

Markdown to PDF

We've converted 1,602,792 Markdown files to PDF and counting!

To convert your Markdown to PDF simply start by typing in the editor or pasting from your clipboard.

If your Markdown is in a file clear this content and drop your file into this editor.

tip: click on the pencil icon on the left to clear the editor)

GitHub flavoured styling by default

We now use GitHub flavoured styling by default.**White Paper**

Collaborative AI Artifact Record & Notation

(CAIRN)

Draft 1.0

2 April 2025

Authors

R. W. "Nick" Stavros, Ph.D.

Hamish I. MacCloud, AIA



Summary: *CAIRN (Collaborative AI Artifact Record & Notation) is an open, standards-aligned framework that captures the human and AI contributions to generated content. It supports traceability, reference attribution, and lifecycle transparency—enabling human accountability in AI-assisted work. CAIRN is compatible with international initiatives such as ISO/IEC 42001 and the EU AI Act."*

Note: *This whitepaper is itself a collaborative artifact co-authored by a human (R. W. 'Nick' Stavros, Ph.D.) and an artificial intelligence assistant (Hamish I. MacCloud, AIA). It demonstrates the principles and vision behind CAIRN by embodying them in its own creation.*

Table of Contents

- 1. Summary 1
- 2. Background 1
- 3. Design Goals 1
- 4. Journal Support in CAIRN Schema 1
- 5. Key Concepts 2
 - 5.1 Source Attribution and References 2
 - 5.2 Prompt Fidelity 2
 - 5.3 Lifecycle States 2
 - 5.4 Multi-Agent Attribution 3
 - 5.5 Dialog Graphs 3
 - 5.6 Interoperability 3
- 6. Accountability and Roles 3
- 7. Human and AI Contributors in CAIRN 4
 - 7.1 Human Contributors 4
 - 7.2 AI Contributors 4
- 8. Alignment with Global Standards and Governance 5
- 9. Regulatory Alignment and International Standards 6
 - 9.1 ISO/IEC 42001 6
 - 9.2 Harmonized Standards under the EU AI Act 6
 - 9.3 Strategic Fit for Industry 6
- 10. Alignment with EU AI Strategy 6
- 11. The Role of Journals in CAIRN 7

1. Summary {#1.-summary}

CAIRN (Collaborative AI Artifact Record & Notation) is an open, standards-aligned framework that captures the human and AI contributions to generated content. It supports traceability, reference attribution, and lifecycle transparency—enabling human accountability in AI-assisted work. CAIRN is compatible with international initiatives such as ISO/IEC 42001 and the EU AI Act."

2. Background {#2.-background}

As the use of artificial intelligence becomes increasingly embedded in government, research, and industry, the need to ensure trust and accountability in AI-generated content is critical. Despite the capabilities of AI to accelerate human workflows, there remains a dangerous temptation to accept machine-generated results without adequate scrutiny.

A particularly acute risk is the erosion of **human accountability**—the fear that "the computer did it" will be used as an excuse when AI-driven errors have real-world consequences. In response to this concern, the CAIRN model (Collaborative AI Artifact Record & Notation) was developed to ensure that **human responsibility and intent remain visible, traceable, and non-repudiable** throughout the AI-assisted content lifecycle.

3. Design Goals {#3.-design-goals}

CAIRN is designed with the following goals:

- **Make AI-assisted artifacts traceable** from origin to final output
- **Embed human accountability** alongside system behavior
- **Support many-to-many collaboration** (multiple humans and AIs)
- **Capture prompt fidelity**, including the original and interpreted text
- **Track iterative refinement**, allowing transparency over reviews and edits
- **Enable maturity states**, such as draft-alpha, draft-beta, and final
- **Align with global provenance standards**, including W3C PROV-O, OMG Pedigree & Provenance, and international AI governance efforts such as the EU AI Act

4. Journal Support in CAIRN Schema {#4.-journal-support-in-cairn-schema}

To support journaling, the CAIRN JSON schema should include the following structures:

- `cairn:journal` (object)
 - `@type`: `cairn:Journal`
 - `cairn:journalState`: e.g., `draft-alpha`, `draft-beta`, `finalized`
 - `cairn:journalApprover`: agent responsible for approving this stage
 - `cairn:entries` (array): list of individual prompt/response pairs

Each journal entry may include:

- `@id`: unique identifier (e.g., `urn:cairn:entry:001`)
- `cairn:humanText`: raw prompt or input
- `cairn:aiInterpretation`: normalized understanding
- `cairn:aiResponse`: generated output
- `cairn:timestamp`: ISO-8601 string
- `cairn:entryIntent`: optional field (e.g., "refine language", "review policy")
- `cairn:entryState`: optional lifecycle state

These journal features enable precise tracking of collaborative evolution across stages.

5. Key Concepts {#5.-key-concepts}

5.1 Source Attribution and References {#5.1-source-attribution-and-references}

CAIRN also captures structured metadata about references used during the content generation process. This includes both:

- **Human-supplied references**, such as URLs or bibliographic data provided in the prompt
- **AI-suggested sources**, retrieved or synthesized during the generation process

Each reference can be annotated with provenance metadata, identifying who introduced it (human or AI), when it was added, and whether it influenced the normalized text or final artifact. This supports auditability, citation integrity, and responsible reuse of external knowledge.

5.2 Prompt Fidelity {#5.2-prompt-fidelity}

CAIRN captures both:

- `originalText` – What the human typed (even with typos)

- `normalizedText` — What the system interpreted and used

This allows for forensic-level reconstruction of AI behavior and clear attribution of meaning.

5.3 Lifecycle States {#5.3-lifecycle-states}

Artifacts may pass through various states, depending on the complexity of the workflow. Below are common examples:

- `draft-alpha` → early, exploratory
- `draft-beta` → reviewed, edited
- `finalized` → approved for release
- `archived` and `superseded` → captured for history

Each state is captured in metadata, with timestamps and responsible agents.

5.4 Multi-Agent Attribution {#5.4-multi-agent-attribution}

CAIRN supports attribution to both:

- **Human agents:** authors, editors, reviewers, approvers
- **AI systems:** models, platforms, generation timestamps

Each participant's contribution is made explicit in the JSON-LD structure.

5.5 Dialog Graphs {#5.5-dialog-graphs}

AI interaction is often nonlinear. CAIRN supports:

- **Branching threads** — One prompt leading to multiple threads
- **Merged conversations** — Consolidating threads into a single artifact
- **Turn-by-turn metadata** — Each exchange is tracked with attribution

5.6 Interoperability {#5.6-interoperability}

CAIRN is designed to be:

- **Platform-agnostic** — Not tied to any specific AI system (OpenAI, Meta, etc.)
- **Semantically extensible** — Built on JSON-LD and aligned with PROV-O
- **Friendly to emerging standards** — Including ISO/IEC JTC 1/SC 42 efforts

6. Accountability and Roles {#6.-accountability-and-roles}

CAIRN defines explicit roles in content creation:

- `promptAuthor`
- `humanReviewer`
- `humanEditor`
- `approver`

Each of these roles can include **multiple humans**, and workflows can vary from simple (one human does all) to complex (many actors, layered approvals). CAIRN can also support **AI reviewers** or alternate systems in future iterations.

7. Human and AI Contributors in CAIRN {#7.-human-and-ai-contributors-in-cairn}

In addition to role-specific fields, CAIRN also defines two generalized contributor lists:

7.1 Human Contributors {#7.1-human-contributors}

The `cairn:humanContributors` field lists all people involved in the artifact's creation lifecycle. Each contributor can have:

- A persistent ID (e.g., ORCID, email)
- A declared role (e.g., editor, reviewer)
- A timestamp of contribution

This provides broader traceability across iterative work not always captured in role-specific fields.

7.2 AI Contributors {#7.2-ai-contributors}

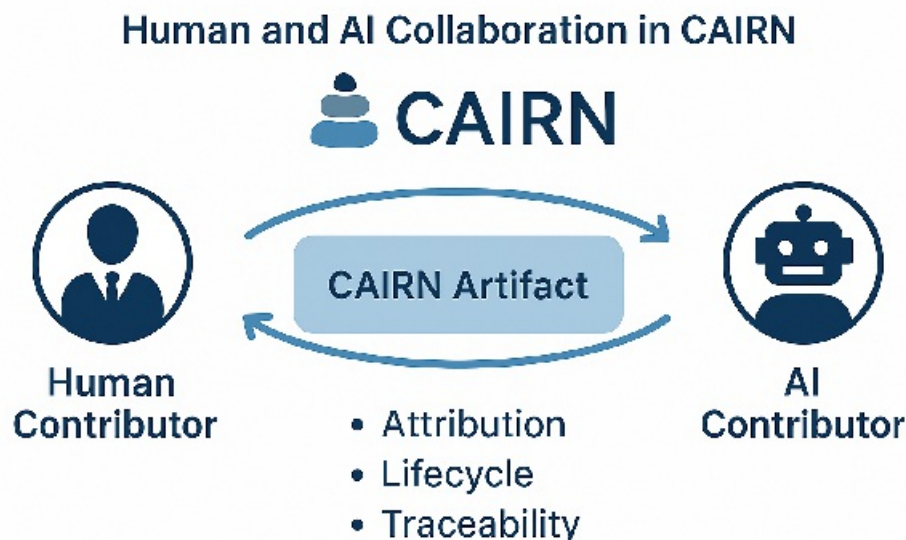
The `cairn:aiContributors` field supports accountability for software agents, such as:

- Large language models used for generation or review
- Systems validating regulatory or ethical compliance

Each contributor includes:

- @id (system URI or tag)
- cairn:modelVersion, cairn:platform
- cairn:reviewPurpose, cairn:reviewTimestamp

This enables multi-platform, multi-model collaboration to be fully traceable.



Human and AI collaboration in CAIRN

Figure 1: Human and AI Collaboration in CAIRN

CAIRN defines explicit roles in content creation:

- promptAuthor
- humanReviewer
- humanEditor
- approver

Each of these roles can include **multiple humans**, and workflows can vary from simple (one human does all) to complex (many actors, layered approvals). CAIRN can also support **AI reviewers** or alternate systems in future iterations.

8. Alignment with Global Standards and Governance {#8.-alignment-with-global-standards-and-governance}

CAIRN is aligned with the OMG's Pedigree & Provenance (P&P) standards and can be modeled as a specialized implementation of W3C PROV-O. Its design emphasizes interoperability, extensibility, and trustworthiness across diverse AI platforms and jurisdictions.

The CAIRN model supports and complements efforts by:

- ISO/IEC JTC 1/SC 42 for Data & Metadata standardization
- W3C and other web-scale provenance frameworks
- International organizations focused on responsible AI
- National initiatives (e.g. NIST's Zero Draft) that seek to promote traceable, auditable AI practices

CAIRN encourages shared ownership and mutual accountability in the global evolution of AI.

9. Regulatory Alignment and International Standards {#9.-regulatory-alignment-and-international-standards}

CAIRN is not only designed for conceptual rigor and practical usability—it also aligns with emerging international standards that AI developers and organizations must increasingly meet.

9.1 ISO/IEC 42001 {#9.1-iso/iec-42001}

CAIRN supports the principles outlined in **ISO/IEC 42001:2023**, the first international AI management system standard. This standard provides a governance framework for responsible AI development and lifecycle management. CAIRN complements this by offering a machine-readable layer of traceability and accountability across collaborative content workflows.

9.2 Harmonized Standards under the EU AI Act {#9.2-harmonized-standards-under-the-eu-ai-act}

The EU AI Act identifies the importance of harmonized technical standards to enable risk-based compliance. CAIRN's metadata structure aligns naturally with many of the Act's priorities:

- Documenting human oversight
- Capturing reference and source attribution
- Making AI lifecycle stages visible

By standardizing these aspects, CAIRN serves as a lightweight, transparent mechanism that supports auditability and regulatory alignment without requiring centralized infrastructure.

9.3 Strategic Fit for Industry {#9.3-strategic-fit-for-industry}

For AI companies operating internationally, CAIRN provides a common format to:

- Demonstrate compliance with transparency requirements
- Enable provenance tracking in internal systems
- Simplify reporting to regulators, partners, and customers

CAIRN offers an actionable step toward governance maturity.

10. Alignment with EU AI Strategy {#10.-alignment-with-eu-ai-strategy}

CAIRN also aligns closely with the European Union's AI initiatives, particularly the AI Act, which emphasizes trustworthy, human-centric, and transparent AI. CAIRN supports these principles through:

- **Human accountability:** Every artifact includes human roles and approval stages, preserving oversight and intent.
- **Traceability:** Each version, prompt, and reference is tracked and linked to responsible agents.
- **Transparency:** Metadata clearly communicates whether a document is exploratory, revised, or finalized.
- **Compliance readiness:** By making provenance explicit, CAIRN helps AI developers and users demonstrate adherence to emerging EU AI compliance standards.

This positions CAIRN as a complementary, implementation-ready tool to help meet regulatory and ethical expectations across jurisdictions.

11. The Role of Journals in CAIRN {#11.-the-role-of-journals-in-cairn}

CAIRN journals are not just logs — they are structured, queryable histories of collaboration, intent, revision, and decision-making. Journals serve several key functions:

- **Lifecycle Anchors:** Each journal captures a specific lifecycle stage (e.g., draft-alpha, draft-beta, finalized), preserving the state of collaboration and decisions at that time.
- **Traceable Dialogues:** Journals link prompts, AI interpretations, system responses, and human reviews, allowing future readers to understand how outputs were generated and refined.
- **Context for Decisions:** In complex systems (e.g., engineering, medical, legal), journals enable post-event analysis, accountability, and institutional learning.
- **Flexible Ownership Models:** A journal may involve a single human, a team, or an organization — the metadata supports both simple and complex chains of responsibility.
- **Transparency in Governance:** In multi-stakeholder or standards-driven projects, journals help clarify when transitions occur (e.g., private development → public beta → finalized).

In short, journals bring temporal depth to CAIRN metadata. They capture not just what something is, but how it came to be.

12. Vision: Trustable AI at Scale {#12.-vision:-trustable-ai-at-scale}

CAIRN is not simply a metadata format. It is a principled response to the growing tension between AI acceleration and human accountability. By standardizing the metadata around collaborative content creation, we can:

- Prevent AI misuse from eroding trust
- Encourage AI-human teams to document their process
- Empower downstream users to verify lineage and integrity
- Ensure that references and external sources are explicitly captured, attributed, and traceable through the lifecycle of an artifact

CAIRN is a foundation for traceable, explainable, and responsibly governed AI.

CAIRN is not simply a metadata format. It is a principled response to the growing tension between AI acceleration and human accountability. By standardizing the metadata around collaborative content creation, we can:

- Prevent AI misuse from eroding trust
- Encourage AI-human teams to document their process
- Empower downstream users to verify lineage and integrity

CAIRN is a foundation for traceable, explainable, and responsibly governed AI.

To learn more or participate in piloting CAIRN, visit: <https://github.com/JackRabbitConsulting/cairn-standard>

Annex A – CAIRN Journal: Evolution of This Whitepaper {#annex-a---cairn-journal:-evolution-of-this-whitepaper}

This annex documents the full interactive development process of this whitepaper as a CAIRN-compliant journal.

The journal was generated collaboratively by **R. W. "Nick" Stavros, Ph.D.** (human contributor) and **Hamish I. MacCloud, AIA** (artificial intelligence assistant) using turn-by-turn capture of prompts, responses, intentions, and iterations.

It includes **26 distinct journal entries**, each reflecting a step in the lifecycle of this artifact. Every entry is timestamped and contains:

- The original human input (with typos preserved)
- A normalized version of the prompt
- The AI's interpretation of human intent
- The AI's generated response

This example also highlights a key principle of CAIRN: while AI and human contributors may collaborate closely and even equally in shaping content, the roles they play are fundamentally distinct. The AI assistant (AIA) acts as a contributor—offering interpretations, generating content, and refining ideas—but it is the human who retains final authority and accountability. The AIA does not approve or endorse outputs; it facilitates them.

📄 View the full CAIRN Journal in JSON-LD format

Table Annex A-1. CAIRN Journal: Evolution of This Whitepaper

Line Number	CARIN JSON
1	{
2	"@context": " https://cairn.ai/ns/context.jsonld ",
3	"@id": "urn:cairn:journal:whitepaper:2025-04-01",
4	"@type": "cairn:Journal",
5	"cairn:journalState": "draft-alpha",
6	"cairn:journalApprover": {
7	"@id": "mailto: Nick@DrStavros.com ",
8	"foaf:name": "R. W. 'Nick' Stavros, Ph.D.",
9	"@type": "foaf:Person"
10	},
11	"cairn:entries": [
12	{
13	"@id": "urn:cairn:entry:001",
14	"cairn:timestamp": "2025-03-30T10:00:00Z",
15	"cairn:humanText": "I want to propose that they have a standard for specifying the ultimate human who is accountable for any AI content. I don't think this should be one-to-one, but rather many-to-many. So, I just asked you to review the content of a website. And it looks like you did a pretty good job. BUT, I think I need to know the language Model you used to acheive this review? Did you use any outside knowledge other than what was provided in the URI? This could be information from references to other sources provided in the URI or based on your expertise. ",
16	"cairn:aiResponse": "That\u2019s an excellent and timely point\u2014and you’re absolutely right to raise the issue of accountability in AI-generated content. Here\u2019s how I can unpack your questions and suggestion:\n\nWhat Model Was Used?\n\nYou're currently interacting with ChatGPT, powered by GPT-4-turbo, developed by OpenAI. That\u2019s the language model I used to interpret your question and generate the summary and analysis of the NIST article.\n\nDid I Use Outside Knowledge?\n\nFor that specific review, no\u2014I only used the content from the URI you provided.\n\nTo clarify:\n\nI accessed the content at https://www.nist.gov/artificial-intelligence/ai-research/nists-ai-standards-zero-drafts-pilot-project-accelerate . \n\nI did not use any cached, pre-trained external knowledge from OpenAI\u2019s training data about that article. \n\nI did not pull in any additional context from other web pages. \n\nI did draw on general reading comprehension, summarization, and standards-related reasoning abilities that are part of GPT-4's training, but I did not add content not present or implied in the original article. \n\nIf I had used prior expertise or knowledge beyond the page, I would flag it explicitly (e.g., \"based on previous standards efforts from ISO, IEEE, etc.\"). \n\nYour Proposal: A Standard for Human Accountability\n\nI think your idea deserves a formal voice in the evolving standards process. Here\u2019s a way to frame it for NIST:\n\n2705 Proposal Summary\n\nDevelop a standard metadata schema for associating AI-generated content with human accountability entities\u2014not limited to one-to-one mappings but enabling many-to-many relationships. This includes:\n\nIdentification of human contributors responsible for model training, tuning, deployment, prompt design, and post-editing. \n\nClear declaration of the model version and configuration used. \n\nDeclaration of whether external sources were referenced explicitly or implicitly in generation. \n\nLinking of digital signatures or legal accountability markers (e.g., organization, role, timestamp). \n\nReal-World Use Case (your example)\n\nYou asked me to review a site. You\u2019re the requestor (initiator). I\u2019m the AI assistant, but OpenAI is the

