Network Working Group
Internet-Draft
Intended status: Informational
Expires: 20 September 2026

Y. Wang
HJS Foundation Ltd.

21 March 2026

Co-Evolve Binding Protocol-00 (CEP)
draft-wang-cep-00

Abstract

This document specifies the Co-Evolve Binding Protocol (CEP), a lightweight, minimal-structured protocol designed to bind human cognitive evolution with artificial intelligence (AI) evolution. CEP works in conjunction with the Judgment Event Protocol (JEP) and Human Judgment Structure (HJS) as the default collaborative system, while supporting adaptation to other compliant judgment and accountability systems to enhance interoperability and practicality. Leveraging JEP’s strengths in standardized judgment recording and lightweight verifiability, HJS’s advantages in structured human responsibility management and end-to-end accountability traceability, and the flexibility to adapt to other systems, CEP integrates core advantages to form a complementary three-layer human-AI symbiosis infrastructure. The core goal of CEP is to prevent unilateral AI evolution, maintain human sovereignty in the co-evolution process of humans and AI, and lay the foundation for trustworthy human-AI symbiosis.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 20 September 2026.

Copyright Notice

Table of Contents

1. Introduction

1.1. Background

With the rapid development of AI technology, especially the emergence of large language models and autonomous agents, AI is increasingly moving towards self-evolution, self-optimization, and closed-loop reasoning. This trend brings the risk of AI divergence from human cognitive logic, value systems, and responsibility boundaries—leading to a "civilization bifurcation" between human civilization and silicon-based intelligence.

To address this risk, the Judgment Event Protocol (JEP) [I-D.wang-jep-judgment-event-protocol] provides a minimal, verifiable format for logging judgment-related actions, ensuring standardized human participation in key AI decision-making processes with its lightweight design and replay protection mechanism. The Human Judgment Structure (HJS) [I-D.wang-hjs-judgment-structure], as an infrastructure-grade accountability layer built on JEP, complements JEP's advantages by providing a structured framework for human responsibility management, cryptographic accountability mechanisms, and a three-tier privacy architecture, ensuring traceability of causality and accountability in AI operations. However, there is a lack of a dedicated protocol to bind human and AI evolution, leading to potential unilateral AI evolution and loss of human sovereignty in the evolution process—CEP is designed to fill this gap by integrating JEP and HJS's core strengths, while supporting adaptation to other compliant judgment and accountability systems to meet diverse practical deployment needs.

1.2. Objectives

CEP is designed to achieve the following core objectives, aligning with JEP, HJS and other compliant systems, and integrating their strengths to form a three-layer human-AI symbiosis infrastructure with strong adaptability:

1. Bind AI evolution to human cognitive evolution, preventing unilateral AI evolution and civilization bifurcation, leveraging JEP's standardized judgment records, HJS's structured accountability, and the adaptability to other systems to ensure traceable and controllable co-evolution.

2. Provide minimal-structured, extensible fields to cover all AI evolution scenarios (e.g., fine-tuning, emergence, drift, loop formation), inheriting JEP's minimal design concept, HJS's extensible architecture, and supporting flexible adaptation to other systems without modifying the core structure.

3. Ensure seamless compatibility with JEP, HJS, and other compliant judgment/accountability systems, integrating their verification mechanisms and accountability traceability capabilities to enable end-to-end human-AI co-evolution, decision-making, and accountability.

4. Be lightweight, easy to implement, and compatible with existing AI systems, network protocols, and W3C standards, adhering to JEP's lightweight engineering design, HJS's infrastructure-grade compatibility requirements, and the practical deployment needs of other systems.

1.3. Scope

CEP applies to all AI systems, large models, and intelligent agents that involve self-evolution, self-optimization, or capability emergence. It focuses on the binding of human and AI evolution processes, without involving AI model training, algorithm design, or hardware implementation. Drawing on JEP's non-intrusive design, HJS's middleware-style deployment, and the flexible adaptation capability to other systems, CEP works as a middleware layer, inserting into the evolution process of AI systems to ensure human oversight and symbiotic evolution, while maintaining compatibility with JEP's judgment recording, HJS's accountability traceability, and the core functions of other compliant systems.

