts, agentic actions, and inference chains do not map cleanly onto artifacts conceived for human authorship. The remainder of the article reviews the relevant frameworks, describes the mapping method, presents the correspondence objective by objective, discusses where coverage holds and where extension is required, and states limitations and directions for further work.

Citation: Paul Gresham, "Mapping C3P Control Objectives to ISO/IEC 42001 Requirements for AI Lifecycle Governance", Universal Library of Innovative Research and Studies, 2026; 3(3): 14-19. DOI: https://doi.org/10.70315/uloap.ulirs.2026.0303003.

BACKGROUND AND RELATED FRAMEWORKS

C3P is treated here as a known framework that the present work builds upon. It defines four load-bearing control objectives for regulated delivery, namely traceability of every deployment to an authorized business change, integrity and provenance such that reviewed code is the code that is built and deployed, immutable history that makes tampering evident, and auditability by default through evidence produced continuously and queried on demand (Gresham, 2026). A further design principle separates the control plane from day-to-day development discretion, so that no single actor holds permission to break the chain of custody. These elements ground the mapping and stay outside the set of findings.

ISO/IEC 42001:2023 supplies the requirements against which those objectives are measured. The standard follows the plan-do-check-act cycle common to management system standards, and it places artificial intelligence risk assessment and risk treatment within its planning clause, system operation within its operation clause, and monitoring, internal audit, and management review within its performance evaluation clause (International Organization for Standardization, 2023). Its annex enumerates controls for data governance, transparency, accountability, and the management of the artificial intelligence system lifecycle. The standard is certifiable, which means an external party can audit conformance, and that auditability raises the stakes for evidence that can be inspected and trusted.

Two adjacent instruments frame the regulatory context and stay outside the objects of the mapping. The NIST AI Risk Management Framework frames risk work as govern, map, measure, and manage, and serves many organizations as a voluntary baselining that is consistent with emerging law (Tabassi, 2023). The European Union Artificial Intelligence Act creates requirements for high-risk systems that include conformity assessment prior to deployment, and monitoring after deployment during the life cycle of the system. Scholarship that reads the act through the lens of auditing considers the feasibility of these requirements for providers (Mökander et al., 2022). What unites these instruments is a reliance on continuous, inspectable evidence that none of them operationalize at the artifact level, which is the space the mapping addresses.

METHODS

The study is an analytical mapping and contains no empirical data, experiments, or quantitative results. The reasoning rests on the normative texts, the published C3P framework, and peer-reviewed literature on artificial intelligence governance, provenance, and auditing. The procedure proceeds through four stages that build on one another.

The first stage decomposes each C3P control objective into a small set of verifiable properties, thereby converting a broad objective into testable statements about the evidence that must exist and the relationships that must hold. The second stage matches each property against the clauses and annex controls of ISO/IEC 42001 that govern the same concern, so that traceability is read against planning and lifecycle controls, provenance against operation and data controls, immutable history against documented information and records requirements, and auditability against performance evaluation. The third stage classifies every correspondence into one of three categories: direct coverage, where the control plane already satisfies the requirement; partial coverage, where it satisfies the requirement for conventional artifacts but not for artificial intelligence-specific ones; and extension, where a new artifact class or a new control must be introduced. The fourth stage records the new managed artifacts introduced by artificial intelligence, since the presence of models, datasets, prompts, agentic actions, and inference chains determines where extension is necessary. Table 1 summarizes the categories that the third stage applies.

Table 1. Classification categories used in the mapping

Category	Meaning in the mapping
Direct coverage	The C3P control plane already satisfies the ISO/IEC 42001 requirement without modification.
Partial coverage	The control plane satisfies the requirement for conventional artifacts but not for artificial intelligence specific artifacts.
Extension	A new managed artifact class or control objective must be introduced to satisfy the requirement for artificial intelligence.

This procedure yields a transparent record of where an existing control plane meets a management system requirement and where artificial intelligence forces the control plane to grow, which is the substance of the sections that follow.

