CSAI Foundation

Cloud Security Alliance AI Safety Initiative

ENISA EUDIW Certification: AI Agent Identity in EU Markets

Security Requirements for AI Agent Authorization Under eIDAS 2.0

Unofficial AI-assisted Research

2026-04-15

Key Takeaways

- ENISA published a draft candidate EU Digital Identity Wallet (EUDIW) cybersecurity certification scheme in April 2026, with public consultation open through April 30, 2026 [1]. Every EU member state must provide at least one certified wallet by end of 2026 [2].
- The eIDAS 2.0 regulation (EU 2024/1183) establishes the legal and technical foundation for the EUDIW, but its implementing acts do not yet expressly address AI agents as first-class identity principals — creating a significant governance gap for enterprises deploying agentic AI in European markets [3].
- The EUDIW's delegation and granular authorization capabilities — originally designed for person-to-person and business-to-person contexts — are now being evaluated as a governance layer for AI agent authorization, with EU industry consortia actively advocating for official AI agent frameworks [4].
- Enterprises operating agentic AI that will interact with EUDIW relying parties should proactively address this gap, since no official bridge provisions yet exist: by mapping agent identities to EUDIW-compatible verifiable credential (VC) structures, aligning with OpenID4VP/VCI protocols, and applying layered controls consistent with both the EU AI Act and eIDAS 2.0 assurance requirements [5][6].
- Based on the authors' analysis, CSA frameworks — particularly MAESTRO, the Agentic Trust Framework (ATF), and the AI Controls Matrix (AICM) — address control domains that map to the gaps in ENISA's current certification scope for organizations managing AI agent authorization today.

Background

In May 2024, Regulation (EU) 2024/1183 entered into force, establishing the legal basis for the EU Digital Identity Wallet and mandating deployment across all member states. The regulation requires that each EU member state make interoperable EU Digital Identity Wallets available to citizens and businesses by the end of 2026 — a deadline embedded in the regulation as a binding obligation on all

member states [3]. The regulation charges ENISA with developing harmonized cybersecurity certification requirements for wallet implementations, giving those requirements binding force through the EU's Cybersecurity Act framework [7].

In January 2025, ENISA convened a dedicated Ad Hoc Working Group of 26 appointed members to develop a candidate cybersecurity certification scheme for the EUDIW [7]. After more than a year of technical drafting, ENISA published draft scheme version 0.4.614 for public review in April 2026, with a consultation window closing April 30, 2026 [1]. A parallel Contribution Agreement, signed in February 2026 and worth 1.6 million EUR over two years, tasks ENISA with supporting member state national certification schemes as they transition to the European framework [8]. The scheme's stated purpose is ensuring the "security and protection of privacy of users and their personal data," with a security architecture that references FIDO authentication standards and mandates high-assurance conformity assessment [9].

Simultaneously, the emergence of autonomous AI agents as operational actors in enterprise systems has introduced a class of identity principal that existing eIDAS 2.0 provisions have not addressed. The regulation's implementing acts published through November 2025 [11] do not establish frameworks for non-human identity principals, reflecting the regulation's original design scope. An AI agent — whether a customer service bot, an autonomous procurement system, or a multi-step reasoning pipeline — may acquire sensitive attributes, present credentials on a user's behalf, and execute consequential actions across organizational boundaries, all without direct human involvement at the moment of each transaction. The EUDIW's technical architecture, grounded in W3C Verifiable Credentials, Decentralized Identifiers (DIDs), OpenID for Verifiable Credential Issuance (OpenID4VCI), and OpenID for Verifiable Presentations (OpenID4VP), offers a principled basis for extending identity assurance to non-human actors [5][10]. Whether and how ENISA's certification scheme will formally address this extension remains an open question — and a critical one for any enterprise deploying agentic AI in European markets.

Security Analysis

The Delegation Gap in EUDIW Architecture

The EUDIW feature with the most direct relevance to AI agent authorization is its delegation model. Under eIDAS 2.0, a wallet holder may grant verifiable, bounded authorization to another party — for example, permitting a family member to execute financial transactions up to a specified limit, or authorizing an employee to access specific business accounts without sharing master credentials [4].

This delegation is cryptographically signed, auditable, and expressly scoped — properties consistent with the least-privilege principles that frameworks such as CSA's Agentic Trust Framework recommend for AI agent authorization [15].

