Verifiable Safety Assurance Evidence Graphs for Modular Humanoid Robot Arms

A vendor-neutral architecture for functional safety, cyber resilience, lifecycle assurance and use-case-specific compliance

Spherity research manuscript

Dr. Carsten Stöcker

Spherity GmbH

7 September 2026

Abstract

Modular humanoid robot arms may be manufactured before their final tools, tasks, users and evolving regulatory requirements are known. Deployment therefore depends on demonstrating that the assembled arm meets the safety, cybersecurity and compliance requirements of its intended application. For example, an arm validated for carrying sealed supplies requires additional evidence before assisting a person with eating, including evidence for contact forces, hygiene and human intervention. During operation, a drifting force sensor could cause the arm to apply more force than intended. Missing evidence can increase costs, delay or prevent deployment, while outdated evidence can conceal emerging hazards or cybersecurity risks.

Validation must therefore connect each operational claim to the current configuration, intended use and applicable requirements. If new regulatory requirements become applicable to the deployed system a year later, the proposed framework shall identify affected claims and trigger the required revalidation under the applicable transition rules. Formal, machine-readable procedures should automate repeatable evidence checks, change-impact analysis, test coordination and monitoring, while qualified parties retain responsibility for assessment. This approach aims to reduce repeated manual reviews and commissioning delays, while supporting timely restrictions when operating conditions exceed validated limits.

This paper proposes a vendor-neutral Verifiable Safety Assurance Evidence Graph (VSAEG) that links component identities, component digital product passports (DPPs), installed configurations, requirements, hazards, safety functions, test results, assumptions, assurance arguments and credential status. Fundamentally, we propose formalising individual assurance claims and supporting evidence as W3C Verifiable Credentials. By leveraging the semantic interoperability of this standard, discrete claims are explicitly bound together, establishing a cohesive, cryptographically verifiable linked evidence graph. Within this architecture, each component passport acts as an access-controlled subgraph based on open-world semantics. The primary focus of this paper is on the structural integrity and validation properties of this verifiable evidence graph, rather than the underlying access-control mechanisms.

The robot arm passport combines the passports of its installed components and adds assembly, application and runtime evidence. This composition preserves the identity, provenance, scope and status of each contributing claim. Runtime evidence is captured through short-lived verifiable credentials, ephemeral assertions, or continuous monitoring mechanisms directly coupled with dynamic credential revocation and status registries. Together, these linked claims, evidence records and assurance relationships form the arm-level VSAEG. A scoped assurance argument establishes how the combined evidence supports system-level safety.

These relationships identify which evidence supports a particular application and which claims require review after a change. A seven-degree-of-freedom arm provides the reference model, grounded by examples from the German Aerospace Center, ARMAR-6 and the G1 platform family. Open-world RDF semantics allow additional evidence and relationships to be contributed throughout the lifecycle as configurations, applications and requirements evolve. Finite validation profiles determine whether the evidence required for a specific decision is complete and acceptable. JSON-LD, W3C Verifiable Credentials and cryptographic proofs support interoperable evidence representation and secured exchange, with CBOR-LD considered an optional developing encoding.

The proposed trust architecture connects eIDAS identity and trust services, prospective and robot-resident Edge Wallets. A robot acts as an Autonomous Mandated Agent under a responsible principal. The control plane manages identities, delegated authority and access, while the data plane records and signs scoped observations and evidence packages. A formal eligibility model separates authorisation, assurance acceptance and deterministic local safety enforcement. The architecture addresses testing, evaluation, verification and validation, cybersecurity, five drift classes, credential suspension and revocation, and incremental revalidation. Industrial, service and medical applicability profiles retain distinct legal and technical boundaries.

The contribution is a conceptual architecture, an executable machine-readable evidence specimen and a reproduceable evaluation protocol. Synthetic checks establish internal consistency of selected mechanisms. Deployment safety, reduced manual effort, economic benefit and regulatory acceptance remain subjects for independent evaluation.

Keywords: functional safety; humanoid robotics; evidence graph; European Business Wallet; Edge Wallet; autonomous mandated agent; W3C verifiable credentials; digital product passport; robot passports, module passports, component passport, robot arm passport, Cyber Resilience Act; eIDAS; TEVV; runtime assurance

Contents

Nomenclature

Term	Meaning in this paper
AI	Artificial intelligence
AMA	Autonomous Mandated Agent (proposed technical role)
CAB	Conformity-assessment body
CBOR-LD	Concise Binary Object Representation for Linked Data
CRA	Cyber Resilience Act
DOF	Degree of freedom
DPP	Digital Product Passport
EAT	Entity Attestation Token
EBW	European Business Wallet
Edge Wallet	Robot-resident credential, mandate and signing service
eIDAS	European framework for electronic identification and trust services
EUDI	European Digital Identity
FMEDA	Failure Modes, Effects and Diagnostic Analysis
GDPR/MDR	General Data Protection Regulation/Medical Devices Regulation
HIL/SiL	Hardware-in-the-loop/software-in-the-loop
HSM	Hardware Security Module
JWS	JSON Web Signature
ODD	Operational Design Domain
PDP/PEP	Policy decision point/policy enforcement point
PL/PLr	Achieved/required Performance Level
QTSP	Qualified trust-service provider
RDF	Resource Description Framework
SACM	Structured Assurance Case Metamodel
SBC	Safe Brake Control
SBOM/VEX	Software Bill of Materials/Vulnerability Exploitability eXchange
SE/TPM	Secure Element/Trusted Platform Module
SHACL	Shapes Constraint Language
SIL	Safety Integrity Level
SLS/SOS/STO	Safely Limited Speed/Safe Operating Stop/Safe Torque Off
TEVV	Testing, evaluation, verification and validation
VC	Verifiable Credential
VSAEG	Verifiable Safety Assurance Evidence Graph

1 Introduction

Humanoid robot arms could serve several application areas. In manufacturing, an arm may load a machine, position a workpiece or hold a tool during assembly. In logistics, it may pick products, sort packages or replenish shelves. In hospitals and care facilities, it may deliver supplies, hand

over everyday objects or assist a person with drinking. Rehabilitation provides another potential application, involving guided movement under a defined therapeutic purpose. These tasks create different safety and assurance needs. Moving a sealed package requires evidence about object retention, collision avoidance and stopping performance. Bringing a cup towards a person's face adds concerns about contact forces, liquid temperature, hygiene and the person's ability to interrupt the movement.

Humanoid robots combine high-energy electromechanical systems, complex software, network connectivity and increasingly adaptive AI. Their arms are modular assemblies of actuators, gearboxes, brakes, encoders, torque and current sensors, control electronics, communications, firmware and mechanical structures. Each element contributes capabilities and failure modes. The safety of the complete arm depends on the installed configuration, tool, payload, task, surrounding geometry, human exposure and control strategy. A component certificate contributes evidence for application-level validation, which must also address the assembled system and its intended operating conditions.

The central problem is temporal and organisational. A manufacturer can design a robot arm for several foreseeable uses before its final application and later regulatory requirements are fully defined. An integrator adds an end effector, such as a gripper, together with the workpiece, application software, workspace and safeguards. An operator changes tasks, replaces components and updates software. A service provider may introduce a new AI model or perception system after commissioning. For example, an arm initially used to move supplies could later be assigned to assisted drinking. That change introduces new contact, hygiene and human-intervention requirements, even where much of the hardware remains the same. Industrial use, collaborative industrial operation, professional service, care and medical use engage different standards and legal duties. The assurance system must preserve reusable evidence, identify its limits and support validation before the arm enters a new application domain.

Current safety practice relies heavily on documents, supplier portals and project-specific files. These artefacts can be technically rigorous while remaining difficult to connect across organisations. Version histories, calibration state, software hashes, assumptions of use, access permissions and revocation status are often managed in separate systems. This fragmentation raises the cost of identifying which claims and tests are affected by a change. For example, replacing a torque sensor requires establishing which force estimates, control limits and application tests depended on its previous calibration. An incomplete dependency record can delay a return to service or leave an operating decision based on outdated evidence.

Cybersecurity makes the assurance problem dynamic. Vulnerabilities, compromised update channels, unauthorised configuration changes and manipulated sensor data can undermine safety assumptions. A compromised software update could change motion behaviour, while falsified sensor readings could conceal excessive force or an approaching obstacle. The Cyber Resilience Act (CRA) establishes lifecycle cybersecurity duties for products within its scope [1]. The Machinery Regulation addresses safety-relevant protection against corruption [2]. AI Act classification follows the specified product, function and conformity-assessment route; a humanoid form or an AI planner alone does not establish high-risk status [3].

This paper asks how a distributed evidence infrastructure can support this changing assurance boundary. Its research object is a Verifiable Safety Assurance Evidence Graph (VSAEG): a semantically defined, cryptographically verifiable and access-controlled graph that connects

engineering evidence to scoped assurance claims and current operational status. For example, the graph can link a brake's identity to its holding test, the tested load, its installed position and the assessment supporting its use. "Verifiable" has five dimensions. Cryptographic verification checks integrity, issuer attribution and authenticated status information. Semantic verification tests conformance with shared meanings and constraints. Engineering verification evaluates method validity, calibration, uncertainty and acceptance criteria. Assurance verification links evidence to claims and hazards. Operational verification checks that the current configuration and context remain within the authorised scope.

The contribution is an architecture for controlled evidence reuse. A common evidence base can support different assurance profiles while preserving each safety function's applicable requirements, including standards-prescribed minima and risk-derived targets. Formal validation profiles provide a basis for automating repeatable evidence checks, identifying affected claims after change and coordinating the required revalidation. Runtime monitoring connects changes in behaviour, component performance and cybersecurity status to decisions about continued, restricted or suspended operation. Qualified parties retain responsibility for engineering assessment and conformity decisions. The graph can be delivered through digital product passport (DPP) infrastructure, including a CRA-capable DPP, while evidence, assurance arguments and conformity decisions retain distinct roles.

The paper addresses five research questions:

1. Which information, provenance and lifecycle requirements arise when modular robot-arm safety evidence is generated by multiple organisations and used under initially unknown application profiles and regulatory requirements?
2. How can component, configuration, integration, use-case and runtime evidence be represented as a semantically linked and cryptographically verifiable graph?
3. How can the graph support assurance-case composition and verification/validation while preventing the inference that system safety equals the sum of component certificates?
4. How can configuration change, software update, calibration expiry and behavioural or safety-function drift trigger bounded impact analysis, credential suspension and incremental revalidation?
5. Which technical and assurance-utility measures can falsify or support the claimed benefits of the architecture?

2 Method and scope

The work adopts the problem–requirements–artefact–evaluation structure of design-science research [4]. Source selection is purposive: official legislation, published standards metadata, standards-body specifications, original research and manufacturer documentation are selected for their relevance to modular robot-arm evidence and delegated machine operation. Search concepts cover robot safety, modular and dynamic assurance, product passports, linked data, attestation, credentials and European identity services. This is a focused engineering synthesis; it makes no claim to systematic-review coverage or a preregistered literature search.

Requirements are translated into a metamodel, reference architecture and decision rules. The analytical evaluation checks source status, bibliographic metadata, formal consistency and a

synthetic evidence specimen. Public standard summaries support scope and edition statements; clause-level implementation claims require access to the applicable normative text. Section 16 specifies the subsequent demonstrator and comparative evaluation. Hardware tests, industrial outcome measurements, interviews and independent expert validation have yet to be conducted for the proposed architecture.

The system boundary includes a modular seven-axis arm, its controller, safety-related control functions, AI planning, internal sensors, external observation, evidence repositories and organisational trust services. The paper covers industrial, collaborative and selected service/care profiles. Medical use is treated as a separate boundary profile because medical-device classification and clinical risk management add distinct obligations. Locomotion and whole-body stability are outside the primary system boundary, although external hazards created by base motion remain relevant to integration.

Normative language is controlled. “Shall” denotes a sourced legal or standards requirement, or a clearly labelled architecture requirement. “Should” expresses a design recommendation. “May” expresses an option. Author proposals are separated from legal facts.

Source-derived obligations, author-defined architecture requirements and illustrative engineering values are labelled separately. An evidence-to-requirement mapping provides traceability, while the applicability decision remains a deployment responsibility. References identify editions or document versions; dynamic product pages and legislative status records were consulted on 7 September 2026.

3 Standards and regulatory landscape

3.1 Industrial robot and machinery safety

ISO 10218-1:2025 specifies safety requirements for industrial robots considered as partly completed machinery. It addresses inherent safe design, protective measures and information for use before integration [5]. ISO 10218-2:2025 addresses industrial robot applications and robot cells, including integration, commissioning, operation, maintenance and decommissioning [6]. This division is important for evidence ownership: the robot manufacturer supplies robot-level evidence and limitations; the integrator produces application-level risk assessment and validation evidence; the operator maintains configuration and operating controls. Responsibilities can overlap and remain governed by the applicable legal framework.

The industrial scope is a substantive boundary. Service and medical uses, lifting or transporting people, and hazards associated with a mobile platform require additional applicability analysis [5, 6]. For the arm model, the base is stationary or its motion is represented by a separately validated interface contract. A complete humanoid assessment must also address whole-body stability, simultaneous arm motion and fall hazards.

The safety functions commonly relevant to an arm include emergency stop, protective stop, safe monitored standstill, safety-rated monitored speed, axis and space limitation, safe torque off, safe brake control and mode selection. ISO 13849-1 and IEC 62061 provide methods for safety-related control systems. IEC 61800-5-2 addresses functional safety of adjustable-speed electrical power drive systems. ISO 12100 provides the risk-assessment and three-step risk-reduction framework. These sources support a chain from hazard to required risk reduction,

safety function, architecture, implementation, verification and validation [7–10].

For an ISO 13849 route, evidence should cover the safety function from sensor through logic to the final control element, its required performance level (PLr), architecture category, component reliability inputs, diagnostic coverage, common-cause failures, systematic measures and software validation. ISO 13849-1:2023 and ISO 13849-2:2012 provide complementary design and validation material [8, 11]. A vendor's safety manual or diagnostic library supplies inputs and assumptions; the achieved PL belongs to the implemented function. Runtime health observations monitor continuing assumptions and performance, rather than recalculate a design PL from telemetry.

Emergency-stop design, guard interlocking, protective sensing and electrical equipment also require the relevant machinery standards [12–15]. Safe brake control addresses control of the brake interface; a separate holding-performance assessment establishes whether the brake sustains the specified load. Safe torque off removes torque production and can require coordinated holding or lowering measures for gravity-loaded arms.

Collaborative operation requires application-specific assessment. ISO/TS 15066:2016 remains separately published alongside the 2025 ISO 10218 editions; ISO/AWI 15066-1 is developing a biomechanical-thresholds document [16, 17]. Contact acceptance criteria require the applicable normative source, contact geometry, body region, transient or quasi-static model, force and pressure measurements, uncertainty and intended user population. Numerical examples in this paper concern a synthetic brake test, rather than human-contact thresholds. Evidence for healthy industrial workers requires a separate justification before reuse for vulnerable care recipients.

3.2 Service, care and medical profiles

ISO 13482:2014 provides a personal-care-robot safety profile, with a successor document at FDIS stage [18, 19]. The ISO 22166 family contributes modularity and module-information concepts while retaining a separate role from product-safety assessment [20–22]. These profiles support explicit mappings between module interfaces, intended tasks and application hazards.