RESULTS

Traceability

The traceability objective requires that every deployment trace to an authorized business change and that every change trace to an approved requirement. Read against ISO/IEC 42001, this property aligns with the planning clause that governs artificial intelligence risk assessment and with the lifecycle controls that require an organization to know what enters a system and why (International Organization for Standardization, 2023). For conventional code changes, the control plane already binds a deployment to its authorizing record, which is direct coverage. The requirement becomes partial once artificial intelligence enters, because a model update or a dataset revision can change system behavior without a corresponding code change, and the authorization

chain must now reach the model and the data. Extension links the business change to model versions, dataset revisions, prompt definitions, and the agentic actions that a system performs on its own initiative.

Integrity and Provenance

The integrity and provenance objective requires that the reviewed artifact be the deployed artifact and that artifacts and approvals be attributed to specific identities. ISO/IEC 42001 addresses this concern through its operational clause and annex controls for data governance and the artificial intelligence system lifecycle (International Organization for Standardization, 2023). The control plane already provides cryptographic or procedural verification for build artifacts, which is direct coverage for code. Coverage turns partial for artificial intelligence because a model carries provenance that a binary does not, including the training data lineage, the training configuration, and the version of the model that served a given inference. Research on machine learning pipelines shows that artifact management systems capture only rudimentary provenance unless lineage is modeled explicitly across data, code, and model events (Schlegel & Sattler, 2025). Extension makes model version, training-data lineage, and prompt and inference provenance first-class links in the artifact graph, so that the chain that once ran from requirement to binary now reaches the model and the data that shaped its outputs.

Immutable History and Tamper-Evidence

The immutable history objective requires that records be retained as journaled versions and that history be resistant to silent rewriting. ISO/IEC 42001 expresses the same demand through its requirements for documented information and record control, which obligate an organization to retain evidence that remains trustworthy over time (International Organization for Standardization, 2023). The control plane satisfies this requirement for change records and requirement references through direct coverage. The requirement becomes partial once a system includes agents, because agents generate a stream of actions, and an inference chain produces intermediate states that a human workflow never creates. An agent can also overwrite its own working memory, thereby erasing the very trail that tamper-evidence depends upon. The literature on auditing autonomous systems argues that every agent action needs a logged trail that records who initiated it and why, and is retained so that outcomes can be traced to the responsible agent (Phiri, 2025). Extension therefore journals agentic actions and inference chains as append-only versions that cannot be silently rewritten.

Auditability by Default

The auditability objective requires that evidence be produced continuously and that exceptions remain explicit and reviewable. This property can be mapped to the performance evaluation clause of ISO/IEC 42001 on monitoring, internal audit, and management review, as organizations can provide evidence of a system's behavior across its lifetime (International Organization for Standardization, 2023). The continuous evidence model of the control plane provides a direct mapping to the structural requirement because evidence is generated as a by-product of delivery. However, coverage is partial, since e.g. model drift, input distribution shift, or agent emergent behavior are not accounted for by any continuing monitoring. In conformity assessment and post-market monitoring of high-risk systems, regulators deem evidence of performance at all stages of a system's lifecycle as necessary, thus raising the bar for what continuous evidence must include (Mökander et al., 2022). The extension adds continuous evidence of artificial intelligence behavior, drift signals, and bounded exceptions for autonomous action to the evidence that the control plane already generates.

Table 2 consolidates the four objectives, the clauses and controls that each addresses, the coverage classification, and the artificial intelligence artifact that extension introduces.

Table 2. Mapping of C3P control objectives to ISO/IEC 42001 requirements

C3P objective	ISO/IEC 42001 clause or annex control	Coverage	AI artifact added by extension
Traceability	Planning and AI risk assessment with lifecycle controls	Direct to partial	Model version, dataset revision, prompt definition, agentic action
Integrity and provenance	Operation with data governance and lifecycle controls	Direct to partial	Training-data lineage, model provenance, inference provenance
Immutable history and tamper-evidence	Documented information and control of records	Direct to partial	Journaled agent actions, journaled inference chains
Auditability by default	Performance evaluation covering monitoring, internal audit, management review	Direct to partial	Drift and behavior evidence, bounded autonomy exceptions

The Extended Artifact Graph

The four extensions converge on a single structural change to the artifact graph that underpins every objective. A graph built for human authorship runs from a business change record through a requirement, a code change, a build artifact, and a deployment, culminating in the evidence pack the pipeline produces. Artificial intelligence inserts new nodes into that chain where models, data, prompts, and autonomous actions shape what a system does, so that the chain that once ran from requirement to binary now reaches the model, the data, and the autonomous actions behind a given output.