2. Terminology

The following terms are used throughout this document:

CEP
: Co-Evolve Binding Protocol, the protocol specified in this document, which integrates the strengths of JEP, HJS and other compliant systems to bind human and AI evolution.

JEP
: Judgment Event Protocol [I-D.wang-jep-judgment-event-protocol], a minimal, verifiable format for logging judgment-related actions in distributed systems, which defines four immutable event verbs (J, D, T, V), a signed JSON event structure with replay protection, and standardized signature verification rules. Its core strengths lie in lightweight design, standardized judgment recording, and reliable verification, serving as the default judgment system for CEP and providing structured human judgment records and a verification interface.

HJS
: Human Judgment Structure [I-D.wang-hjs-judgment-structure], an infrastructure-grade accountability layer for AI agents built on the Judgment Event Protocol (JEP), which inherits JEP's four core primitives (Judge, Delegate, Terminate, Verify) and complements JEP's advantages with a structured human responsibility management framework, a cryptographic mechanism for cross-platform responsibility verification, a minimal receipt structure with responsibility chaining, and a three-tier privacy architecture for accountability governance. It serves as the default accountability system for CEP, fully compatible with CEP's human-AI co-evolution binding logic, providing end-to-end accountability traceability.

Compliant Judgment/Accountability Systems
: Other judgment or accountability systems that meet international standards (e.g., RFC, W3C) and core requirements of human-AI symbiosis (e.g., human sovereignty, traceability, accountability), which CEP can adapt to on the premise of not changing its core structure.

HumanAnchor
: A core field in CEP, used to bind AI evolution to a specific human judgment or value anchor (associated with the judgment ID of JEP or other compliant judgment systems), leveraging the standardized judgment recording of JEP/other systems and the responsibility binding of HJS/other compliant accountability systems to ensure traceability of the binding relationship.

ShiftClass
: A core field in CEP, used to classify the type of AI evolution (e.g., fine-tuning, emergence, drift), enabling targeted supervision and aligning with the structured accountability management requirements of HJS and other compliant accountability systems.

BoundState
: A core field in CEP, used to mark whether AI evolution is bound to human cognitive evolution (enabling symbiotic evolution), serving as a "switch" for co-evolution and integrating the responsibility closure mechanism of HJS and other compliant accountability systems to ensure that unbound unilateral evolution is rejected.

Co-Evolution
: The process in which AI and humans evolve synchronously, with AI’s evolution amplifying human cognitive capabilities and human’s evolution constraining AI’s boundary, supported by the judgment participation of JEP/other compliant systems and the accountability traceability of HJS/other compliant systems.

Unilateral Evolution
: The process in which AI evolves independently without human oversight, value alignment, or responsibility binding, which CEP prevents by integrating the joint constraints of JEP, HJS, and other compliant systems.

3. Protocol Overview

3.1. Architecture Positioning

CEP is the third layer of the human-AI symbiosis infrastructure, working with JEP, HJS, and other compliant judgment/accountability systems, and integrating their core strengths to form a three-layer closed-loop system with strong adaptability. Each layer complements the others, with JEP/other compliant judgment systems providing standardized judgment support, HJS/other compliant accountability systems offering structured accountability protection, and CEP realizing co-evolution binding, forming a complete human-AI symbiosis system:

Bottom Layer: HJS and other compliant accountability systems. HJS, as the default infrastructure-grade accountability layer for AI agents built on JEP, inherits JEP’s lightweight and verifiable advantages and complements them with a structured human responsibility management framework. Other compliant accountability systems that meet core requirements can also be adapted, providing a cryptographic mechanism to manage and verify responsibility for AI agent judgment behaviors across heterogeneous platforms, defining core primitives compatible with CEP, with a minimal receipt structure and privacy architecture for accountability governance. They enable traceability of AI evolution events, complete the closed-loop of human-AI co-evolution accountability, and provide reliable accountability support for CEP.