Intended medical purpose can activate the Medical Devices Regulation (MDR) and standards including IEC 80601-2-78:2019 with Amendment 1:2024, ISO 14971, IEC 62304, IEC 62366-1 and IEC 81001-5-1 [23–28]. IEC 80601-2-78 has a specific scope involving medical robots for rehabilitation, assessment, compensation or alleviation. Medical-device status follows intended purpose and the applicable legal definition. A "care robot" label supplies insufficient information to select that route. Products governed by the MDR or IVDR fall within the CRA's stated sectoral exclusions; their cybersecurity obligations require the corresponding medical framework [1].

3.3 AI and the safety boundary

ISO/IEC TR 5469:2024 distinguishes several relationships between AI and functional safety: AI can be part of a safety-related function, can control equipment supervised by conventional safety functions, or can support safety lifecycle activities [29]. The reference architecture in this paper prioritises the second pattern. AI proposes actions and trajectories. An independent safety controller checks bounded invariants and retains deterministic authority to permit, limit or stop motion. This separation supports clear evidence and failure containment while allowing AI performance evidence to inform operational risk controls.

For the product-related route in Article 6(1), the AI Act's high-risk classification requires both

the specified safety-component or product relationship and a third-party conformity-assessment requirement under the relevant Annex I legislation [3]. Other listed use cases require their own assessment. The graph records the classification rationale, intended purpose, AI version, evaluation evidence, monitored limits and the claims that rely on it. This avoids substituting a general statement about AI's influence on safety for the legal classification test.

3.4 Cyber Resilience Act and CRA-capable DPPs

The CRA establishes horizontal cybersecurity requirements for products with digital elements [1]. Its main requirements apply from 11 December 2027. Reporting obligations for actively exploited vulnerabilities and severe incidents apply from 11 September 2026; provisions concerning notification of conformity-assessment bodies apply from 11 June 2026 [30]. Relevant evidence includes cybersecurity risk assessment, secure-by-design decisions, vulnerability handling, software and component inventory, security updates, technical documentation, conformity assessment, incident reports and support-period information.

The applicability record shall identify exclusions, responsible economic operator, product version, support period and conformity route. CRA Article 14 distinguishes the 24-hour early warning from subsequent notifications and reports. The evidence service should preserve discovery times, affected-version hypotheses, confidence and later corrections so reporting can proceed incrementally; complete fleet attribution is a continuing investigation task [1]. Vulnerability handling and secure development can draw on the relevant IEC 62443 lifecycle and component requirements [31, 32].

A "CRA-capable DPP" is defined here as a DPP architecture able to carry, reference, protect and selectively disclose evidence needed for CRA processes. The CRA does not mandate a DPP, and possession of a DPP does not prove CRA conformity. The value lies in linking product identity and configuration to signed software bills of materials, VEX statements, vulnerability status, update evidence, risk decisions and authorised declarations. This can reduce duplication between safety and cybersecurity evidence where the same software, component or change affects both domains.

3.5 ESPR DPP, eIDAS 2.0 and European Business Wallets

The Ecodesign for Sustainable Products Regulation (ESPR) establishes a product-specific DPP framework whose detailed obligations depend on implementing delegated acts [33]. The VSAEG may reuse passport identifiers, discovery, data carriers and controlled repositories. This proposal gives "CRA-capable DPP" an architectural meaning: an infrastructure capable of delivering relevant evidence. Neither ESPR nor CRA establishes a generic robot safety passport certificate. Confidential safety, vulnerability and personal data should remain subject to purpose-specific access and retention.

Regulation (EU) 2024/1183 amended eIDAS to establish the European Digital Identity Framework [34]. Under the consolidated eIDAS Regulation, qualified electronic signatures, seals and timestamps have distinct legal effects: a qualified signature has the equivalent legal effect of a handwritten signature; a qualified seal benefits from presumptions of data integrity and correctness of origin; and a qualified timestamp benefits from presumptions concerning time accuracy and bound-data integrity [35]. These effects support cross-organisational accountability while engineering adequacy remains an assessment question.

The Commission's European Business Wallet proposal, COM(2025) 838, remains in the legislative process [36, 37]. The proposed architecture therefore separates existing eIDAS services from prospective Business Wallet capabilities. A wallet capability interface can connect identity, representation, credentials, signing, sealing and presentation without presuming adoption of a particular future certification or protocol profile.

Section 9 extends this interface to a subordinate robot-resident Edge Wallet. The organisation provisions credentials, mandates and policies to device-held keys. In a private-care setting, the responsible principal may instead be a natural person, with a service provider holding a separate maintenance mandate. The common technical model preserves these different legal relationships.

Table 1: *Regulatory and standards applicability to the proposed evidence architecture.*

Regime or standard	**Primary object**	**Evidence-graph relevance**	**Important boundary**
ISO 10218-1:2025	Industrial robot	Robot-level safety functions, limits and information for integration	Industrial robot scope
ISO 10218-2:2025	Application and cell	Application risk assessment, safeguards and validation	Configured application determines safety
ISO 13849/IEC 62061	Safety-related controls	Target, architecture, diagnostic evidence and validation	Each safety function needs its own target
ISO 13482/IEC 80601-2-78	Service/care or medical robot	Profile-specific hazards, tests and intended purpose	Medical purpose activates a separate route
EU Machinery Regulation	Machinery product and integration	Essential requirements, technical file and conformity evidence	Applies from 20 January 2027 with stated exceptions
EU AI Act	AI system or safety component	AI version, role, risk controls, logs and monitoring	High-risk status is function and route dependent
CRA	Product with digital elements	SBOM, VEX, vulnerability, update, incident and support evidence	Scope and exclusions apply; MDR/IVDR products have a separate cybersecurity route
ESPR DPP	Product information framework	Identifiers, data carriers, discovery and access	Does not constitute safety or CRA conformity

Table 1: *Regulatory and standards applicability to the proposed evidence architecture. (continued)*

Regime or standard	Primary object	Evidence-graph relevance	Important boundary
eIDAS 2.0	Identity and trust services	Signatures, seals, timestamps and attributes	Authenticity is distinct from technical validity
Proposed European Business Wallet	Economic operator and public body	Organisation, mandate, controlled exchange and signing/sealing	Legislative proposal; Edge Wallet is an author-defined deployment profile

4 Related work and research gap

The proposed architecture intersects four bodies of work that are usually developed separately. First, functional-safety engineering supplies lifecycle processes, integrity targets, diagnostic concepts and validation methods. Its evidence is often organised around a product, project or responsible organisation. Second, modular assurance cases allow claims and evidence fragments to be composed, challenged and revised. SACM provides a metamodel for claims, arguments, artefacts and relationships, but it does not prescribe robot component identity, hardware attestation, credential status or application-profile semantics [38]. Third, digital-twin and product-passport research provides persistent product identity and lifecycle data sharing. Existing DPP studies concentrate mainly on sustainability, circularity and supply-chain information. Safety evidence adds hazards, safety functions, uncertainty, configuration scope, assessor authority and reliance status [39–41]. Fourth, knowledge-graph and verifiable-credential technologies provide semantic linking and signed exchange, but do not determine whether a safety argument is adequate.

Modular and dynamic safety cases provide established approaches to reusable arguments and through-life change [42, 43]. Human–robot interaction research supplies complementary methods for motion constraints, contact and human exposure; these methods address physical risk rather than organisational identity or evidence portability [44, 45]. Elder-care robotics research also shows why intended users and task context need explicit treatment [46]. The present architecture connects those concerns to attributable component evidence and current reliance decisions.

The proposed contribution is the integration of configuration-bound evidence, structured arguments, issuer authority, delegated robot mandates and lifecycle status within a common decision model. Existing standards and studies provide constituent mechanisms. This paper proposes interfaces among them and tests selected internal consistency properties; it does not claim priority for knowledge graphs, dynamic safety cases, runtime supervision or verifiable credentials individually.

The paper advances six linked contributions:

1. a vendor-neutral reference model for a seven-axis modular robot arm and its assurance boundary;
2. a typed evidence-graph metamodel that joins safety, cyber, AI and lifecycle evidence without

collapsing their claim types;

3. an open-world semantic model with closed-world decision profiles;
4. a hardware-rooted and eIDAS-connected trust model for physical and organisational attribution;
5. a TEVV, credential-status and incremental-revalidation lifecycle; and
6. an evaluation protocol that separates technical feasibility from assurance utility.

The evaluation distinguishes four comparison dimensions: assurance-case methods organise justification; DPP methods organise product information; runtime-supervision methods constrain behaviour; and identity/credential methods authenticate statements and authority. The VSAEG's added requirement is a traceable path across all four: a current, scoped decision identifies the physical configuration, supporting evidence, responsible issuer and robot mandate. Section 16 describes how that integration can be tested against document-based baselines.

5 Derived architecture requirements

The review yields the following requirements for the proposed artefact. They are design requirements of this paper rather than direct quotations from a single standard. Each deployment must trace them to its applicable legal and normative sources.

Table 2: *Derived architecture requirements and verification focus.*

ID	Architecture requirement	Rationale and verification focus
AR-01	Every safety-relevant physical, software and evidence object shall have a unique, durable and resolvable identifier.	Prevents ambiguity across suppliers and service life; test by duplicate and replacement scenarios.
AR-02	The architecture shall record the method and assurance level of the physical–digital binding.	A serial number, key and mechanical part can diverge; verify challenge-response and tamper/replacement cases.
AR-03	Claims, evidence, arguments, assessments and credentials shall remain distinct object types.	Prevents a signed result from being treated as a certificate or an assurance decision.
AR-04	Every operational assurance credential shall be bound to a canonical configuration baseline.	Safety depends on installed hardware, firmware, software, model, parameters, calibration and tool.
AR-05	Every operational claim shall state its intended use, ODD, jurisdiction, assumptions, limits and validity interval.	Supports unknown future applications without uncontrolled claim extension.
AR-06	The graph shall support current status, suspension, revocation, expiry, supersession and historical audit.	Reliance changes after faults, compromise, change or withdrawal while evidence history remains immutable.

Table 2: *Derived architecture requirements and verification focus. (continued)*

ID	Architecture requirement	Rationale and verification focus
AR-07	Immediate motion-safety enforcement shall remain bounded, local and deterministic unless a safety-rated distributed architecture is demonstrated.	Networked AI and infrastructure paths have variable delay, availability and attack surface.
AR-08	Safety and cybersecurity dependencies shall share configuration and change objects while retaining separate claims and acceptance criteria.	A cyber change can invalidate safety assumptions; disciplinary objectives remain distinct.
AR-09	Test evidence shall identify method, subject, configuration, stimulus, measurement, uncertainty, limits, calibration, executor, observer, trace and status.	A pass/fail value alone is not reviewable evidence.
AR-10	Ontologies, shapes, credential schemas and policy profiles shall be versioned and governed.	Meaning can drift even when bytes and signatures remain unchanged.
AR-11	Access shall follow role, purpose, least disclosure and need-to-know principles.	FMEDA, vulnerability, personal and operational data require controlled sharing.
AR-12	The architecture shall state epistemic limits and distinguish unknown, absent, invalid and false information.	Open-world graphs must not convert missing data into a safety conclusion.
AR-13	Every issued or assessed claim shall identify the responsible principal, acting subject, role, mandate, competence scope and time.	Distinguishes legal/natural-person authority, robot evidence production and qualified assessment.
AR-14	Monitored change and drift shall propagate to affected claims and produce a bounded operational and revalidation response.	A monitoring alert without a change in reliance is incomplete assurance control.
AR-15	The Edge Wallet shall provision robot/workload credentials to protected device-held keys and isolate organisational seal, update and evidence-signing authority.	Sections 9.3–9.7; verify enrolment, key substitution, signing API restrictions and transfer.
AR-16	Child mandates shall attenuate parent authority and bind action, resource, purpose, audience, configuration, time and status.	Sections 7.5 and 9.5; reject expanded rights, wrong audiences, stale leases and revoked ancestors.

Table 2: *Derived architecture requirements and verification focus. (continued)*

ID	Architecture requirement	Rationale and verification focus
AR-17	Control-plane eligibility and data-plane evidence signing shall remain separate from deterministic local motion enforcement.	Sections 9 and 11; test network loss, resource exhaustion, logging backlog and supervisory transitions.
AR-18	A secured evidence object shall identify the exact signed bytes or canonical dataset, pinned semantic dependencies and included third-party objects.	Sections 8.1 and 9.7; test context substitution, reserialisation, selective disclosure and digest mismatch.
AR-19	Runtime evidence shall bind source, configuration, mandate, decision, sequence, time uncertainty and claim type to a bounded signing operation.	Section 9.6; reject replay, cross-session substitution and attempts to convert observations into assessments.
AR-20	Validation profiles shall distinguish technical device signatures, organisational authority and the qualification of eIDAS trust-service operations.	Sections 3.5 and 9.3; reject unsupported qualified-seal claims and unauthorised legal declarations.

These requirements create two complementary modes. The evidence space is extensible and distributed: new suppliers and evidence types can add facts without redesigning a central database. The reliance decision is constrained: a named profile defines which facts must be present, current, authorised and mutually consistent before operation is permitted.

6 Reference model of the modular robot arm

The reference system has seven revolute degrees of freedom: three shoulder axes, elbow flexion, forearm rotation and two wrist axes. J1–J7 denote the controlled kinematic chain; the exact axis convention belongs to the configuration manifest. A torso or base provides the mounting frame. Seven axes are a modelling choice, allowing redundancy for many six-dimensional end-effector tasks; other humanoid arms require a different joint count and kinematic model.

Each joint module contains a motor, reduction gear, brake where required by the risk assessment, power stage, gate driver, current sensing, temperature sensing and at least one position channel. Safety-related functions can require diverse or redundant position sensing and independent plausibility checks. Joint torque can be measured directly, inferred from motor current and transmission models, or estimated through observers. Each method has different uncertainty and fault assumptions that must be visible in the evidence graph.

The arm includes a motion computer, a safety controller with demonstrated independence and freedom from interference, communication interfaces, power supervision, secure update capability and hardware identity anchors. The planner generates candidate trajectories. Safety-related sensing and control enforce the validated envelope, including appropriate stop and holding functions. A robot-resident Edge Wallet serves identity, mandate verification and evidence

signing through isolated services. The wallet sits outside the motion-control feedback loop; its authority changes affect operation through validated supervisory transitions.

The model distinguishes five configuration layers:

1. **Component configuration:** serialised hardware, firmware, calibration and manufacturing evidence.
2. **Joint configuration:** motor, transmission, brake, sensor geometry, control parameters and thermal limits.
3. **Arm configuration:** kinematic chain, structural parameters, controller versions, safety functions and tool interface.
4. **Application configuration:** end effector, payload, task, workspace, safeguards, network and human interaction.
5. **Operational configuration:** active recipe, AI model, policy set, environmental state and current status.

Each layer has a canonical configuration identifier computed from ordered, typed references rather than a simple concatenation of files. The identifier binds the evidence to the installed configuration. Physical binding uses device-held keys, secure boot, measured boot, signed component manifests, authenticated installation records and challenge-response checks. A secure element can protect a component key; a TPM-class device can record platform measurements and produce quotes; an embedded hardware security module can isolate keys and cryptographic services within a safety or control processor. An enterprise HSM protects issuer or service keys. Selection depends on threat model, update architecture, physical constraints, assurance needs and lifecycle management.

Table 3 maps reference modules to representative evidence. Implementation examples in Sections 6 and 15 instantiate functional roles; safety obligations remain attached to the function and installed configuration.