This baseline mapping shows that C3P meets the general

requirements of ISO/IEC 42001 for a conventional, human-in-the-loop deployment, and it exposes the limit of that reading, since the control points were written for delivery that a person authorizes, reviews, and releases. A general deployment exercises only part of what the framework can govern, and the case where it earns its keep is the one where autonomy removes the human from each step, which is the subject of the next section.

GENERALIZING C3P TO AGENTIC AI

The baseline mapping treats artificial intelligence as one more artifact inside a pipeline that a person still drives. Agentic systems break that assumption, because a software agent pursues an objective, takes actions, and produces output with little or no human intervention at each step. The control questions a reviewer answers by hand, namely what the work was for, what was done, whether the result is correct, and whether it is safe to release, now have to be answered by the system about its own behavior. The four control points of C3P translate into that setting without modification, and the same evidence that satisfies a human reviewer satisfies an ISO/IEC 42001 auditor. The strongest case for the framework is therefore the governance of autonomous agents, where the four points generalize into a reusable pattern whose documented evidence maps onto the clauses an audit examines.

The Four Control Points Generalized for Agents

The first point is a defined, agreed purpose. An agent operates against an explicit objective and a bounded mandate that fix what it may attempt and where its authority ends. This corresponds to the objectives an organization sets under Clause 6.2, the operational planning and control of Clause 8.1, and the scoping of intended use and impact under Clause 8.4, with the responsible-development aims of Annex A and the intended-use control of A.9 framing what a legitimate mandate contains (International Organization for Standardization, 2023). A purpose recorded before execution becomes the reference against which every later check is measured.

The second point is that the agent is tracked and auditable. Every action is recorded in a tamper-evident trail, version-controlled wherever the medium allows, so that the sequence of decisions can be reconstructed after the fact. This evidentiary backbone of certification corresponds to the monitoring and measurement of Clause 9.1 and to the life-cycle records, traceability, and change control of Annex A.6, which feed the performance evaluation of Clause 9 and the improvement activity of Clause 10. An agent that overwrites its own working state without journaling it destroys the record on which this point depends, which is why append-only logging is the operative mechanism.

The third point is verification that the agent delivered what was expected. A check that is automated, human, or both confirms that the output matches the agreed purpose before the work counts as complete, which corresponds to the risk treatment of Clause 8.3, the impact assessment on change of Clause 8.4, and the Annex A controls on verification, validation, model evaluation, and human oversight under A.9. Verification closes the loop opened by the first point, since the agreed purpose supplies the criterion the check applies.

The fourth point is safe delivery of the resulting output. The output is released through controlled channels with appropriate oversight and a path to roll back, which corresponds to the controlled deployment of Clause 8.1, responsible use under A.9, the provision of information for interested parties under A.8, the management of third-party and customer relationships under A.10, and the incident-handling controls that govern a release. Safe delivery ensures that the autonomy granted by the first point does not extend to an uncontrolled effect on the world.

AI Coding Agents

The clearest intersection is the coding agent, a system that writes or modifies source code on its own initiative. Code is an artifact whose purpose, audit trail, verification, and safe delivery are non-negotiable in a regulated setting, which are precisely the four control points. Wrapping a coding agent in C3P automates the controls a human engineer would otherwise apply by discipline. The agreed purpose binds a change to an authorized objective before the agent writes a line. The audit trail captures each edit as a version-controlled record, which source control already supports and which yields tamper-evident history at no extra cost. Verification runs the tests, reviews, and evaluations that confirm the change does what the purpose required, with a human gate where the risk warrants one. Safe delivery routes the result through the same controlled pipeline that governs human contributions, including the ability to revert. The agent inherits the control plane that already governs the repository, and the evidence it generates is the evidence an ISO/IEC 42001 audit expects for a change to a managed system.