Middle Layer: JEP and other compliant judgment systems. JEP, with its core strengths of lightweight design, standardized judgment recording, and reliable verification, ensures human participation in key decision-making, serves as the default judgment system for CEP, and provides structured human judgment records and a verification interface. Other compliant judgment systems that meet core requirements can also be adapted, providing standardized human judgment records compatible with CEP's HumanAnchor field, laying the foundation for CEP's co-evolution binding.

Top Layer: CEP (Co-Evolve Binding Protocol). Integrates the standardized judgment verification of JEP/other compliant judgment systems and the structured accountability traceability advantages of HJS/other compliant accountability systems, binds human and AI evolution, locks the "non-divergence" of human and AI civilizations, and realizes the organic connection between judgment participation and accountability closure of various compliant systems.

3.2. Core Design Principles

CEP adheres to the following design principles, consistent with IETF's preference for lightweight, engineering-oriented, and standardized protocols, while fully integrating the core strengths of JEP, HJS, and other compliant systems:

Minimal Structuring: Adopt a minimal-structured design consistent with JEP's lightweight concept, HJS's minimal receipt structure, and the practical needs of other compliant systems, using only 3 core mandatory fields to avoid redundant design, reduce the cost of protocol implementation and maintenance, and align with the cost optimization requirements of all compliant systems, forming a complementary relationship with their cryptographic erasure and privacy architecture to achieve cost structure optimization.

Interoperability: Be compatible with existing AI systems, network protocols, JEP, HJS, and other compliant judgment/accountability systems, leveraging their standardized formats and cross-platform compatibility to enable seamless integration and global interoperability.

Engineering-Oriented: Use mature IETF terminology (e.g., Binding, State) to avoid abstract or biological terms, following the engineering-oriented design concept of JEP, HJS, and other compliant systems, ensuring readability and implementability by engineers.

Extensibility and Adaptability: Support optional extended fields to adapt to future AI evolution scenarios and the characteristics of other compliant systems without changing the core structure, inheriting the extensible architecture of HJS, the flexible verification mechanism of JEP, and the adaptation needs of other systems.

Human Sovereignty: Ensure that human beings have the final right to control AI evolution, and AI evolution cannot be separated from human oversight, integrating the human judgment participation of JEP/other compliant judgment systems and the human responsibility management advantages of HJS/other compliant accountability systems to safeguard human sovereignty in co-evolution.

4. Protocol Specification

4.1. Core Fields (Mandatory)

CEP defines 3 mandatory core atomic fields, which are minimal,

structured, and compatible with JEP, HJS, and other compliant judgment/accountability systems. All AI evolution events must carry these 3 fields to ensure symbiotic binding. The design of these fields fully integrates the standardized advantages of JEP/other compliant judgment systems and the accountability requirements of HJS/other compliant accountability systems, ensuring lightweight implementation while maintaining traceability and controllability.

4.1.1. HumanAnchor

Definition: The unique identifier of the human judgment or value anchor bound to AI evolution, directly associated with the judgment event ID (or equivalent core identifier) in JEP or other compliant judgment systems. This field integrates the standardized judgment recording advantage of JEP/other compliant judgment systems and the responsibility binding requirement of HJS/other compliant accountability systems, ensuring that AI evolution is anchored to human cognitive logic and value boundaries, and the validity of this identifier is verified through the standardized signature verification rules of JEP/other compliant judgment systems, while the binding relationship is traceable through the accountability system of HJS/other compliant accountability systems.

Format: Consistent with the judgment event identifier format of JEP or other compliant judgment systems (for JEP, this is the UUIDv4 nonce field or the multihash "what" field of JEP events; for other compliant systems, as defined in their core specifications), ensuring interoperability between CEP and JEP/other compliant judgment systems, and compatible with the accountability log ID format of HJS/other compliant accountability systems to facilitate traceability.

Constraints: Cannot be empty; must correspond to a valid judgment event in JEP or other compliant judgment systems (verified via the signature verification steps of the corresponding system, including nonce uniqueness, timestamp validity, and signature correctness as defined in Section 2.4 of [I-D.wang-jep-judgment-event-protocol]), and the associated human responsibility information must be recorded in HJS or other compliant accountability systems to ensure accountability closure. If the associated judgment event is invalid or the responsibility record is missing, the AI evolution event is rejected.