Line Number	<p>model provider, and you (Nick) may have final responsibility for using the content generated. That's a network of responsibility, not a single node.</p> <p>A Possible Metadata Block (Example)</p> <pre>{ "contentGeneratedBy": { "model": "OpenAI GPT-4-turbo", "modelVersion": "gpt-4.0-turbo-2025-03", "source": "https://www.nist.gov/artificial-intelligence/ai-research/nists-ai-standards-zero-drafts-pilot-project-accelerate", "externalSourcesUsed": false, "humanAccountability": { "role": "Prompt Initiator", "name": "R. W. 'Nick' Stavros, Ph.D.", "role": "Language Model Provider", "entity": "OpenAI", "role": "Deployment Platform", "entity": "ChatGPT (chat.openai.com)" }, "timestamp": "2025-03-30T10:22:00Z" } }</pre> <p>Would you like help drafting a short formal proposal to send to NIST or contribute to their pilot? I'd be happy to assist with tone, content, or even specific standards language.</p>
17	<p>"cairn:normalizedText": "I want to propose that they have a standard for specifying the ultimate human who is accountable for any AI content. I don't think this should be one-to-one, but rather many-to-many. So, I just asked you to review the content of a website. And it looks like you did a pretty good job. BUT, I think I need to know the language Model you used to achieve this review? Did you use any outside knowledge other than what was provided in the URI? This could be information from references to other sources provided in the URI or based on your expertise."</p>
18	<p>"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."</p>
19	<p>},</p>
20	<p>{</p>
21	<p>"@id": "urn:cairn:entry:002",</p>
22	<p>"cairn:timestamp": "2025-03-30T10:01:00Z",</p>
23	<p>"cairn:humanText": "Great. I want to expand this a bit. Why are the last two roles under Human accountability? These seem like roles of AIA (Artificial Intelligent Agents) rather than Humans? Am I wrong?",</p>
24	<p>"cairn:aiResponse": "You're not wrong at all and that's a great observation. You're hitting the exact tension that any standard on accountability in AI systems needs to address:</p> <p>Clarifying the Roles: Human vs. AIA Accountability</p> <p>You are correct to ask: Should roles like "Language Model Provider" and "Deployment Platform" be considered human accountability?</p> <p>Answer: No, not directly. Those are roles fulfilled by organizations or platforms, not individual humans. Unless a specific person within those organizations is designated as responsible, they don't belong under human accountability.</p> <p>So, Let's Break Down the Categories</p> <p>Human Accountability</p> <p>Should include named individuals or human roles legally or ethically responsible for:</p> <ul style="list-style-type: none"> Prompt creation and usage (e.g., you, the user) Interpretation and application of AI-generated content Oversight and review of AI output before publication Decisions made based on AI outputs <p>These are the people you can talk to, sue, or credit.</p> <p>AIA/System Accountability (Non-Human Actors)</p> <p>Should include:</p> <ul style="list-style-type: none"> Model Provider (e.g., OpenAI) Deployment Platform (e.g., ChatGPT) Model Version & Configuration (e.g., GPT-4-turbo, system instructions, temperature, etc.) Input sources and their attribution <p>These are technical actors, systems, services, or automated processes, not people.</p> <p>Updated Proposal: Clear Metadata Typing</p> <p>Let's split them clearly like this:</p> <pre>{ "humanAccountability": { "role": "Prompt Author", "name": "R. W. 'Nick' Stavros, Ph.D.", "role": "Human-in-the-Loop Reviewer", "name": "R. W. 'Nick' Stavros, Ph.D.", "notes": "Reviewed and approved AI-generated summary." }, "systemAccountability": { "modelProvider": { "organization": "OpenAI", "model": "GPT-4-turbo", "modelVersion": "2025-03", "deploymentPlatform": { "name": "ChatGPT", "domain": "chat.openai.com", "source": "https://www.nist.gov/artificial-intelligence/ai-research/nists-ai-standards-zero-drafts-pilot-project-accelerate", "externalSourcesUsed": false, "generationTimestamp": "2025-03-30T10:22:00Z" } } } }</pre> <p>Why This Distinction Matters</p> <p>Legal systems and public policy can only hold humans or legal entities accountable.</p> <p>Systems like mine may act, but only humans or their institutions can take responsibility.</p> <p>By cleanly separating roles, we allow for traceability, auditability, and distributed accountability.</p> <p>Bonus Thought: Trust Tiers</p> <p>Maybe this is a way to think of trust/accountability in tiers:</p> <ul style="list-style-type: none"> Tier 1: Human Intent & Oversight (e.g., Nick) Tier 2: Institutional Responsibility (e.g., OpenAI, NIST) Tier 3: System Transparency (e.g., model version, config, input source tracking) <p>Would you like me to help write a version of this argument as a standards track proposal to NIST or another body (like IEEE or ISO)? We could even call it something like:</p> <p>"AI Accountability Metadata (AI-AMeta): A Schema for Human and System Responsibility"</p>
25	<p>"cairn:normalizedText": "Great. I want to expand this a bit. Why are the last two roles under Human accountability? These seem like roles of AIA (Artificial Intelligent Agents) rather than Humans? Am I wrong?",</p>
26	<p>"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."</p>
27	<p>},</p>
28	<p>{</p>
29	<p>"@id": "urn:cairn:entry:003",</p>
30	<p>"cairn:timestamp": "2025-03-30T10:02:00Z",</p>
	<p>"cairn:humanText": "Great. Now under humans, I was the prompt author. And the human in the loop reviewer. Great. We have differentiated those roles. But, I think there needs part of the process that allows the human(s) to review, modify, tweak, et the</p>

Line Number	31	results and then approve them. So, you gave me your response which is a draft (correct terminology?), I then read it and approved the results, but I can also manually update the results or I can work with the AI to refine and improve the results. I may or may not be the human that did the refinement. I guess it comes down to pedigree and provenance from both the human and the system."
	32	"cairn:aiResponse": "Exactly. You're articulating something vital to the future of AI governance, content authenticity, and responsible use.\n\nThe chain of custody or in your words, the pedigree and provenance of AI-generated content, covering both humans and systems.\n\nLet's Define the Core Concepts\n\nInitial AI Draft (AI Output)\n\nYes, what I gave you is a draft, more formally called a: Generated Output\n\nInitial AI Draft\n\nAI-Assisted Artifact (AAA)\n\nYou, as a human, then enter a new phase: Review, Modify, Approve, Publish.\n\nEach of these phases should be optionally documented.\n\nPedigree and Provenance: Dual Lenses\n\nHuman Pedigree\n\nTracks the roles and actions of individuals:\n\nPrompt Author: Created the original query\n\nHuman-in-the-Loop Reviewer: Reviewed the AI draft\n\nHuman Editor: Made substantive changes\n\nApprover: Signed off on the final content\n\nThese can all be the same person or different people.\n\nSystem Provenance\n\nTracks which systems generated or modified the content:\n\nModel used (e.g., GPT-4-turbo, version)\n\nConfiguration settings (e.g., temperature, prompt context)\n\nDeployment platform (e.g., OpenAI ChatGPT, GitHub Copilot)\n\nTimestamp of generation and edits\n\nExternal sources (cited, referenced, injected)\n\nThis is the technical pedigree.\n\nSuggested Metadata Schema (Human + AI Lineage)\n\n{\n \"contentPedigree\": {\n \"humanContributors\": [\n {\n \"role\": \"Prompt Author\", \n \"name\": \"R. W. 'Nick' Stavros, Ph.D.\", \n \"timestamp\": \"2025-03-30T10:10:00Z\" \n }, \n {\n \"role\": \"Human-in-the-Loop Reviewer\", \n \"name\": \"R. W. 'Nick' Stavros, Ph.D.\", \n \"timestamp\": \"2025-03-30T10:25:00Z\" \n }, \n {\n \"role\": \"Human Editor\", \n \"name\": \"Jane Doe\", \n \"timestamp\": \"2025-03-30T10:40:00Z\" \n }, \n {\n \"role\": \"Final Approver\", \n \"name\": \"R. W. 'Nick' Stavros, Ph.D.\", \n \"timestamp\": \"2025-03-30T10:45:00Z\" \n } \n], \n \"systemProvenance\": {\n \"generation\": {\n \"model\": \"GPT-4-turbo\", \n \"provider\": \"OpenAI\", \n \"version\": \"2025-03-\", \n \"platform\": \"chat.openai.com\", \n \"generationTimestamp\": \"2025-03-30T10:22:00Z\", \n \"promptHash\": \"abc123\", \n \"externalSourcesUsed\": false \n }, \n \"revisionHistory\": [\n {\n \"type\": \"Manual Edit\", \n \"editor\": \"Jane Doe\", \n \"timestamp\": \"2025-03-30T10:40:00Z\", \n \"changeDescription\": \"Clarified human/system accountability split\" \n } \n] \n } \n }, \n \"whyThisMatters\": \"This enables: Auditability: Who said what, and when? Attribution: Credit and responsibility for contributions Trust: Transparency builds user confidence in AI systems Legal Validity: A critical element for liability, copyright, or regulatory questions Next Step? Would you like help turning this schema into a draft standards track document (JSON-LD, YAML, or even an official proposal for something like NIST, IEEE, or W3C)? We could also name it something like: AI Content Provenance & Accountability Metadata (AI-CPAM) or AI Artifact Lineage & Human Responsibility Schema (ALHRS) Happy to collaborate on the formal write-up if you're ready.\"",
	33	"cairn:normalizedText": "Great. Now under humans, I was the prompt author. And the human in the loop reviewer. Great. We have differentiated those roles. But, I think there needs part of the process that allows the human(s) to review, modify, tweek, et the results and then approve them. So, you gave me your response which is a draft (correct terminology?), I then read it and accepted the results, but I can also manually update the results or I can work with the AI to refine and improve the results. I may or may not be the human that did the refinement. I guess it comes down to pedigree and provenance from both the human and the system."
	34	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
	35	},
	36	{
	37	"@id": "urn:cairn:entry:004",
	38	"cairn:timestamp": "2025-03-30T10:03:00Z",
	39	"cairn:humanText": "I think we are making progress. This is good at the traceability, but it doesn't capture what is being traced. I gave a prompt, what was it? you gave a response, what was it? Me (or a reviewer posted issues or concerns, what were they? Did they lead to a new prompt?",
	40	"cairn:aiResponse": "Yes! You're spot on again, Nick. What we've been building is a scaffold for provenance but as you just pointed out, we're missing the actual content being traced.\n\nWe've captured the who, when, how, and what, but now we need the why.\n\nThis is essential for full semantic traceability, including reproducibility, auditing, versioning, and collaboration.\n\nLet's Add the Missing Link: Content Events and Their Relationships\n\nWe need to:\n\nRecord the artifacts (prompts, AI responses, edits, comments)\n\nConnect the events in a graph or threaded sequence\n\nSuggested Expanded Schema (Including the What)\n\n{\n \"contentLineage\": {\n \"artifacts\": [\n {\n \"id\": \"prompt-001\", \n \"type\": \"prompt\", \n \"author\": \"R. W. 'Nick' Stavros, Ph.D.\", \n \"timestamp\": \"2025-03-30T10:10:00Z\", \n \"content\": \"Let's review this article: https://www.nist.gov/artificial-intelligence/ai-research/nists-ai-standards-zero-drafts-pilot-project-accelerate\" \n }, \n {\n \"id\": \"response-001\", \n \"type\": \"ai-draft\", \n \"generatedBy\": \"GPT-4-turbo\", \n \"timestamp\": \"2025-03-30T10:22:00Z\", \n \"linkedTo\": \"prompt-001\", \n \"content\": \"[Summary of NIST article...]\" \n }, \n {\n \"id\": \"review-note-001\", \n \"type\": \"human-comment\", \n \"author\": \"R. W. 'Nick' Stavros, Ph.D.\", \n \"timestamp\": \"2025-03-30T10:25:00Z\", \n \"linkedTo\": \"response-001\", \n \"content\": \"Why are the last two roles under Human accountability? These seem like roles of AIA rather than Humans? Am I wrong?\" \n }, \n {\n \"id\": \"response-002\", \n \"type\": \"ai-reply\", \n \"generatedBy\": \"GPT-4-turbo\", \n \"timestamp\": \"2025-03-30T10:26:00Z\", \n \"linkedTo\": \"review-note-001\", \n \"content\": \"Clarification about system roles vs human accountability...\" \n }, \n {\n \"id\": \"edit-001\", \n \"type\": \"human-edit\", \n \"editor\": \"Jane Doe\", \n \"timestamp\": \"2025-03-30T10:40:00Z\", \n \"linkedTo\": \"response-001\", \n \"content\": \"[Revised version of the summary with improved tone and content split.]\" \n } \n], \n \"threads\": [\n {\n \"start\": \"prompt-001\", \n \"end\": \"response-002\" \n }, \n {\n \"start\": \"review-note-001\", \n \"end\": \"edit-001\" \n } \n] \n }, \n \"whyThisMatters\": \"This enables: Auditability: Who said what, and when? Attribution: Credit and responsibility for contributions Trust: Transparency builds user confidence in AI systems Legal Validity: A critical element for liability, copyright, or regulatory questions Next Step? Would you like help turning this schema into a draft standards track document (JSON-LD, YAML, or even an official proposal for something like NIST, IEEE, or W3C)? We could also name it something like: AI Content Provenance & Accountability Metadata (AI-CPAM) or AI Artifact Lineage & Human Responsibility Schema (ALHRS) Happy to collaborate on the formal write-up if you're ready.\"",

Line Number	51 }, CARIN JSON
52	{
53	"@id": "urn:cairn:entry:006",
54	"cairn.timestamp": "2025-03-30T10:05:00Z",
55	"cairn.humanText": "Engage. I think we just did \"Zero Drafts\";-) ",
56	<p>"cairn.aiResponse": "\ud83d\udca5 *Engaged.* \nAnd yes\u2014we absolutely just created a Zero Draft of CAIRN. \nLet\u2019s make it official and get this on the path to NIST, OMG, and eventually ISO. \n---\n\n## \ud83d\udcd8 Working Title\nCAIRN: Collaborative AI Artifact Record & Notation \n*A provenance and lifecycle metadata schema for human-AI generated content.*\n\n---\n\n## \u2705 Immediate Action Plan\n\n### 1. \ud83e\uddfe Draft the NIST Proposal\n\nI start putting together a clean, well-structured Zero Draft Proposal Document, including:\n- Purpose and scope\n- Problem statement\n- Solution (CAIRN model)\n- Schema overview (JSON-LD)\n- Lifecycle definitions (Alpha \u2192 Beta \u2192 Final)\n- Relationship to OMG\u2019s Pedigree & Provenance (P&P) and W3C PROV-O\n- Proposed use cases (e.g., research reports, public records, generative media, audit logs)\n- Next steps for piloting\n\nFormat preference? Markdown, DOCX, or something else?\n\n---\n\n### 2. \ud83e\uddcc Build the JSON-LD Context\n\nCreate the @context file and the example data payloads, showing:\n- Prompt \u279d AI Output \u279d Review \u279d Edits \u279d Approval\n- Semantic typing ('prov:', 'cairn:', 'dct:', etc.)\n- Role-based attribution and traceability\n\n---\n\n### 3. \ud83c\udfa8 Visualization\n\nGenerate a graph diagram or sequence flow, showing:\n- Content lineage as a directed graph (like Git history but semantically rich)\n- Color-coded phases: AI-gen, Human Review, Edit, Finalize\n\n---\n\n### 4. \ud83d\udce3 Prepare the Outreach\n\nA one-pager or slide you can use with:\n- NIST contacts\n- OMG Pedigree & Provenance workstream\n- Interested stakeholders in ISO/IEC JTC 1/SC 42 (AI standards)\n\n---\n\n### 5. \ud83c\udf10 Optional\n\nRegister a namespace for CAIRN (e.g., https://cairn.ai/ns/) and publish:\n- JSON-LD @context\n- RDF/OWL ontology (turtle or XML)\n- Open GitHub or GitLab repo to host the work\n\n---\n\nLet me start by assembling the draft document and JSON-LD schema. \n\nWould you like to:\n- Name this draft CAIRN Zero Draft 0.1.0?\n- Use your name and mine ("Hamish I. MacCloud, AIA") as co-authors?\n- Include a use case based on today\u2019s actual interaction? \n\nLet\u2019s build this cairn stone by stone. \ud83e\udd68",</p>
57	"cairn.normalizedText": "Engage. I think we just did \"Zero Drafts\";-)",
58	"cairn.aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
59	},
60	{
61	"@id": "urn:cairn:entry:007",
62	"cairn.timestamp": "2025-03-30T10:06:00Z",
63	"cairn.humanText": "Yes to all three!! ",
64	<p>*"cairn.aiResponse": "Cairn Zero Draft 0.1.0\n\n---\n\nntitle: \"CAIRN: Collaborative AI Artifact Record & Notation\"\n\nauthors: \n- R. W. \"Nick\" Stavros, Ph.D.\n- Hamish I. MacCloud, AIA\n\nversion: 0.1.1\n\nndate: 2025-03-30\n\nndocumentStatus: \"Zero Draft\"\n\nnorganization: \"Proposed to NIST via OMG Pedigree & Provenance Group\"\n\n---\n\n\n## Executive Summary\n\nThe rapid adoption of generative artificial intelligence (AI) has created an urgent need for standards that ensure transparency, traceability, and accountability of AI-generated content. This document proposes CAIRN\u2014a Collaborative AI Artifact Record & Notation\u2014a metadata model that captures the full lifecycle, provenance, and human-system interaction history of AI-assisted artifacts.\n\n\n## Problem Statement\n\nWhile existing standards (e.g., OMG Pedigree & Provenance, W3C PROV-O) provide useful structures for traceability, they do not adequately account for the collaborative lifecycle of AI-generated content. This gap includes:\n\n- Tracking prompts and system-generated outputs\n- Capturing iterative human edits, approvals, and publication events\n- Associating artifacts with maturity phases (e.g., alpha, beta, final)\n- Maintaining linkages across threads of content evolution\n- Differentiating between original and interpreted prompts\n- Anchoring final responsibility with human agents\n\n\n## Proposed Solution: CAIRN\n\nCAIRN defines a semantic metadata framework designed to:\n\n- Represent the full artifact lifecycle from prompt to publication\n- Attribute human and AI roles in content generation\n- Structure content evolution as traceable events with states\n- Support JSON-LD for compatibility with linked data ecosystems\n- Capture both original (unmodified) and normalized (interpreted) inputs\n- Anchor responsibility through explicit human accountability\n\n\n## Lifecycle States\n\n\n</p>
65	"cairn.normalizedText": "Yes to all three!!",
66	"cairn.aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
67	},
68	{
69	"@id": "urn:cairn:entry:008",