Agentic Workflows

Beyond coding, the same four points generalize to any autonomous task execution, where an agent plans and carries out a sequence of steps toward a goal. Here the control points become a reusable governance pattern that no longer depends on a code-specific discipline. The agreed purpose and bounded mandate define the task and the limits of the agent's authority, the audit trail journals the actions and intermediate states, verification confirms that the result meets the goal, and safe delivery releases the outcome through a controlled channel with oversight and rollback. A workflow built this way carries its governance with it, and the pattern is reusable because it depends on the structure of autonomous execution and not on the particular task, which lets an organization apply one control model across a fleet of agents.

Across both cases the result is the same. The four control points produce tamper-proof evidence at each stage, and that evidence assembles into a chain of custody for the agent, whether it writes code or executes a general workflow. Figure 1 shows how the generalized control points anchor to the clauses of ISO/IEC 42001 and how the evidence each point emits collects into a tamper-evident ledger that yields the chain of custody an audit examines.

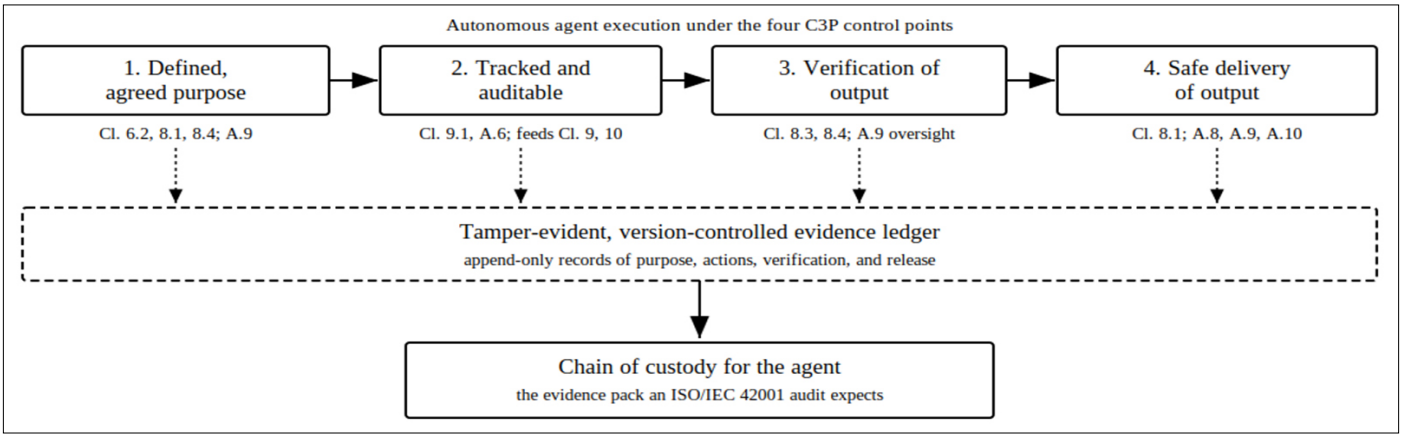

Fig. 1. Chain of custody for an autonomous agent governed by the four generalized C3P control points

The chain of custody converts a claim of governance into an artifact an auditor can inspect. Each control point contributes a record, the records are bound together in order, and the resulting evidence pack answers the questions of purpose, action, verification, and release for the agent's entire run. This is the concrete contribution of the framework to agentic governance, since execution itself produces the documentation that ISO/IEC 42001 certification requires.

DISCUSSION

The mapping shows a consistent pattern across all four objectives. For a general deployment the control plane meets the structural requirements of ISO/IEC 42001, and the value it adds is to convert point-in-time attestation into a standing output that answers the performance evaluation clause with evidence that exists by construction. The agentic setting of the previous section is where this property becomes decisive, because an autonomous agent offers no human reviewer whose sign-off could stand in for the evidence, and the control plane is the only thing that produces a record of purpose, action, verification, and release.