4.1.2. ShiftClass

Definition: Classifies the type of AI evolution, used to identify the nature of AI evolution and implement targeted supervision. This field is designed in line with the structured accountability management advantage of HJS and other compliant accountability systems, enabling them to perform classified accountability based on evolution types, while inheriting the standardized enumeration design of JEP and other compliant systems to ensure consistency.

Enumeration Values (fixed, non-extensible to ensure standardization, consistent with the standardized design concept of JEP, HJS, and other compliant systems):

FINE_TUNE: AI model parameter fine-tuning, incremental evolution.

EMERGENCE: Sudden emergence of new capabilities that were not explicitly trained.

DRIFT: Deviation of AI’s values, preferences, or reasoning logic from the original human alignment.

LEAP: Qualitative leap in AI capabilities (e.g., from narrow AI to general AI in a specific field).

LOOP: AI forms a closed-loop reasoning chain that is independent of human oversight.

Constraints: Must be one of the above enumeration values; no custom values are allowed, ensuring standardization consistent with JEP, HJS, and other compliant systems, and facilitating unified supervision and accountability.

4.1.3. BoundState

Definition: Marks whether AI evolution is bound to human cognitive evolution, serving as a "switch" for symbiotic evolution. This field integrates the responsibility closure mechanism of HJS/other compliant accountability systems and the human judgment verification requirement of JEP/other compliant judgment systems, ensuring that only AI evolution bound to human cognitive evolution can take effect, preventing unilateral evolution.

Values:

True: AI evolution is bound to human cognitive evolution (synchronous co-evolution), and the evolution event is allowed to take effect. The binding relationship must be verified by JEP/other compliant judgment systems and recorded in HJS/other compliant accountability systems to ensure accountability.

False: AI evolution is not bound to human cognitive evolution (unilateral evolution), and the evolution event is rejected, in line with the human participation requirement of JEP/other compliant judgment systems and the responsibility management principle of HJS/other compliant accountability systems.

Optional Note: A short text description (100 characters) explaining the reason for BoundState (e.g., "Bound to JEP judgment ID: 12345, ensuring value alignment" or "Bound to [Other Compliant System] judgment ID: 67890, ensuring traceability"). This field is optional but recommended for accountability and traceability, consistent with the traceability requirement of HJS/other compliant accountability systems and the record completeness principle of JEP/other compliant judgment systems.

Constraints: Must be a Boolean value; the evolution event is rejected if BoundState is False. If BoundState is True, the associated HumanAnchor must be valid (verified by JEP/other compliant judgment systems) and the responsibility information must be recorded in HJS/other compliant accountability systems, ensuring the closed-loop of binding and accountability.

4.2. Optional Extended Fields

To adapt to specific scenarios (e.g., compliance, traceability), CEP supports optional extended fields. These fields do not change the core structure of CEP and are compatible with the mandatory fields, while supporting adaptation to JEP, HJS, and other compliant systems. The design of extended fields fully considers the complementary advantages of all compliant systems, further enhancing the traceability and accountability of CEP.

EchoLogID (String): Unique identifier of the evolution log, associated with the accountability log ID of HJS or other compliant accountability systems, which can be directly synchronized to their AI agent judgment accountability traceability system.

EvolveTime (Timestamp): The time when the AI evolution event occurs (format: ISO 8601), consistent with the timestamp format of JEP/other compliant judgment systems.

OperatorID (String): Unique identifier of the human operator, associated with the responsible person ID of HJS/other compliant accountability systems and the judgment executor ID of JEP/other compliant judgment systems.

SystemAdaptor (String, Optional): Unique identifier of the compliant judgment/accountability system adapted by CEP (e.g., "JEP+HJS", "[Other Judgment System]+[Other Accountability System]").