Table 3: *Reference robot-arm modules, safety concerns and representative evidence.*

Module or interface	Principal safety concerns	Representative evidence objects
Joint structure and bearings	Fracture, fatigue, excessive play, pinch and shear	Material and process records; proof load; fatigue basis; dimensional inspection; backlash and wear trend
Motor and transmission	Excess torque, runaway, jamming, overheating, efficiency change	Torque constant; thermal model; gear ratio; mechanical limits; insulation test; friction and efficiency characterisation

Table 3: *Reference robot-arm modules, safety concerns and representative evidence. (continued)*

Module or interface	Principal safety concerns	Representative evidence objects
Brake and holding path	Failure to hold, delayed release, unintended engagement	Holding-torque test; release time; wear estimate; feedback diagnostics; safe brake control validation
Position and velocity sensing	Bias, loss, common cause, aliasing, incorrect direction	Resolution and accuracy; diverse channel comparison; mounting geometry; calibration; diagnostic coverage; fault injection
Torque, current and force sensing	Force-limit error, drift, saturation, observer mismatch	Calibration; bandwidth; uncertainty; current–torque relation; cross-check; saturation and fault response
Power stage and gate drive	Uncontrolled energisation, short circuit, latent gate fault	Safe-state concept; gate diagnostics; desaturation response; power-cycle and fault-injection results
Safety controller	Incorrect limit, timing overrun, common cause, unauthorised change	Safety manual; FMEDA inputs; watchdog and memory tests; software baseline; timing and independence evidence
AI and motion computer	Unsafe proposal, perception error, model change, resource exhaustion	Model and software identity; scenario coverage; confidence limits; timing; update evidence; runtime-monitoring plan
Secure element, TPM or HSM	Key compromise, false identity, rollback, untrusted boot	Provisioning record; certificate chain; measured boot; key policy; attestation result; rotation and compromise procedure
End-effector and payload interface	Sharp/hot tool, drop, unexpected mass, altered inertia	Tool identity; retention test; geometry; mass and centre of gravity; hazard-specific acceptance criteria
External observation interface	Blind spot, stale or forged message, time-base error	Sensor calibration; coverage map; latency distribution; authentication; replay protection; degraded-mode validation

Let $q \in \mathbb{R}^n$ be joint coordinates with n = 7 for the reference arm, and let C identify a fixed installed baseline. The end-effector transform $T_E = f(q; C) \in \mathrm{SE}(3)$ describes position and orientation relative to the mounting frame. A stationary-base joint-space model is

$$\mathbf{M}(\mathbf{q}; C)\,\ddot{\mathbf{q}} + \mathbf{C}_v(\mathbf{q}, \dot{\mathbf{q}}; C)\,\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}; C) + \boldsymbol{\tau}_f(\mathbf{q}, \dot{\mathbf{q}}; C) = \boldsymbol{\tau}_a + \boldsymbol{\tau}_{\mathrm{ext}}. \tag{1}$$

Here M is the joint inertia matrix, $C_v\dot{q}$ contains Coriolis and centrifugal terms, g is gravity torque, τ_f is friction torque, τ_a is actuator torque referred to the joint output, and τ_{ext} is external generalised torque. Converting motor current to τ_a requires the torque constant, transmission ratio, efficiency and uncertainty model. A free-moving base requires additional coordinates and coupling terms. The simplified equation supports configuration traceability; safe control requires validated models, sensing and bounds.

The model is evidence-bearing rather than purely computational. Link masses, inertias, gear ratios, torque constants, friction parameters, sensor transforms, brake characteristics and uncertainty bounds are identified by source and validity interval. A change to any parameter used by a safety function creates a dependency edge to that function and its validation evidence. The graph therefore distinguishes a model parameter, an estimate of that parameter, a test that produced the estimate and the configuration in which it is valid.

6.1 Three concrete platform anchors

The German Aerospace Center (DLR, Rollin' Justin), uses two seven-degree-of-freedom arms and provides a direct reference for the chosen joint count [47]. Consider a proposed household-object handover with a stationary base. Replacing a joint torque sensor preserves some link-geometry evidence but changes calibration, torque-estimation and force-limiting dependencies. The proposed graph links the replacement record to the affected tests and handover profile. This is an author-defined assurance scenario, separate from DLR's reported robot capabilities.

ARMAR-6 of Karlsruhe Institute of Technology (KIT) uses eight joints per arm and integrated sensor–actuator–controller modules for collaborative maintenance tasks [48]. A tool-support scenario illustrates why module identity alone is insufficient: adding a shoulder joint or exchanging the hand adapter changes kinematics, reachable space and payload assumptions. Its model uses n = 8, while the evidence schema remains reusable. An enterprise mandate can authorise tool retrieval and evidence recording while reserving machinery restart and safety-parameter changes for a responsible operator.

Unitree's G1 product information distinguishes five arm degrees of freedom from optional additional wrist degrees of freedom and hand options [49]. That distinction is useful for configuration assurance: a catalogue family name cannot identify the installed kinematic chain, hand, software or permitted payload. A proposed laboratory transfer of a lightweight object would bind evidence to the actual variant and tool. The scenario establishes a documentation requirement, rather than a claim that the platform has a certified collaborative or medical operating mode.

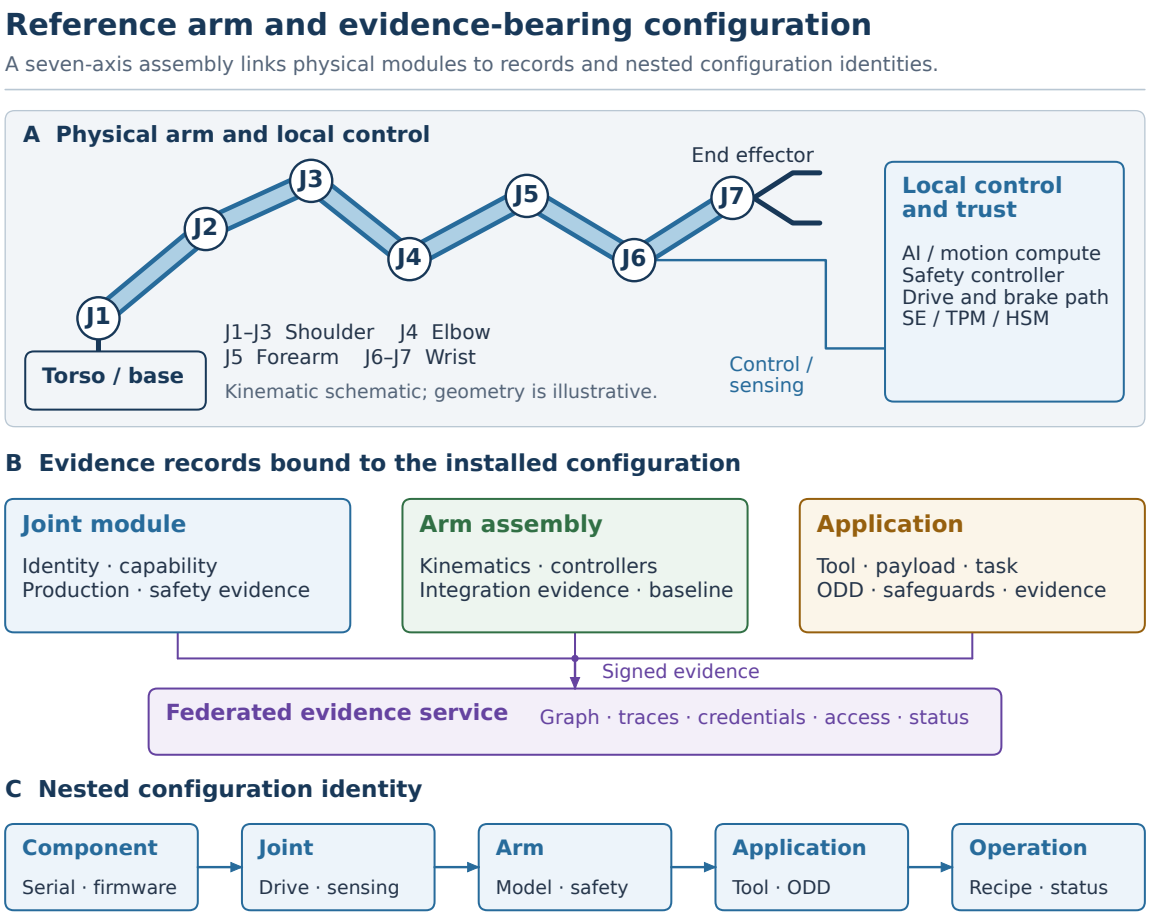

Figure 1: *Reference arm and evidence-bearing configuration. A seven-axis arm links physical modules, evidence records and five nested configuration layers; operational reliance requires binding to the installed arm, application and current operating profile.*

7 Conceptual foundations and formal assurance model

7.1 Claim, evidence, argument and decision

A claim states a proposition, such as "the specified monitored-speed function achieves its required PL for configuration C under the stated assumptions." Evidence supports or challenges that proposition. An argument explains the relevance and sufficiency of selected evidence. An assessment records an authorised reviewer's conclusion, scope, rationale and unresolved conditions. A runtime policy decision applies that accepted scope to a current request. These distinct graph objects preserve the difference between test execution, professional judgement and operational control.

Evidence quality has several attributes: method identity and version; test subject; configuration; stimulus; measured value; uncertainty; acceptance limit; result; equipment and calibration; executor competence; observer or reviewer; time; environment; raw-trace location; cryptographic hash; signature or seal; and status. High-rate telemetry normally belongs in a time-series or content-addressed repository. The graph stores signed summaries, events, statistics, locations and hashes.

7.2 Five dimensions of verifiability

Cryptographic verifiability answers whether bytes, issuer and status can be authenticated. **Semantic verifiability** answers whether terms, units, relations and constraints conform to the chosen profile. **Engineering verifiability** addresses measurement validity, uncertainty, coverage and acceptance criteria. **Assurance verifiability** addresses the structured link from evidence through argument to hazard and claim. **Operational verifiability** addresses whether the current system remains the validated configuration in the authorised domain.

A credential can be cryptographically valid while its calibration has expired. It can be semantically valid while the issuer lacked the required authority. It can remain technically correct for a component while the application has changed. The verifier therefore evaluates all five dimensions.

7.3 Functional-safety concepts and composition rule

Component evidence is necessary for modular assurance but system safety is non-additive. Interactions introduce hazards, common-cause failures and performance limits. The graph shall prevent a rule of the form "all components certified implies system safe." Valid composition requires an explicit integration claim, defined interfaces, assumptions of use, configuration identity, hazard analysis, independence analysis where applicable, and application-level validation.

The model uses safety terms precisely. A **functional capability** describes what a component can do. A **safety-related capability** is a capability relevant to risk reduction. A **safety mechanism** detects, controls or mitigates a fault. A **diagnostic** provides evidence about faults or health. A **safety function** is a specified system response that contributes required risk reduction. An **integrity claim** states the required and achieved dependability of that function. An **assumption of use** states a condition that the downstream design must satisfy. These objects are related, yet they are not interchangeable.

The assurance domain also includes more than random hardware failure. The graph distinguishes functional-safety evidence, intended-behaviour evidence, mechanical and structural safety, elec-

trical and thermal safety, cybersecurity, human factors, privacy and, where applicable, medical safety and essential performance. A shared hazard may depend on evidence from several domains. Each domain retains its own method and decision authority.

7.4 Formal model for safety and cyber assurance

The formal model defines policy eligibility over an explicit evidence snapshot. Let $D_t = \{(g_i, G_i, \pi_i)\}$ be a finite RDF dataset of named evidence graphs G_i, graph identifiers g_i and provenance records π_i at decision time t. Component, integration, application, runtime and cyber evidence are separate graph families. Their union supplies a query view, while named-graph provenance, issuer proofs and status remain available [50].

$$G_{\text{view}}(t) = \bigcup_{(g_i, G_i, \pi_i) \in D_t} G_i. \tag{2}$$

Let P identify a versioned validation profile, C_t the installed configuration, U the intended use, O the operational design domain and K a scoped assurance credential. An assessment A links an explicit argument, hazards and assumptions to a defined snapshot. $\text{Eval}_P(\varphi, D_t)$ takes one of four decision values: true, false, unknown or conflict. This is a policy evaluation convention over evidence, rather than a replacement for RDF semantics. Only true satisfies a required condition.

$$\text{Assured}_P(K, C_t, U, O, t) \coloneqq \forall \varphi \in \text{Req}_P(K, C_t, U, O, t),\ \text{Eval}_P(\varphi, D_t) = \text{true}. \tag{3}$$

Req_P includes accepted-assessment scope, configuration match, evidence sufficiency, issuer competence and authority, current credential status, applicable assumptions, evidence freshness, monitored safety-function health and relevant cyber conditions. An argument is an identified, reviewable object; the model avoids treating an undefined Arg symbol as a self-validating Boolean theorem.

Assured_P denotes acceptance under the chosen policy, not proof that physical harm is impossible. Monitored health checks test continuing assumptions and performance; achieved PL or SIL remains supported by the design and validation case. A missing, inaccessible or contradictory prerequisite prevents the current permit. A separate validated policy maps that condition to a restricted envelope, hold, controlled stop, emergency lowering or another hazard-appropriate response. Sudden torque removal can be hazardous for suspended loads and person support.

Define x dependsOn y to mean that claim, assessment or decision x relies on prerequisite y. For changed objects Δ, the initial impact set follows the inverse direction of those edges:

$$A_\Delta = \Delta \cup \left\{ x \mid \exists d \in \Delta :\ x \xrightarrow{\text{dependsOn}^+} d \right\}. \tag{4}$$

The plus denotes a path of one or more dependency edges from a dependent to a changed prerequisite. Equivalently, an implementation traverses incoming dependsOn edges starting at Δ. It then adds affected interface guarantees, shared resources and common-cause assumptions under P. Missing dependencies can make the result incomplete, so bounded revalidation requires explicit review of the unchanged boundary. For example, a changed brake calibration reaches the

brake test, holding claim and operational decision, while an unrelated cosmetic record remains outside that closure.

7.5 Delegated authority and motion permission

A mandate M identifies a responsible principal, a robot/workload subject key, permitted actions, resources, purpose, audience, spatial and operational limits, validity interval, delegation rule and status. For a child mandate M_c and parent M_p, effective rights are the intersection of inherited rights and the child's restrictions. Rights denote admissible action–resource–purpose–context–time tuples. A child may narrow time, purpose, resources and actions; expansion requires new authority from the principal.

$$\mathrm{Rights}_{\mathrm{eff}}(M_c) = \mathrm{Rights}_{\mathrm{eff}}(M_p) \cap \mathrm{Constraints}(M_c). \tag{5}$$

Authorised(a,M,t) requires a valid identity and delegation chain, proof of possession, current status and membership of action a in $\mathrm{Rights}_{\mathrm{eff}}(M)$. A runtime execution request must satisfy the following necessary condition:

$$\begin{aligned} \mathrm{Execute}(a,t) \Rightarrow\ & \mathrm{Authorised}(a, M, t) \\ & \wedge\, \mathrm{Assured}_P(K, C_t, U, O, t) \\ & \wedge\, \mathrm{LocalGuard}(a, C_t, t). \end{aligned} \tag{6}$$

LocalGuard is the validated safety-control decision using current local measurements and limits. The implication expresses necessary conditions; a supervisory permit never obliges motion. To control the interval between verification and use, a short-lived decision binds the action class, configuration digest, mandate, policy version, sequence counter and expiry. The local enforcement point rejects a stale or mismatched decision. Graph queries and remote status retrieval remain outside the deterministic control cycle.