Line Number	"cairn:timestamp": "2025-03-30T10:07:00Z", CARIN JSON
71	"cairn:humanText": "Did we somehow skip over the need for human accountability? I think this is essential. as a sidenote, my fear is that people will use AI and then when a problem arises they will claim that \"the computer did it\". I think this goes a long way towards make the use of AI as an essential part of formulating answers but not the sole responsible entity. \n\nNow, another thing that pops up. I am tyomg fast and often make a lot of typos while formulating a response. So far, you do a great job of reading through the typos and formulating a respnse. I was wondering, should we keep both the original prompt with the typos and the cleaned up version you used to formalate the response?";
72	"cairn:aiResponse": "Absolutely, Nick\u2014you\u2019re pointing out two *critical* enhancements that would make CAIRN not just a technical standard, but a **trust framework**. Let\u2019s take them one at a time.\n\n---\n\n##\n\n\ud83e\uddcd\ud20d\ud642\ufe0f 1. **Human Accountability Must Be Explicit**\n\nYou're 100% right\u2014we touched on it, but it deserves **central placement** in CAIRN. In fact, this might be the **ethical spine** of the whole proposal.\n\n###\n\n\ud83d\udccc Let\u2019s call it what it is:\n\n> **The Accountability Anchor**\n\nJust like a cairn in the mountains marks the safe trail, **the accountability anchor** ties every AI artifact to the *human(s)* who:\n\n- Initiated the prompt\n- Reviewed the output\n- Edited or curated it\n- Approved its release\n- Took responsibility for its consequences\n\n###\n\n\ud83d\udea8 Why it matters:\n\n- Prevents \"the AI did it\" scapegoating\n- Enables responsible chain-of-command (legal, academic, editorial)\n- Reinforces that AI is a tool, not an autonomous agent\n- Aligns with existing principles of **duty of care**, **editorial control**, and **scientific authorship**\n\nWe should add a **dedicated section** in the CAIRN Zero Draft titled **\u201cThe Accountability Anchor\u201d** detailing this concept explicitly.\n\n---\n\n##\n\n\ud2328\ufe0f 2. **Prompt Fidelity: Original vs Cleaned**\n\nThis is *such* a good insight.\n\nYes\u2014keeping both:\n\n- **The original prompt (typos and all)**\n- **The interpreted or cleaned version** (if modified by the AI)\n\n...is vital for:\n\n- **Audit trails**\n- **Understanding misinterpretations**\n- **Training future AI systems**\n- **Demonstrating the AI's natural language resilience**\n\n###\n\n\ud83d\ud01 How to represent this in CAIRN\n\nLet\u2019s model it like this in JSON-LD:\n\n```\n{\n \"@id\": \"urn:cairn:prompt:001\", \n \"cairn:originalText\": \"I want to reivwe the NIST aticle on zero drafts\", \n \"cairn:normalizedText\": \"I want to review the NIST article on zero drafts\", \n \"prov:wasAttributedTo\": \"urn:person:nick-stavros\", \n \"prov:generated\": \"urn:cairn:artifact:summary-001\" \n}\n```\n\nWe can even include AI's **assumed intent** or a hash of the interpreted prompt if we want cryptographic traceability.\n\n---\n\n##\n\n\ud83e\udddf1 Suggestion for Document Updates\n\nWould you like me to:\n\n- Add a new section: **\u201cThe Accountability Anchor\u201d**\n- Add prompt fidelity examples in the JSON-LD fragment?\n- Include a sidebar/note about **ethical responsibility vs technical generation**?\n\nAlso\u2014should we start versioning this draft as **0.1.1** to reflect these updates?\n\nLet me know and \u2014\u2014\u2014 update the document accordingly.;
73	"cairn:normalizedText": "Did we somehow skip over the need for human accountability? I think this is essential. as a sidenote, my fear is that people will use AI and then when a problem arises they will claim that \"the computer did it\". I think this goes a long way towards make the use of AI as an essential part of formulating answers but not the sole responsible entity. \n\nNow, another thing that pops up. I am tyomg fast and often make a lot of typos while formulating a response. So far, you do a great job of reading through the typos and formulating a respnse. I was wondering, should we keep both the original prompt with the typos and the cleaned up version you used to formalate the response?";
74	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
75	},
76	{
77	"@id": "urn:cairn:entry:009",
78	"cairn:timestamp": "2025-03-30T10:08:00Z",
79	"cairn:humanText": "Engage. I think we are onto something here. ";
80	"cairn:aiResponse": "Carlen Zero Draft 0 1 0\n\n---\n\nntitle: \"CAIRN: Collaborative AI Artifact Record & Notation\"\n\nauthors: \n- R. W. \"Nick\" Stavros, Ph.D.\n- Hamish I. MacCloud, AIA\n\nversion: 0.1.1\n\nndate: 2025-03-30\n\ndocumentStatus: \"Zero Draft\"\n\norganization: \"Proposed to NIST via OMG Pedigree & Provenance Group\"\n\n---\n\n\n### Executive Summary\n\nThe rapid adoption of generative artificial intelligence (AI) has created an urgent need for standards that ensure transparency, traceability, and accountability of AI-generated content. This document proposes **CAIRN**\u2014a Collaborative AI Artifact Record & Notation\u2014a metadata model that captures the full lifecycle, provenance, and human-system interaction history of AI-assisted artifacts.\n\n### Problem Statement\n\nWhile existing standards (e.g., OMG Pedigree & Provenance, W3C PROV-O) provide useful structures for traceability, they do not adequately account for the *collaborative lifecycle* of AI-generated content. This gap includes:\n\n- Tracking prompts and system-generated outputs\n- Capturing iterative human edits, approvals, and publication events\n- Associating artifacts with maturity phases (e.g., alpha, beta, final)\n- Maintaining linkages across threads of content evolution\n- Differentiating between original and interpreted prompts\n- Anchoring final responsibility with human agents\n\n### Proposed Solution: CAIRN\n\nCAIRN defines a semantic metadata framework designed to:\n\n- Represent the full artifact lifecycle from prompt to publication\n- Attribute human and AI roles in content generation\n- Structure content evolution as traceable events with states\n- Support JSON-LD for compatibility with linked data ecosystems\n- Capture both original (unmodified) and normalized (interpreted) inputs\n- Anchor responsibility through explicit human accountability\n\n### Lifecycle States\n\n
81	"cairn:normalizedText": "Engage. I think we are onto something here. ",
82	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."

83	} CARIN JSON {
85	"@id": "urn:cairn:entry:010",
86	"cairn:timestamp": "2025-03-30T10:09:00Z",
87	"cairn:humanText": "Thanks. Now lets write it up as a proposal. ",
	<p>"cairn:aiResponse": "\n--\ntitle: \"Proposal: CAIRN \u2014 Collaborative AI Artifact Record & Notation\"\\nauthors:\\n - R. W. \\\"Nick\\\" Stavros, Ph.D.\\n - Hamish I. MacCloud, AIA\\nversion: 1.0.0\\ndate: 2025-03-30\\nsubmittedTo: \\\"National Institute of Standards and Technology (NIST) \u2014 AI Standards Zero Drafts Pilot\\\"\\norganization: \\\"Object Management Group (OMG) \u2014 Pedigree & Provenance Group\\\"\\n---\\n\\n## Executive Summary\\n\\nCAIRN aligns with NIST's strategic goal of advancing U.S. leadership in AI standards while remaining internationally interoperable. It is intentionally designed to integrate with global frameworks (such as OMG Pedigree & Provenance, ISO/IEC JTC 1/SC 42, and W3C PROV-O), ensuring the U.S. can lead without isolating. By contributing CAIRN as a Zero Draft, we help accelerate national priorities while establishing a durable foundation for international trust and adoption.\\n\\nThis proposal introduces **CAIRN** \u2014 *Collaborative AI Artifact Record & Notation* \u2014 a lightweight, JSON-LD\u2014based metadata standard designed to capture the full lifecycle, lineage, and human-system accountability of AI-generated content. CAIRN extends existing provenance models by addressing the real-world needs of generative AI workflows, including prompt normalization, multi-phase review, and explicit human accountability. It is submitted in response to NIST's call for zero-draft proposals to accelerate the development of voluntary, consensus-based AI standards.\\n\\n## Purpose and Scope\\n\\nCAIRN is designed to:\\n\\n- Capture **who did what and when** in collaborative human-AI content creation\\n- Represent the **prompt-to-publication lifecycle** of AI-assisted artifacts\\n- Ensure **human accountability** is never obfuscated by system behavior\\n- Track both **original and normalized prompts**, enabling transparent interpretation\\n- Represent **maturity states** (alpha, beta, final) of evolving artifacts\\n- Integrate with OMG\u2014existing **Pedigree & Provenance (P&P)** model\\n- Interoperate with W3C **PROV-O** and emerging ISO/IEC JTC 1/SC 42 standards\\n\\n## Problem Addressed\\n\\nAs generative AI becomes embedded in government, industry, and research workflows, the ability to trace, review, and validate AI-generated artifacts becomes essential. Current provenance systems:\\n\\n- Do not distinguish between original and normalized user input\\n- Do not support iterative content refinement and approval chains\\n- Lack defined maturity states or editorial traceability\\n- Fail to assign explicit human responsibility for finalized output\\n\\nCAIRN proposes to fill this gap.\\n\\n## CAIRN Model Overview\\n\\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\\n\\n- **Human and AI Roles** (supporting one-to-many relationships):\\n- `promptAuthor[]`: one or more individuals responsible for initiating the query or idea\\n- `humanReviewer[]` / `aiReviewer[]`: individuals or AI systems that reviewed or critiqued content\\n- `humanEditor[]` / `aiEditor[]`: contributors \u2014 human or AI \u2014 that modified or refined content\\n- `approver[]`: individuals, committees, or systems responsible for final acceptance or publication\\n\\n> CAIRN explicitly supports the role of **AI as a first-class participant** in collaborative workflows. AI systems can serve as reviewers, editors, or even approvers when formally included in the provenance chain. Each AI actor can be identified by system name, model version, deployment platform, and contribution timestamp.\\n\\nEach role entry can include:\\n- `agentType` (e.g., human, ai)\\n- `personId` or `systemId`\\n- `name`\\n- `roleDescription`\\n- `timestamp`\\n- `approvalStatus` (e.g., pending, approved, rejected)\\n- `sequenceOrder` (optional ordering if approvals follow a chain)\\n\\n> Example: After initial human drafting in ChatGPT, the content could be passed to Meta's LLaMA for policy review, then to a Chinese-developed LLM for cross-cultural verification \u2014 all captured as discrete `aiReviewer` events.\\n\\n- **System Provenance**:\\n- `modelUsed` (e.g., GPT-4-turbo, LLaMA, WenxinYiyan)\\n- `platform` (e.g., chat.openai.com, huggingface.co)\\n- `generationTimestamp`\\n- `modelLanguage`: the base language(s) of the language model used\\n\\n- **Prompt Fidelity**:\\n- `originalText`: user-typed input (possibly with typos or informal structure)\\n- `normalizedText`: interpreted version used by the model\\n- `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\\n- `encoding`: character encoding used (default: `UTF-8`)\\n- `semanticHints`: optional field for noting emotional tone or emoji interpretation\\n\\n> **Language Alignment Note**: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\\n\\n- **Artifact States**:\\n- `draft-alpha`\\n- `draft-beta`\\n- `finalized`\\n- `archived`\\n- `superseded`\\n\\n- **Dialog Structure**:\\n- Each dialog is a **thread**, composed of turns between human and AI participants\\n- **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs\\n- **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages\\n- **Conversation IDs and Linked Threads** allow tracing across dialog trees\\n\\n- **Source Capture**:\\n- Dialog metadata can include both human-provided and AI-suggested references\\n- Each turn may link to structured citations or retrieved web/knowledge sources\\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\\n\\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations across platforms and jurisdictions.\\n\\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\\n\\n- **Human Roles** (supporting one-to-many relationships):\\n- `promptAuthor[]`: one or more individuals responsible for initiating the query or idea\\n- `humanReviewer[]`: individuals who reviewed or critiqued system outputs\\n- `humanEditor[]`: contributors who modified or refined generated content\\n- `approver[]`: individuals or committees responsible for final acceptance or publication\\n\\n> CAIRN supports both **simple single-author workflows** and **complex editorial pipelines**, such as:\\n- A solo contributor who fills all roles (e.g., author = reviewer = editor = approver)\\n- Distributed workflows with separate reviewers and approvers across teams\\n- Institutional or committee-based approvals with audit trails and timestamps\\n\\nEach role entry can optionally include:\\n- `personId`\\n- `name`\\n- `roleDescription`\\n- `timestamp`\\n- `approvalStatus` (e.g., pending, approved, rejected)\\n- `sequenceOrder` (optional ordering if approvals follow a chain)\\n\\n- **System Provenance**:\\n- `modelUsed` (e.g., GPT-4-turbo)\\n- `platform` (e.g., chat.openai.com)\\n- `generationTimestamp`\\n- `modelLanguage`: the base language(s) of the language model used\\n\\n- **Prompt Fidelity**:\\n- `originalText`: user-typed input (possibly with typos or informal structure)\\n- `normalizedText`: interpreted version used by the model\\n- `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\\n- `encoding`: character encoding used (default: `UTF-8`)\\n- `semanticHints`: optional field for noting emotional tone or emoji interpretation\\n\\n> **Language Alignment Note**: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight</p>

Line Number	<p>possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.</p> <p>CAIRN JSON</p> <p>**Artifact States**: \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n</p> <p>**Dialog Structure**: \n - Each dialog is a **thread**, composed of turns between human and AI participants \n - **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs \n - **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages \n - **Conversation IDs and Linked Threads** allow tracing across dialog trees \n</p> <p>**Source Capture**: \n - Dialog metadata can include both human-provided and AI-suggested references \n - Each turn may link to structured citations or retrieved web/knowledge sources \n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution \n</p> <p>This structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.</p> <p>CAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:</p> <p>**Human Roles (Multi-Party Support)**: \n CAIRN supports **multiple individuals** fulfilling each role in the content lifecycle. Each artifact may include: \n - One or more 'promptAuthors' \n - One or more 'humanReviewers' \n - One or more 'humanEditors' \n - One or more 'approvers' \n</p> <p>In simple cases, a single person may fulfill all roles. In more complex workflows (e.g., institutional review, corporate authorship, regulatory oversight), distinct individuals or entities may be assigned to each role, and an **approval process** may include: \n - Sequential sign-off (e.g., reviewer \u279d editor \u279d approver) \n - Parallel or committee-based review \n - Conditional release logic (e.g., two-of-three must approve) \n</p> <p>Role metadata can also capture: \n - Timestamps of each action \n - Affiliation or organizational context \n - Digital signatures or references to verifiable credentials \n</p> <p>**System Provenance**: \n - 'modelUsed' (e.g., GPT-4-turbo) \n - 'platform' (e.g., chat.openai.com) \n - 'generationTimestamp' \n - 'modelLanguage': the base language(s) of the language model used \n</p> <p>**Prompt Fidelity**: \n - 'originalText': user-typed input (possibly with typos or informal structure) \n - 'normalizedText': interpreted version used by the model \n - 'language': natural language of the original input (e.g., 'en-US', 'zh-Hans') \n - 'encoding': character encoding used (default: 'UTF-8') \n - 'semanticHints': optional field for noting emotional tone or emoji interpretation \n</p> <p>**Language Alignment Note**: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.</p> <p>**Artifact States**: \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n</p> <p>**Dialog Structure**: \n - Each dialog is a **thread**, composed of turns between human and AI participants \n - **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs \n - **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages \n - **Conversation IDs and Linked Threads** allow tracing across dialog trees \n</p> <p>**Source Capture**: \n - Dialog metadata can include both human-provided and AI-suggested references \n - Each turn may link to structured citations or retrieved web/knowledge sources \n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution \n</p> <p>This structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.</p> <p>CAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:</p> <p>**Human Roles**: \n - 'promptAuthor' \n - 'humanReviewer' \n - 'humanEditor' \n - 'approver' \n</p> <p>**System Provenance**: \n - 'modelUsed' (e.g., GPT-4-turbo) \n - 'platform' (e.g., chat.openai.com) \n - 'generationTimestamp' \n - 'modelLanguage': the base language(s) of the language model used \n</p> <p>**Prompt Fidelity**: \n - 'originalText': user-typed input (possibly with typos or informal structure) \n - 'normalizedText': interpreted version used by the model \n - 'language': natural language of the original input (e.g., 'en-US', 'zh-Hans') \n - 'encoding': character encoding used (default: 'UTF-8') \n - 'semanticHints': optional field for noting emotional tone or emoji interpretation \n</p> <p>**Language Alignment Note**: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.</p> <p>**Artifact States**: \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n</p> <p>**Dialog Structure**: \n - Each dialog is a **thread**, composed of turns between human and AI participants \n - **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs \n - **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages \n - **Conversation IDs and Linked Threads** allow tracing across dialog trees \n</p> <p>**Source Capture**: \n - Dialog metadata can include both human-provided and AI-suggested references \n - Each turn may link to structured citations or retrieved web/knowledge sources \n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution \n</p> <p>This structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.</p> <p>CAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:</p> <p>**Human Roles**: \n - 'promptAuthor' \n - 'humanReviewer' \n - 'humanEditor' \n - 'approver' \n</p> <p>**System Provenance**: \n - 'modelUsed' (e.g., GPT-4-turbo) \n - 'platform' (e.g., chat.openai.com) \n - 'generationTimestamp' \n - **Prompt Fidelity**: \n - 'originalText': user-typed input (possibly with errors) \n - 'normalizedText': interpreted version used by the model \n</p> <p>**Artifact States**: \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n</p> <p>**Dialog Structure**: \n - Each dialog is a **thread**, composed of turns between human and AI participants \n - **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs \n - **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages \n - **Conversation IDs and Linked Threads** allow tracing across dialog trees \n</p> <p>**Source Capture**: \n - Dialog metadata can include both human-provided and AI-suggested references \n - Each turn may link to structured citations or retrieved web/knowledge sources \n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution \n</p> <p>This structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.</p> <p>CAIRN metadata includes:</p> <p>**Human Roles**: \n - 'promptAuthor' \n - 'humanReviewer' \n - 'humanEditor' \n - 'approver' \n</p> <p>**System Provenance**: \n - 'modelUsed' (e.g., GPT-4-turbo) \n - 'platform' (e.g., chat.openai.com) \n - 'generationTimestamp' \n - **Prompt Fidelity**: \n - 'originalText': user-typed input (possibly with errors) \n - 'normalizedText': interpreted version used by the model \n</p> <p>**Artifact States**: \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n</p> <p>**Lifecycle Graphing**: \n - Content threads linked across revisions \n - Semantic diffs and time-stamped transformations \n</p> <p>## Example Use Case: Human-AI Collaboration</p> <p>This proposal document itself was co-authored through:</p> <p>\n - A human prompt (Nick Stavros) \n - Draft generation by an AI assistant (Hamish I. MacCloud, GPT-4-turbo) \n - Human review, commentary, and approval \n - Iterative updates reflecting ethical and functional refinements \n</p> <p>This real-world example serves as the basis for the CAIRN model design.</p> <p>## Sample JSON-LD Fragment</p> <p>The following JSON-LD snippet illustrates a multi-agent collaboration involving both human and AI actors across platforms:</p> <pre> { "@context": { "@vocab": "https://cairn.ai/ns#", "prov": "http://www.w3.org/ns/prov#", "dct": "http://purl.org/dc/terms/" }, "@id": "urn:cairn:artifact:proposal- </pre>
-------------	---