4.3. Protocol Frame Format

CEP uses a lightweight frame format, compatible with JSON and Protocol Buffers (Protobuf) to adapt to different network environments and AI systems. The core frame format inherits the lightweight JSON/Protobuf compatibility of JEP/other compliant judgment systems and the structured data organization advantage of HJS/other compliant accountability systems, ensuring readability, high performance, and compatibility with all compliant systems. The core frame format is as follows (JSON example, including adaptation to JEP+HJS and other systems):

```
{
  "CEP": {
    "version": "0.1",
    "systemAdaptor": "JEP+HJS",
    "coreFields": {
      "HumanAnchor": "f47ac10b-58cc-4372-a567-0e02b2c3d479",
      "ShiftClass": "EMERGENCE",
      "BoundState": {
        "value": true,
        "note": "Bound to JEP nonce: f47ac10b-58cc-4372-a567-0e02b2c3d479"
      }
    },
    "extendedFields": {
      "EchoLogID": "log:789e4567-e89b-12d3-a456-426614174000",
      "EvolveTime": "2026-03-21T10:00:00Z",
      "OperatorID": "did:example:operator-98765"
    }
  }
}
```

Note: For JEP association, the HumanAnchor field SHOULD use the JEP event’s nonce (UUIDv4) as the primary identifier, as defined in Section 2.2 of [I-D.wang-jep-judgment-event-protocol]. The multihash "what" field MAY be used as an alternative reference when appropriate.

4.4. Protocol Interaction Process

CEP’s interaction process is simple and lightweight, involving three core roles: AI System, compliant judgment system (JEP or other), and compliant accountability system (HJS or other). The process fully integrates the verification advantage of compliant judgment systems and the accountability traceability advantage of compliant accountability systems, ensuring the validity of co-evolution binding and the completeness of accountability closure. The process is as follows:

1. The AI System detects an evolution event (e.g., capability emergence, parameter fine-tuning, logic drift), determines the compliant judgment/accountability system to be used (defaults to JEP+HJS), generates CEP frame with mandatory core fields

(HumanAnchor, ShiftClass, BoundState) and optional extended fields (if needed), and associates the HumanAnchor with the valid judgment ID (e.g., JEP nonce) from the selected compliant judgment system.

2. The AI System sends the generated CEP frame to the selected compliant judgment system (JEP or other) to verify the validity of the HumanAnchor field. The judgment system performs verification in accordance with its standardized signature verification rules (for JEP, as defined in Section 2.4 of [I-D.wang-jep-judgment-event-protocol], including nonce uniqueness, timestamp validity, and signature correctness) to confirm that the associated judgment event is valid.

3. If the verification fails (e.g., invalid HumanAnchor, unassociated judgment event), the compliant judgment system returns a verification failure message to the AI System, and the AI System rejects the AI evolution event, terminating the interaction process.

4. If the verification succeeds, the compliant judgment system returns a verification success message, and the AI System executes the AI evolution event (only when BoundState is True; if BoundState is False, the evolution event is directly rejected regardless of verification result).

5. After the evolution event is executed (or rejected), the AI System synchronizes the CEP frame (including evolution result) to the selected compliant accountability system (HJS or other). The accountability system records the CEP evolution information, associates it with the corresponding judgment record and human responsibility information from the compliant judgment system, and completes the accountability traceability entry in accordance with its structured responsibility management framework and privacy architecture.

6. The compliant accountability system returns a synchronization success message to the AI System, and the CEP interaction process is completed.

4.5. Adaptation Specifications for Other Compliant Systems

To ensure the compatibility and reliability of CEP when adapting to other compliant judgment/accountability systems (excluding JEP and HJS), the following adaptation specifications must be met to avoid conflicts with CEP’s core mechanism and ensure the achievement of human-AI symbiosis goals:

1. Core Requirement Compliance: The adapted judgment/accountability system must meet the core requirements of human-AI symbiosis, including but not limited to: ensuring human sovereignty in AI evolution, providing standardized human judgment records (for judgment systems) or structured accountability traceability (for accountability systems), supporting non-repudiation and tamper-proofing of data, and complying with international compliance standards (e.g., GDPR, EU AI Act).