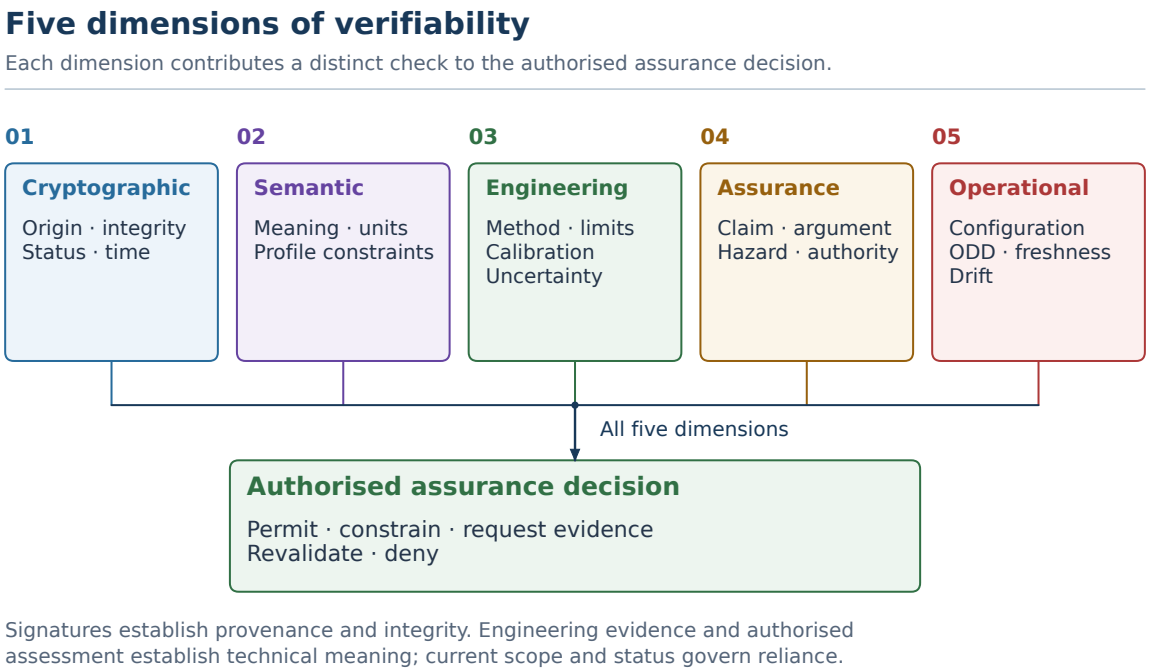

Figure 2: *Five dimensions of verifiability. Cryptographic validity is necessary and remains insufficient without semantic, engineering, assurance and operational validity.*

8 Semantic evidence graph and assurance-case connection

8.1 RDF, JSON-LD and CBOR-LD representations

RDF provides a graph data model and an associated semantic framework. JSON-LD supplies a JSON representation whose contexts map terms to IRIs; it can represent RDF statements without prescribing one database [50, 51]. CBOR-LD compresses linked data using CBOR, the Concise Binary Object Representation. Its specification is a W3C Working Draft, so use requires a pinned version, codec and context/dictionary profile [52]. These formats describe and transport data. Cryptographic proofs, trusted identifiers, provenance and verification policies supply verifiability.

These technologies enable verifiable evidence **graphs** because a signed evidence object can contain or reference globally identified entities and relations rather than an isolated document. A brake test can identify the physical brake, installed joint, arm configuration, method, calibration, executor, observer, hazard and claim. The verifier can follow these links across separately issued credentials while checking each proof and status.

RDF/OWL's open-world interpretation allows additional facts to exist beyond the available graph [53]. Absence of a disclosed calibration record therefore leaves calibration status unknown. This supports modular supply chains: a supplier publishes component facts before an integrator knows the future application; later actors add tool, user, jurisdiction and context statements without rewriting the supplier's signed claims. The graph name alone proves neither publisher identity nor completeness. Inference rules, provenance and access policies must be governed explicitly.

An operational decision evaluates a bounded snapshot and a named validation profile. SHACL can check datatypes, cardinalities, class relations and required fields [54]. Temporal, cryptographic and legal-authority checks require additional validators or explicitly specified extensions. Counting two sensor records does not establish sensor independence; that requires engineering evidence. An absent or undisclosed required object leaves the decision incomplete. Closed-world validation applies to this finite decision scope while the wider evidence model remains open to extension.

Signing also requires an exact boundary. A JOSE-secured VC signs the selected payload bytes; changes in whitespace or serialisation produce different bytes. An RDF-based Data Integrity suite uses its specified canonicalisation and transformation rules [55–57]. JSON-LD-to-CBOR-LD conversion preserves an existing proof only when the selected profile defines and validates that preservation. Contexts and dictionaries should be version-pinned and integrity-checked, with an allowlisted loader to prevent semantic substitution and uncontrolled network retrieval. A signed finite graph manifest can reference independently signed evidence without changing the original issuers' claims.

8.2 Core metamodel

The core metamodel contains Actor, Principal, RobotAgent, EdgeWallet, Role, Mandate, Product, Component, Assembly, Configuration, SoftwareRelease, AIModel, OperatingDomain, IntendedUse, Hazard, RiskEstimate, SafetyFunction, SafetyMechanism, Requirement, Assumption, TestMethod, Calibration, EvidenceRecord, EvidenceBatch, Claim, Argument, Assessment, Credential, StatusEvent, Vulnerability, Update, Incident and PolicyDecision.

Important relations include: *contains*, *installedIn*, *hasConfiguration*, *implements*, *mitigates*, *constrainedBy*, *requires*, *assumes*, *testedBy*, *produced*, *observedBy*, *supports*, *challenges*, *issuedBy*, *authorisedBy*, *validFor*, *supersedes*, *suspendedBy*, *revokedBy* and *affectedByChange*. Relations carry provenance and validity intervals where appropriate.

The metamodel uses PROV-O for entities, activities and agents, VC Data Model 2.0 for portable issuer statements, and SACM 2.3 for structured arguments and assessments [38, 58, 59]. RATS separates attestation evidence, appraisal and relying-party decisions; EAT provides a claims format for attested characteristics [60, 61]. These standards contribute compatible roles and representations, while application-specific safety and mandate profiles define their combined use.

Example competency questions include:

- Which installed components and software releases formed the arm configuration during test T?
- Which hazards and assumptions are addressed by evidence package E?
- Is the issuer authorised for this claim type, product class, jurisdiction and time interval?
- Which safety claims depend on a calibration expiring within 30 days?
- Which active credentials depend on a vulnerable software component?
- Does the current tool, payload and operating domain remain within the validated scope?

- Which delta tests are required after replacement of encoder channel B?
- Which evidence supports the decision to continue, degrade or stop operation after a drift event?
- Which principal and mandate authorised the robot workload to sign observation E?
- Which decisions remain eligible after a parent mandate is suspended or a device key is compromised?

The graph maintains immutable historical evidence and append-only status events. Correction creates a new version linked by *supersedes*. Suspension marks a credential temporarily unacceptable for new reliance. Revocation ends acceptance under the governing policy. Historical auditability remains available subject to access rights and retention duties.

8.3 Evidence, assurance and credential layers

The model separates three related layers. The **evidence layer** contains measured or observed artefacts and provenance. The **assurance layer** contains hazards, requirements, claims, arguments, counterclaims, assumptions and assessments. The **credential layer** contains portable issuer statements, cryptographic proofs, validity intervals and status references. An EvidenceRecord can support or challenge one or more Claims. An Argument uses those relations to justify an Assessment. An authorised issuer may then create a scoped Credential that materialises the assessment outcome. Credential validity never bypasses the argument.

The graph supports negative and contradictory evidence. A later test can challenge a prior claim without deleting the original record. Conflicting observations retain their sources, times, methods and uncertainties. A policy defines whether the conflict blocks operation, triggers a conservative envelope or requires an assessor decision. This is especially important for external perception and drift monitoring, where uncertainty and disagreement are expected.

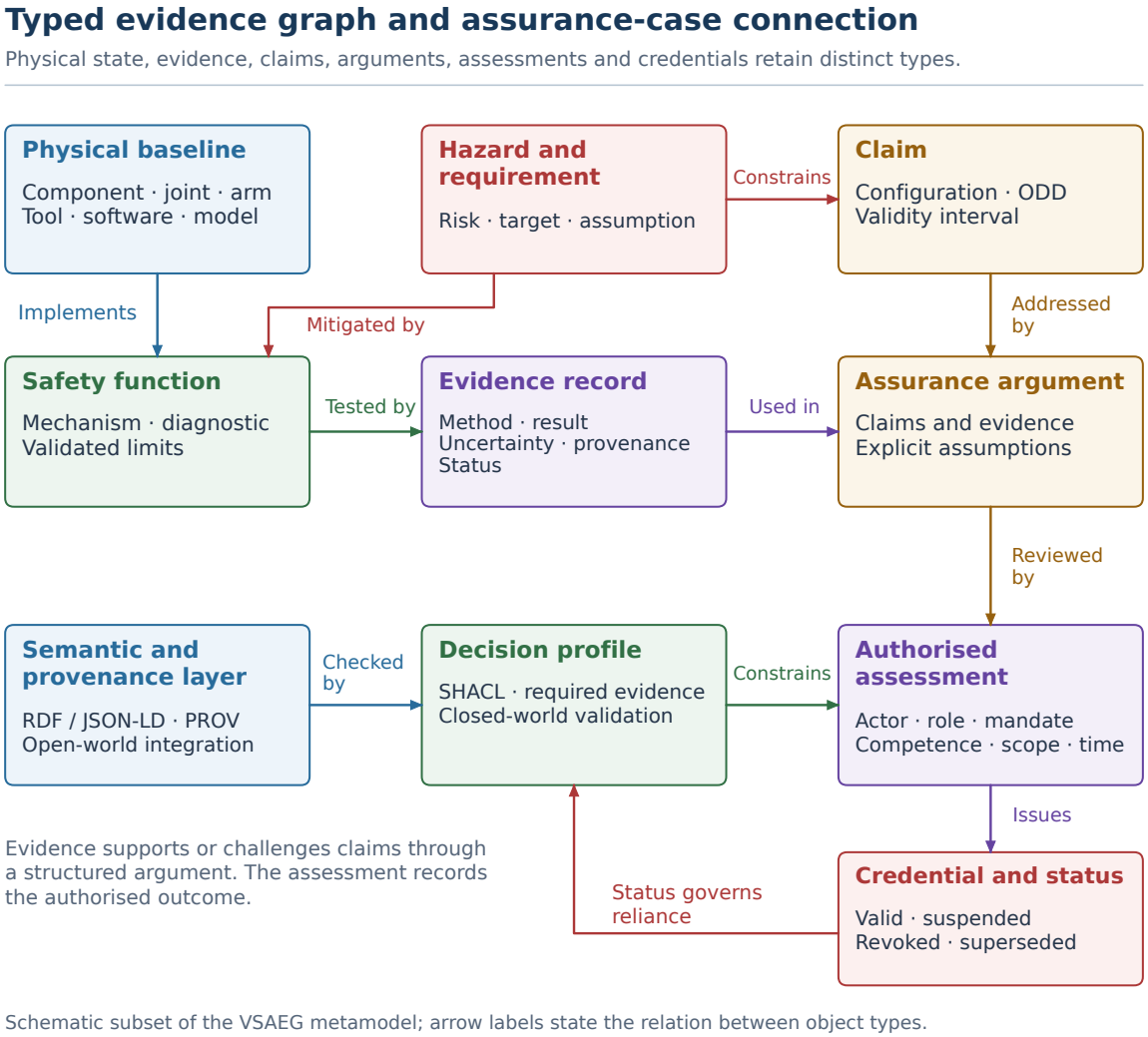

Figure 3: *Typed evidence graph and assurance-case connection. Physical state, evidence, claims, arguments, authorised assessment and credential status remain distinct objects linked by explicit dependencies.*

9 Trust, identity and access-control architecture

9.1 Threat model and trust roles

The threat model includes component substitution, evidence forgery, replay, rollback, issuer-key compromise, false authority, stale status, unauthorised disclosure and denial of service. An autonomous planner also creates a confused-deputy risk: malicious instructions or environmental content can attempt to invoke a privileged signing or actuation service. The architecture assumes attackers can control ordinary network traffic and some non-safety workloads. Physical binding, trusted sensing, approved firmware and issuer competence are explicit dependencies, each requiring separate evidence.

The following four roles separate hardware identity, platform state, controller-local signing and organisational issuance. They can be combined only when the implementation demonstrates the required isolation and lifecycle controls [62].

1. A **secure element** or protected device identity stores a component key and supports challenge-response with a small attack surface.
2. A **TPM-class platform root** measures boot state, protects keys and produces attestation quotes tied to freshness challenges.
3. An **embedded controller HSM** isolates safety-controller keys, secure-boot material, counters and signed controller-local evidence.
4. An **enterprise HSM or qualified trust-service system** protects organisational issuer, seal, timestamp and status-service keys.

Key separation is mandatory at the design level. Device identity, software signing, evidence signing, remote administration and organisational sealing use distinct keys and policies. Credentials record algorithms, key identifiers and verification methods. Long service life requires rotation, revocation, compromise recovery, archival validation and cryptographic agility.

9.2 Cyber-physical identity binding

A product identifier identifies the described object; a cryptographic device identity proves control of a key associated with hardware. Binding requires a controlled process. At manufacturing or onboarding, an issuer verifies the device, provisions or reads a protected key, records relevant measurements and creates a signed binding credential. At installation, an authorised integrator records parent assembly, slot, component identity, configuration and time. Runtime challenge-response and attestation can demonstrate continued key possession and approved platform state.

The binding process has ten controlled steps:

1. the device key is generated in protected hardware or securely provisioned;
2. a manufacturer or approved onboarding authority issues a device identity credential linking the public key to product identity and production context;

3. a durable DataMatrix, NFC carrier or protected electronic identifier links field service to the same subject;
4. the integrator performs a fresh challenge-response and verifies the manufacturer chain;
5. an installation credential links component identity, arm identity, joint or slot, orientation, interface and time;
6. a canonical arm configuration manifest binds all installed components, firmware, software, AI model, safety parameters and calibrations;
7. measured boot or secure boot evidence establishes the approved executable baseline;
8. TEVV records include the configuration identifier, method, calibration and trace hash;
9. service events update installation, firmware, calibration and status through authorised transactions; and
10. compromise, replacement or withdrawal creates a status event and triggers dependency-based impact analysis.

The binding remains probabilistic in the broader engineering sense. A protected key can authenticate electronics while a mechanical component has been replaced. Tamper evidence, secure assembly procedures, physical unclonable features, serialisation, installation inspection and sensor consistency checks strengthen the cyber-physical link. The evidence graph records the binding method, assurance level, inspector and residual risk rather than using a binary "trusted" label.

9.3 Organisational authority and eIDAS 2.0

Evidence must be attributed to organisations and authorised persons or systems. A manufacturer can issue production and design evidence. A calibration laboratory can issue calibration results within its scope. An integrator can issue installation and application-validation records. A conformity-assessment body can issue only the statements it is legally entitled to issue. The operator can report runtime events and maintenance. Wallet-held role and mandate credentials allow a verifier to evaluate who acted for which legal person, under which authority and at what time.

eIDAS matters because relying parties need accountable origin, recognised identity and appropriate evidence of time across organisational boundaries. Under Articles 25, 35 and 41, qualified signatures, seals and timestamps have different legal effects [35]. An electronic-signature signatory is a natural person; a seal creator is a legal person. The robot is a technical subject acting for a principal. Its device signature can authenticate evidence origin within the trust framework, while qualification as an eIDAS service or creation device requires the applicable regulated conditions. A VC signature or a Business Wallet label alone establishes neither those conditions nor a right to make a safety or CRA declaration.