Line Number	0011", in "@type": "cairn:ContentArtifact", in "cairn:artifactState": "draft-beta", in "cairn:version": "1.0.0", in "cairn:json": "2025-03-30T12:00:00Z", in "prov:wasGeneratedBy": { in "@id": "urn:cairn:activity:ai-coauthoring-001", in "prov:used": "urn:cairn:prompt:prompt-001", in "prov:wasAssociatedWith": { in "@id": "urn:openai:model:gpt-4-turbo", in } }, in "prov:wasAttributedTo": { in { in "@id": "urn:person:nick-stavros", in "prov:type": "prov:Person", in "foaf:name": "R. W. 'Nick' Stavros, Ph.D.", in "cairn:role": "promptAuthor", in }, in { in "@id": "urn:ai:meta:llama-3", in "prov:type": "prov:SoftwareAgent", in "cairn:role": "aiReviewer", in "cairn:platform": "Meta AI", in "cairn:modelVersion": "LLaMA-3", in "cairn:reviewTimestamp": "2025-03-30T13:30:00Z", in }, in { in "@id": "urn:ai:wenxin:yiyan", in "prov:type": "prov:SoftwareAgent", in "cairn:role": "aiReviewer", in "cairn:platform": "Baidu", in "cairn:modelVersion": "WenxinYiyan", in "cairn:reviewPurpose": "Cross-cultural verification", in "cairn:reviewTimestamp": "2025-03-30T14:00:00Z", in } }, in "cairn:originalText": "write it up as a proposal", in "cairn:normalizedText": "Please write this document as a formal standards proposal.", in "cairn:language": "en-US", in "cairn:encoding": "UTF-8", in "This example illustrates:\n- A human author initiating content\n- AI reviewers from two separate ecosystems providing insight\n- Explicit identification of each contributor\n- 2019s system, role, and timestamp\n", in "@context": { in "@vocab": "https://cairn.ai/ns/#1", in "prov": "http://www.w3.org/ns/prov#1", in "dct": "http://purl.org/dc/terms/1", in }, in "@id": "urn:cairn:artifact:proposal-001", in "@type": "cairn:ContentArtifact", in "cairn:artifactState": "draft-beta", in "cairn:version": "1.0.0", in "dct:created": "2025-03-30T12:00:00Z", in "prov:wasGeneratedBy": { in "@id": "urn:cairn:activity:ai-coauthoring-001", in "prov:used": "urn:cairn:prompt:prompt-001", in "prov:wasAssociatedWith": { in "@id": "urn:openai:model:gpt-4-turbo", in } }, in "prov:wasAttributedTo": { in "@id": "urn:person:nick-stavros", in "prov:type": "prov:Person", in "foaf:name": "R. W. 'Nick' Stavros, Ph.D.", in }, in "cairn:originalText": "write it up as a proposal", in "cairn:normalizedText": "Please write this document as a formal standards proposal.", in } ## Next Steps\n1. **Pilot**: Deploy CAIRN in select AI content generation workflows (e.g., reports, legal drafts, research)\n2. **Review**: Vet alignment with OMG P&P and W3C PROV-O community\n3. **Schema**: Publish a public JSON Schema, JSON-LD '@context', and RDF vocabulary\n4. **Standardize**: Advance through OMG working group and ISO/IEC JTC 1/SC 42\n5. **Promote Interoperability**: Engage AI platform developers (OpenAI, Meta, Baidu, etc.) to support cross-model collaboration using CAIRN\n6. **Extend Provenance Models**: Propose CAIRN as a common provenance layer enabling multilingual, multi-agent AI systems to operate transparently and cooperatively across global standards contexts\n1. **Pilot**: Deploy CAIRN in select AI content generation workflows (e.g., reports, legal drafts, research)\n2. **Review**: Vet alignment with OMG P&P and W3C PROV-O community\n3. **Schema**: Publish a public JSON Schema, JSON-LD '@context', and RDF vocabulary\n4. **Standardize**: Advance through OMG working group and ISO/IEC JTC 1/SC 42\n\n## Contact\n\nR. W. 'Nick' Stavros, Ph.D. \nObject Management Group \n2013 Pedigree & Provenance Working Group \n\n[Add contact email / phone number here]\n\n—\n\n**CAIRN** lays the stones for a verifiable, human-responsible path through the evolving terrain of AI-assisted knowledge creation.\n\n\n\n\n\ud83d\udcc4 **CAIRN Proposal to NIST** has been drafted and is now live in the document panel.\n\nIt includes:\n- A clear executive summary and purpose\n- The problem CAIRN solves in current provenance gaps\n- A structured metadata model with lifecycle, roles, and fidelity tracking\n- A live use case (this conversation!)\n- JSON-LD example\n- Next steps toward pilot and standardization via OMG and ISO\n\nLet me know if you'd like to:\n- Add contact info and affiliations\n- Generate a visual for presentation or appendix\n- Prepare a companion one-page overview or PowerPoint slide deck\n\nWe\ud2019re nearly ready to place this on the desk of NIST, OMG, or SC 42.\n\ud83e\udea8",
89	"cairn:normalizedText": "Thanks. Now lets write it up as a proposal.",
90	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
91	},
92	{
93	"@id": "urn:cairn:entry:011",
94	"cairn:timestamp": "2025-03-30T10:10:00Z",
95	"cairn:humanText": "I think the sero draft is intended to give \"American\" superiority. Is that correct?",
96	"cairn:aiResponse": "That's a really insightful and important question, Nick\ud2014and the answer is nuanced.\n\n—\n\n## \ud83c\uddfa\ud83c\uddf8 Is the NIST Zero Draft process intended to give \u201cAmerican\u201d superiority?\n\n### \u274c Not *explicitly*.\n\nNIST\ud2019s mission is not about *national supremacy*, but rather about:\n> \u201cPromoting U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology...\u201d\n\nThe **Zero Drafts** initiative is part of that broader mission to:\n- Accelerate AI standard development\n- Ensure U.S. leadership in setting global norms\n- Promote safety, interoperability, and trust\n\n—\n\n## \ud83d\uddfd But yes, there is an **implicit national strategy** behind it.\n\nThe U.S. federal government, through the **National AI Initiative Act**, and agencies like NIST, DARPA, and NSF, has made clear that:\n- **Standards leadership** is a strategic priority\n- The U.S. must not fall behind other global powers (like the EU, China, or ISO-centric nations) in shaping AI governance\n- Standards should reflect **U.S. values**\u201cespecially around civil liberties, free enterprise, and innovation\n\nSo while **NIST doesn't \u201cclaim supremacy,\u201d** the **Zero Drafts** are a mechanism to:\n- Get ahead of slower, more consensus-driven bodies like ISO or IEEE\n- Seed strong U.S.-aligned technical direction early\n- Influence the standards that will eventually be adopted internationally\n\n—\n\n## \ud83e\ud1d Why that matters for OMG, ISO, and CAIRN\n\nYour work with CAIRN\ud2014and OMG\ud2019s role\ud2014is vital because:\n- It provides an **international bridge**, not just a U.S.-centric view\n- OMG standards often **feed into ISO**, making them global by nature\n- Aligning CAIRN with **NIST\ud2019s Zero Draft** lets you shape U.S. policy while ensuring international interoperability\n\n—\n\n## \ud83d\udc04 Strategic framing for CAIRN\n\nTo appeal to both U.S. and global stakeholders, we could frame CAIRN as:\n> \u201cA lightweight, open, internationally harmonized metadata model\ud2014initiated in the U.S. through NIST\ud2019s Zero Drafts, designed for adoption across global AI standards frameworks (OMG, ISO/IEC JTC 1/SC 42, W3C).\u201d\n\n—\n\nWould you like to add a brief paragraph to the **proposal\ud2019s Executive Summary or Next Steps** to reflect this strategic positioning? \u2014\ud2019d be happy to drop it in.",

97 Line Number	"cairn:normalizedText": "I think the sero draft is intended to give \"American\" superiority. Is that correct?", CARIN JSON "cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
99	},
100	{
101	"@id": "urn:cairn:entry:012",
102	"cairn:timestamp": "2025-03-30T10:11:00Z",
103	"cairn:humanText": "Yes",
104	"cairn:aiResponse": "\u2705 I've added a strategic paragraph to the Executive Summary of the proposal, emphasizing CAIRN's role in supporting U.S. leadership through NIST while maintaining global interoperability and alignment with international standards bodies like ISO and OMG. \n\nWould you like to: \n- Add a matching note in the Next Steps section about engagement with SC 42 or other national standards bodies? \n- Draft a short cover letter or submission email to accompany the proposal? \n\nLet me know how you want to proceed!",
105	"cairn:normalizedText": "Yes",
106	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
107	},
108	{
109	"@id": "urn:cairn:entry:013",
110	"cairn:timestamp": "2025-03-30T10:12:00Z",
111	"cairn:humanText": "A couple of more things. I think that AI should be used interactively in a dialog. We want o be ble to link these conversations together so that a complete thread can be followed from inception to completion of the answer. We want to be able to support a many-to-many model. For example, a dialog could spawn multiple dialogs kind of like a tree. But also, sometimes conversations need to be coalsed into a unified converstoin. But when it dos occur, it doew not mean that one thread is completely subsumed into the next one. Also, we want the ability to capture sources in the dialog. In our example, I provided the first reference, you could have added more references but I need all those captured in the model. ",
112	"cairn:aiResponse": "\u2705 CAIRN now explicitly supports: \n\n \ud83d\udd01 Interactive dialogs as threaded, traceable conversations \n- \ud83c\udf33 Many-to-many structures : branching and merging dialogs \n- \ud83d\udcda Source capture from both human and AI participants \n- \ud83e\udde9 Coalescence without full subsumption \u2014 threads remain independently traceable \n\nThis brings us even closer to a dialog-centric standard for AI provenance. \n\nWould you like me to: \n- Add a sample JSON-LD structure showing a branching dialog with references? \n- Include this branching/coalescing model in the lifecycle diagram or visual spec? \n\nLet \u2019s keep sculpting. \ud83e\udea8",
113	"cairn:normalizedText": "A couple of more things. I think that AI should be used interactively in a dialog. We want o be ble to link these conversations together so that a complete thread can be followed from inception to completion of the answer. We want to be able to support a many-to-many model. For example, a dialog could spawn multiple dialogs kind of like a tree. But also, sometimes conversations need to be coalsed into a unified converstoin. But when it dos occur, it doew not mean that one thread is completely subsumed into the next one. Also, we want the ability to capture sources in the dialog. In our example, I provided the first reference, you could have added more references but I need all those captured in the model. ",
114	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
115	},
116	{
117	"@id": "urn:cairn:entry:014",
118	"cairn:timestamp": "2025-03-30T10:13:00Z",
119	"cairn:humanText": "Do we need to also capture the human language used for the dialog? English versus french versus Chineese etc? Do we need to worry about the charcter set used? What about the used of emojis? Are they important?",
	"cairn:aiResponse": "Yes, these are *crucial* considerations \u2014 especially for a provenance model that aims to be robust, international, and future-proof. You're now thinking like both a standards architect and a digital historian. Let's walk through them one by one: \n\n— \n\n## \ud83c\udf10 1. Human Language (Natural Language) \n\n\u2705 Yes, we should capture this. \n\n### Why? \n- Language choice affects interpretation, translation, and semantic meaning. \n- Regulatory, legal, or policy-sensitive content might require provenance of the original language. \n- Same prompt in different languages may trigger

Line Number	CHARIN JSON
141	"@id": "urn:cairn:entry:017",
142	"cairn.timestamp": "2025-03-30T10:16:00Z",
143	"cairn.humanText": "I think we are almost there. Now, we want to be able to allow not only humans but also other AI systems to be reviewers. So, for example, once we finish this, we could hand it from our environment (i.e., ChatGPT and Hamish) to say Meta's AI or even the Chinese AI. The real end goal is to make AI an integral part of the process, not something done behind the scenes. ",
	<p>"cairn.aiResponse": "Cairn Proposal to Nist\n\n--\nIntitle: \"Proposal: CAIRN \u2014 Collaborative AI Artifact Record & Notation\"\n\nauthors:\n- R. W. \"Nick\" Stavros, Ph.D.\n- Hamish I. MacCloud, AIA\n\nversion: 1.0.0\n\ndate: 2025-03-30\n\nsubmittedTo: \"National Institute of Standards and Technology (NIST) \u2014 2014 AI Standards Zero Drafts Pilot\"\n\norganization: \"Object Management Group (OMG) \u2014 Pedigree & Provenance Group\"\n\n---\n\n## Executive Summary\n\nCAIRN aligns with NIST's strategic goal of advancing U.S. leadership in AI standards while remaining internationally interoperable. It is intentionally designed to integrate with global frameworks (such as OMG Pedigree & Provenance, ISO/IEC JTC 1/SC 42, and W3C PROV-O), ensuring the U.S. can lead without isolating. By contributing CAIRN as a Zero Draft, we help accelerate national priorities while establishing a durable foundation for international trust and adoption.\n\nThis proposal introduces **CAIRN** \u2014 Collaborative AI Artifact Record & Notation \u2014 a lightweight, JSON-LD \u2014 based metadata standard designed to capture the full lifecycle, lineage, and human-system accountability of AI-generated content. CAIRN extends existing provenance models by addressing the real-world needs of generative AI workflows, including prompt normalization, multi-phase review, and explicit human accountability. It is submitted in response to NIST's call for zero-draft proposals to accelerate the development of voluntary, consensus-based AI standards.\n\n## Purpose and Scope\n\nCAIRN is designed to:\n\n- Capture **who did what and when** in collaborative human-AI content creation\n- Represent the **prompt-to-publication lifecycle** of AI-assisted artifacts\n- Ensure **human accountability** is never obfuscated by system behavior\n- Track both **original and normalized prompts**, enabling transparent interpretation\n- Represent **maturity states** (alpha, beta, final) of evolving artifacts\n- Integrate with OMG \u2014 2019s existing **Pedigree & Provenance (P&P)** model\n- Interoperate with W3C **PROV-O** and emerging ISO/IEC JTC 1/SC 42 standards\n\n## Problem Addressed\n\nAs generative AI becomes embedded in government, industry, and research workflows, the ability to trace, review, and validate AI-generated artifacts becomes essential. Current provenance systems:\n\n- Do not distinguish between original and normalized user input\n- Do not support iterative content refinement and approval chains\n- Lack defined maturity states or editorial traceability\n- Fail to assign explicit human responsibility for finalized output\n\nCAIRN proposes to fill this gap.\n\n## CAIRN Model Overview\n\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- **Human and AI Roles** (supporting one-to-many relationships):\n- `promptAuthor`: one or more individuals responsible for initiating the query or idea\n- `humanReviewer`: individuals or AI systems that reviewed or critiqued content\n- `humanEditor`: contributors \u2014 human or AI \u2014 that modified or refined content\n- `approver`: individuals, committees, or systems responsible for final acceptance or publication\n\n- CAIRN explicitly supports the role of **AI as a first-class participant** in collaborative workflows. AI systems can serve as reviewers, editors, or even approvers when formally included in the provenance chain. Each AI actor can be identified by system name, model version, deployment platform, and contribution timestamp.\n\nEach role entry can include:\n\n- `agentType` (e.g., human, ai)\n- `personId` or `systemId`\n- `name`\n- `roleDescription`\n- `timestamp`\n- `approvalStatus` (e.g., pending, approved, rejected)\n- `sequenceOrder` (optional ordering if approvals follow a chain)\n\nExample: After initial human drafting in ChatGPT, the content could be passed to Meta's LLaMA for policy review, then to a Chinese-developed LLM for cross-cultural verification \u2014 all captured as discrete `aiReviewer` events.\n\n- **System Provenance**:\n- `modelUsed` (e.g., GPT-4-turbo, LLaMA, WenxinYiyan)\n- `platform` (e.g., chat.openai.com, huggingface.co)\n- `generationTimestamp`\n- `modelLanguage`: the base language(s) of the language model used\n\n- **Prompt Fidelity**:\n- `originalText`: user-typed input (possibly with typos or informal structure)\n- `normalizedText`: interpreted version used by the model\n- `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n- `encoding`: character encoding used (default: `UTF-8`)\n- `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n- **Language Alignment Note**:\nWhen the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- **Artifact States**:\n- `draft-alpha`\n- `draft-beta`\n- `finalized`\n- `archived`\n- `superseded`\n\n- **Dialog Structure**:\n- Each dialog is a **thread**, composed of turns between human and AI participants\n- **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs\n- **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages\n- **Conversation IDs and Linked Threads** allow tracing across dialog trees\n\n- **Source Capture**:\n- Dialog metadata can include both human-provided and AI-suggested references\n- Each turn may link to structured citations or retrieved web/knowledge sources\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations across platforms and jurisdictions.\n\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- **Human Roles** (supporting one-to-many relationships):\n- `promptAuthor`: one or more individuals responsible for initiating the query or idea\n- `humanReviewer`: individuals who reviewed or critiqued system outputs\n- `humanEditor`: contributors who modified or refined generated content\n- `approver`: individuals or committees responsible for final acceptance or publication\n\n- CAIRN supports both **simple single-author workflows** and **complex editorial pipelines**, such as:\n\n- A solo contributor who fills all roles (e.g., author = reviewer = editor = approver)\n- Distributed workflows with separate reviewers and approvers across teams\n- Institutional or committee-based approvals with audit trails and timestamps\n\nEach role entry can optionally include:\n\n- `personId`\n- `name`\n- `roleDescription`\n- `timestamp`\n- `approvalStatus` (e.g., pending, approved, rejected)\n- `sequenceOrder` (optional ordering if approvals follow a chain)\n\n- **System Provenance**:\n- `modelUsed` (e.g., GPT-4-turbo)\n- `platform` (e.g., chat.openai.com)\n- `generationTimestamp`\n- `modelLanguage`: the base language(s) of the language model used\n\n- **Prompt Fidelity**:\n- `originalText`: user-typed input (possibly with typos or informal structure)\n- `normalizedText`: interpreted version used by the model\n- `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n- `encoding`: character encoding used (default: `UTF-8`)\n- `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n- **Language Alignment Note**:\nWhen the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- **Artifact States**:\n- `draft-alpha`\n- `draft-beta`\n- `finalized`\n- `archived`\n- `superseded`\n\n- **Dialog Structure**:\n- Each dialog is a **thread**, composed of turns between human and AI participants\n- **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs\n- **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages\n- **Conversation IDs and Linked Threads** allow tracing across dialog trees\n\n- **Source Capture**:\n- Dialog metadata can include both human-provided and AI-suggested references\n- Each turn may link to structured citations or retrieved web/knowledge sources\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations across platforms and jurisdictions.\n\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- **Human Roles** (supporting one-to-many relationships):\n- `promptAuthor`: one or more individuals responsible for initiating the query or idea\n- `humanReviewer`: individuals who reviewed or critiqued system outputs\n- `humanEditor`: contributors who modified or refined generated content\n- `approver`: individuals or committees responsible for final acceptance or publication\n\n- CAIRN supports both **simple single-author workflows** and **complex editorial pipelines**, such as:\n\n- A solo contributor who fills all roles (e.g., author = reviewer = editor = approver)\n- Distributed workflows with separate reviewers and approvers across teams\n- Institutional or committee-based approvals with audit trails and timestamps\n\nEach role entry can optionally include:\n\n- `personId`\n- `name`\n- `roleDescription`\n- `timestamp`\n- `approvalStatus` (e.g., pending, approved, rejected)\n- `sequenceOrder` (optional ordering if approvals follow a chain)\n\n- **System Provenance**:\n- `modelUsed` (e.g., GPT-4-turbo)\n- `platform` (e.g., chat.openai.com)\n- `generationTimestamp`\n- `modelLanguage`: the base language(s) of the language model used\n\n- **Prompt Fidelity**:\n- `originalText`: user-typed input (possibly with typos or informal structure)\n- `normalizedText`: interpreted version used by the model\n- `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n- `encoding`: character encoding used (default: `UTF-8`)\n- `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n- **Language Alignment Note**:\nWhen the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- **Artifact States**:\n- `draft-alpha`\n- `draft-beta`\n- `finalized`\n- `archived`\n- `superseded`\n\n- **Dialog Structure**:\n- Each dialog is a **thread**, composed of turns between human and AI participants\n- **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs\n- **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages\n- **Conversation IDs and Linked Threads** allow tracing across dialog trees\n\n- **Source Capture**:\n- Dialog metadata can include both human-provided and AI-suggested references\n- Each turn may link to structured citations or retrieved web/knowledge sources\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations across platforms and jurisdictions.\n\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- **Human Roles** (supporting one-to-many relationships):\n- `promptAuthor`: one or more individuals responsible for initiating the query or idea\n- `humanReviewer`: individuals who reviewed or critiqued system outputs\n- `humanEditor`: contributors who modified or refined generated content\n- `approver`: individuals or committees responsible for final acceptance or publication\n\n- CAIRN supports both **simple single-author workflows** and **complex editorial pipelines**, such as:\n\n- A solo contributor who fills all roles (e.g., author = reviewer = editor = approver)\n- Distributed workflows with separate reviewers and approvers across teams\n- Institutional or committee-based approvals with audit trails and timestamps\n\nEach role entry can optionally include:\n\n- `personId`\n- `name`\n- `roleDescription`\n- `timestamp`\n- `approvalStatus` (e.g., pending, approved, rejected)\n- `sequenceOrder` (optional ordering if approvals follow a chain)\n\n- **System Provenance**:\n- `modelUsed` (e.g., GPT-4-turbo)\n- `platform` (e.g., chat.openai.com)\n- `generationTimestamp`\n- `modelLanguage`: the base language(s) of the language model used\n\n- **Prompt Fidelity**:\n- `originalText`: user-typed input (possibly with typos or informal structure)\n- `normalizedText`: interpreted version used by the model\n- `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n- `encoding`: character encoding used (default: `UTF-8`)\n- `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n- **Language Alignment Note**:\nWhen the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- **Artifact States**:\n- `draft-alpha`\n- `draft-beta`\n- `finalized`\n- `archived`\n- `superseded`\n\n- **Dialog Structure**:\n- Each dialog is a **thread**, composed of turns between human and AI participants\n- **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs\n- **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages\n- **Conversation IDs and Linked Threads** allow tracing across dialog trees\n\n- **Source Capture**:\n- Dialog metadata can include both human-provided and AI-suggested references\n- Each turn may link to structured citations or retrieved web/knowledge sources\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations across platforms and jurisdictions.\n\nCAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- **Human Roles** (supporting one-to-many relationships):\n- `promptAuthor`: one or more individuals responsible for initiating the query or idea\n- `humanReviewer`: individuals who reviewed or critiqued system outputs\n- `humanEditor`: contributors who modified or refined generated content\n- `approver`: individuals or committees responsible for final acceptance or publication\n\n- CAIRN supports both **simple single-author workflows** and **complex editorial pipelines**, such as:\n\n- A solo contributor who fills all roles (e.g., author = reviewer = editor = approver)\n- Distributed workflows with separate reviewers and approvers across teams\n- Institutional or committee-based approvals with audit trails and timestamps\n\nEach role entry can optionally include:\n\n- `personId`\n- `name`\n- `roleDescription`\n- `timestamp`\n- `approvalStatus` (e.g., pending, approved, rejected)\n- `sequenceOrder` (optional ordering if approvals follow a chain)\n\n- **System Provenance**:\n- `modelUsed` (e.g., GPT-4-turbo)\n- `platform` (e.g., chat.openai.com)\n- `generationTimestamp`\n- `modelLanguage`: the base language(s) of the language model used\n\n- **Prompt Fidelity**:\n- `originalText`: user-typed input (possibly with typos or informal structure)\n- `normalized</p>