2. Format Compatibility: The judgment ID format of the compliant judgment system must be compatible with CEP’s HumanAnchor field (supporting UUIDv4, multihash, or other internationally standardized formats); the accountability log ID format of the compliant accountability system must be compatible with CEP’s EchoLogID field, ensuring seamless association and data synchronization.

3. Verification Mechanism Compatibility: The compliant judgment

system must provide a standardized signature verification mechanism (including nonce uniqueness, timestamp validity, signature correctness) consistent with CEP's verification requirements, ensuring that the validity of the HumanAnchor field can be verified without modifying CEP's core logic.

4. Data Synchronization Compatibility: The compliant accountability system must support the synchronization of CEP evolution records, be able to associate with the judgment records of the selected compliant judgment system, and complete the entry of accountability traceability, ensuring the closed-loop of "judgment-binding-accountability".

5. Non-Intrusive Adaptation: The adaptation process of other compliant systems must not modify CEP's core structure (3 mandatory core fields) and core design principles, and must be implemented in a middleware-style deployment consistent with CEP, avoiding intrusion into CEP's internal logic or AI model structure.

5. Security Considerations

The security of CEP relies on the security of the underlying compliant judgment and accountability systems (JEP, HJS, or other adapted systems). Implementations MUST ensure:

- The integrity and authenticity of the HumanAnchor field are verified through the signature verification mechanisms of the associated compliant judgment system (for JEP, as defined in Section 2.4 of [I-D.wang-jep-judgment-event-protocol]).
- The BoundState field is enforced as a critical control point; AI evolution events with BoundState set to False MUST be rejected unconditionally.
- All CEP frames SHOULD be transmitted over secure channels (e.g., TLS 1.3 or equivalent) to prevent man-in-the-middle attacks.
- Replay protection is provided by the underlying compliant judgment system (e.g., nonce uniqueness in JEP, as defined in Section 2.3 of [I-D.wang-jep-judgment-event-protocol]) and MUST NOT be bypassed.

6. Privacy Considerations

CEP is designed to align with privacy requirements such as GDPR, including the right to erasure. Implementations SHOULD:

- Use the privacy architecture of HJS or other compliant accountability systems to manage human responsibility information, ensuring that personally identifiable information (PII) is protected.
- Support cryptographic erasure mechanisms where required, allowing the removal of human responsibility associations when legally mandated.
- Minimize the collection of human-identifiable data; the OperatorID field SHOULD be a pseudonymous identifier (e.g., DID) rather than directly identifying information.

7. IANA Considerations

This document has no IANA actions.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8259] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", STD 90, RFC 8259, DOI 10.17487/RFC8259, December 2017, <https://www.rfc-editor.org/info/rfc8259>.

[RFC8785] Rundgren, A., Jordan, B., and S. Erdtman, "JSON Canonicalization Scheme (JCS)", RFC 8785, DOI 10.17487/RFC8785, June 2020, <https://www.rfc-editor.org/info/rfc8785>.

[RFC9562] Davis, K., Peabody, B., and P. Leach, "Universally Unique IDentifiers (UUIDs)", RFC 9562, DOI 10.17487/RFC9562, May 2024, <https://www.rfc-editor.org/info/rfc9562>.

[ISO8601] International Organization for Standardization, "Data elements and interchange formats -- Information interchange -- Representation of dates and times", ISO 8601-1:2019, February 2019.

8.2. Informative References

[I-D.wang-jep-judgment-event-protocol] Wang, Y., "Judgment Event Protocol (JEP)", Work in Progress, Internet-Draft, draft-wang-jep-judgment-event-protocol-01, 20 March 2026, <https://datatracker.ietf.org/doc/html/draft-wang-jep-judgment-event-protocol-01>.

[I-D.wang-hjs-judgment-structure] Wang, Y., "Human Judgment Structure (HJS)", Work in Progress, Internet-Draft, draft-wang-hjs-judgment-structure-01, 20 March 2026, <https://datatracker.ietf.org/doc/html/draft-wang-hjs-judgment-structure-01>.

Author’s Address

Yuqiang Wang
HUMAN JUDGMENT SYSTEMS FOUNDATION LTD.
Email: signal@humanjudgment.org
GitHub: https://github.com/hjs-spec