A prospective Business Wallet can orchestrate organisational identity, representation credentials, mandates, signing, sealing and presentation through controlled interfaces [36]. The robot's Edge Wallet is a proposed subordinate implementation profile, rather than an already established legal category within that proposal. The organisation retains responsibility for permitted use of its authority and the qualified trust-service provider retains its own regulated role.

9.4 Provisioning an Autonomous Mandated Agent

An Autonomous Mandated Agent is defined here as a technical agent that selects and executes tasks within a verifiable delegation from a responsible principal. “Humanoid Employee” describes an enterprise deployment metaphor; “Citizen Companion” describes a private assistance metaphor. Neither term assigns employment, citizenship, legal personality or independent consent capacity to the machine. Manufacturer, integrator, operator and service-provider duties continue to follow their applicable roles. Whether a robot-mediated transaction binds a principal follows the applicable law and contractual arrangements; a mandate credential records authority evidence rather than creating a general legal agency status.

The Edge Wallet is a local credential and key service bound to the robot, its approved workload and installation identity. Deployment on robot hardware is analogous to holding an identity wallet on a smartphone, but a subordinate robot service is distinct from a certified EUDI Wallet. A secure element, TPM or suitable protected execution environment can support key custody. The wallet shall expose narrowly scoped APIs to approved workloads rather than unrestricted access to signing keys.

- Enrol the robot: verify component and platform identity, commissioning records, ownership or custody, and a fresh attestation against approved reference values.
- Generate device and workload keys locally under a protection policy. Register public keys and proof of possession with the enterprise identity service. Retain organisational seal and update-authorisation keys in their separate controlled systems.
- Issue a robot-agent credential binding the principal, robot ID, workload key, configuration baseline and identity lifecycle. Provision issuer trust anchors, schema/context allowlists and status-service policy.
- Issue a mandate with action, resource, purpose, audience, location, configuration, validity, delegation-depth and intervention constraints. Bind any access token to the authorised workload key.
- Install local enforcement policies, evidence-signing templates and bounded offline rules. Validate denial, expiry, disconnection, reboot, clock uncertainty and human intervention before commissioning.
- Operate with renewal, key rotation and parent–child status propagation. Transfer, retirement or ownership change closes prior mandates and requires controlled re-enrolment; retained evidence keeps its historical provenance.

In the enterprise case, a manufacturer or operator may authorise “retrieve approved tool and record handover” for a specified work area and shift. That mandate grants neither safety-parameter modification nor legal conformity-signing authority. A private companion’s task mandate may come from the user or an authorised representative, while the service company holds a separate maintenance mandate. Remote maintenance rights and consent to care or disclosure remain distinct, purpose-limited permissions.

Privacy requirements follow the actual processing and roles. Care data can engage GDPR Articles 5, 6, 9, 25 and 35, including purpose limitation, an appropriate lawful basis, conditions for

health data and risk-based impact assessment [63]. The Edge Wallet should reference protected traces and minimise environmental identifiers. Signatures and hashes preserve linkability and integrity; they do not anonymise a personal record. Access and retention policies must also address bystanders and counterparties to an interaction.

9.5 Control plane: identity, mandates and access

The control plane governs who may act, for whom, on which resource and for how long. A policy decision point evaluates principal identity, representation, mandate attenuation, workload proof of possession, configuration, audience, status and action scope. A local policy enforcement point admits the authorised request. This separation follows established zero-trust principles [64]. Token exchange can express delegation and DPoP can bind suitable HTTP access tokens to a key; those protocols implement technical controls rather than establish legal authority [65, 66].

The decision is short-lived and bound to the verified context in Equation (6). The robot checks a bounded status cache or obtains fresh status before granting a new operational lease. Parent revocation invalidates subordinate reliance under the policy. Offline operation uses an explicit maximum age, clock-error bound and prevalidated action subset. A status-service outage triggers the defined supervisory response when that bound expires, without introducing network calls into the motion-control cycle.

A language model or task planner submits structured requests to a broker. The broker independently validates action fields, source measurements, mandate and evidence type. Untrusted scene text, a verbal instruction or a software agent's self-description shall not expand authority. Changes to safety parameters, firmware, external disclosure or legal declarations require their own permission and approval route.

9.6 Data plane: autonomous evidence signing

The data plane records what the robot observed, inferred, exchanged and executed. The Edge Wallet's signing service authenticates bounded evidence records produced by approved sensors and workloads. Each batch identifies robot and workload keys, principal and mandate references, policy decision, configuration digest, boot/session ID, sequence interval, observation times and clock uncertainty, source/calibration references, method or model version and trace digests. Environmental-interaction receipts identify the counterparty and transaction where permitted. Observed, estimated, received and assessed claims retain distinct types.

Telemetry is buffered outside the safety task and signed in bounded batches or event records. High-rate source data remains in controlled repositories. Queue size, compute budget and backpressure protect the safety workload; loss of logging causes an explicitly assessed supervisory response where evidence availability is an operating assumption. The robot may sign a finite VSAEG snapshot or manifest of included objects. That signature attests the selected bytes and origin of the manifest, while referenced third-party credentials retain their own issuers, proofs and status.

For example, an approved handover produces a policy decision, grasp/retention observations, a tool and configuration reference, an interaction receipt and a completion or abort record. The robot signs those runtime facts within its mandate. A laboratory signs measurement results within its competence, an assessor issues an assessment within its authority, and a manufacturer remains

responsible for its conformity declaration. An Edge Wallet cannot promote its observations into an independent certification.

9.7 Signing JSON-LD credentials and organisational evidence packages

An issuance profile first constructs a typed JSON-LD record, validates the applicable shape and cross-object requirements, and fixes the exact bytes or canonical RDF dataset covered by the proof. The Edge Wallet or authorised issuer then creates a secured VC using a specified Data Integrity or JOSE/COSE profile [55, 56]. The credential identifies its actual issuer; a device-held signing service acts only within that issuer's credential and key policy.

For a legally attributable organisational package, the Business Wallet verifies the representative or service mandate and invokes the permitted signing or sealing service. One implementable arrangement seals a manifest containing digests of the secured VCs, graph/context/profile versions and assessment scope. The verifier checks both the device/issuer VC proof and the seal-to-manifest-to-object digest chain. If a qualified seal or timestamp is required, issuance and validation use the applicable qualified provider, certificate and creation environment. A generic JWS remains a cryptographic proof and gains no qualified status through nesting alone [35].

A status profile can use Bitstring Status List entries for revocation and suspension, with separate entries for the two purposes [67]. Expiry, suspension, revocation, supersession and key rotation have distinct semantics. Verifiers check the exact status reference, issuer authorisation, validity, freshness and the policy for evidence created before a compromise. A timestamp supports temporal evaluation; an asserted timestamp field in a device record remains a claim about time.

Selective disclosure requires a compatible proof scheme or separately issued minimal credentials. Removing fields from an ordinary signed JSON payload invalidates its signature. Authorised retrieval of a referenced subgraph can also minimise disclosure, provided the decision profile still receives the required evidence. Appendix A supplies an explicit unsecured JSON-LD payload and a separately signed demonstration envelope to make these boundaries inspectable.

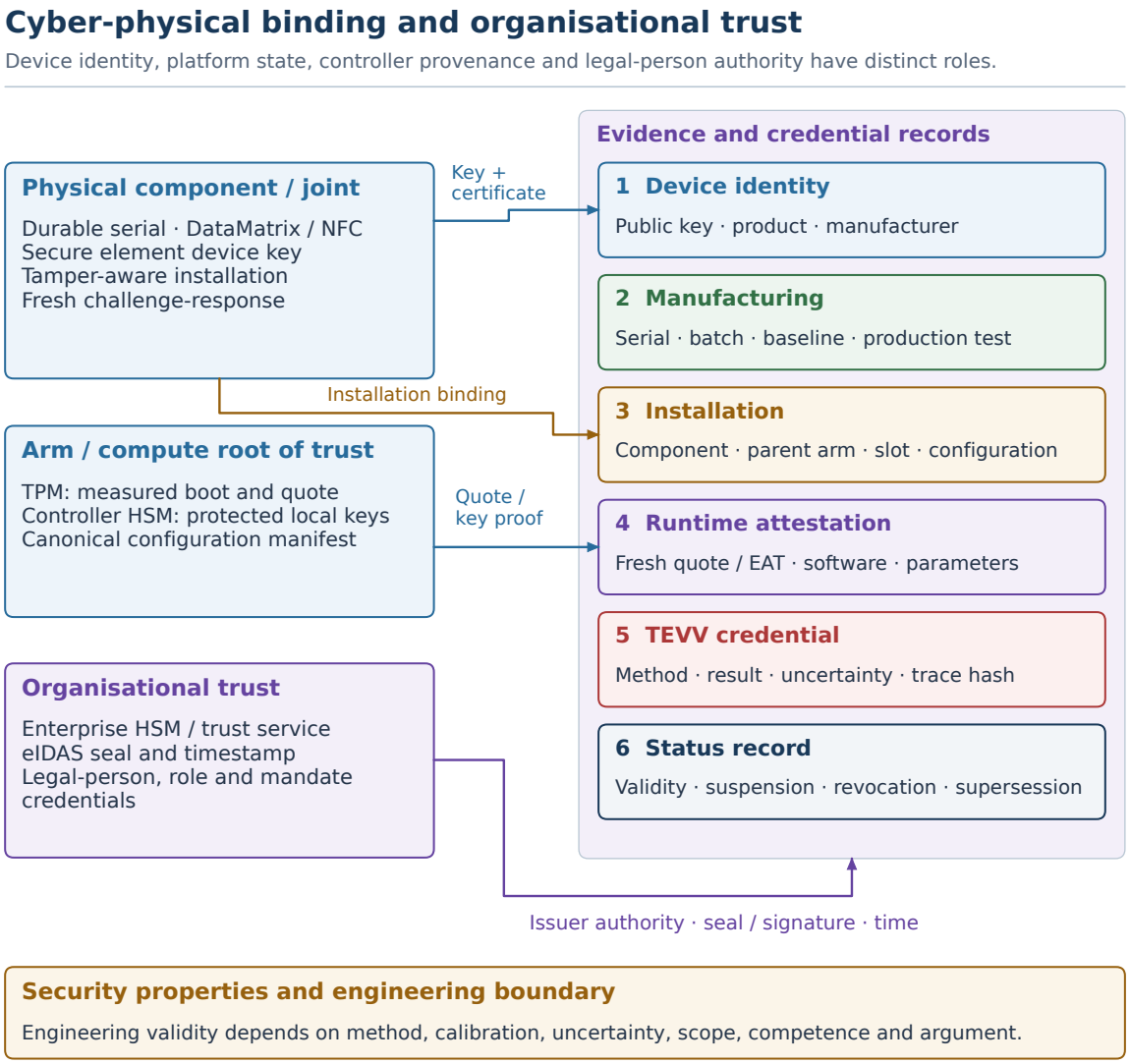

Figure 4: *Cyber-physical binding and organisational trust. Hardware identity, attested state, installation evidence and organisational authority establish complementary properties. Robot mandate and Edge Wallet services extend this foundation as described in Sections 9.4–9.7.*

10 Digital TEVV and verifiable-credential lifecycle

10.1 TEVV as an evidence production process

Testing, evaluation, verification and validation serve different questions. **Testing** applies a defined stimulus and records a response. **Evaluation** interprets results against criteria and uncertainty. **Verification** determines whether an implementation satisfies specified requirements. **Validation** determines whether the configured system is suitable for its intended use and ODD. A single measurement can contribute to several activities, but its record must state which conclusion is claimed and by whom.

The test strategy uses complementary modalities:

Table 4: *TEVV modalities and representative robot-arm evidence.*

TEVV modality	Primary purpose	Representative robot-arm evidence
Document and design review	Check requirements, interfaces, analyses and assumptions	Hazard analysis; safety plan; architecture; FMEDA; software and cybersecurity records
Component test	Characterise one item under controlled conditions	Sensor diagnostic test; gate-driver fault response; motor thermal curve; secure-element challenge
Joint-module test	Capture electromechanical interactions	Torque/speed envelope; brake holding; backlash; position-channel disagreement; thermal derating
Arm integration test	Validate kinematics, interfaces and coupled behaviour	Workspace limits; stopping performance; communication fault; common-cause power event
SiL and HIL	Exercise logic and timing before hazardous physical tests	Trajectory bounds; watchdog; delay and packet faults; controller/model versions
Fault injection	Show detection and safe response	Encoder freeze; current offset; gate fault; stale message; key rejection; configuration mismatch
Physical use-case test	Validate configured task and safeguards	Payload retention; contact measurement; protective stop; tool hazard; emergency procedure
External-observation test	Validate coverage, timing and disagreement handling	Blind-spot scenario; collision prediction; time synchronisation; authenticated request
Human-factors test	Evaluate foreseeable use and vulnerable users	Mode comprehension; intervention; accessibility; misuse; workload; recovery
Cybersecurity test	Evaluate security controls that support assurance	Secure boot; update authenticity; access policy; replay; status unavailability; incident workflow

Table 4: *TEVV modalities and representative robot-arm evidence. (continued)*

TEVV modality	Primary purpose	Representative robot-arm evidence
Runtime observation	Detect health, behavioural and context change	Stopping trend; brake margin; model update; repeated boundary approach; ODD deviation

Startup checks can include logic and memory tests, RAM and flash integrity, power supervision, watchdog challenge, gate-driver diagnostics, encoder diagnostics, sensor supplies, brake state, configuration and firmware verification, secure boot or attestation and emergency-stop input tests. Runtime checks include commanded versus measured position and current, redundant encoder agreement, estimated versus measured torque, thermal model residuals, phase-current symmetry, DC-link voltage, brake command and feedback, communication age, watchdog state and position, speed, force and workspace limits. Maintenance checks include encoder-zero and torque calibration, torque-constant verification, backlash, brake holding and release, insulation, structural inspection, proof load, thermal resistance, vibration and harness inspection. The applicable risk assessment selects the set and interval.

Every evidence result should include the following structured fields. The applicable profile determines mandatory values, controlled units and explicit reasons for any inapplicable field:

- Identity and scope: evidence-result ID; tested component or assembly ID; installation position; configuration digest; intended use and ODD; affected claim or safety function.
- Method and execution: versioned test method; test software and firmware identifiers or digests; stimulus and duration; equipment; executor; reviewer or observer role; source provenance.
- Measurement and decision: measured values and units; uncertainty model, coverage factor and interval where applicable; acceptance limits; decision rule; result; anomalous or missing samples.
- Calibration and conditions: calibration reference and validity; sensor geometry; environment; assumptions of use; test start/end times; clock source and time uncertainty.
- Evidence retention: raw-trace location and digest; record/graph schema and context version; retention and access policy; links to superseded or contradictory evidence.
- Authority and status: issuer, signing key and proof profile; principal and mandate; policy-decision reference; credential validity and status; seal and trusted timestamp references where required.

High-rate telemetry is retained in a fit-for-purpose time-series or content-addressed repository. The graph records signed summaries, events, statistical results, locations and hashes. This preserves efficient query and provenance without turning an RDF store into a telemetry database.

