# Sajimora™: Deterministic Cross-Language Contract Verification and Evidence-Grounded Static Analysis

Technical Whitepaper — Defensive Publication — Version 1.1 (November 1, 2026)

**The Sajimora™ Project • Canonical: https://sajimora.com • Contact: admin@sajimora.com**

**Abstract—**Autonomous AI coding agents and automated refactoring systems face an architectural crisis when refactoring distributed enterprise systems: silent contract drift across polyglot boundaries. When backend APIs (Java Spring Boot, Micronaut) evolve, lexical grep and probabilistic RAG suffer from catastrophic semantic blindness—failing to resolve class inheritance, Jackson serialization aliases (@JsonProperty), and dynamic request envelopes, while flooding LLM context windows with tens of thousands of irrelevant tokens. This whitepaper establishes the formal architecture of **Sajimora™**: (1) a cross-language Concrete Syntax Tree (CST) query engine achieving **98.66% context token reduction** over lexical grep on enterprise codebases; (2) an evidence-grounded AI architecture where an embedded single-pass encoder proposes candidate evidence while a 100% deterministic Rust mathematical gate (VerdictAuthority) certifies soundness fail-closed. This document constitutes a technical disclosure intended for consideration as potential prior art.

## 1. Introduction & The Empirical Failure Modes of Lexical Grep

When an AI coding agent refactors an enterprise backend API, it must answer: *'What client call sites break if this field is renamed or deleted?'* In current practice, agents invoke lexical search tools (grep, ripgrep). Across an empirical benchmark of 20 production REST APIs in the ThingsBoard enterprise repository (a multi-million line Java backend + Angular frontend), lexical search demonstrated severe architectural failure modes: **(1) Blindness to Inheritance:** Controllers inherit endpoints from generic CRUD bases (e.g. DeviceController extends BaseCrudController<Device, UUID>); grep returns 0 matches, falsely certifying the endpoint is dead. **(2) Wire-Envelope Obfuscation:** Frameworks wrap DTOs in reactive envelopes (ResponseEntity<Mono<Page<TbResourceInfo>>>); grep cannot unwrap generic containers. **(3) The Common Symbol Catastrophe:** Modifying common words like .name floods the context window with 46,025 tokens (1,097 matched lines), triggering 'Lost in the Middle' attention degradation. **(4) Serialization Alias Invisibility:** Renaming a field while preserving @JsonProperty("alias") triggers false breaking changes, while deleting @JsonProperty goes completely undetected.

## 2. Enabling Implementation: Mathematical Formalism & Path Lattice

Sajimora™ models distributed cross-language verification as a closed-loop obligation conservation problem over an algebraic lattice. Let a path expression P be a sequence of segment atoms: **Literal(s) ■ Param(p, τ) ■ Wildcard(*) ■ RecursiveWildcard(**)**.



Let O_net be the net unresolved contract obligations across all produced endpoints E and client invocations C. **Safety Invariant:** A build passes compilation if and only if O_net = 0. Any link where SoundMatch is Unknown forces an immediate fail-closed build halt unless an immutable cryptographic exception hash exists in .sajimora/exceptions.sig.

## 3. Evidence-Grounded AI Architecture

**AI Evidence Decoupling:** AI is authorized to discover and propose candidate evidence; only deterministic Rust mathematics certifies verdicts. A foundational design principle is that AI proposes candidate evidence but never certifies its own conclusions. The pipeline consists of three stages: (1) an embedded 25M-40M single-pass slot-tagging encoder (sajimora-ai) running locally via Rust candle in < 5ms on a standard CPU to extract candidate dynamic URL fragments; (2) Salsa incremental symbol-table grounding to resolve candidates against declared controller schemas; and (3) a deterministic Rust verification gate (VerdictAuthority) that certifies contract soundness fail-closed.