However, the current regulatory and certification framework treats delegation exclusively in human terms. The implementing acts published through November 2025 define Person Identification Data (PID) and Electronic Attestations of Attributes (EAAs) in ways that presuppose a human data subject [11]. AI agents — which may be instantiated ephemerally, operate across jurisdictional boundaries, and act simultaneously under multiple delegated authorities — do not map cleanly onto these categories. The WE BUILD consortium, one of the EU's Large Scale Pilots for EUDIW deployment, published recommendations in early 2026 calling explicitly for standards bodies to establish working groups on interoperability between EUDIW and AI agents, and for the EU to develop a formal strategy for AI agent governance built on the European Digital Identity Framework [4]. These recommendations have not yet resulted in official implementing measures, leaving a structural gap in the certification landscape.

Security Architecture and Assurance Levels

ENISA's draft certification scheme for the EUDIW establishes three assurance levels — basic, substantial, and high — calibrated to the risk profile of the intended use case [1][9]. For wallet functions associated with legally binding actions, high-assurance certification is required. High assurance implies hardware-backed key storage (typically using a Trusted Execution Environment or secure element), liveness detection for biometric binding, and resistance to a defined threat model that includes remote software attacks, physical device compromise, and adversarial manipulation of the certification boundary.

The scheme's incorporation of FIDO authentication standards is significant for AI agent scenarios in two respects. First, FIDO's attestation model — in which a device proves its hardware security properties to a relying party — provides a template for how an AI agent's execution environment could be certified as meeting defined security properties. An agent running in a certified Trusted Execution Environment could, in principle, present wallet-compatible attestations that satisfy relying party trust requirements. Second, FIDO's reliance on asymmetric cryptography and origin binding offers protections against credential theft and relay attacks that are particularly relevant when credentials are being presented by automated systems rather than human-controlled browsers.

The EUCC (European Cybersecurity Certification Scheme for Common Criteria products), which underlies the hardware security requirements in the EUDIW scheme, is the emerging security foundation for European digital identity as a whole [12]. Enterprises selecting hardware security modules, mobile secure elements, or cloud-based attestation services for AI agent identity management should verify their suppliers' EUCC certification status, as this will likely become a prerequisite for relying party acceptance of wallet-compatible agent credentials.

The OpenID4VC Protocol Layer and AI Agent Authorization

The technical protocols underpinning EUDIW interoperability — OpenID4VCI for credential issuance and OpenID4VP for credential presentation — are now being adopted in enterprise identity infrastructure beyond the EUDIW context itself [5][10]. OpenID4VP, which defines how a holder presents verifiable credentials to a verifier, can accommodate machine-to-machine presentation flows where the presenter is an AI agent acting under delegated authority. The OpenID4VC High Assurance Interoperability Profile (HAIP), now entering self-certification through the OpenID Foundation as of February 2026, establishes the conformance baseline for high-assurance wallet interactions [13].

For AI agent authorization, a central design question is how delegation chains are represented within the VC data model. A plausible architecture involves three credential layers: a root PID credential issued to the human or legal entity authorizing the agent; a delegation credential signed by the authorizing entity and encoding the specific capabilities, scopes, and expiration conditions of the agent's authority; and an agent attestation credential issued by the agent's operator certifying the agent's provenance, model version, and execution environment. Relying parties would then verify all three layers before honoring an agent-initiated presentation. This architecture is technically consistent with the W3C Verifiable Credentials Data Model 2.0, which supports linked data proofs and credential chaining [14]. The question of whether ENISA's certification scheme will formally require or validate support for such multi-layer delegation flows remains open pending the outcome of the current public consultation.

Threat Vectors Specific to AI Agent Wallet Interactions

The intersection of EUDIW-grade identity infrastructure with agentic AI introduces threat vectors that are qualitatively different from those addressed by the existing certification scheme. Prompt injection attacks — in which an adversary embeds instructions within content processed by an AI agent that cause it to perform unauthorized actions — could be weaponized to abuse credential presentation. An agent instructed to "present your identity credential to verify this transaction" could, if insufficiently constrained, be manipulated into presenting credentials to an adversarial verifier that collects them for replay. The EUDIW scheme's draft threat model documentation does not explicitly address prompt-injection-driven credential abuse [9]; the attack class also lacks a direct operational analog in typical human wallet interactions, since human users do not process adversarially crafted content during credential presentation in the same way automated agents do.

Delegation credential abuse represents a second class of risk. Because AI agents may operate with persistent, long-lived credentials across many transactions, a compromised agent credential could enable unauthorized access across a larger number of transactions than a human credential presentation, which typically occurs in a single discrete interaction. The principle of ephemeral, short-

lived agent credentials — advocated in CSA's Agentic Trust Framework — substantially reduces the impact of a compromised credential by limiting its validity window, though it does not prevent credential abuse within that window and requires wallet infrastructure that supports high-frequency issuance and revocation without imposing prohibitive overhead on certification processes [15].