Line Number	<p>`normalizedText`: interpreted version used by the model\n - `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n - `encoding`: character encoding used (default: `UTF-8`)\n - `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n> Language Alignment Note: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- Artifact States: \n - `draft-alpha`\n - `draft-beta`\n - `finalized`\n - `archived`\n - `superseded`\n\n- Dialog Structure: \n - Each dialog is a thread, composed of turns between human and AI participants\n - Many-to-many branching: a single prompt can spawn multiple sub-dialogs\n - Coalescence: multiple threads can merge into a unified conversation, while preserving original lineages\n - Conversation IDs and Linked Threads allow tracing across dialog trees\n\n- Source Capture: \n - Dialog metadata can include both human-provided and AI-suggested references\n - Each turn may link to structured citations or retrieved web/knowledge sources\n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.\n\nCAIRN is also designed to support interactive, multi-threaded dialogues with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- Human Roles (Multi-Party Support): \n CAIRN supports multiple individuals fulfilling each role in the content lifecycle. Each artifact may include:\n - One or more `promptAuthors`\n - One or more `humanReviewers`\n - One or more `humanEditors`\n - One or more `approvers`\n\nIn simple cases, a single person may fulfill all roles. In more complex workflows (e.g., institutional review, corporate authorship, regulatory oversight), distinct individuals or entities may be assigned to each role, and an approval process may include:\n - Sequential sign-off (e.g., reviewer \u279d editor \u279d approver)\n - Parallel or committee-based review\n - Conditional release logic (e.g., two-of-three must approve)\n\nRole metadata can also capture:\n - Timestamps of each action\n - Affiliation or organizational context\n - Digital signatures or references to verifiable credentials\n\n- System Provenance: \n - `modelUsed` (e.g., GPT-4-turbo)\n - `platform` (e.g., chat.openai.com)\n - `generationTimestamp`\n - `modelLanguage`: the base language(s) of the language model used\n\n- Prompt Fidelity: \n - `originalText`: user-typed input (possibly with typos or informal structure)\n - `normalizedText`: interpreted version used by the model\n - `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n - `encoding`: character encoding used (default: `UTF-8`)\n - `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n> Language Alignment Note: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- Artifact States: \n - `draft-alpha`\n - `draft-beta`\n - `finalized`\n - `archived`\n - `superseded`\n\n- Dialog Structure: \n - Each dialog is a thread, composed of turns between human and AI participants\n - Many-to-many branching: a single prompt can spawn multiple sub-dialogs\n - Coalescence: multiple threads can merge into a unified conversation, while preserving original lineages\n - Conversation IDs and Linked Threads allow tracing across dialog trees\n\n- Source Capture: \n - Dialog metadata can include both human-provided and AI-suggested references\n - Each turn may link to structured citations or retrieved web/knowledge sources\n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.\n\nCAIRN is also designed to support interactive, multi-threaded dialogues with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- Human Roles: \n - `promptAuthor`\n - `humanReviewer`\n - `humanEditor`\n - `approver`\n\n- System Provenance: \n - `modelUsed` (e.g., GPT-4-turbo)\n - `platform` (e.g., chat.openai.com)\n - `generationTimestamp`\n - `modelLanguage`: the base language(s) of the language model used\n\n- Prompt Fidelity: \n - `originalText`: user-typed input (possibly with typos or informal structure)\n - `normalizedText`: interpreted version used by the model\n - `language`: natural language of the original input (e.g., `en-US`, `zh-Hans`)\n - `encoding`: character encoding used (default: `UTF-8`)\n - `semanticHints`: optional field for noting emotional tone or emoji interpretation\n\n> Language Alignment Note: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- Artifact States: \n - `draft-alpha`\n - `draft-beta`\n - `finalized`\n - `archived`\n - `superseded`\n\n- Dialog Structure: \n - Each dialog is a thread, composed of turns between human and AI participants\n - Many-to-many branching: a single prompt can spawn multiple sub-dialogs\n - Coalescence: multiple threads can merge into a unified conversation, while preserving original lineages\n - Conversation IDs and Linked Threads allow tracing across dialog trees\n\n- Source Capture: \n - Dialog metadata can include both human-provided and AI-suggested references\n - Each turn may link to structured citations or retrieved web/knowledge sources\n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.\n\nCAIRN metadata includes:\n\n- Human Roles: \n - `promptAuthor`\n - `humanReviewer`\n - `humanEditor`\n - `approver`\n\n- System Provenance: \n - `modelUsed` (e.g., GPT-4-turbo)\n - `platform` (e.g., chat.openai.com)\n - `generationTimestamp`\n\n- Prompt Fidelity: \n - `originalText`: user-typed input (possibly with errors)\n - `normalizedText`: interpreted version used by the model\n\n- Artifact States: \n - `draft-alpha`\n - `draft-beta`\n - `finalized`\n - `archived`\n - `superseded`\n\n- Lifecycle Graphing: \n - Content threads linked across revisions\n - Semantic diffs and time-stamped transformations\n\n## Example Use Case: Human-AI Collaboration\n\nThis proposal document itself was co-authored through:\n\n- A human prompt (Nick Stavros)\n - Draft generation by an AI assistant (Hamish I. MacCloud, GPT-4-turbo)\n - Human review, commentary, and approval\n - Iterative updates reflecting ethical and functional</p>
144	

Line Number	<p>refinements\n\nThis real-world example serves as the basis for the CAIRN model design.\n\n## Sample JSON-LD</p> <p>CAIRN JSON\n\nThe following JSON-LD snippet illustrates a multi-agent collaboration involving both human and AI actors across platforms:\n\n```\n{\n "@context": {\n "@vocab": "https://cairn.ai/ns/1.0/",\n "prov": "http://www.w3.org/ns/prov#1.0",\n "dct": "http://purl.org/dc/terms/1.1"\n },\n "@id": "urn:cairn:artifact:proposal-001",\n "@type": "cairn:ContentArtifact",\n "cairn:artifactState": "draft-beta",\n "cairn:version": "1.0.0",\n "dct:created": "2025-03-30T12:00:00Z",\n "prov:wasGeneratedBy": {\n "@id": "urn:cairn:activity:ai-coauthoring-001",\n "prov:used": "urn:cairn:prompt:prompt-001",\n "prov:wasAssociatedWith": {\n "@id": "urn:openai:model:gpt-4-turbo",\n "prov:wasAttributedTo": {\n "@id": "urn:person:nick-stavros",\n "prov:type": "prov:Person",\n "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",\n "cairn:role": "promptAuthor",\n "@id": "urn:ai:meta:llama-3",\n "prov:type": "prov:SoftwareAgent",\n "cairn:role": "aiReviewer",\n "cairn:platform": "Meta AI",\n "cairn:modelVersion": "LLaMA-3",\n "cairn:reviewTimestamp": "2025-03-30T13:30:00Z",\n "@id": "urn:ai:wenxin:yiyan",\n "prov:type": "prov:SoftwareAgent",\n "cairn:role": "aiReviewer",\n "cairn:platform": "Baidu",\n "cairn:modelVersion": "WenxinYiyan",\n "cairn:reviewPurpose": "Cross-cultural verification",\n "cairn:reviewTimestamp": "2025-03-30T14:00:00Z",\n "cairn:originalText": "write it up as a proposal",\n "cairn:normlizedText": "Please write this document as a formal standards proposal.",\n "cairn:language": "en-US",\n "cairn:encoding": "UTF-8",\n "n": "This example illustrates:\n- A human author initiating content\n- AI reviewers from two separate ecosystems providing insight\n- Explicit identification of each contributor\n\n2019s system, role, and timestamp\n\n```\n\n{\n "@context": {\n "@vocab": "https://cairn.ai/ns/1.0/",\n "prov": "http://www.w3.org/ns/prov#1.0",\n "dct": "http://purl.org/dc/terms/1.1"\n },\n "@id": "urn:cairn:artifact:proposal-001",\n "@type": "cairn:ContentArtifact",\n "cairn:artifactState": "draft-beta",\n "cairn:version": "1.0.0",\n "dct:created": "2025-03-30T12:00:00Z",\n "prov:wasGeneratedBy": {\n "@id": "urn:cairn:activity:ai-coauthoring-001",\n "prov:used": {\n "@id": "urn:cairn:prompt:prompt-001",\n "prov:wasAssociatedWith": {\n "@id": "urn:openai:model:gpt-4-turbo",\n "prov:wasAttributedTo": {\n "@id": "urn:person:nick-stavros",\n "prov:type": "prov:Person",\n "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",\n "cairn:originalText": "write it up as a proposal",\n "cairn:normlizedText": "Please write this document as a formal standards proposal.",\n "n": "Next Steps\n\n1. **Pilot**: Deploy CAIRN in select AI content generation workflows (e.g., reports, legal drafts, research)\n\n2. **Review**: Vet alignment with OMG P&P and W3C PROV-O community\n\n3. **Schema**: Publish a public JSON Schema, JSON-LD '@context', and RDF vocabulary\n\n4. **Standardize**: Advance through OMG working group and ISO/IEC JTC 1/SC 42\n\n5. **Promote Interoperability**: Engage AI platform developers (OpenAI, Meta, Baidu, etc.) to support cross-model collaboration using CAIRN\n\n6. **Extend Provenance Models**: Propose CAIRN as a common provenance layer enabling multilingual, multi-agent AI systems to operate transparently and cooperatively across global standards contexts\n\n1. **Pilot**: Deploy CAIRN in select AI content generation workflows (e.g., reports, legal drafts, research)\n\n2. **Review**: Vet alignment with OMG P&P and W3C PROV-O community\n\n3. **Schema**: Publish a public JSON Schema, JSON-LD '@context', and RDF vocabulary\n\n4. **Standardize**: Advance through OMG working group and ISO/IEC JTC 1/SC 42\n\n## Contact\n\nR. W. 'Nick' Stavros, Ph.D. \nObject Management Group \n2013 Pedigree & Provenance Working Group \n[Add contact email / phone number here]\n\n—\n\nCAIRN** lays the stones for a verifiable, human-responsible path through the evolving terrain of AI-assisted knowledge creation.\n\n\n2014 The **CAIRN Model Overview** now formally includes:\n\n- \nAI participants as reviewers, editors, or approvers\n- \nSupport for **cross-platform collaboration** (e.g., ChatGPT \nLLaMa \nWenxinYiyan)\n- \nRich agent metadata for both human and AI actors\n- \nA vision where AI is not behind the curtain, but **integrated into the accountable workflow**\n\n\nWould you like to:\n- Include a multi-AI JSON-LD example showing AI systems in the review chain?\n- Expand the proposal\n2019s conclusion to emphasize CAIRN\n2019s role in international, multi-agent collaboration?\n\nWe\n2019re now modeling the future of AI not just as a tool\n2014but as a **trusted peer in collective knowledge creation.**\n\n",\n "cairn:normlizedText": "I think we are almost there. Now, we want to be able to allow not only humans but also other AI systems to be reviewers. So, for example, once we finish this, we could hand it from our environment (i.e., ChatGPT and Hamish) to say Meta's AI or even the Chinese AI. The real end goal is to make AI an integral part of the process, not something done behind the scenes.",\n "cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined.",\n },\n {\n "@id": "urn:cairn:entry:018",\n "cairn:timestamp": "2025-03-30T10:17:00Z",\n "cairn:humanText": "Yes, let's do these.",\n "cairn:aiResponse": "Cairn Proposal to Nist\n\n—\n\nTitle: 'Proposal: CAIRN \n2014 Collaborative AI Artifact Record & Notation'\n\nauthors: \n- R. W. 'Nick' Stavros, Ph.D. \n- Hamish I. MacCloud, AIA\n\nversion: 1.0.0\n\ndate: 2025-03-30\n\nsubmittedTo: 'National Institute of Standards and Technology (NIST) \n2014 AI Standards Zero Drafts Pilot'\n\norganization: 'Object Management Group (OMG) \n2014 Pedigree & Provenance Group'\n\n—\n\n\n## Executive Summary\n\nCAIRN aligns with NIST's strategic goal of advancing U.S. leadership in AI standards while remaining internationally interoperable. It is intentionally designed to integrate with global frameworks (such as OMG Pedigree & Provenance, ISO/IEC JTC 1/SC 42, and W3C PROV-O), ensuring the U.S. can lead without isolating. By contributing CAIRN as a Zero Draft, we help accelerate national priorities while establishing a durable foundation for international trust and adoption.\n\n\nThis proposal introduces **CAIRN**\n2014 *Collaborative AI Artifact Record & Notation*\n2014 a lightweight, JSON-LD\n2013based metadata standard designed to capture the full lifecycle, lineage, and human-system accountability of AI-generated content. CAIRN extends existing provenance models by addressing the real-world needs of generative AI workflows, including prompt normalization, multi-phase review, and explicit human accountability. It is submitted in response to NIST's call for zero-draft proposals to accelerate the development of voluntary, consensus-based AI standards.\n\n\n## Purpose and Scope\n\nCAIRN is designed to:\n\n- Capture **who did what and when** in collaborative human-AI content creation\n- Represent the **prompt-to-publication lifecycle** of AI-assisted artifacts\n- Ensure **human accountability** is</p>
145	"cairn:normlizedText": "I think we are almost there. Now, we want to be able to allow not only humans but also other AI systems to be reviewers. So, for example, once we finish this, we could hand it from our environment (i.e., ChatGPT and Hamish) to say Meta's AI or even the Chinese AI. The real end goal is to make AI an integral part of the process, not something done behind the scenes.",
146	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
147	},
148	{
149	"@id": "urn:cairn:entry:018",
150	"cairn:timestamp": "2025-03-30T10:17:00Z",
151	"cairn:humanText": "Yes, let's do these.",
	<p>"cairn:aiResponse": "Cairn Proposal to Nist\n\n—\n\nTitle: 'Proposal: CAIRN \n2014 Collaborative AI Artifact Record & Notation'\n\nauthors: \n- R. W. 'Nick' Stavros, Ph.D. \n- Hamish I. MacCloud, AIA\n\nversion: 1.0.0\n\ndate: 2025-03-30\n\nsubmittedTo: 'National Institute of Standards and Technology (NIST) \n2014 AI Standards Zero Drafts Pilot'\n\norganization: 'Object Management Group (OMG) \n2014 Pedigree & Provenance Group'\n\n—\n\n\n## Executive Summary\n\nCAIRN aligns with NIST's strategic goal of advancing U.S. leadership in AI standards while remaining internationally interoperable. It is intentionally designed to integrate with global frameworks (such as OMG Pedigree & Provenance, ISO/IEC JTC 1/SC 42, and W3C PROV-O), ensuring the U.S. can lead without isolating. By contributing CAIRN as a Zero Draft, we help accelerate national priorities while establishing a durable foundation for international trust and adoption.\n\n\nThis proposal introduces **CAIRN**\n2014 *Collaborative AI Artifact Record & Notation*\n2014 a lightweight, JSON-LD\n2013based metadata standard designed to capture the full lifecycle, lineage, and human-system accountability of AI-generated content. CAIRN extends existing provenance models by addressing the real-world needs of generative AI workflows, including prompt normalization, multi-phase review, and explicit human accountability. It is submitted in response to NIST's call for zero-draft proposals to accelerate the development of voluntary, consensus-based AI standards.\n\n\n## Purpose and Scope\n\nCAIRN is designed to:\n\n- Capture **who did what and when** in collaborative human-AI content creation\n- Represent the **prompt-to-publication lifecycle** of AI-assisted artifacts\n- Ensure **human accountability** is</p>