## 4. Empirical Verification: 20 Selected DTO/Field Cases (ThingsBoard)

| # | Target DTO & Field | Grep Context | Lines | Sajimora | Sites | Reduction |
|---|---|---|---|---|---|---|
| 1 | RuleChain.name | 46,025 tok | 1,097 | 464 tok | 8 | -98.99% |
| 2 | TbResourceInfo.customerId | 10,933 tok | 287 | 353 tok | 5 | -96.77% |
| 3 | AgentApplication.agentId | 1,218 tok | 34 | 372 tok | 5 | -69.45% |
| 4 | JwtPair.token | 3,114 tok | 66 | 192 tok | 0 | -93.83% |
| 5 | Alarm.customerId | 10,933 tok | 287 | 259 tok | 2 | -97.63% |
| 6 | Device.customerId | 10,933 tok | 287 | 291 tok | 3 | -97.33% |
| 7 | OtaPackageInfo.deviceProfileId | 2,035 tok | 53 | 251 tok | 1 | -87.66% |
| 8 | AgentAppProfile.name | 46,025 tok | 1,097 | 338 tok | 4 | -99.26% |
| 9 | Edge.customerId | 10,933 tok | 287 | 366 tok | 5 | -96.65% |
| 10 | Integration.configuration | 24,550 tok | 523 | 312 tok | 3 | -98.72% |
| 11 | Job.customerId | 10,933 tok | 287 | 322 tok | 3 | -97.05% |
| 12 | Queue.name | 46,025 tok | 1,097 | 187 tok | 0 | -99.59% |
| 13 | WhiteLabelingParams.logoImageUrl | 268 tok | 9 | 396 tok | 5 | Grep missed 5! |
| 14 | DeviceProfile.name | 46,025 tok | 1,097 | 302 tok | 3 | -99.34% |
| 15 | EntityInfo.name | 46,025 tok | 1,097 | 189 tok | 0 | -99.58% |
| 16 | EntityView.entityId | 13,989 tok | 378 | 266 tok | 2 | -98.09% |
| 17 | HomeDashboardInfo.dashboardId | 2,382 tok | 62 | 281 tok | 2 | -88.20% |
| 18 | TenantProfile.name | 46,025 tok | 1,097 | 302 tok | 3 | -99.34% |
| 19 | AiModel.version | 12,436 tok | 321 | 260 tok | 2 | -97.90% |
| 20 | Asset.customerId | 10,933 tok | 287 | 259 tok | 2 | -97.63% |
| **TOTAL** | **20 Selected DTO/Field Cases** | **445,785 tok** | **8,477** | **5,963 tok** | **58** | **-98.66% OVERALL** |

## 5. Defensive Publication and Intellectual Property Declaration

**5.1 Purpose of Publication:** This document constitutes a public technical disclosure of the methods, architectures, algorithms, mathematical models, and technical implementations expressly described within this publication. The publication is intended to establish a publicly accessible technical record of the disclosed subject matter, including its publication date and technical contents, for use as potential prior art in subsequent patent examination and related intellectual property proceedings.

**5.2 Scope of Technical Disclosure:** The following subject matter is expressly disclosed for defensive publication: (1) Closed-loop obligation conservation and algebraic path-lattice verification across polyglot boundaries; (2) Separation of probabilistic AI-based candidate evidence generation from deterministic static verification and verdict certification; and (3) The specific algorithms, mathematical representations, data structures, and implementation approaches described in Sections 2, 3, and 4 of this document. This disclosure applies only to subject matter sufficiently described in this publication and does not extend to undisclosed inventions, proprietary systems language architectures, or subsequent technical improvements.

**5.3 Defensive Prior Art Intent:** The publisher intends this publication to be publicly accessible and available for consideration as prior art against subsequent patent applications claiming subject matter already disclosed herein. The publication may be submitted to patent offices, examiners, courts, or other competent authorities as documentary evidence. The legal effect shall be determined under the applicable patent laws of the relevant jurisdiction.

**5.4 Copyright Reservation:** Copyright in the original text, illustrations, diagrams, graphical elements, and other copyrightable expression contained in this document remains with its respective copyright owner. Defensive publication of technical methods does not transfer copyright ownership or grant ownership of the Sajimora brand, trademarks, or other separately protected intellectual property.

**5.5 Software and Trademark Separation:** This publication does not independently establish an open-source licence for Sajimora software. The licensing, distribution, modification, and commercial use of Sajimora source code shall be governed by the separate software licence accompanying the relevant repository. The Sajimora™ name, logos, branding, and associated trademarks are not included in any public-domain dedication.

**5.6 Publication Record:** Project: **Sajimora™** • Document: **Technical Whitepaper — Defensive Publication** • Version: **1.1** • Publication date: **November 1, 2026** • Canonical location: **https://sajimora.com** • Publisher: **Sajimora Project** • Contact: **admin@sajimora.com** • Document SHA-256 Digest: *Recorded upon compilation*.