The point of genuine novelty is that the four control points written for human-led delivery carry over to autonomous agents without modification, and that satisfying them yields the documented evidence an ISO/IEC 42001 audit expects. A coding agent and a general workflow agent differ in what they produce, yet both are governed by the same four points and both emit the same chain of custody. This is a stronger claim than the baseline mapping, since it shows the framework doing work a human process cannot do for itself once the human is removed from the loop, and it is the result that carries practical weight for organizations deploying agents under a management system standard.

Two principles from observability practice in artificial intelligence infrastructure illustrate the same logic without standing as empirical evidence. Open telemetry projects that surface low-level system signals demonstrate that behavior must be made visible before it can be governed, and that separating the plane that records signals from the plane that produces them protects the integrity of the record. The reference implementation for C3P demonstrates that commodity tooling can implement control-plane separation without proprietary infrastructure (Gresham, 2026). These examples connect the abstract requirement for continuous evidence to concrete mechanisms that already exist.

The analysis carries limitations that bound its claims. It is a conceptual mapping without empirical validation, so the categories in Table 2 represent reasoned correspondence. The reading of ISO/IEC 42001 relies on the published clause and annex structure, and a certified implementation might surface correspondences that a textual reading does not. The practitioner experience that informs the source framework comes from globally regulated financial institutions, and the specific regulatory requirements of other regulated sectors differ enough that the mapping would need adjustment before transfer.

CONCLUSION

This article mapped the four control objectives of the Continuous Compliance Control Protocol onto the lifecycle requirements of ISO/IEC 42001 and then carried the mapping into the agentic setting, where the same four points govern autonomous coding agents and general agentic workflows. The baseline mapping establishes that C3P meets the requirements of the standard for a conventional deployment. The agentic generalization is the substantive contribution, since the four points apply without modification to systems that act on their own initiative, and the tamper-evident evidence they produce assembles into a chain of custody that supplies exactly the documentation a certification audit examines.

Several directions remain open. Empirical validation against a certified ISO/IEC 42001 implementation would test the mapping with measured outcomes. A detailed treatment

of coding-agent controls would document the nuances of wrapping a repository pipeline around an autonomous author, and a treatment of multi-agent orchestration would address the diffusion of responsibility that arises when many agents coordinate on one task. A mapping onto the European Union Artificial Intelligence Act and the NIST AI Risk Management Framework would situate the pattern within the wider regulatory landscape. Each direction extends the central claim that durable, continuous evidence is the foundation on which the governance of autonomous artificial intelligence must rest.

REFERENCES

1. Gresham, P. (2026). C3P: A Continuous Compliance Control Protocol for regulated software delivery. *International Journal of Advanced Engineering, Management and Science, 12*(3), 245–253. https://doi.org/10.22161/ijaems.123.22
2. Gueorguiev, T. (2025). An approach to integrate Artificial Intelligence in ISO 9001-based quality management systems. *Measurement: Sensors, 38*, Article 101787. https://doi.org/10.1016/j.measen.2024.101787
3. International Organization for Standardization. (2023). *Information technology — Artificial intelligence — Management system* (ISO/IEC 42001:2023). https://www.iso.org/standard/42001
4. Mökander, J., Axente, M., Casolari, F., & Floridi, L. (2022). Conformity assessments and post-market monitoring: A guide to the role of auditing in the proposed European AI regulation. *Minds and Machines, 32*(2), 241–268. https://doi.org/10.1007/s11023-021-09577-4
5. Phiri, C. C. (2025). Creating characteristically auditable agentic AI systems. In *Proceedings of the Intelligent Robotics FAIR 2025 (IntRob '25)*. https://doi.org/10.1145/3759355.3759356
6. Schlegel, M., & Sattler, K.-U. (2025). Capturing end-to-end provenance for machine learning pipelines. *Information Systems, 132*, Article 102495. https://doi.org/10.1016/j.is.2024.102495
7. Tabassi, E. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (Report No. NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1