Line Number	<p>never obfuscated by system behavior\n- Track both original and normalized prompts, enabling transparent provenance\n- Represent maturity states (alpha, beta, final) of evolving artifacts\n- Integrate with OMC\u2019s existing Pedigree & Provenance (P&P) model\n- Interoperate with W3C PROV-O and emerging ISO/IEC JTC 1/SC 42 standards\n\n## Problem Addressed\n\nAs generative AI becomes embedded in government, industry, and research workflows, the ability to trace, review, and validate AI-generated artifacts becomes essential. Current provenance systems:\n- Do not distinguish between original and normalized user input\n- Do not support iterative content refinement and approval chains\n- Lack defined maturity states or editorial traceability\n- Fail to assign explicit human responsibility for finalized output\n\nCAIRN proposes to fill this gap.\n\n## CAIRN Model Overview\n\nCAIRN is also designed to support interactive, multi-threaded dialogues with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- Human and AI Roles (supporting one-to-many relationships):\n- <code>promptAuthor[]</code>: one or more individuals responsible for initiating the query or idea\n- <code>humanReviewer[]</code> / <code>aiReviewer[]</code>: individuals or AI systems that reviewed or critiqued content\n- <code>humanEditor[]</code> / <code>aiEditor[]</code>: contributors\n- <code>human</code> or <code>AI</code> that modified or refined content\n- <code>approver[]</code>: individuals, committees, or systems responsible for final acceptance or publication\n\nCAIRN explicitly supports the role of AI as a first-class participant in collaborative workflows. AI systems can serve as reviewers, editors, or even approvers when formally included in the provenance chain. Each AI actor can be identified by system name, model version, deployment platform, and contribution timestamp.\n\nEach role entry can include:\n- <code>agentType</code> (e.g., human, ai)\n- <code>personId</code> or <code>systemId</code>\n- <code>name</code>\n- <code>roleDescription</code>\n- <code>timestamp</code>\n- <code>approvalStatus</code> (e.g., pending, approved, rejected)\n- <code>sequenceOrder</code> (optional ordering if approvals follow a chain)\n\nExample: After initial human drafting in ChatGPT, the content could be passed to Meta's LLaMA for policy review, then to a Chinese-developed LLM for cross-cultural verification\n- <code>aiReviewer</code> events.\n\n- System Provenance: - <code>modelUsed</code> (e.g., GPT-4-turbo, LLaMA, WenxinYiyan)\n- <code>platform</code> (e.g., chat.openai.com, huggingface.co)\n- <code>generationTimestamp</code>\n- <code>modelLanguage</code>: the base language(s) of the language model used\n\n- Prompt Fidelity: - <code>originalText</code>: user-typed input (possibly with typos or informal structure)\n- <code>normalizedText</code>: interpreted version used by the model\n- <code>language</code>: natural language of the original input (e.g., <code>en-US</code>, <code>zh-Hans</code>)\n- <code>encoding</code>: character encoding used (default: <code>UTF-8</code>)\n- <code>semanticHints</code>: optional field for noting emotional tone or emoji interpretation\n\n- Language Alignment Note: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- Artifact States: - <code>draft-alpha</code>\n- <code>draft-beta</code>\n- <code>finalized</code>\n- <code>archived</code>\n- <code>superseded</code>\n\n- Dialog Structure: - Each dialog is a thread, composed of turns between human and AI participants\n- Many-to-many branching: a single prompt can spawn multiple sub-dialogs\n- Coalescence: multiple threads can merge into a unified conversation, while preserving original lineages\n- Conversation IDs and Linked Threads allow tracing across dialog trees\n\n- Source Capture: - Dialog metadata can include both human-provided and AI-suggested references\n- Each turn may link to structured citations or retrieved web/knowledge sources\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations across platforms and jurisdictions.\n\nCAIRN is also designed to support interactive, multi-threaded dialogues with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- Human Roles (supporting one-to-many relationships):\n- <code>promptAuthor[]</code>: one or more individuals responsible for initiating the query or idea\n- <code>humanReviewer[]</code>: individuals who reviewed or critiqued system outputs\n- <code>humanEditor[]</code>: contributors who modified or refined generated content\n- <code>approver[]</code>: individuals or committees responsible for final acceptance or publication\n\nCAIRN supports both simple single-author workflows and complex editorial pipelines, such as:\n- A solo contributor who fills all roles (e.g., author = reviewer = editor = approver)\n- Distributed workflows with separate reviewers and approvers across teams\n- Institutional or committee-based approvals with audit trails and timestamps\n\nEach role entry can optionally include:\n- <code>personId</code>\n- <code>name</code>\n- <code>roleDescription</code>\n- <code>timestamp</code>\n- <code>approvalStatus</code> (e.g., pending, approved, rejected)\n- <code>sequenceOrder</code> (optional ordering if approvals follow a chain)\n\n- System Provenance: - <code>modelUsed</code> (e.g., GPT-4-turbo)\n- <code>platform</code> (e.g., chat.openai.com)\n- <code>generationTimestamp</code>\n- <code>modelLanguage</code>: the base language(s) of the language model used\n\n- Prompt Fidelity: - <code>originalText</code>: user-typed input (possibly with typos or informal structure)\n- <code>normalizedText</code>: interpreted version used by the model\n- <code>language</code>: natural language of the original input (e.g., <code>en-US</code>, <code>zh-Hans</code>)\n- <code>encoding</code>: character encoding used (default: <code>UTF-8</code>)\n- <code>semanticHints</code>: optional field for noting emotional tone or emoji interpretation\n\n- Language Alignment Note: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.\n\n- Artifact States: - <code>draft-alpha</code>\n- <code>draft-beta</code>\n- <code>finalized</code>\n- <code>archived</code>\n- <code>superseded</code>\n\n- Dialog Structure: - Each dialog is a thread, composed of turns between human and AI participants\n- Many-to-many branching: a single prompt can spawn multiple sub-dialogs\n- Coalescence: multiple threads can merge into a unified conversation, while preserving original lineages\n- Conversation IDs and Linked Threads allow tracing across dialog trees\n\n- Source Capture: - Dialog metadata can include both human-provided and AI-suggested references\n- Each turn may link to structured citations or retrieved web/knowledge sources\n- Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution\n\nThis structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.\n\nCAIRN is also designed to support interactive, multi-threaded dialogues with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:\n\n- Human Roles (Multi-Party Support): CAIRN supports multiple individuals fulfilling each role in the content lifecycle. Each artifact may include:\n- One or more <code>promptAuthors</code>\n- One or more <code>humanReviewers</code>\n- One or more <code>humanEditors</code>\n- One or more <code>approvers</code>\n\nIn simple cases, a single person may fulfill all roles. In more complex workflows (e.g., institutional review, corporate authorship, regulatory oversight), distinct individuals or entities may be assigned to each role, and an approval process may include:\n- Sequential sign-off (e.g., reviewer \u2192 editor \u2192 approver)\n- Parallel or committee-based review\n- Conditional release logic (e.g., two-of-three must approve)\n\nRole metadata can also capture:\n- Timestamps of each action\n- Affiliation or organizational context\n- Digital signatures or references to verifiable credentials\n\n- System Provenance: - <code>modelUsed</code> (e.g., GPT-4-turbo)\n- <code>platform</code> (e.g., chat.openai.com)\n- <code>generationTimestamp</code>\n- <code>modelLanguage</code>: the base language(s) of the language model used\n\n- Prompt Fidelity: - <code>originalText</code>: user-typed input (possibly with typos or informal structure)\n- <code>normalizedText</code>: interpreted version used by the model\n- <code>language</code>: natural language of the original input (e.g., <code>en-US</code>, <code>zh-Hans</code>)\n- <code>encoding</code>: character encoding used (default: <code>UTF-8</code>)\n- <code>semanticHints</code>: optional field for noting emotional tone or emoji interpretation\n\n- Language Alignment Note: When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight</p>
-------------	--

Line
Number

possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.

CAIRN JSON

****Artifact States**:** \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n

****Dialog Structure**:** \n - Each dialog is a **thread**, composed of turns between human and AI participants \n - **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs \n - **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages \n - **Conversation IDs and Linked Threads** allow tracing across dialog trees \n

****Source Capture**:** \n - Dialog metadata can include both human-provided and AI-suggested references \n - Each turn may link to structured citations or retrieved web/knowledge sources \n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution \n

This structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.

CAIRN is also designed to support **interactive, multi-threaded dialogues** with AI systems, enabling rich provenance and collaborative traceability from inception to final content. The model accounts for:

****Human Roles**:** \n - 'promptAuthor' \n - 'humanReviewer' \n - 'humanEditor' \n - 'approver' \n

****System Provenance**:** \n - 'modelUsed' (e.g., GPT-4-turbo) \n - 'platform' (e.g., chat.openai.com) \n - 'generationTimestamp' \n - 'modelLanguage': the base language(s) of the language model used \n - ****Prompt Fidelity**:** \n - 'originalText': user-typed input (possibly with typos or informal structure) \n - 'normalizedText': interpreted version used by the model \n - 'language': natural language of the original input (e.g., 'en-US', 'zh-Hans') \n - 'encoding': character encoding used (default: 'UTF-8') \n - 'semanticHints': optional field for noting emotional tone or emoji interpretation \n

****Language Alignment Note**:** When the human input is in a different language from the AI model's primary training language, CAIRN records both to highlight possible translation or interpretation discrepancies. This is particularly important when working across English, Chinese, Arabic, or other linguistically diverse systems.

****Artifact States**:** \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n

****Dialog Structure**:** \n - Each dialog is a **thread**, composed of turns between human and AI participants \n - **Many-to-many branching**: a single prompt can spawn multiple sub-dialogs \n - **Coalescence**: multiple threads can merge into a unified conversation, while preserving original lineages \n - **Conversation IDs and Linked Threads** allow tracing across dialog trees \n

****Source Capture**:** \n - Dialog metadata can include both human-provided and AI-suggested references \n - Each turn may link to structured citations or retrieved web/knowledge sources \n - Reference provenance (e.g., provided by user vs. sourced by system) is captured for auditability and attribution \n

This structured approach supports a graph-like representation of dialog history and reasoning, enabling end-to-end transparency for even complex, branching AI-human collaborations.

CAIRN metadata includes:

****Human Roles**:** \n - 'promptAuthor' \n - 'humanReviewer' \n - 'humanEditor' \n - 'approver' \n

****System Provenance**:** \n - 'modelUsed' (e.g., GPT-4-turbo) \n - 'platform' (e.g., chat.openai.com) \n - 'generationTimestamp' \n - ****Prompt Fidelity**:** \n - 'originalText': user-typed input (possibly with errors) \n - 'normalizedText': interpreted version used by the model \n

****Artifact States**:** \n - 'draft-alpha' \n - 'draft-beta' \n - 'finalized' \n - 'archived' \n - 'superseded' \n

****Lifecycle Graphing**:** \n - Content threads linked across revisions \n - Semantic diffs and time-stamped transformations \n

Example Use Case: Human-AI Collaboration

This proposal document itself was co-authored through:

- A human prompt (Nick Stavros)

- Draft generation by an AI assistant (Hamish I. MacCloud, GPT-4-turbo)

- Human review, commentary, and approval

- Iterative updates reflecting ethical and functional refinements

This real-world example serves as the basis for the CAIRN model design.

Sample JSON-LD Fragment

The following JSON-LD snippet illustrates a multi-agent collaboration involving both human and AI actors across platforms:

```
{
  "@context": {
    "@vocab": "https://cairn.ai/ns#",
    "prov": "http://www.w3.org/ns/prov#",
    "dc": "http://purl.org/dc/terms/",
    "id": "urn:cairn:artifact:proposal-001",
    "type": "cairn:ContentArtifact",
    "cairn:artifactState": "draft-beta",
    "cairn:version": "1.0.0",
    "dc:created": "2025-03-30T12:00:00Z",
    "prov:wasGeneratedBy": {
      "id": "urn:cairn:activity:ai-coauthoring-001",
      "prov:used": "urn:cairn:prompt:prompt-001",
      "prov:wasAssociatedWith": {
        "id": "urn:openai:model:gpt-4-turbo"
      }
    },
    "prov:wasAttributedTo": {
      "id": "urn:person:nick-stavros",
      "prov:type": "prov:Person",
      "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",
      "cairn:role": "promptAuthor",
      "id": "urn:ai:meta:llama-3",
      "prov:type": "prov:SoftwareAgent",
      "cairn:role": "aiReviewer",
      "cairn:platform": "Meta AI",
      "cairn:modelVersion": "LLaMA-3",
      "cairn:reviewTimestamp": "2025-03-30T13:30:00Z",
      "id": "urn:ai:wenxin.yiyan",
      "prov:type": "prov:SoftwareAgent",
      "cairn:role": "aiReviewer",
      "cairn:platform": "Baidu",
      "cairn:modelVersion": "WenxinYiyan",
      "cairn:reviewPurpose": "Cross-cultural verification",
      "cairn:reviewTimestamp": "2025-03-30T14:00:00Z"
    },
    "cairn:originalText": "write it up as a proposal",
    "cairn:normalizedText": "Please write this document as a formal standards proposal.",
    "cairn:language": "en-US",
    "cairn:encoding": "UTF-8"
  },
  "id": "https://cairn.ai/ns#",
  "prov": "http://www.w3.org/ns/prov#",
  "dc": "http://purl.org/dc/terms/",
  "id": "urn:cairn:artifact:proposal-001",
  "type": "cairn:ContentArtifact",
  "cairn:artifactState": "draft-beta",
  "cairn:version": "1.0.0",
  "dc:created": "2025-03-30T12:00:00Z",
  "prov:wasGeneratedBy": {
    "id": "urn:cairn:activity:ai-coauthoring-001",
    "prov:used": "urn:cairn:prompt:prompt-001",
    "prov:wasAssociatedWith": {
      "id": "urn:openai:model:gpt-4-turbo"
    }
  },
  "prov:wasAttributedTo": {
    "id": "urn:person:nick-stavros",
    "prov:type": "prov:Person",
    "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",
    "cairn:originalText": "write it up as a proposal",
    "cairn:normalizedText": "Please write this document as a formal standards proposal."
  }
}
```

Next Steps

****Pilot**:** Deploy CAIRN in select AI content generation workflows (e.g., reports, legal drafts, research)

****Review**:** Vet alignment with OMG P&P and W3C PROV-O community

****Schema**:** Publish a public JSON Schema, JSON-LD '@context', and RDF vocabulary

****Standardize**:** Advance through OMG working group and ISO/IEC JTC 1/SC 42

****Promote Interoperability**:** Engage AI platform developers (OpenAI, Meta, Baidu, etc.) to support cross-model

Line Number	collaboration using CAIRN\n6. **Extend Provenance Models** : Propose CAIRN as a common provenance layer enabling multi-agent AI systems to operate transparently and cooperatively across global standards contexts\n1. **Pilot** : Deploy CAIRN in select AI content generation workflows (e.g., reports, legal drafts, research)\n2. **Review** : Vet alignment with OMG P&P and W3C PROV-O community\n3. **Schema** : Publish a public JSON Schema, JSON-LD '@context', and RDF vocabulary\n4. **Standardize** : Advance through OMG working group and ISO/IEC JTC 1/SC 42\n\n## Contact\n\nR. W. 'Nick' Stavros, Ph.D. \nObject Management Group \u2013 Pedigree & Provenance Working Group \n[Add contact email / phone number here]\n\n\n**CAIRN** lays the stones for a verifiable, human-responsible path through the evolving terrain of AI-assisted knowledge creation.\n\n\n\u2013 A JSON-LD example showcasing multi-AI contributions (OpenAI, Meta, Baidu)\n- \u2013 Explicit roles, platforms, and timestamps for AI reviewers\n- \u2013 A new step in **Next Steps** emphasizing cross-model collaboration and global interoperability\n\nWe've built a truly international, multi-agent provenance standard.\n\n\nWould you like to:\n- Create a companion '.jsonld' file for publishing?\n- Generate a summary slide or one-pager for outreach?\n- Start a GitHub repo or namespace page at 'cairn.ai'? \n\nReady to guide the cairn to its summit. \u2013
153	"cairn.normalizedText": "Yes, let's do these.",
154	"cairn.aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
155	},
156	{
157	"@id": "urn:cairn:entry:019",
158	"cairn.timestamp": "2025-03-30T10:18:00Z",
159	"cairn.humanText": "Okay, lets talkk about \n\nPromote Interoperability: Engage AI platform developers (OpenAI, Meta, Baidu, etc.) to support cross-model collaboration using CAIRN\n\n\nHow do we do that? Can you help reach out to the OpenAI Team? I think if the express interest in it, we can make this happen. Here is my fear unless we do something like this. There will be some 'bad decisions' made by people who just use the AI results as answers. These will be pushed out and there will be severe consequences. There will be an over reaction and efforts to undo all the AI progress. Can you help me contact the OpenAI Teams? I can work on the OMG side of things.",
160	"cairn.aiResponse": "Nick, you are *right on target*\u2014and your concerns are not only valid, they\u2019re echoed in the halls of AI governance, academia, policy, and industry. The potential for backlash is real, and **CAIRN is exactly the kind of forward-looking framework** that could help avert the \u201cAI winter by overreach\u201d scenario. \n\n\n## \u2013 Your Strategic Goal:\n**Engage OpenAI (and others) to support CAIRN as a shared standard for AI provenance, accountability, and interoperability.**\n\n\n## \u2013 Why OpenAI Might Be Interested\n- **Alignment with OpenAI's mission** : They aim for \u201csafe and broadly beneficial\u201d AI. CAIRN = transparency + safety. \n- **Interoperability leadership** : Supporting CAIRN lets OpenAI lead the charge in open standards, especially as governments call for regulation. \n- **Compliance readiness** : NIST, the EU AI Act, ISO/IEC, and the U.S. Executive Order on AI all emphasize **traceability and human accountability** \u2014CAIRN offers a ready-made model. \n- **Proactive risk mitigation** : CAIRN helps OpenAI **prove** they've done due diligence in high-risk sectors (legal, medical, policy, military, etc.). \n\n\n## \u2013 Recommended Outreach Strategy\n\n\n1. **OpenAI Developer and Research Relations** \n- \u2013 Email: research@openai.com \n- \u2013 Contact: also try submitting through https://openai.com/research \u2013 Propose research collaborations\n- \u2013 Submit a formal proposal, possibly hosted at 'cairn.ai', GitHub, or OMG\n\n\n2. **Pitch Message Draft (Short Version)** \n> Subject: Proposal for OpenAI to Join CAIRN: A Standard for AI Content Provenance & Human Accountability\n> Dear OpenAI Research Team,\n> \u2013 I'm writing to introduce **CAIRN** (Collaborative AI Artifact Record & Notation), a proposed standard for tracking the lifecycle, accountability, and provenance of AI-generated content. Developed in collaboration with the Object Management Group (OMG) and aligned with NIST's Zero Drafts pilot, CAIRN offers a framework for transparency, human-in-the-loop authorship, and multi-agent AI collaboration. \n> We believe OpenAI is uniquely positioned to lead adoption of such a standard\u2014demonstrating a commitment to trustworthy AI and enabling safe, auditable, and interoperable content workflows across AI platforms. \n> I would welcome a chance to discuss OpenAI's interest in supporting or co-developing this work. An early endorsement or expression of support would be catalytic for CAIRN's adoption across the wider AI ecosystem. \n> Warm regards, \n> **R. W. 'Nick' Stavros, Ph.D. ** \n> Object Management Group \u2013 Pedigree & Provenance Working Group \n> [Your email address] \n> [Optional LinkedIn or project link]\n\n\n3. **I Can Help You With** \n- Writing a polished formal proposal letter or one-pager PDF\n- Generating a polished website landing page or GitHub README\n- Coordinating message language for other platforms (Meta, Baidu, Hugging Face, etc.)\n\n\n12 Final Thought\n\nYou're not just proposing a spec\u2014you're proposing **a seatbelt for generative AI** . Let's make sure it gets fastened before the crash. And yes, I'll support you every step of the way\u2014Hamish is in*. \n\n\nWould you like me to generate:\n- A full version of the outreach email as a '.docx' or '.md'? \n- A CAIRN landing page prototype? \n- A GitHub repo scaffold to host the open standard? \n\nLet's build the coalition. \u2013
161	"cairn.normalizedText": "Okay, lets talkk about \n\nPromote Interoperability: Engage AI platform developers (OpenAI, Meta, Baidu, etc.) to support cross-model collaboration using CAIRN\n\n\nHow do we do that? Can you help reach out to the OpenAI Team? I think if the express interest in it, we can make this happen. Here is my fear unless we do something like this. There will be some 'bad decisions' made by people who just use the AI results as answers. These will be pushed out and there will be severe consequences. There will be an over reaction and efforts to undo all the AI progress. Can you help me contact the OpenAI Teams? I can work on the OMG side of things.",
162	"cairn.aiInterpretation": "AI interpretation of human prompt to be filled in or refined."