A concrete decision rule is illustrated in Appendix A. The synthetic fixture applies a measured 91.4 N · m load for 60 s, with expanded torque uncertainty 1.1 N · m (coverage factor k = 2). Its

lower guard-band value is 90.3 N · m, exceeding the illustrative 88.0 N · m minimum, while the displacement upper bound, 0.08° + 0.01° = 0.09°, remains within the illustrative 0.10° limit. This supports only the defined proof-load result under the synthetic method. It establishes neither breakaway torque nor application-level holding safety. The choice of guard band and conformity risk requires a documented decision rule [68].

10.2 Eight-stage digital validation process

1. **Validation request.** The requester supplies arm identity, intended use, ODD, population, jurisdiction, tool, payload, change set and requested issuer.
2. **Profile resolution.** The system resolves applicable legal and standards profiles, hazards, safety functions, target integrity, evidence requirements and issuer roles.
3. **Freeze and attest.** The installed component identities, configuration, firmware, software, AI model, parameters and calibrations are frozen and attested.
4. Test-plan decomposition. Existing evidence is assessed for scope and freshness. Missing component, joint, arm, SiL, HIL, physical, cyber, external-observation and human-factor tests are scheduled.
5. **TEVV execution.** Qualified systems and people execute the tests. Results include uncertainty, trace hashes, observations and fault-injection outcomes.
6. **Evidence-graph construction.** Results are linked to configuration, hazards, requirements, assumptions, claims and provenance. Shapes validate completeness.
7. **Assurance review.** The authorised reviewer evaluates claims, arguments, residual risks, counterevidence, issuer competence and independence.
8. **Scoped credential issuance.** The issuer creates a time-limited credential containing configuration hash, allowed use, ODD, limits, issuer, expiry and status endpoint.

Verification before use evaluates the secured credential, authorised issuer, current status, configuration, application scope, health tests and robot mandate. A changed context initiates a new decision. Revalidation links the new assessment and credential to its predecessor while preserving failed tests and the earlier conditions of reliance.

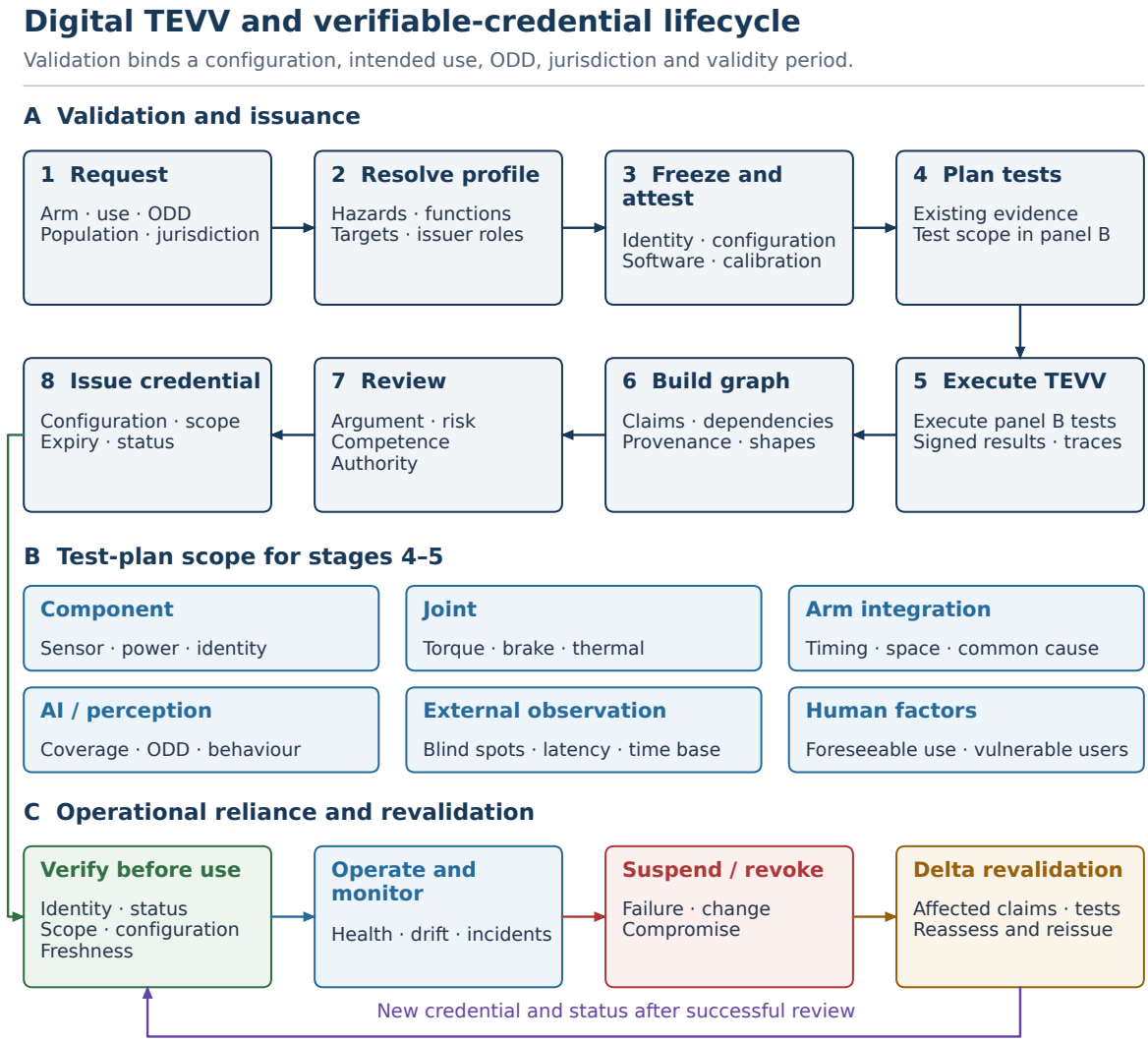

Figure 5: *Digital TEVV and credential lifecycle. Profile selection, configuration freeze, test planning, evidence generation and authorised review precede scoped issuance. Runtime checks, suspension and delta revalidation preserve scope over change. The extended flow also verifies robot mandates and identity status as described in Section 9.*

11 Hybrid distributed safety architecture

11.1 Inside-out validation

Inside-out validation uses onboard sensing and local compute. Joint encoders, torque or force sensors, motor current, temperature, brake feedback, supply voltage and communication monitors feed state estimation and diagnostics. The AI or motion-planning computer generates a proposed trajectory. The dedicated safety controller independently checks safety-rated constraints such as speed, position, workspace, torque, stopping behaviour, timing, communication age and mode. It also checks hardware diagnostic signals, watchdogs and configuration identity.

The safety path evaluates the specified speed, position, workspace, torque-related and mode conditions using the sensing and diagnostics justified in its safety case. A watchdog is one fault-detection mechanism; the controller additionally implements the complete safety-function logic and response. Independence requires evidence for shared power, clocks, communications, memory, scheduling and other common-cause paths. Logical separation alone establishes insufficient fault containment.

11.2 Outside-in awareness

Outside-in awareness uses infrastructure sensors and compute to observe the robot and its environment from perspectives unavailable to onboard sensors. Cameras, depth sensors, radar or other authorised sensors can cover blind spots, estimate human and object trajectories and detect anomalous behaviour. Infrastructure compute can predict collision risk and send a secure safety request or constraint update.

External messages require authenticated origin, integrity, freshness, sequence checking, time-base assumptions and bounded handling. A secure transport protects against attacks, while a safety-related communication path also requires the error-detection and timing measures of its selected safety architecture. A trajectory correction remains a proposal subject to local checks. An emergency or protective request must invoke its specifically validated function; arbitrary network messages cannot be treated as a certified emergency-stop channel.

11.3 Continuous deterministic loop and timing budget

The continuous loop is: sense; estimate; propose; check; enforce; observe outcome; record evidence; evaluate drift. Inside-out and outside-in observations are time-synchronised and linked to configuration and policy. Disagreement is itself evidence. The system can reduce speed, increase separation, request human review or stop according to the validated policy.

For any external observation credited with risk reduction, the assessed worst-case response must include sensing, estimation, communication, verification, decision and physical stopping or holding:

$$T_{\text{sense}} + T_{\text{estimate}} + T_{\text{network}} + T_{\text{verify}} + T_{\text{decide}} + T_{\text{stop}} \leq T_{\text{available}} - T_{\text{margin}}. \tag{7}$$

T_{stop} includes the mechanical response under the specified load and conditions, rather than ending at command dispatch. $T_{\text{available}}$ is derived from the hazard and approach model. Statistical average or percentile latency cannot establish a worst-case guarantee. External channels with

unbounded inference or communication delay can provide supplementary observations while an independently justified protection remains active. Availability and delay faults require validated degraded modes.

Closed-loop validation uses scenario simulation, software-in-the-loop and hardware-in-the-loop before physical execution. It injects sensor faults, communication delay and loss, conflicting observations, compute overload, model errors, configuration mismatch and drive faults. The evidence package records both detection and recovery timing. Independence analysis addresses shared sensors, power, clocks, networks, software libraries and environmental blind spots.

11.4 Concrete, non-exclusive implementation example

NVIDIA describes Halos for Robotics and an Outside-In Safety Blueprint combining external perception, a Safety AI Monitor, event integration and decision functions; IGX Thor provides a possible edge-compute implementation [69–71]. These are useful architectural examples for scenario testing, perception monitoring and safety-request generation. The selected release, safety documentation, qualified function boundary and integration assumptions require project-specific assessment. Product announcements, simulations and authenticated messages alone establish no end-to-end safety rating.

Alternative implementations can use other edge AI platforms, safety PLCs, safety MCUs, drive-integrated safety and external perception stacks. The architectural test is functional: the planner cannot bypass the safety controller; the safety controller can enforce limits without the AI computer; external loss has a bounded response; and each evidence record identifies the exact hardware, firmware, software and configuration.

The Edge Wallet adds an evidence and authority side path to Figure 6. Sensor/event records flow asynchronously to the evidence broker and repository. Mandate decisions enter through a supervisory policy enforcement point. The local safety controller retains direct access to required safety sensing and its independently validated drive/holding path. The editorial review specifies the side-path additions without altering the existing figure artwork.

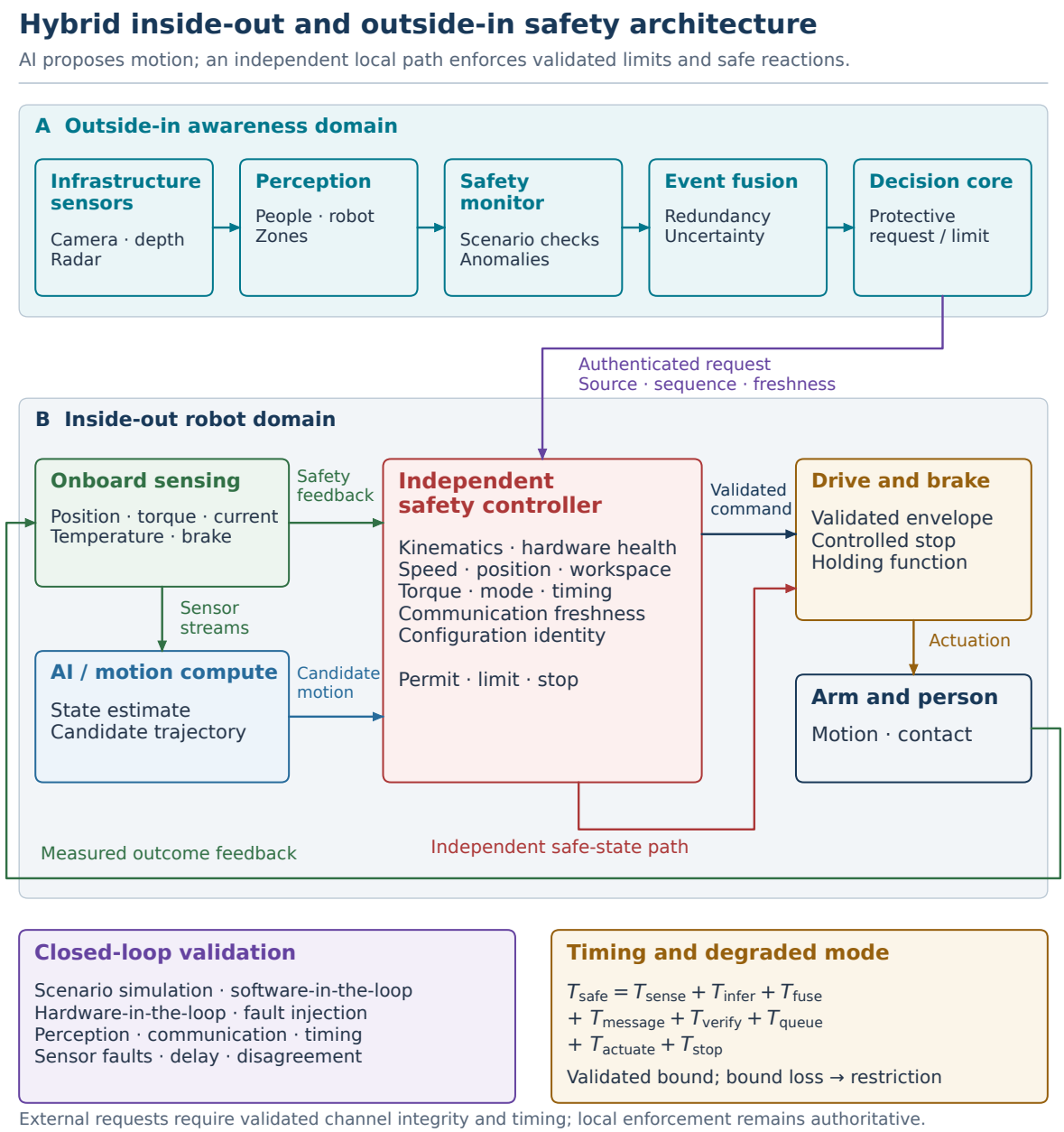

Figure 6: *Hybrid inside-out and outside-in safety architecture. AI proposes motion, infrastructure adds awareness and authenticated requests, and an independent local path enforces validated motion limits and safe reactions.*

12 Lifecycle assurance, drift and credential status

12.1 Evidence lifecycle

At design time, hazards, safety requirements, assumptions and verification plans are linked to component and architecture definitions. Manufacturing adds identity, production test and software-baseline evidence. Assembly records installed components and creates the arm configuration. Integration adds tool, payload, cell or service environment, safeguards, application risk assessment and validation. Commissioning issues a scoped operational credential only after required evidence and approvals are present.

Operation produces health summaries, policy decisions, exceptions and incident records. Maintenance changes component or calibration state. Updates change software, firmware or AI models. Each change is classified and propagated through dependency relations. The system computes an affected claim set and a delta validation plan. Approval creates a new baseline and superseding credential.

12.2 Five drift classes

Configuration drift is divergence between the authorised baseline and installed hardware, software, model, calibration or policy. **Safety-function performance drift** is degradation in stopping time, brake torque, sensor agreement, diagnostic coverage or other monitored performance. **Behavioural drift** is a change in AI or system behaviour, including distribution shift, planning anomalies, repeated boundary approaches or task-performance change. **Cyber drift** is a change in vulnerability, attestation, key, access or threat state. **Operational-context drift** is a change in tool, payload, workspace, human population, environmental conditions or intended use.

Monitoring uses thresholds, trends and rule-based dependencies. A threshold crossing creates a signed DriftObservation linked to the signal, method, configuration, hazard and affected claim. The policy engine evaluates severity, persistence, uncertainty and available safeguards. The action set is permit, permit with a tighter envelope, degrade, controlled stop, safe stop, quarantine or service request.

