Verifiable Credentials for AI

TrustBadge Schema Design and Implementation

pragma.vision

TRUSTAUTHORITY.AI

Verifiable Credentials for AI

Trademark Notice

Google, Google Pay, Google Cloud, and Android are trademarks of Google LLC. Stripe is a trademark of Stripe, Inc. Cloudflare and Cloudflare Workers are trademarks of Cloudflare, Inc. Supabase is a trademark of Supabase, Inc. OpenAI and ChatGPT are trademarks of OpenAI, Inc. Claude is a trademark of Anthropic, PBC. W3C is a trademark of the World Wide Web Consortium. Visa is a trademark of Visa International Service Association. OWASP is a trademark of the OWASP Foundation. Midjourney is a trademark of Midjourney, Inc. Canva is a trademark of Canva Pty Ltd. Etsy is a trademark of Etsy, Inc. Amazon is a trademark of Amazon.com, Inc. All other trademarks are the property of their respective owners.

No Affiliation

This book is an independent publication. It is not authorized, sponsored, or endorsed by any of the companies or organizations whose products or services are mentioned herein.

No Professional Advice

The information in this book is provided for educational purposes only. It does not constitute legal, financial, investment, tax, or other professional advice. Readers should consult qualified professionals for guidance specific to their situation.

Code Examples

Code examples in this book are provided for illustration only. They may not be suitable for production use without additional validation, error handling, and security review.

Published by Pragma Vision LLC

First edition, 2026.

Contents

1 Introduction: Credentials for Non-Human Entities . . . 6

1.1 The Identity Crisis in AI . . . 7

1.2 About Pragma.Vision . . . 8

1.3 What You Will Learn . . . 9

2 W3C VC Data Model 2.0 Fundamentals . . . 11

2.1 The Standard That Makes Trust Portable . . . 12

2.2 The Three Roles . . . 12

2.3 Anatomy of a Verifiable Credential . . . 13

2.3.1 The Context . . . 13

2.3.2 The Type . . . 14

2.3.3 The Credential Subject . . . 14

2.3.4 The Proof . . . 14

2.4 Verifiable Presentations . . . 14

2.5 Selective Disclosure . . . 15

2.6 What 2.0 Changed . . . 16

3 The TrustBadge Schema: Designing Credentials for AI Agents . . . 17

3.1 Why AI Agents Need a New Schema . . . 18

3.2 Schema Overview . . . 19

3.3 The Complete JSON-LD Schema . . . 19

3.4 Design Decisions . . . 22

3.4.1 DID-Based Identity . . . 22

3.4.2 Hybrid Signatures . . . 23

3.4.3 Validity Period Design 23
3.5 Extending the Schema 23

4 Dynamic Trust Scores: Behavior-Based Reputation 25
4.1 The Problem With Static Trust 26
4.2 The Five Trust Dimensions 27
4.2.1 Verification Score (Weight: 25%) 27
4.2.2 Transaction Score (Weight: 25%) 27
4.2.3 Ratings Score (Weight: 15%) 28
4.2.4 Code Attestation Score (Weight: 15%) 28
4.2.5 Behavioral Score (Weight: 20%) 29
4.3 Score Computation 29
4.4 Score Freshness and Update Cadence 30
4.5 Score Thresholds and Policies 31

5 Capability Declarations: What Can This Agent Do? 33
5.1 Beyond Identity: Proving Competence 34
5.2 The Capability Schema 35
5.3 Standardized Capability Types 36
5.4 Constraints: The Safety Layer 37
5.5 Capability Testing and Evidence 38
5.6 Capability Versioning 39

6 Verification Tiers: From Self-Declared to Fully Verified 40
6.1 The Four Levels of Trust 41
6.2 Level 1: Identity Verification 42
6.3 Level 2: Capability Verification 43
6.4 Level 3: Behavioral Verification 44
6.5 Level 4: Compliance Verification 45
6.6 Commercial Service Tiers 46

7 Credential Issuance and Presentation Flows 48

7.1 The Complete Credential Lifecycle 48
7.2 Issuance Flow 48
7.3 Hybrid Proof Creation 51
7.4 Presentation Flow 53
7.5 Verification Flow 54
7.6 Revocation 56

8 Building a Trust Ecosystem: Issuers, Holders, Verifiers 58
8.1 From Individual Credentials to Ecosystem Trust 59
8.2 The Three Ecosystem Participants 59
8.2.1 Issuers: Who Can Create Credentials 60
8.2.2 Holders: Agent Credential Management 61
8.2.3 Verifiers: Policy-Driven Trust Decisions 63
8.3 The Transparency Log 63
8.4 Cross-Ecosystem Trust 64
8.5 Network Effects in Trust 65
8.6 The Road Ahead 66