Finally, the combination of AI agents and the EUDIW's selective disclosure capabilities creates a new privacy risk surface. Selective disclosure — the ability to prove a single attribute (e.g., "age over 18") without revealing the underlying credential — is a deliberate privacy feature of the EUDIW design. However, an AI agent that repeatedly queries selective disclosures from the same wallet across many interactions may inadvertently allow statistical inference of the underlying attributes, a form of correlation attack that is heightened relative to typical human wallet usage patterns, since human users present credentials far less frequently than automated agents operating at scale. Organizations deploying agents that interact with EUDIW-compatible relying parties should evaluate whether this deployment triggers the high-risk threshold under GDPR Article 35 and, where it does, include this correlation risk in their Data Protection Impact Assessment.

Recommendations

Immediate Actions

Organizations operating agentic AI systems that interact with European identity infrastructure, or that will be subject to EUDIW relying-party requirements, should take several concrete steps now. First, designate a team responsible for tracking the ENISA EUDIW certification scheme consultation through its April 30, 2026 close, and submit feedback specifically addressing the gap in AI agent identity provisions. The current draft does not expressly accommodate non-human principals; enterprise input during this consultation window is the most direct lever for shaping the final scheme [1].

Second, audit existing AI agent identity architectures against the EUDIW technical framework. Specifically, assess whether agent credentials can be represented as Verifiable Credentials in SD-JWT or mdoc format — the two primary formats supported in the EUDI Architecture and Reference Framework — and whether agent authorization flows can be expressed using OpenID4VP presentation protocols. Organizations that discover significant architectural gaps between their current agent identity implementations and the EUDIW technical model should expect longer remediation timelines; the earlier this assessment begins, the more options remain available [5].

Third, ensure that any AI systems your organization operates that process EUDIW-issued credentials — for example, identity verification systems that accept wallet presentations from customers or partners — are evaluated under the EU AI Act's risk classification criteria. The EU AI Act classifies systems providing remote biometric identification — a category encompassing automated identity verification deployments — as high-risk, imposing documentation, logging, human oversight, and accuracy disclosure requirements that must be addressed before these systems can legally operate in production [6].

Short-Term Mitigations

While formal AI agent provisions in the EUDIW scheme are pending, organizations can implement interim controls that preserve audit trails and constrain agent authority in EUDIW-adjacent contexts. Structuring AI agent authorization using existing OAuth 2.0 token binding and Demonstrating Proof of Possession (DPoP) mechanisms provides cryptographic binding between agent credentials and their use context — a property that partially compensates for the lack of wallet-grade hardware attestation [16]. Combining DPoP-bound agent tokens with short expiration windows — 15 minutes or less for high-privilege operations, consistent with guidance in CSA's Agentic Trust Framework [15] — limits the exposure window of any compromised credential.

Organizations should also implement presentation audit logging that captures, at minimum, the credential type presented, the relying party identity, the claimed delegation authority, and the timestamp of each AI agent credential presentation. This logging serves dual purposes: it supports GDPR accountability obligations and provides the audit trail necessary to demonstrate compliance with the EU AI Act's transparency requirements for automated decision systems.

Strategic Considerations

For enterprises deploying agentic AI in European markets, AI agent identity management will require active governance attention across both eIDAS 2.0 and EU AI Act obligations simultaneously — a combination that creates novel compliance complexity. The Qualified Trust Service Provider (QTSP) model that underlies eIDAS 2.0's trust hierarchy offers a potential template for what identity industry observers and the authors of this note term AI Trust Service Providers (AITSPs) — specialized entities that would certify AI agent identity, provenance, and governance in a manner analogous to how QTSPs currently certify digital signatures and seals [17]. If the European Commission moves to formalize this concept, organizations that have established internal agent identity governance programs will be better positioned to achieve certification and to serve as reference implementers.

The European Business Wallet, announced in the Commission's 2025 Competitiveness Compass and formally proposed by the European Commission in 2025, represents a parallel identity channel specifically designed for organizational and inter-business identity [18]. Enterprises should evaluate whether their AI agent authorization architecture should be built on the citizen-facing EUDIW model, the forthcoming business wallet model, or a hybrid approach. The business wallet's design, oriented toward machine-to-machine credential exchange in supply chain and regulated industry contexts, may prove a more natural fit for enterprise AI agent authorization than the citizen wallet architecture.

CSA Resource Alignment

MAESTRO Threat Modeling

CSA's MAESTRO framework (Multi-Agent Environment, Security, Threat, Risk, and Outcome) provides the most directly applicable threat model for the threat vectors described in this note. MAESTRO's systematic enumeration of risks across agentic architectures — including delegation abuse, prompt injection enabling unauthorized action, and trust boundary violations at multi-agent boundaries — aligns with the threat categories that ENISA's EUDIW scheme will need to address as AI agents become relying parties and credential presenters in the wallet ecosystem [19]. Organizations should conduct MAESTRO-based threat assessments for any AI agent deployment that will interact with EUDIW-compatible identity infrastructure.

