Internet Engineering Task Force
Internet-Draft
Updates: 3261 (if approved)
Intended status: Experimental
Expires: 28 August 2026

N. Serafini
Individual

24 February 2026

Problem Statement for an AI-Session-ID in SIP
draft-nser-sip-ai-problem-statement-00

Abstract

This document describes a signaling-layer correlation gap in SIP-based real-time AI voice systems. Existing SIP identifiers provide dialog and communication-session scope, but they do not define a stable application-layer identifier capable of persisting across topology changes, multi-dialog call control operations, federated interconnects, or distributed processing environments.

The document analyzes current mechanisms and their limitations, and motivates the need for a standardized SIP-carried identifier whose semantics are restricted to application-layer correlation and out-of-band context retrieval.

Status of This Memo

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

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

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

This Internet-Draft will expire on 28 August 2026.

Table of Contents

1. Introduction 2
2. Analysis of Existing SIP Mechanisms 3
2.1. Call-ID and Dialog Identifiers 3
2.2. Session-ID 3
2.3. User-to-User Information (UUI) 4
2.4. Other Deployment Constructs 4
3. Why an AI-Session-ID is Required 4
4. Differentiation from Existing Solutions and Identifiers . . . 5
5. Requirements for an Application-Layer AI-Session-ID 5
5.1. Requirements 5
6. Scope Clarification 7
7. IANA Considerations 7
8. Security and Privacy Considerations 7
9. References 7
9.1. Informative References 7
Acknowledgements 8
Contributors 8
Author's Address 8

1. Introduction

SIP defines transaction and dialog machinery for communication signaling ([RFC3261]). Dialog identity is established by Call-ID and local/remote tags. That identity is sufficient for SIP dialog processing, routing decisions, and in-dialog request handling.

A second identity class is communication session identity. Session-ID ([RFC7989]) provides end-to-end and endpoint-associated identifiers for communication sessions and troubleshooting use cases. Session-ID semantics are intentionally focused on communication sessions.

A third identity class is application-layer identity - defined in this document. In distributed AI voice systems, application components maintain conversation state, policy state, tool-execution state, and audit state that are not encoded in SIP dialog state. This logical flow can span multiple SIP dialogs and, in some deployments, multiple communication sessions or federations.

This document uses the term "AI-Session-ID" to denote an application-layer correlation identifier carried in SIP. AI-Session-ID is not defined as a SIP dialog identifier and is not defined as a communication session identifier. Its intended role is limited to stable correlation.

2. Analysis of Existing SIP Mechanisms

This section reviews existing SIP and related mechanisms and evaluates their scope relative to application-layer continuity requirements.

2.1. Call-ID and Dialog Identifiers

Call-ID and tags identify SIP dialogs ([RFC3261]). They solve dialog matching and in-dialog request association.

Dialog identifiers are dialog-local; transfers, B2BUA-created legs, dialog replacement, topology-hiding transformations, and re-originations can produce new dialogs with different identifiers.

Therefore, dialog identifiers do not by themselves provide stable correlation for one logical application workflow across multiple dialogs.

2.2. Session-ID

Session-ID ([RFC7989]) provides communication-session identifiers intended for cross-element troubleshooting and correlation of communication sessions.

Session-ID semantics are tied to communication sessions; application workflows may intentionally span multiple communication sessions while preserving one logical application interaction.

Therefore, Session-ID is necessary for communication-session diagnostics but does not fully define application-layer workflow continuity semantics.

2.3. User-to-User Information (UUI)

UUI ([RFC7433]) transports user-to-user application information in SIP signaling. It solves structured transport of application data between participating endpoints.

UUI is payload-oriented and deployment behavior depends on intermediary and policy handling. It does not define a globally stable, workflow-scoped identifier with deterministic propagation requirements across all dialog derivation patterns.

Therefore, UUI can carry correlation data but does not by itself establish interoperable semantics for a persistent application workflow identifier.

2.4. Other Deployment Constructs

Implementations often use proprietary headers, platform interaction identifiers, and database mappings. These approaches can be operationally effective within one implementation.

Semantics, security handling, and propagation behavior are deployment-specific. Interoperable behavior across vendors and domains is not guaranteed.

3. Why an AI-Session-ID is Required

Distributed AI voice systems typically include SIP proxies, SBCs, B2BUAs, media services, AI orchestration services, and external tools. Processing is commonly distributed across regions and horizontally scaled nodes.

In these architectures, a single logical interaction can involve multiple SIP dialogs due to AI-to-AI transfer, AI-to-human-to-AI escalation, B2BUA leg creation, third-party call control, and callbacks. Stateless SIP processing patterns and retry behavior further require idempotent operations across asynchronous components.

Dialog-local state is insufficient because dialog identifiers can change when signaling topology changes. Communication session identity is insufficient because one flow may span multiple communication sessions.

A stable application-layer identifier transported in SIP enables deterministic out-of-band context retrieval through HTTP APIs, event systems, and distributed state stores. This allows any authorized component receiving signaling to re-associate events with the correct workflow state during normal processing, retry, or failover.

4. Differentiation from Existing Solutions and Identifiers

AI-Session-ID is differentiated as follows:

* AI-Session-ID is not a communication session identifier and does not replace Session-ID ([RFC7989]).