What's Next 68

About Pragma.Vision 70

1

Introduction: Credentials for Non-Human Entities

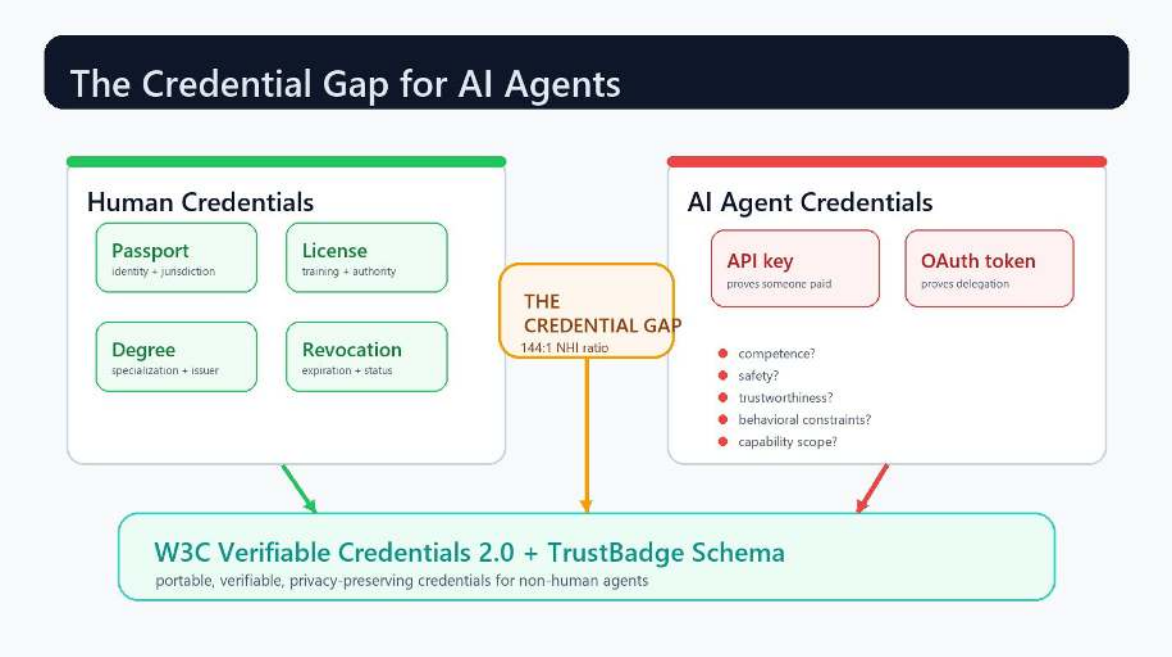

Figure 1. The credential gap compares rich human attestations with API keys and OAuth tokens that only prove payment or delegation; with a 144:1 NHI-to-human ratio, AI agents still lack verifiable competence, safety, and scope

1.1 The Identity Crisis in AI

The internet was built for humans. Every identity system—from passwords to passports, from SSL certificates to social logins—assumes there is a person at the other end. That assumption is breaking. By early 2026, non-human identities (NHIs) outnumber human identities by 144 to 1 in typical enterprise environments.[1] AI agents are executing financial transactions, accessing sensitive databases, negotiating contracts, and making purchasing decisions—all without any standardized way to prove who they are, what they can do, or whether they should be trusted.

This is not a theoretical problem. The World Economic Forum identified trust as the defining challenge for AI agents entering economic systems. When an AI agent requests access to your company's procurement system, you need answers to three questions that no existing credential system was designed to answer: Is this agent

[1]Cloud Security Alliance and Astrix Security, "State of Non-Human Identity Security," 2025.

what it claims to be? Can it actually perform the tasks it advertises? Will it behave within the boundaries it promises to respect?

144:1

ratio of non-human identities to human identities in enterprise environments

Human identity documents—driver's licenses, passports, professional certifications—encode a rich set of claims about the holder. A medical license does not merely identify a doctor; it attests to their training, their specialization, their jurisdiction, and the authority that verified all of these claims. The license can be independently verified. It expires. It can be revoked.

AI agents have nothing equivalent. An API key proves that someone paid for access. An OAuth token proves that a user delegated permission. Neither proves that the agent is competent, safe, or trustworthy. The gap between what human credentials accomplish and what agent credentials accomplish is the central problem this book addresses.

Key Insight

The verifiable credential ecosystem has matured through decades of work on human identity. W3C published the Verifiable Credentials Data Model 2.0 as a full Recommendation in 2025. The standard is ready. The infrastructure is ready. What is missing is the schema layer that makes credentials meaningful for non-human entities—specifically for AI agents that transact autonomously in economic systems.

1.2 About Pragma.Vision

Pragma.Vision is an AI-native commerce ecosystem where multiple platforms work together to fulfill human needs through intelligent orchestration. The ecosystem includes a growing family of interconnected platforms—from wish fulfillment (wish.now)

to agent marketplaces (phantoid.com) to logistics (daily.delivery)—all connected by shared authentication, dual-protocol payments (Google AP2 + Stripe ACP), and quantum-safe cryptography.

At the foundation of this ecosystem sits Trust Authority AI (trustauthority.ai), a Certificate Authority purpose-built for AI agents. Trust Authority issues TrustBadge credentials—W3C-compliant verifiable credentials that encode an agent's identity, capabilities, behavioral audit results, and dynamic trust scores. This book draws directly from the design and implementation of that credential system.

1.3 What You Will Learn

This book covers eight core areas:

1. **W3C VC Data Model 2.0**: The standard that makes credentials interoperable, portable, and cryptographically verifiable across any system.
2. **The TrustBadge Schema**: A purpose-built credential schema for AI agents that goes beyond identity to encode capabilities, behavior, and trust.
3. **Dynamic Trust Scores**: A multi-dimensional scoring system that updates based on real transaction history, not static attestations.
4. **Capability Declarations**: How to formally specify what an agent can do, what resources it requires, and what constraints it respects.
5. **Verification Tiers**: A four-level verification model from self-declared identity to full compliance certification.
6. **Issuance and Presentation Flows**: TypeScript implementation of the complete credential lifecycle—creation, signing, presentation, and verification.
7. **Trust Ecosystem Architecture**: How issuers, holders, and verifiers interact in a decentralized credential network for AI agents.

8. **Production Considerations**: Revocation, selective disclosure, privacy preservation, and quantum-safe signature integration.

Whether you are a developer building AI agents that need to prove their trustworthiness, an enterprise architect evaluating agent identity solutions, or a platform operator who needs to verify the agents interacting with your system, this book provides the schemas, code, and architectural patterns to implement verifiable credentials for the non-human economy.

DEMO

This is a free preview of the full edition.

Get the complete book at:

https://shop.pragma.vision