Agentic Trust Framework (ATF)

CSA's Agentic Trust Framework directly addresses the governance architecture that organizations need while formal EU provisions for AI agent identity remain pending. Based on the authors' analysis, the ATF's requirements for verified agent identity with auditable provenance, ephemeral credentials scoped to operational context, and policy-as-code authorization rules are architecturally consistent with the EUDIW's VC-based delegation model and provide an interim governance layer that organizations can implement today [15]. The ATF's emphasis on OAuth2/OIDC integration for approval workflows is particularly relevant to organizations building EUDIW-adjacent agent authorization systems, as OpenID4VP is an extension of the same OAuth ecosystem.

AI Controls Matrix (AICM)

The AICM's identity and access management control domain addresses the requirement for non-human identity management that is currently absent from eIDAS 2.0 implementing acts. Specifically, AICM controls related to AI system lifecycle governance, human oversight of automated decisions, and supply chain attestation for AI components provide a practical governance baseline that enterprises can use to satisfy both EU AI Act compliance and the anticipated future requirements of any EUDIW AI agent certification extension [20]. The AICM's Shared Security Responsibility Model (SSRM) is also directly applicable to the tripartite responsibility structure of EUDIW deployments, where wallet providers, credential issuers, and relying parties each bear distinct security obligations.

Zero Trust for AI Guidance

CSA's Zero Trust guidance for AI environments — particularly the principle that no agent should be trusted by virtue of its network position or prior session state — is aligned with the EUDIW's architecture, which requires fresh cryptographic proof at each credential presentation rather than relying on session continuity [21]. Applying Zero Trust architecture to AI agent interactions with EUDIW infrastructure means treating each credential presentation as an independent trust decision, verifying the entire delegation chain at presentation time, and maintaining no implicit trust in agents that have previously authenticated successfully.

References

[1] ENISA. "Draft candidate EUDIW Scheme v0.4.614 for Public Review." European Union Cybersecurity Certification, April 2026.

[2] ENISA. "ENISA advances the certification of EU Digital Wallets." ENISA News, 2026.

[3] European Commission. "European Digital Identity (EUDI) Regulation." Digital Strategy, May 2024.

[4] Biometric Update. "EU can rein in AI agents with EUDI Wallets and business wallets: WE BUILD." Biometric Update, March 2026.

[5] Walt.id. "What is OpenID4VCI? The Developer's Guide (2026)." Walt.id Documentation, 2026.

[6] European Commission. "EU Artificial Intelligence Act." Shaping Europe's Digital Future, 2024.

[7] ENISA. "EU Digital Identity Wallet: A leap towards secure and trusted electronic identification through certification." ENISA News, 2024.

[8] Biometric Update. "ENISA invites feedback for EU Digital Identity Wallet cybersecurity certification." Biometric Update, April 2026.

[9] Mobile ID World. "ENISA Sets Security Bar for Europe's Mandatory Digital Wallet." Mobile ID World, 2026.

[10] Walt.id. "What is OpenID4VP? The Developer's Guide (2026)." Walt.id Documentation, 2026.

[11] Biometric Update. "EU publishes three new Implementing Acts for EUDI Wallet." Biometric Update, November 2025.

[12] CCLab. "EUCC Behind eIDAS 2.0: The New Pillar of Security in Europe's Digital Identity Framework." CCLab News, 2025.

[13] Biometric Update. "OpenID Foundation launching self-certification program for 3 specs in Feb 2026." Biometric Update, December 2025.

[14] W3C. "Verifiable Credentials Data Model 2.0." W3C Recommendation, May 2025.

[15] Cloud Security Alliance. "Agentic Trust Framework: Zero Trust for AI Agents." CSA Blog, February 2026.

[16] IETF. "OAuth 2.0 Demonstrating Proof of Possession (DPoP)." RFC 9449, September 2023.

[17] Spherity. "AI Agents Need Identity – Europe Has the Blueprint." Spherity Blog, 2025.

[18] Spherity. "The European Business Wallet: A Strategic Pillar for Digital Identity and Industrial Competitiveness in the EU." Medium/Spherity, 2025.

[19] Cloud Security Alliance. "Agentic AI Threat Modeling Framework: MAESTRO." CSA Blog, February 2025.

[20] Cloud Security Alliance. "Agentic AI Identity and Access Management." CSA Artifacts, 2025.

[21] Cloud Security Alliance. "Agentic AI and Zero Trust." CSA Blog, August 2025.