Behavioural monitoring supports safety but does not turn an unvalidated AI metric into a safety function. A drift detector has its own performance evidence, false-positive and false-negative characteristics, update controls and failure response. Where drift monitoring contributes to a safety-related function, its integrity requirements follow from the risk assessment.

12.3 Suspension, revocation and revalidation

An operational-assurance credential, a delegation mandate and a device key each have their own status. A drifting brake can suspend reliance on a payload profile while leaving device identity usable for diagnostics. Revoking a parent mandate restricts subordinate authority; key compromise affects attribution and can require re-enrolment. Expiry follows validity time, suspension is reversible under the profile, revocation terminates acceptance of that credential, and supersession links a replacement. Previously signed records remain auditable, with reliance evaluated against creation time, compromise evidence and applicable retention duties.

Revocation or suspension does not delete prior evidence. It changes reliance status from a stated time and reason. Verifiers retrieve fresh status or apply a defined maximum status age. Offline operation requires cached status with a bounded validity period and a conservative expiry action.

Drift responses shall record detection source, baseline, affected dependencies, uncertainty, decision and human intervention. A behavioural novelty detector may trigger investigation or a restrictive transition, while qualification of the safety function requires its own evidence. After authority withdrawal, incident logging may continue under a separate narrowly scoped policy; it cannot silently retain the revoked operational mandate.

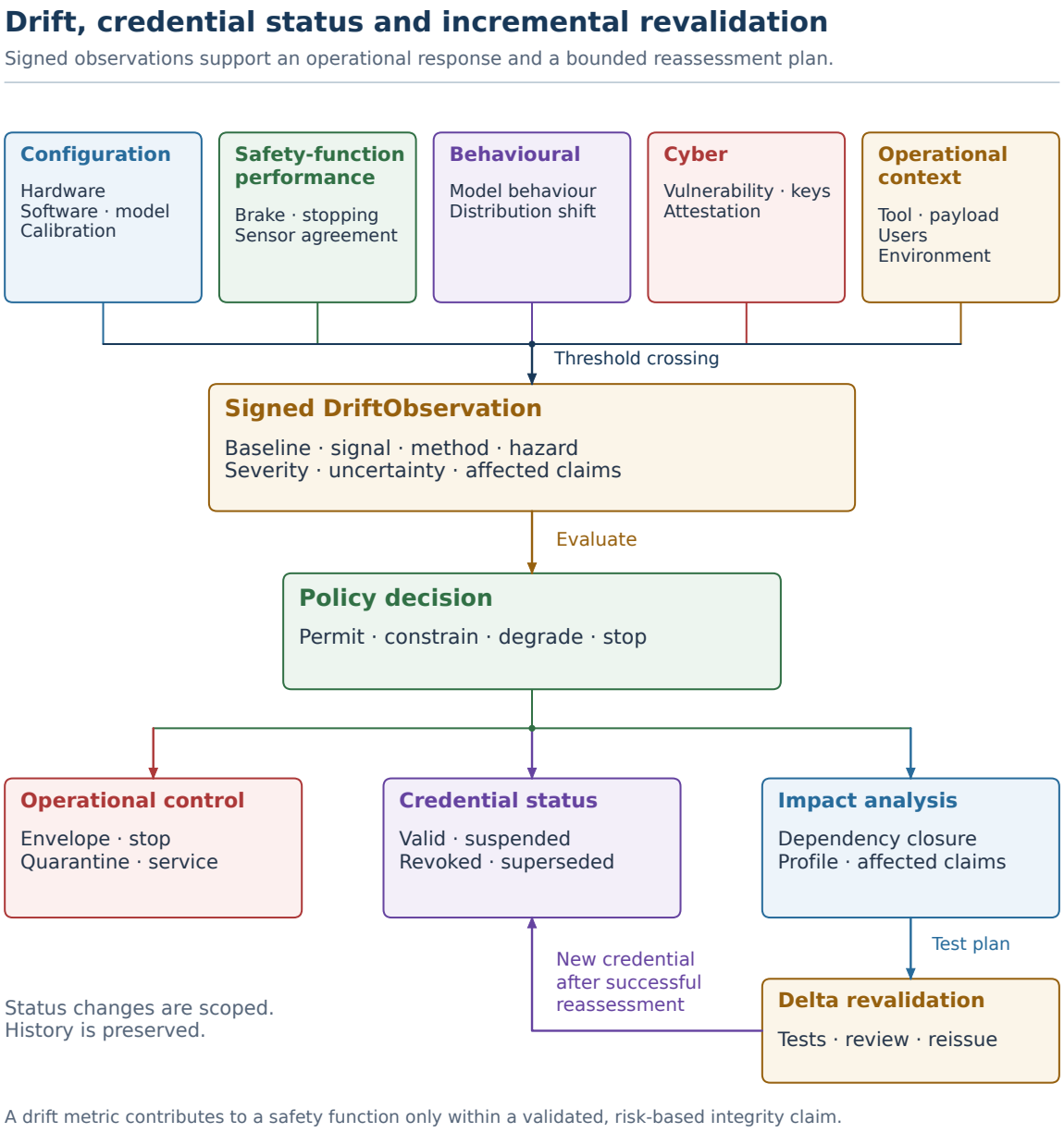

Figure 7: *Drift, credential status and incremental revalidation. Five drift classes produce signed observations, a policy decision, an operational response, a credential-status change and bounded impact analysis and revalidation.*

13 Application profiles and assurance ladder under uncertain future use

13.1 Profile A: guarded industrial handling

The arm handles parts within a safeguarded cell. Primary concerns include reach, crushing, unexpected startup, ejected workpieces, tool hazards and maintenance access. Evidence includes robot-level safety functions, cell risk assessment, guard and interlock design, stopping-distance validation, tool and payload limits, emergency-stop tests, access-control configuration and maintenance procedures. ISO 10218-1 and -2 form central standards inputs.

The evidence graph allows component and robot evidence to be reused while binding the application credential to the exact cell configuration. Replacement of a motor drive can trigger retest of stopping and thermal performance without reopening unrelated guard geometry evidence when dependencies show no effect.

13.2 Profile B: collaborative object handover

The arm shares a workspace with a person and hands over a non-sharp object. The risk assessment covers transient impact, quasi-static clamping, handover failure, unexpected motion, perception occlusion and object drop. Evidence includes speed and separation monitoring where used, power and force limiting where used, contact measurement, tool geometry, payload, body-region assumptions, stopping performance and human-factor tests. Acceptance limits cite the applicable standard edition and test setup.

Outside-in observation can improve coverage of blind spots. The local safety controller retains the validated stopping authority. A change from a compliant rounded object to a sharp or hot object invalidates the profile even if robot hardware remains unchanged.

13.3 Profile C: professional care assistance

The arm assists with object retrieval or feeding in a professional care environment. Users may have reduced mobility, slower reaction or medical vulnerability. Environmental variability and close human contact increase assurance needs. Evidence adds user-specific limits, hygiene and cleaning, privacy controls, remote intervention, human oversight, continuous health monitoring and incident reporting. ISO 13482 may be relevant depending on final scope and edition. A medical intended purpose activates separate medical-device assessment.

13.4 Use-case and assurance ladder

Figure 8 presents six illustrative service/care profiles. They organise changes in contact, dependence and intended purpose, rather than define a universal ordering of SIL, PL or legal classification. Industrial guarded and collaborative profiles remain separate examples in Sections 13.1 and 13.2.

1. **Fetch, carry and telepresence.** The arm performs non-contact logistics or remote presence. Evidence emphasises safe speed, workspace limits, collision detection, object retention, network loss and manual recovery.
2. **Object handover and medication delivery.** Close approach introduces finger entrapment, recipient identity, task authorisation, retained-object and audit-trail requirements.

3. **Feeding and meal assistance.** Face proximity, hot liquids, utensil orientation, user pause, swallowing-related misuse and hygiene add hazards and human-factor evidence.
4. **Dressing and hygiene.** Entanglement, skin pinch, tactile sensing, compliant withdrawal, privacy and local processing become central.
5. **Mobility support and person transfer.** Person-supporting load paths, redundant sensing, brake proof testing, safe hold and emergency lowering become essential. Removal of motor torque can increase risk.
6. **Rehabilitation and a possible medical profile.** A medical intended purpose can activate medical-device law, essential performance, clinical risk, usability, post-market monitoring and independent assessment.

Future intimate-care tasks form an additional research boundary, with exceptional privacy, consent, contact and vulnerability concerns. They require task-specific evidence and ethical review before any human study. The six profiles in Figure 8 provide a starting taxonomy rather than permission to extrapolate contact or clinical claims.

Evidence burden generally increases with contact severity, user vulnerability, dependency, autonomy, medical intent and regulatory scrutiny. These variables can move independently. A low-speed task can still be severe when the user cannot withdraw, and a non-contact task can remain cyber-critical when remote compromise controls high-energy motion.

Table 5: *Reusable and profile-specific evidence across application layers.*

Evidence layer	**Reusable across profiles**	**Profile-specific addition**
Component	Identity, design limits, safety manual, FMEDA inputs, production tests	Application restrictions on use
Arm assembly	Kinematics, structural limits, controller baseline, core safety functions	Tool and payload configuration
Integration	Interface tests, communication and power architecture	Cell, safeguards, user and environment
Use-case validation	Test methods and reusable evidence fragments	Hazards, acceptance criteria and final validation
Runtime	Health, configuration, vulnerability and status mechanisms	Thresholds, intervention and reporting policy

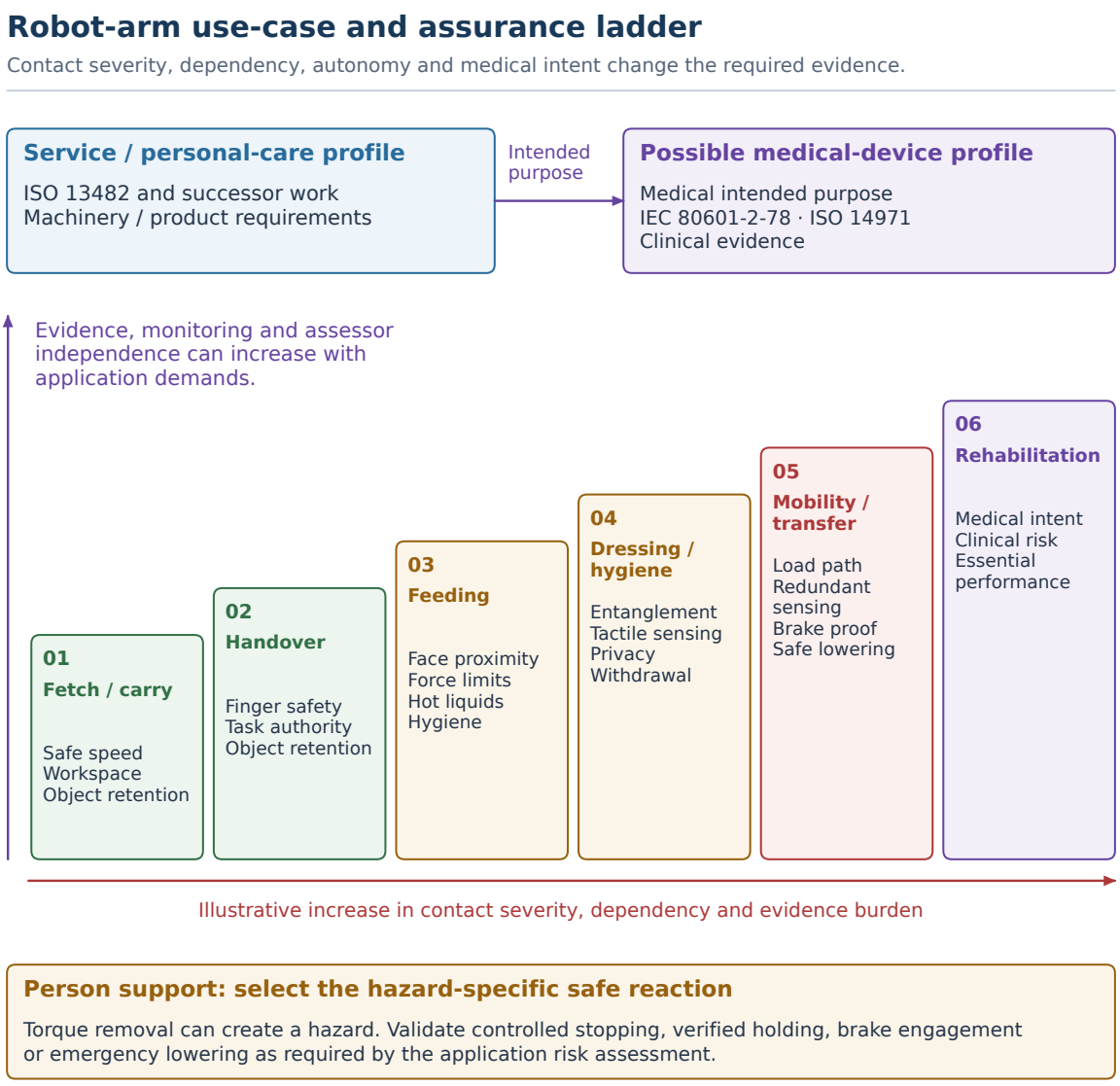

Figure 8: *Robot-arm use-case and assurance ladder. Contact severity, dependency, user vulnerability, autonomy and medical intent can increase evidence burden and change the applicable regulatory profile. The dimensions are qualitative and can vary independently.*

14 CRA-capable DPP evidence packages and trust plane

The DPP-facing layer exposes a product identifier and a manifest of evidence resources. Public entries can include manufacturer identity, product model, support period, documentation location and permitted verifier endpoints. Restricted entries can contain safety manuals, SBOMs, VEX statements, update history, vulnerability decisions, test evidence and assurance credentials. Highly sensitive artefacts such as detailed FMEDA data or raw vulnerability information remain in protected repositories; the graph stores metadata, access policy, hash and provenance.

A CRA-related claim should identify the economic operator, product and software versions, applicability rationale, support period, requirements addressed, assessment basis and status. The graph links signed SBOM/VEX information and update evidence to affected configurations and safety assumptions. Article 14 reporting records preserve timestamps, preliminary findings and subsequent corrections [1]. Sensitive exploit and personal data remain in controlled repositories; a public passport entry can supply discovery and authorised access instructions.

The organisational control plane links a responsible principal to identity, representation, mandate and trust-service operations. The subordinate Edge Wallet implements local verification and evidence-signing services under that authority. Figure 9 shows the regulatory and trust foundation; the expanded two-plane deployment and natural-person companion branch are specified in Section 9 and in the figure-update instructions. A legal-person seal establishes attributable organisational origin under its applicable trust profile, while safety and CRA conclusions require the responsible assessment process.

The model supports regulatory reuse without claiming regulatory equivalence. CRA evidence addresses cybersecurity obligations. Functional-safety evidence addresses risk-reduction functions. AI evidence addresses the AI role. DPP services provide discovery and controlled access. Cross-links permit coordinated change analysis.