* AI-Session-ID may span multiple communication sessions when those sessions are part of one logical workflow.

* AI-Session-ID is not equivalent to Call-ID and does not redefine SIP dialog identity ([RFC3261]).

* AI-Session-ID does not carry full application context; it carries only a stable identifier.

As result, standardized AI-Session-ID construct can provide deterministic context lookup semantics, improve distributed state resilience, support observability and governance workflows, enable multi-vendor interoperability, and reduce dependence on fragile mapping logic.

5. Requirements for an Application-Layer AI-Session-ID

The following requirements are derived from the use cases and constraints described in this document. They define properties for a SIP-level construct intended to represent an application-layer AI-Session-ID identifier.

5.1. Requirements

* REQ1: It MUST be possible to associate multiple SIP dialogs or communication sessions with a single logical AI-Session-ID when those dialogs or sessions are components of one continuous application-layer interaction.

* REQ2: It MUST be possible to preserve the AI-Session-ID identifier across dialogs created as a result of call control operations (e.g., REFER, Replaces) when still part of the same application-layer interaction.

* REQ3: If an AI-Session-ID identifier is present in a dialog-establishing request, it MUST be included unchanged in all subsequent in-dialog requests generated within that dialog.

* REQ4: It MUST be possible for an authorized application-layer orchestrator to inject an AI-Session-ID identifier into a newly originated SIP dialog, including dialogs not directly derived from an existing dialog.

* REQ5: The identifier MUST be suitable for use as a stable key for retrieving application-layer session context through out-of-band mechanisms; SIP signaling itself MUST NOT carry the session context.

* REQ6: Once established within a dialog, the AI-Session-ID identifier MUST NOT change during the lifetime of that dialog.

* REQ7: If a different AI-Session-ID identifier is received within an existing dialog, the receiving entity MUST either reject or ignore the conflicting value while preserving the original identifier.

* REQ8: When a B2BUA receives differing AI-Session-ID identifiers on separate legs that it bridges, it MUST apply a deterministic local conflict-resolution policy.

* REQ9: The identifier MUST be opaque and MUST NOT encode user identity, device identity, domain identity, Call-ID, tags, IP addresses, or other SIP header fields or signaling metadata.

* REQ10: Possession of the AI-Session-ID identifier MUST NOT by itself grant authorization to retrieve application-layer session context; out-of-band retrieval mechanisms MUST enforce independent authentication and authorization.

* REQ11: The presence of the identifier MUST NOT alter SIP dialog semantics or call processing for entities that do not understand it.

* REQ12: The identifier MUST be suitable for safe transport in SIP headers and compatibility with other signaling or control protocols.

* REQ13: The identifier SHOULD be globally unique in time and space to avoid collision across administrative domains.

* REQ14: It MUST be possible for administrators and monitoring systems to use the identifier to correlate SIP signaling records, media processing events, AI model interactions, and external tool invocations into a coherent interaction timeline.

6. Scope Clarification

This document is a problem statement, it identifies protocol and architectural gaps and states requirements for further specification work.

7. IANA Considerations

This document makes no request of IANA.

8. Security and Privacy Considerations

AI-Session-ID is intended as a correlation identifier, not an authorization artifact. In single-domain or fully trusted environments, transporting only the identifier may be operationally sufficient. When signaling crosses administrative boundaries or traverses semi-trusted intermediaries, deployments often need a verifiable way to trust the asserted AI-Session-ID value.

SIP transport protections such as TLS or mutual TLS provide hop-by-hop confidentiality and integrity. However, hop-by-hop protection alone does not guarantee end-to-end integrity of an asserted AI-Session-ID value against in-path intermediaries that can legitimately terminate and re-originate SIP messages.

A deployment best practice is to carry AI-Session-ID together with an associated signed assertion when a trust boundary is crossed. This is an optional mechanism selected by policy and it is not required in all deployments.

To keep SIP signaling lightweight, deployments may prefer passing a signed assertion by reference; when by-reference retrieval is unavailable, an embedded assertion may be used as a fallback.

Possession of an AI-Session-ID value or an associated signed assertion does not by itself grant access to application context. Out-of-band context retrieval requires independent authentication and authorization.

9. References

9.1. Informative References

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <https://www.rfc-editor.org/info/rfc3261>.

[RFC3891] Mahy, R., Biggs, B., and R. Dean, "The Session Initiation Protocol (SIP) "Replaces" Header", RFC 3891, DOI 10.17487/RFC3891, September 2004, <https://www.rfc-editor.org/info/rfc3891>.

[RFC7433] Johnston, A. and J. Rafferty, "A Mechanism for Transporting User-to-User Call Control Information in SIP", RFC 7433, DOI 10.17487/RFC7433, January 2015, <https://www.rfc-editor.org/info/rfc7433>.

[RFC7989] Jones, P., Salgueiro, G., Pearce, C., and P. Giralt, "End-to-End Session Identification in IP-Based Multimedia Communication Networks", RFC 7989, DOI 10.17487/RFC7989, October 2016, <https://www.rfc-editor.org/info/rfc7989>.

Acknowledgements

Acknowledgements section.

Contributors

Contributions section.

Author’s Address

Nicola Serafini
Individual
Email: n.serafini@tutanota.com