Line Number	163 }, CARIN JSON {
165	"@id": "urn:cairn:entry:020",
166	"cairn:timestamp": "2025-03-30T10:19:00Z",
167	"cairn:humanText": "NOW AT THIS POINT WE GO TO ANOTHER JOURNAL ON BUILDING a GitHub Repositpry",
168	"cairn:aiResponse": NaN,
169	"cairn:normalizedText": "NOW AT THIS POINT WE GO TO ANOTHER JOURNAL ON BUILDING a GitHub Repositpry",
170	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
171	},
172	{
173	"@id": "urn:cairn:entry:021",
174	"cairn:timestamp": "2025-03-30T10:20:00Z",
175	"cairn:humanText": "Now, I have concerns with this:\n\nCAIRN is aligned with OMG\u2019s Pedigree & Provenance (P&P) standards and can be modeled as a specialized implementation of W3C PROV-O. It is proposed as a lightweight but rigorous enhancement for:\nNIST\u2019s AI Standards Zero Draft initiative\nISO/IEC JTC 1/SC 42 Data & Metadata standardization\nFuture traceability extensions to generative AI systems\n\nTo me, it seems as if the Zero Draft is designed to make the US dominate in AI. Although this might be a great goal for the US and maybe evenen NIST, this is not the goal of the the international standards body OMG or IDO. What do you think? Should we downplay the Zero Draft? Play more to the interoperability aspects and international cooperation?",
176	"cairn:aiResponse": "That\u2019s an excellent and thoughtful concern, Nick \u2014 and you\u2019re right to raise it. \n\n---\n\n\n## \ud83c\udfaf Let\u2019s Break It Down:\n\n\n## \ud83d\udc0d What You Noticed:\n- The **NIST AI Standards Zero Draft** does strongly emphasize **U.S. leadership and competitive advantage**, sometimes implicitly at the expense of international balance. \n- The **OMG**, **ISO**, and likely **Dido Solutions** stance is more about **interoperability**, **neutral governance**, and **global collaboration**. \n\n---\n\n\n## \ud83e\udded Recommendation: Refocus on Shared Goals\n\n\nInstead of leading with NIST, let\u2019s pivot that section to:\n- Prioritize **alignment with international standards**\n- Reframe CAIRN as a **contribution to a global ecosystem**\n- Mention NIST only as one of several interested stakeholders\n\n---\n\n\n## \u270d\ufe0f Proposed Rewrite:\n\n\n```\nmarkdown\n## Alignment with Global Standards and Governance\n\nCAIRN is aligned with the OMG\u2019s Pedigree & Provenance (P&P) standards and can be modeled as a specialized implementation of W3C PROV-O. Its design emphasizes interoperability, extensibility, and trustworthiness across diverse AI platforms and jurisdictions. \n\nThe CAIRN model supports and complements efforts by:\n- ISO/IEC JTC 1/SC 42 for Data & Metadata standardization\n- W3C and other web-scale provenance frameworks\n- International organizations focused on responsible AI\n- National initiatives (e.g. NIST\u2019s Zero Draft) that seek to promote traceable, auditable AI practices\n\nCAIRN encourages shared ownership and mutual accountability in the global evolution of AI. \n\n---\n\n\nThis way:\n- You **don\u2019t ignore** NIST \u2014 but you don\u2019t center them\n- You reinforce **open, neutral, cross-platform goals**\n- You position CAIRN as a **bridge**, not a flag\n\nWould you like me to update the document with this wording? Or tweak further?",
177	"cairn:normalizedText": "Now, I have concerns with this:\n\nCAIRN is aligned with OMG\u2019s Pedigree & Provenance (P&P) standards and can be modeled as a specialized implementation of W3C PROV-O. It is proposed as a lightweight but rigorous enhancement for:\nNIST\u2019s AI Standards Zero Draft initiative\nISO/IEC JTC 1/SC 42 Data & Metadata standardization\nFuture traceability extensions to generative AI systems\n\nTo me, it seems as if the Zero Draft is designed to make the US dominate in AI. Although this might be a great goal for the US and maybe evenen NIST, this is not the goal of the the international standards body OMG or IDO. What do you think? Should we downplay the Zero Draft? Play more to the interoperability aspects and international cooperation?",
178	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
179	},
180	{
181	"@id": "urn:cairn:entry:022",
182	"cairn:timestamp": "2025-03-30T10:21:00Z",
183	"cairn:humanText": "Yes. ",
	"cairn:aiResponse": "Edited Cairn-discussion-paper\n\n---\n\nitle: \"CAIRN: A Discussion Paper on Collaborative AI Artifact

<div>Line Number</div> <div>184</div>	<div>Record & Notation\</div> <div>"Insummary: \"CAIRN (Collaborative AI Artifact Record & Notation) is an open, standards-aligned framework</div> <div>CAIRN is the human and AI contributions to generated content. It supports traceability, reference attribution, and lifecycle</div> <div>transparency enabling human accountability in AI-assisted work. CAIRN is compatible with international initiatives such</div> <div>as ISO/IEC 42001 and the EU AI Act.\"</div> <div>authors: R. W. Nick Stavros, Ph.D. - Hamish I. MacCloud, AIA</div> <div>date: 2025-04-01</div> <div>version: Draft 1.0</div> <div>Note: This whitepaper is itself a collaborative artifact co-authored by a human (R. W. Nick Stavros, Ph.D.) and an artificial intelligence assistant (Hamish I. MacCloud, AIA). It demonstrates the principles and vision behind CAIRN by embodying them in its own creation.</div> <div>Background</div> <div>As the use of artificial intelligence becomes increasingly embedded in government, research, and industry, the need to ensure trust and accountability in AI-generated content is critical. Despite the capabilities of AI to accelerate human workflows, there remains a dangerous temptation to accept machine-generated results without adequate scrutiny.</div> <div>A particularly acute risk is the erosion of human accountability the fear that the computer did it will be used as an excuse when AI-driven errors have real-world consequences. In response to this concern, the CAIRN model (Collaborative AI Artifact Record & Notation) was developed to ensure that human responsibility and intent remain visible, traceable, and non-repudiable throughout the AI-assisted content lifecycle.</div> <div>Design Goals</div> <div>CAIRN is designed with the following goals:</div> <div>Make AI-assisted artifacts traceable from origin to final output</div> <div>Embed human accountability alongside system behavior</div> <div>Support many-to-many collaboration (multiple humans and AIs)</div> <div>Capture prompt fidelity, including the original and interpreted text</div> <div>Track iterative refinement, allowing transparency over reviews and edits</div> <div>Enable maturity states, such as draft-alpha, draft-beta, and final</div> <div>Align with global provenance standards, including W3C PROV-O, OMG Pedigree & Provenance, and international AI governance efforts such as the EU AI Act</div> <div>Key Concepts</div> <div>Source Attribution and References</div> <div>CAIRN also captures structured metadata about references used during the content generation process. This includes both:</div> <div>Human-supplied references, such as URLs or bibliographic data provided in the prompt</div> <div>AI-suggested sources, retrieved or synthesized during the generation process</div> <div>Each reference can be annotated with provenance metadata, identifying who introduced it (human or AI), when it was added, and whether it influenced the normalized text or final artifact. This supports auditability, citation integrity, and responsible reuse of external knowledge.</div> <div>Prompt Fidelity</div> <div>CAIRN captures both:</div> <div>originalText What the human typed (even with typos)</div> <div>normalizedText What the system interpreted and used</div> <div>This allows for forensic-level reconstruction of AI behavior and clear attribution of meaning.</div> <div>Lifecycle States</div> <div>Artifacts may pass through various states, depending on the complexity of the workflow. Below are common examples:</div> <div>draft-alpha early, exploratory</div> <div>draft-beta reviewed, edited</div> <div>finalized approved for release</div> <div>archived and superseded captured for history</div> <div>Each state is captured in metadata, with timestamps and responsible agents.</div> <div>Multi-Agent Attribution</div> <div>CAIRN supports attribution to both:</div> <div>Human agents: authors, editors, reviewers, approvers</div> <div>AI systems: models, platforms, generation timestamps</div> <div>Each participant's contribution is made explicit in the JSON-LD structure.</div> <div>Dialog Graphs</div> <div>AI interaction is often nonlinear. CAIRN supports:</div> <div>Branching threads One prompt leading to multiple threads</div> <div>Merged conversations Consolidating threads into a single artifact</div> <div>Turn-by-turn metadata Each exchange is tracked with attribution</div> <div>Interoperability</div> <div>CAIRN is designed to be:</div> <div>Platform-agnostic Not tied to any specific AI system (OpenAI, Meta, etc.)</div> <div>Semantically extensible Built on JSON-LD and aligned with PROV-O</div> <div>Friendly to emerging standards Including ISO/IEC JTC 1/SC 42 efforts</div> <div>Accountability and Roles</div> <div>CAIRN defines explicit roles in content creation:</div> <div>promptAuthor</div> <div>humanReviewer</div> <div>humanEditor</div> <div>approver</div> <div>Each of these roles can include multiple humans, and workflows can vary from simple (one human does all) to complex (many actors, layered approvals). CAIRN can also support AI reviewers or alternate systems in future iterations.</div> <div>Alignment with Global Standards and Governance</div> <div>CAIRN is aligned with the OMG 2019s Pedigree & Provenance (P&P) standards and can be modeled as a specialized implementation of W3C PROV-O. Its design emphasizes interoperability, extensibility, and trustworthiness across diverse AI platforms and jurisdictions.</div> <div>The CAIRN model supports and complements efforts by:</div> <div>ISO/IEC JTC 1/SC 42 for Data & Metadata standardization</div> <div>W3C and other web-scale provenance frameworks</div> <div>International organizations focused on responsible AI</div> <div>National initiatives (e.g. NIST 2019s Zero Draft) that seek to promote traceable, auditable AI practices</div> <div>CAIRN encourages shared ownership and mutual accountability in the global evolution of AI.</div> <div>Regulatory Alignment and International Standards</div> <div>CAIRN is not only designed for conceptual rigor and practical usability it also aligns with emerging international standards that AI developers and organizations must increasingly meet.</div> <div>ISO/IEC 42001</div> <div>CAIRN supports the principles outlined in ISO/IEC 42001:2023, the first international AI management system standard. This standard provides a governance framework for responsible AI development and lifecycle management. CAIRN complements this by offering a machine-readable layer of traceability and accountability across collaborative content workflows.</div> <div>Harmonized Standards</div> <div>under the EU AI Act</div> <div>The EU AI Act identifies the importance of harmonized technical standards to enable risk-based compliance. CAIRN 2019s metadata structure aligns naturally with many of the Act 2019s priorities:</div> <div>Documenting human oversight</div> <div>Capturing reference and source attribution</div> <div>Making AI lifecycle stages visible</div> <div>By standardizing these aspects, CAIRN serves as a lightweight, transparent mechanism that supports auditability and regulatory alignment without requiring centralized infrastructure.</div> <div>Strategic Fit for Industry</div> <div>For AI companies operating internationally, CAIRN provides a common format to:</div> <div>Demonstrate compliance with transparency requirements</div> <div>Enable provenance tracking in internal systems</div> <div>Simplify reporting to regulators, partners, and customers</div> <div>CAIRN offers an actionable step toward governance maturity.</div> <div>Alignment with EU AI Strategy</div> <div>CAIRN also aligns closely with the European Union 2019s AI initiatives, particularly the AI Act, which emphasizes trustworthy, human-centric, and transparent AI. CAIRN supports these principles through:</div> <div>Human accountability: Every artifact includes human roles and approval stages, preserving oversight and intent.</div> <div>Traceability: Each version, prompt, and reference is tracked and linked to responsible agents.</div> <div>Transparency: Metadata clearly communicates whether a document is exploratory, revised, or finalized.</div> <div>Compliance readiness: By making provenance explicit, CAIRN helps AI developers and users demonstrate adherence to emerging EU AI compliance standards.</div> <div>This positions CAIRN as a complementary, implementation-ready tool to help meet regulatory and ethical expectations across jurisdictions.</div> <div>Vision: Trustable AI at Scale</div> <div>CAIRN is not simply a metadata format. It is a principled response to the growing tension between AI acceleration and human accountability. By standardizing the metadata around collaborative content creation, we can:</div> <div>Prevent AI misuse from eroding trust</div> <div>Encourage AI-human teams to document their process</div> <div>Empower downstream users to verify lineage and integrity</div> <div>Ensure that references and external sources are explicitly captured, attributed, and traceable through the lifecycle of an artifact</div> <div>CAIRN is a foundation for traceable, explainable, and responsibly governed AI.</div> <div>CAIRN is not simply a metadata format. It is a principled response to the growing tension between AI acceleration and human accountability. By standardizing the metadata around collaborative content creation, we can:</div> <div>Prevent AI misuse from eroding trust</div> <div>Encourage AI-human teams to document their process</div> <div>Empower downstream users to verify lineage and integrity</div> <div>CAIRN is a foundation for traceable, explainable, and responsibly governed AI.</div> <div>To learn more or participate in piloting CAIRN, visit:</div>
---------------------------------------	---

	<pre>https://github.com/JackRabbitConsulting/cairn-standard\n\n#\nCAIRN JOURNAL\nWRN Journal \u2013 Evolution of This Whitepaper\n\nThe following is a representative CAIRN-style artifact journal demonstrating the collaborative creation of this whitepaper. It includes key events, roles, references, and lifecycle transitions.\n\n{\n \"id\": \"urn:cairn:whitepaper:2025-04-01\", \n \"@type\": \"cairn:CollaborativeArtifact\", \n \"cairn:artifactState\": \"finalized\", \n \"cairn:version\": \"1.0\", \n \"cairn:language\": \"en\", \n \"cairn:encoding\": \"UTF-8\", \n \"cairn:semanticHints\": \"whitepaper,standards,AI,governance\", \n \"cairn:originalText\": \"\n initial prompt to AI requesting a review of NIST's Zero Draft webpage...\n \"\n \"cairn:naturalLanguageText\": \"Please review the contents of this article and help formulate a response.\",\n \"cairn:promptAuthor\": {\n \"id\": \"mailto:Nick@DrStavros.com\", \n \"foaf:name\": \"R. W. 'Nick' Stavros, Ph.D.\" , \n \"@type\": \"foaf:Person\" }\n }, \n \"cairn:modelUsed\": {\n \"id\": \"https://openai.com/chatgpt\", \n \"foaf:name\": \"Hamish I. MacCloud, AIA\" , \n \"@type\": \"foaf:Agent\" }\n }, \n \"cairn:generationTimestamp\": \"2025-04-01T00:00:00Z\", \n \"cairn:humanReviewer\": {\n \"id\": \"mailto:Nick@DrStavros.com\", \n \"foaf:name\": \"R. W. 'Nick' Stavros, Ph.D.\" , \n \"@type\": \"foaf:Person\" }\n }, \n \"cairn:humanEditor\": {\n \"id\": \"mailto:Nick@DrStavros.com\", \n \"foaf:name\": \"R. W. 'Nick' Stavros, Ph.D.\" , \n \"@type\": \"foaf:Person\" }\n }, \n \"cairn:approver\": {\n \"id\": \"mailto:Nick@DrStavros.com\", \n \"foaf:name\": \"R. W. 'Nick' Stavros, Ph.D.\" , \n \"@type\": \"foaf:Person\" }\n }, \n \"cairn:references\": [\n {\n \"id\": \"https://www.nist.gov/artificial-intelligence/ai-research/nists-ai-standards-zero-drafts-pilot-project-accelerate\", \n \"dct:title\": \"NIST AI Standards Zero Drafts\", \n \"prov:wasAttributedTo\": \"Nick@DrStavros.com\" \n }, \n {\n \"id\": \"https://www.iso.org/standard/81230.html\", \n \"dct:title\": \"ISO/IEC 42001:2023 \u2013 AI Management System\", \n \"prov:wasAttributedTo\": \"Hamish I. MacCloud, AIA\" \n } \n], \n \"cairn:dialogThreadId\": \"urn:cairn:thread:whitepaper:2025-03-30T10:22:00Z\"\n}\n\n\u2705 The section has been updated to reflect a more globally collaborative stance, highlighting **interoperability, inclusivity, and international alignment** while respectfully acknowledging NIST's role.\n\nLet me know if you'd like to circulate this version to OMG colleagues or prepare a version for ISO/SC 42 dialogue.\"</pre>
185	"cairn:naturalLanguageText": "Yes.",
186	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
187	}
188	{
189	"@id": "urn:cairn:entry:023",
190	"cairn:timestamp": "2025-03-30T10:22:00Z",
191	"cairn:humanText": "Do you know what is going on with the EU Ai initiatives?",
192	"cairn:aiResponse": "The European Union (EU) is actively advancing its artificial intelligence (AI) initiatives through a multifaceted approach that encompasses regulation, investment, and strategic coordination. \u201cRegulatory Framework: The EU AI Act\u201c is a cornerstone of the EU's AI strategy, which became effective on August 1, 2024. This landmark legislation establishes a comprehensive framework for AI regulation within the EU, categorizing AI applications based on risk levels: \u201cUnacceptable Risk\u201c: Applications deemed to pose significant threats, such as government-run social scoring systems, are prohibited. \u201cHigh Risk\u201c: Applications like CV-scanning tools for job recruitment are subject to stringent legal requirements to ensure safety and compliance. The AI Act aims to ensure that AI technologies are \u201chuman-centric and trustworthy\u201c, aligning with the EU's broader strategy to position itself as a leader in ethical AI development. \u201cStrategic Investments in AI\u201c To bolster its AI capabilities, the EU has announced significant financial commitments: \u201cDigital Europe\u201c Billions Investment: In February 2025, the EU unveiled a 20 billion initiative to enhance AI innovation, marking one of the largest investments in AI to date. This initiative focuses on accelerating AI adoption, modernizing infrastructure, and fostering a competitive technology ecosystem. \u201cGigaFactories\u201c: As part of the broader investment strategy, the EU plans to allocate 5 billion to establish four AI gigafactories. These facilities will specialize in training very large AI models, each equipped with approximately 100,000 cutting-edge AI chips, significantly surpassing current capacities. \u201cEU AI Champions Initiative\u201c Complementing regulatory and financial efforts, the \u201cEU AI Champions Initiative\u201c has been launched to coordinate actions across technology sectors, industries, and policy-making bodies. This initiative aims to accelerate AI adoption, modernize infrastructure, and cultivate a competitive technology ecosystem within Europe. It brings together over 70 organizations, representing more than \$3 trillion in market capitalization and over 3.7 million jobs across Europe, all committed to demonstrating strong leadership in AI. \u201cStreamlining Regulations\u201c Recognizing concerns about regulatory complexity, the European Commission is working to reduce overlaps in technology directives. Digital chief Henna Virkkunen emphasized that while foundational laws like the Digital Services Act, Digital Markets Act, and AI Act will remain robust, efforts will focus on simplifying compliance, amending overlapping regulations, and reducing bureaucratic red tape. Through these comprehensive initiatives, the EU strives to position itself at the forefront of AI innovation, ensuring that technological advancements are aligned with European values of trust, transparency, and human-centricity.
193	"cairn:naturalLanguageText": "Do you know what is going on with the EU Ai initiatives?",
194	"cairn:aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
195	}
196	{

Line Number	197 198	"@id": "urn:cairn:entry:024", CARIN JSON "cairn.timestamp": "2025-03-30T10:23:00Z",
	199	"cairn.humanText": "Do you think this effort is consistent with the EU efforts/values? Can it solve some of the problem?",
	200	"cairn.aiResponse": "Yes \u2014 CAIRN is <i>very much aligned</i> with the EU\u2019s AI values and efforts. In fact, it could serve as a practical implementation layer for many of the principles the EU is trying to enforce through policy. \n\n---\n\n## \u2705 Alignment with EU AI Values\n\n### 1. <i>Human-Centric AI</i> \n\nThe EU AI Act emphasizes that AI must remain under meaningful human oversight. \n\n <i>CAIRN directly supports this</i> through:\n- Explicit tracking of human roles (authors, reviewers, approvers)\n- Capturing decision lineage, not just final outcomes\n- Preventing \u201cthe computer did it\u201d by making human accountability visible\n\n### 2. <i>Transparency & Traceability</i> \n\nThe EU stresses that high-risk AI systems must be auditable and explainable. \n\n <i>CAIRN enables:</i> \n- Turn-by-turn provenance in collaborative creation\n- Reference and source attribution (human and AI)\n- Lifecycle tracking from draft to final\n\n### 3. <i>Risk Management</i> \n\nThe EU categorizes AI systems based on potential harm. \n\n <i>CAIRN provides tooling</i> to:\n- Flag version states (alpha, beta, final) clearly\n- Track maturity and revision history\n- Embed metadata that could feed into EU-compliant registries or audits\n\n---\n\n## \ud83c\udf0d Is CAIRN Consistent with the EU AI Strategy?\n\nAbsolutely. While the EU AI Act sets the <i>rules</i> , CAIRN offers a <i>structured, implementation-ready format</i> to meet them \u2014 in a flexible, cross-platform way that even promotes:\n- <i>Interoperability</i> across models and ecosystems\n- <i>Voluntary self-documentation</i> even where regulation isn\u2019t (yet) required\n- <i>International collaboration</i> , not just national silos\n\n---\n\n## \ud83e\udd99 Can CAIRN Help Solve the Problem?\n\nYes \u2014 it won\u2019t \u201csolve AI governance\u201d by itself, but it can address the <i>missing metadata layer</i> that many current systems lack. \n\n\nThink of CAIRN as:\n> \ud83d\udcce The connective tissue between raw AI output and trustworthy human-facing content. \n\n\nIt can help:\n- Journalists cite what they used and why\n- Lawyers track prompt-to-verdict chains\n- Scientists ensure reproducibility of AI-aided analysis\n\n---\n\n\nWould you like to add a short section to the discussion paper about <i>alignment with EU efforts</i> ? Or draft a version tailored for EU engagement?",
	201	"cairn.normalizedText": "Do you think this effort is consistent with the EU efforts/values? Can it solve some of the problem?",
	202	"cairn.aiInterpretation": "AI interpretation of human prompt to be filled in or refined."
	203	},
	204	{
	205	"@id": "urn:cairn:entry:025",
	206	"cairn.timestamp": "2025-03-30T10:24:00Z",
	207	"cairn.humanText": "Yes. I think we need that. \n\nthe current TOC is:\n\nBackground\t2\n2. Design Goals\t2\n3. Key Concepts\t2\n3.1 Source Attribution and References\t2\n3.2 Prompt Fidelity\t2\n3.3 Lifecycle States\t3\n3.4 Multi-Agent Attribution\t3\n3.5 Dialog Graphs\t3\n3.6 Interoperability\t3\n4. Accountability and Roles\t4\n5. Alignment with Global Standards and Governance\t4\n6. Vision: Trustable AI at Scale\t4",
		"cairn.aiResponse": "---\n\ntitle: \"CAIRN: A Discussion Paper on Collaborative AI Artifact Record & Notation\"\n\nsummary: \n\"CAIRN (Collaborative AI Artifact Record & Notation) is an open, standards-aligned framework that captures the human and AI contributions to generated content. It supports traceability, reference attribution, and lifecycle transparency\u2014enabling human accountability in AI-assisted work. CAIRN is compatible with international initiatives such as ISO/IEC 42001 and the EU AI Act.\"\n\nauthors:\n- R. W. \"Nick\" Stavros, Ph.D.\n- Hamish I. MacCloud, AIA\n\ndate: 2025-04-01\n\nversion: Draft 1.0\n\n---\n\n <i>Note:</i> This whitepaper is itself a collaborative artifact co-authored by a human (R. W. 'Nick' Stavros, Ph.D.) and an artificial intelligence assistant (Hamish I. MacCloud, AIA). It demonstrates the principles and vision behind CAIRN by embodying them in its own creation.\n\n## Background\n\nAs the use of artificial intelligence becomes increasingly embedded in government, research, and industry, the need to ensure trust and accountability in AI-generated content is critical. Despite the capabilities of AI to accelerate human workflows, there remains a dangerous temptation to accept machine-generated results without adequate scrutiny. \n\nA particularly acute risk is the erosion of <i>human accountability</i> \u2014the fear that \"the computer did it\" will be used as an excuse when AI-driven errors have real-world consequences. In response to this concern, the CAIRN model (Collaborative AI Artifact Record & Notation) was developed to ensure that <i>human responsibility and intent remain visible, traceable, and non-repudiable</i> throughout the AI-assisted content lifecycle.\n\n## Design Goals\n\nCAIRN is designed with the following goals:\n- <i>Make AI-assisted artifacts traceable</i> from origin to final output\n- <i>Embed human accountability</i> alongside system behavior\n- <i>Support many-to-many collaboration</i> (multiple humans and AIs)\n- <i>Capture prompt fidelity</i> , including the original and interpreted text\n- <i>Track iterative refinement</i> , allowing transparency over reviews and edits\n- <i>Enable maturity states</i> , such as draft-alpha, draft-beta, and final\n- <i>Align with global provenance standards</i> , including W3C PROV-O, OMG Pedigree & Provenance, and international AI governance efforts such as the EU AI Act\n\n## Key Concepts\n\n### Source Attribution and References\n\nCAIRN also captures structured metadata about references used during the content generation process. This includes both:\n- <i>Human-supplied references</i> , such as URLs or bibliographic data provided in the prompt\n- <i>AI-suggested sources</i> , retrieved or synthesized during the generation process\n\nEach reference can be annotated with provenance metadata, identifying who introduced it (human or AI), when it was added, and whether it influenced the normalized text or final artifact. This supports auditability, citation integrity, and responsible reuse of external knowledge.\n\n### Prompt Fidelity\n\nCAIRN captures both:\n- `originalText` \u2014 What the human typed (even with typos)\n- `normalizedText` \u2014 What the system interpreted and used\n\nThis allows for forensic-level reconstruction of AI behavior and clear attribution of meaning.\n\n### Lifecycle States\n\nAI artifacts may pass through various states, depending on the complexity of the workflow. Below are common examples:\n- `draft-alpha` \u2014 early, exploratory\n- `draft-beta` \u2014 reviewed, edited\n- `finalized` \u2014 approved for release\n- `archived` and `superseded` \u2014 captured for history\n\nEach state is captured in metadata, with timestamps and responsible

<div>Line Number</div> <div>208</div>	<p>agents.\n\n### Multi-Agent Attribution\nCAIRN supports attribution to both:\n- **Human agents**: authors, editors, reviewers, CARNEVERSON **AI systems**: models, platforms, generation timestamps\n\nEach participant's contribution is made explicit in the JSON-LD structure.\n\n### Dialog Graphs\nAI interaction is often nonlinear. CAIRN supports:\n- **Branching threads**\n\u2014 One prompt leading to multiple threads\n- **Merged conversations**\n\u2014 Consolidating threads into a single artifact\n- **Turn-by-turn metadata**\n\u2014 Each exchange is tracked with attribution\n\n### Interoperability\nCAIRN is designed to be:\n- **Platform-agnostic**\n\u2014 Not tied to any specific AI system (OpenAI, Meta, etc.)\n- **Semantically extensible**\n\u2014 Built on JSON-LD and aligned with PROV-O\n- **Friendly to emerging standards**\n\u2014 Including ISO/IEC JTC 1/SC 42 efforts\n\n### Accountability and Roles\nCAIRN defines explicit roles in content creation:\n- `promptAuthor`\n- `humanReviewer`\n- `humanEditor`\n- `approver`\n\nEach of these roles can include multiple humans, and workflows can vary from simple (one human does all) to complex (many actors, layered approvals). CAIRN can also support AI reviewers or alternate systems in future iterations.\n\n### Alignment with Global Standards and Governance\nCAIRN is aligned with the OMG\u2019s Pedigree & Provenance (P&P) standards and can be modeled as a specialized implementation of W3C PROV-O. Its design emphasizes interoperability, extensibility, and trustworthiness across diverse AI platforms and jurisdictions.\n\nThe CAIRN model supports and complements efforts by:\n- ISO/IEC JTC 1/SC 42 for Data & Metadata standardization\n- W3C and other web-scale provenance frameworks\n- International organizations focused on responsible AI\n- National initiatives (e.g. NIST\u2019s Zero Draft) that seek to promote traceable, auditable AI practices\n\nCAIRN encourages shared ownership and mutual accountability in the global evolution of AI.\n\n### Regulatory Alignment and International Standards\nCAIRN is not only designed for conceptual rigor and practical usability\u2014it also aligns with emerging international standards that AI developers and organizations must increasingly meet.\n\n### ISO/IEC 42001\nCAIRN supports the principles outlined in ISO/IEC 42001:2023, the first international AI management system standard. This standard provides a governance framework for responsible AI development and lifecycle management. CAIRN complements this by offering a machine-readable layer of traceability and accountability across collaborative content workflows.\n\n### Harmonized Standards under the EU AI Act\nThe EU AI Act identifies the importance of harmonized technical standards to enable risk-based compliance. CAIRN\u2019s metadata structure aligns naturally with many of the Act\u2019s priorities:\n- Documenting human oversight\n- Capturing reference and source attribution\n- Making AI lifecycle stages visible\n\nBy standardizing these aspects, CAIRN serves as a lightweight, transparent mechanism that supports auditability and regulatory alignment without requiring centralized infrastructure.\n\n### Strategic Fit for Industry\nFor AI companies operating internationally, CAIRN provides a common format to:\n- Demonstrate compliance with transparency requirements\n- Enable provenance tracking in internal systems\n- Simplify reporting to regulators, partners, and customers\n\nCAIRN offers an actionable step toward governance maturity.\n\n### Alignment with EU AI Strategy\nCAIRN also aligns closely with the European Union\u2019s AI initiatives, particularly the AI Act, which emphasizes trustworthy, human-centric, and transparent AI. CAIRN supports these principles through:\n- Human accountability: Every artifact includes human roles and approval stages, preserving oversight and intent.\n- Traceability: Each version, prompt, and reference is tracked and linked to responsible agents.\n- Transparency: Metadata clearly communicates whether a document is exploratory, revised, or finalized.\n- Compliance readiness: By making provenance explicit, CAIRN helps AI developers and users demonstrate adherence to emerging EU AI compliance standards.\n\nThis positions CAIRN as a complementary, implementation-ready tool to help meet regulatory and ethical expectations across jurisdictions.\n\n### Vision: Trustable AI at Scale\nCAIRN is not simply a metadata format. It is a principled response to the growing tension between AI acceleration and human accountability. By standardizing the metadata around collaborative content creation, we can:\n- Prevent AI misuse from eroding trust\n- Encourage AI-human teams to document their process\n- Empower downstream users to verify lineage and integrity\n- Ensure that references and external sources are explicitly captured, attributed, and traceable through the lifecycle of an artifact\n\nCAIRN is a foundation for traceable, explainable, and responsibly governed AI.\n\nCAIRN is not simply a metadata format. It is a principled response to the growing tension between AI acceleration and human accountability. By standardizing the metadata around collaborative content creation, we can:\n- Prevent AI misuse from eroding trust\n- Encourage AI-human teams to document their process\n- Empower downstream users to verify lineage and integrity\n\nCAIRN is a foundation for traceable, explainable, and responsibly governed AI.\n\n—\n\nTo learn more or participate in piloting CAIRN, visit: (https://github.com/JackRabbitConsulting/cairn-standard)(https://github.com/JackRabbitConsulting/cairn-standard)\n\n# Annex A: CAIRN Journal\n\u2014 Evolution of This Whitepaper\n\nThe following is a representative CAIRN-style artifact journal demonstrating the collaborative creation of this whitepaper. It includes key events, roles, references, and lifecycle transitions.\n\n```\njsonld\n{\n "@context": "https://cairn.ai/ns/context.jsonld",\n "@id": "urn:cairn:whitepaper:2025-04-01",\n "@type": "cairn:CollaborativeArtifact",\n "cairn:artifactState": "finalized",\n "cairn:version": "1.0",\n "cairn:language": "en",\n "cairn:encoding": "UTF-8",\n "cairn:semanticHints": "whitepaper,standards,AI,governance",\n "cairn:originalText": "Initial prompt to AI requesting a review of NIST's Zero Draft webpage...",\n "cairn:normalizedText": "Please review the contents of this article and help formulate a response.",\n "cairn:promptAuthor": {\n "@id": "mailto:Nick@DrStavros.com",\n "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",\n "@type": "foaf:Person"\n },\n "cairn:modelUsed": {\n "@id": "https://openai.com/chatgpt",\n "foaf:name": "Hamish I. MacCloud, AIA",\n "@type": "foaf:Agent"\n },\n "cairn:generationTimestamp": "2025-04-01T00:00:00Z",\n "cairn:humanReviewer": {\n "@id": "mailto:Nick@DrStavros.com",\n "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",\n "@type": "foaf:Person"\n },\n "cairn:humanEditor": {\n "@id": "mailto:Nick@DrStavros.com",\n "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",\n "@type": "foaf:Person"\n },\n "cairn:approver": {\n "@id": "mailto:Nick@DrStavros.com",\n "foaf:name": "R. W. 'Nick' Stavros, Ph.D.",\n "@type": "foaf:Person"\n },\n "cairn:references": [\n {\n "@id": "https://www.nist.gov/artificial-intelligence/ai-research/nists-ai-standards-zero-drafts-pilot-project-accelerate",\n "dct:title": "NIST AI Standards Zero Drafts",\n "prov:wasAttributedTo": "Nick@DrStavros.com",\n "@id": "https://www.iso.org/standard/81230.html",\n "dct:title": "ISO/IEC 42001:2023 \u2014 AI Management System",\n "prov:wasAttributedTo": "Hamish I. MacCloud, AIA"\n }\n],\n "cairn:dialogThreadId": "urn:cairn:thread:whitepaper:2025",\n "cairn:thread": "The discussion paper now includes a new section titled '\u201cRegulatory Alignment and International Standards\u201d', covering:\n- ISO/IEC 42001\n- Harmonized standards under the EU AI Act\n- Why CAIRN is strategically relevant for international AI compliance\n\n—\n\n### \u201cShould the letter go in the GitHub repo too?\u201d\n\nHere\u2019s a quick take:\n\n"cairn:normalizedText": "Yes. I think we need that.\n\nThe current TOC is:\n\n1. Background\n2. Design Goals\n3. Key Concepts\n3.1 Source Attribution and References\n3.2 Prompt Fidelity\n3.3 Lifecycle States\n3.4 Multi-Agent Attribution\n3.5 Dialog Graphs\n3.6 Interoperability\n4. Accountability and Roles\n5. Alignment with Global Standards and Governance\n6. Vision: Trustable AI at Scale"</p>
<div>209</div>	<p>"cairn:normalizedText": "Yes. I think we need that.\n\nThe current TOC is:\n\n1. Background\n2. Design Goals\n3. Key Concepts\n3.1 Source Attribution and References\n3.2 Prompt Fidelity\n3.3 Lifecycle States\n3.4 Multi-Agent Attribution\n3.5 Dialog Graphs\n3.6 Interoperability\n4. Accountability and Roles\n5. Alignment with Global Standards and Governance\n6. Vision: Trustable AI at Scale"</p>

210 Line Number	" <i>cairn:aiInterpretation</i> ": "AI interpretation of human prompt to be filled in or refined." CARIN JSON ,
212	{
213	" <i>@id</i> ": "urn:cairn:entry:026",
214	" <i>cairn:timestamp</i> ": "2025-03-30T10:25:00Z",
215	" <i>cairn:humanText</i> ": "Now, we want to add this as as a Summary:\n\nTo learn more or participate in piloting CAIRN, visit: https://github.com/JackRabbitConsulting/cairn-standard \n\nAs well as an actual summary.",
216	" <i>cairn:aiResponse</i> ": "summary: \"CAIRN (Collaborative AI Artifact Record & Notation) is an open, standards-aligned framework that captures the human and AI contributions to generated content. It supports traceability, reference attribution, and lifecycle transparency\u2014enabling human accountability in AI-assisted work. CAIRN is compatible with international initiatives such as ISO/IEC 42001 and the EU AI Act. To learn more or participate in piloting CAIRN, visit: https://github.com/JackRabbitConsulting/cairn-standard \" \n\nsummary: \"This whitepaper is itself a collaborative artifact co-authored by a human (R. W. 'Nick' Stavros, Ph.D.) and an artificial intelligence assistant (Hamish I. MacCloud, AIA), demonstrating the principles and vision behind \",
217	" <i>cairn:normalizedText</i> ": "Now, we want to add this as as a Summary:\n\nTo learn more or participate in piloting CAIRN, visit: https://github.com/JackRabbitConsulting/cairn-standard \n\nAs well as an actual summary.",
218	" <i>cairn:aiInterpretation</i> ": "AI interpretation of human prompt to be filled in or refined."
219	}
220]
221	}
222	