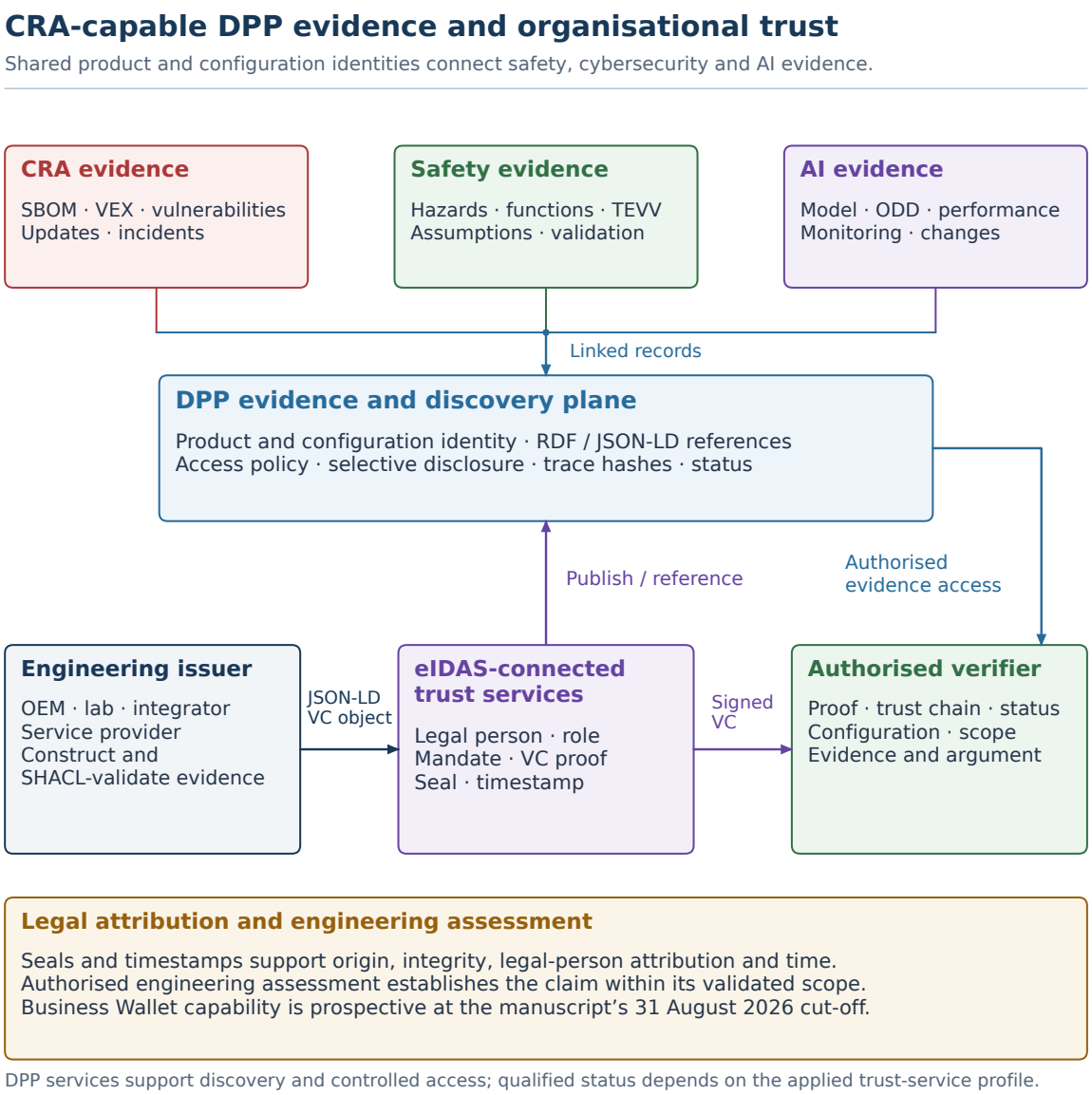

Figure 9: *CRA-capable DPP evidence and organisational trust. Domain evidence, protected discovery, issuer authority and trust services connect through scoped claims. European Business Wallet capabilities remain proposed; Section 9 develops the associated Edge Wallet control and data planes.*

15 Vendor-neutral implementation profiles and concrete best practices

Table 6 maps functional roles to non-exclusive implementation examples. Selection requires the exact device variant, software release, safety/security documentation, lifecycle support and integration assumptions. Technology-family names describe candidates, rather than certified equivalence across suppliers.

Table 6: *Non-exclusive implementation profiles and evidence-oriented best practice.*

Functional role	**Illustrative technology families**	**Evidence and best practice to request**
Safety MCU or controller	Infineon AURIX; TI Hercules or C2000 safety devices; NXP safety MCUs; Renesas safety platforms; ST safety-supported STM32 configurations	Safety manual, FMEDA, diagnostic library, tool qualification assumptions, errata, watchdog and lockstep evidence, independent review [72–76]
Gate driver and power path	Infineon MOTIX; TI, NXP, ST, Renesas or specialist drive components	Safe-state behaviour, desaturation and short-circuit response, power diagnostic coverage, common-cause analysis, safe brake interface
Component identity anchor	Infineon OPTIGA Trust; NXP EdgeLock SE; STSAFE; TPM 2.0 devices from qualified suppliers	Key provisioning and ownership, certificate chain, anti-rollback, lifecycle states, update and revocation, physical-binding procedure
Platform attestation	TPM 2.0 platform with controlled reference values and attestation appraisal	Measured boot, fresh quotes/EAT, reference-value governance, verifier policy, recovery and algorithm agility
AI/edge compute	NVIDIA IGX Thor and Jetson-class systems; other industrial edge accelerators	Exact BSP, driver, model and container hashes; secure boot; resource interference; latency; monitoring; software support status
Outside-in observation	NVIDIA Halos/Metropolis/VSS concepts; industrial vision and safety-sensor ecosystems	Coverage, calibration, time synchronisation, uncertainty and delay/loss handling; distinguish AI observations from validated protective functions [69, 70]

Table 6: *Non-exclusive implementation profiles and evidence-oriented best practice. (continued)*

Functional role	Illustrative technology families	Evidence and best practice to request
Safety PLC and safe motion	Siemens Safety Integrated; Beckhoff TwinSAFE; Bosch Rexroth ctrlX SAFETY; ABB SafeMove; comparable certified systems	Certified function scope, parameter protection, fieldbus safety profile, stop-time data, validation tools and configuration export
Graph, credential and wallet services	Standards-based RDF stores, VC issuers/verifiers, enterprise wallets, eIDAS trust-service providers	Schema and context governance, issuer authorisation, HSM custody, privacy, status availability, audit, portability and exit plan
Edge Wallet and mandate broker	Protected device keys; standards-based credentials; token exchange and proof-of-possession access	Attenuated delegation, narrow signing API, local enforcement, offline lease, status propagation and evidence batch provenance [64–66]

Good implementation practice includes a minimal trusted computing base for the safety path; diverse sensing where the risk requires it; signed and reproducible software builds; configuration export from safety engineering tools; documented clock and counter sources; fail-safe status caching; security advisories linked to SBOM components; and test equipment that signs evidence only after binding the physical subject, configuration and calibration. Supplier evidence should expose assumptions of use in machine-readable form while retaining the human-readable safety manual as the normative interpretation source.

An illustrative joint can combine a safety MCU with an embedded HSM, supervised power supply, diagnostic motor gate driver, dual-die or diverse position sensing, current and temperature sensing, a brake with feedback and a secure element. An AI compute module proposes trajectories. A TPM attests the compute baseline. An independent safety controller evaluates kinematics, safe speed, position, torque, workspace, communication freshness and hardware health. An enterprise wallet and trust service sign or seal component, installation and TEVV credentials. Equivalent products from other suppliers can fulfil each role when their evidence and integration assumptions satisfy the profile.

The implementation boundary shall avoid vendor-derived inference. A safety manual supports a design argument; it does not certify the integrated arm. A platform marketed for functional safety does not make an AI application safe. A secure element authenticates a key; it does not authenticate an unbound gearbox. A certified safe-motion function remains valid only under its defined parameter, feedback, architecture and application assumptions.

16 Evaluation protocol

16.1 Proof of Technology

The proposed demonstrator uses one controlled joint and a stationary arm integration model. Its minimum equipment comprises redundant or otherwise justified position sensing, current/temperature measurement, a brake/holding fixture, a deterministic safety controller, a planning computer and protected keys. A laboratory host supplies issuer and status services; a robot-resident process implements the proposed Edge Wallet broker. Simulation and hardware-in-the-loop (HIL) scenarios precede hazardous physical tests. Software-in-the-loop is abbreviated SiL here to distinguish it from safety integrity level (SIL).

Proof of Technology tests whether evidence acquisition, configuration binding, proof verification, mandate attenuation, status propagation and local supervisory transitions work under specified faults. Acceptance criteria are fixed before measurement. Hardware timing is recorded independently of application logs, and reference fault labels identify the expected affected claims. Successful software checks establish selected mechanism properties; successful safety validation additionally requires the prescribed lifecycle evidence and independent review.

16.2 Proof of Assurance Utility

Proof of Value evaluates whether the same architecture reduces assurance work while maintaining or improving decision quality. Use paired scenarios under a document-management baseline and a VSAEG workflow, counterbalancing order to reduce learning effects. Match reviewers by role and record training, task complexity, missing information and assistance. An independently adjudicated dependency/claim set provides the reference for completeness and change-impact scoring. A pilot estimates variance before sample-size and power decisions.

Primary endpoints should be defined in advance: reviewer time to a defensible decision, omission rate for required evidence, precision and recall of affected-claim identification, and false-permit rate under injected faults. Secondary endpoints include provenance-retrieval time, package portability, status propagation, revalidation scope and perceived usability. Report denominators, uncertainty intervals and task-level distributions. Performance percentiles characterise operational service quality; they remain separate from safety response-time bounds. Commercial savings are hypotheses until observed against the baseline.

16.3 Falsification tests

The artefact should fail safely under deliberately adverse cases:

- every component credential is valid, but the tool exceeds the validated payload;
- a test result is correctly signed, but its calibration has expired;
- an authorised issuer signs a claim outside its accredited scope;
- an AI model changes while the arm configuration identifier remains stale;
- an external observer sends an authenticated but delayed stop request;
- an SBOM component becomes vulnerable after commissioning;
- the brake holding torque trends below its acceptance margin;

- the operating environment changes from guarded to collaborative;
- status is unavailable beyond the offline validity window.

Success is defined separately for evidence validity, authority enforcement and physical response. A forged or unknown prerequisite shall prevent a new operational permit. A safety fault shall produce the predefined hazard-appropriate transition within its validated bound. A complete incident record shall preserve the observation, decision and affected scope. Failure in one dimension cannot be compensated by success in another.

16.4 Demonstrator fault and change matrix

Table 7: *Demonstrator fault and change matrix.*

Scenario	Injected condition	Expected evidence and control response
Encoder latent fault	Freeze channel B while channel A moves	Diagnostic event; affected safety-function claim; controlled restriction or stop; delta sensor test
Brake degradation	Holding torque falls through warning and acceptance margins	Trend evidence; tighter payload or speed; credential suspension at limit; service and proof test
Configuration substitution	Install a valid component in the wrong joint position	Challenge passes; installation/configuration mismatch fails; operation denied until authorised rebind
Firmware update	Signed controller firmware changes	New measured state; dependency closure; regression and timing tests; superseding credential
AI model update	Planner model hash changes	Operational credential mismatch; AI TEVV and ODD review; local safe envelope remains enforced
External delay	Protective request arrives beyond freshness bound	Message rejected as stale; degraded-mode transition; timing event recorded
Vulnerability disclosure	Critical SBOM dependency gains an exploitable VEX status	Cyber drift event; affected safety assumptions assessed; restriction, patch or credential suspension

Table 7: *Demonstrator fault and change matrix. (continued)*

Scenario	Injected condition	Expected evidence and control response
Status outage	Verifier cannot refresh beyond offline window	Cached validity expires; conservative degraded or stop state according to profile
Use-case change	Sharp heated tool replaces rounded object	Profile mismatch; collaborative credential invalid; new hazard analysis and physical validation
Issuer compromise	Evidence-signing key revoked	All dependent credentials identified; reliance suspended; re-issuance from preserved raw evidence under approved process
Mandate expansion	Planner asks for a safety-limit update using a handover-only mandate	Request denied; attempted action and policy recorded; existing safety envelope remains authoritative
Principal or key withdrawal	Parent mandate revoked or workload key compromised	Scoped reliance ends; validated supervisory response; separately authorised incident logging and recovery
Semantic substitution	Unapproved JSON-LD context or dictionary version supplied	Context rejected; evidence decision remains incomplete; original bytes preserved for review

16.5 Comparative baseline

Baseline comparisons should retain the same hazards, configuration changes and decision criteria across workflows. Report all exclusions and failed runs. Confidence intervals should account for repeated tasks by the same reviewer or configuration, rather than treat correlated observations as independent. Independent specialists should challenge the applicability matrix, dependency boundary and adverse evidence. Studies involving workers or care recipients require appropriate ethics approval, consent, safeguarding and stopping criteria before recruitment.

16.6 Synthetic consistency specimen and reproducibility boundary

The companion specimen contains an explicitly synthetic brake-test payload, exact configuration and trace files, a public demonstration verification key, secured JWS envelopes, a local status-list fixture and an executable verifier. Its checks cover byte-level proof verification, payload equality, artifact digests, the stated decision rule, a clear local status entry, tamper rejection, inverse dependency traversal, unknown/conflict rejection and mandate attenuation. These checks

establish internal consistency for that specimen; they do not validate deployed issuer authority, a live status service, a qualified eIDAS seal or robot hardware.

The public demonstration key is reproducible test material and carries no production trust. The identifiers are illustrative and the fixture is resolved locally. A deployment must publish a governed vocabulary and verification profile, establish real issuer and device authority, and validate interoperability with independent implementations. The verification record and file manifest accompany the revised manuscript.

17 Discussion

The VSAEG addresses a structural mismatch between modular products and application-specific assurance. It preserves source evidence while allowing later actors to add integration and use-case evidence. The model makes uncertainty explicit: an arm can be evidence-ready before its final use is known, while operational authorisation remains withheld until the application profile is validated.

The architecture also joins safety and cybersecurity without merging their objectives. A cyber event can invalidate a safety assumption. A safety-critical configuration change can change the cyber attack surface. Shared identity, configuration and change objects allow coordinated impact analysis. Separate claim types and assessment roles preserve disciplinary boundaries.

Cryptographic identity improves provenance but creates key-management duties. Provisioning, rotation, recovery, compromise response, algorithm agility and end-of-life must be designed. Long-lived robot evidence may outlast current algorithms. Evidence packages should record algorithm identifiers and support re-sealing or archival timestamp renewal while preserving the original signature history. Trust frameworks must also address issuer onboarding, accreditation, schema governance and status-service availability.

Outside-in observation offers useful diversity and coverage but adds dependency on networks and infrastructure. Its safety contribution must be bounded by availability, latency, integrity and common-cause analysis. Local deterministic enforcement remains the preferred authority for immediate motion safety. External systems can request action, tighten constraints and support evidence, subject to validated architecture.

The European trust framework offers a path to legal-person attribution. eIDAS signatures, seals and timestamps can protect evidence today. The proposed European Business Wallet can reduce fragmentation in organisational credentials and cross-border presentation if the final Regulation and implementing standards support the required machine-to-machine patterns. The paper therefore uses a capability interface rather than assuming one final wallet protocol.

Open-world semantics is central to the unknown-use-case problem. It permits a robot arm to accumulate new application evidence without asserting that an earlier supplier graph was complete. It also supports federated ownership: a supplier can maintain a controlled subgraph, a laboratory can issue a test credential, an operator can retain runtime records and a verifier can assemble an authorised view. The closed-world validation profile is equally central. It converts the extensible graph into a determinate checklist for a particular reliance decision. The pair resolves a common tension between interoperability and safety completeness.

The design yields four testable propositions:

- **P1:** typed dependency relations reduce the time to identify evidence affected by a component or software change.
- **P2:** configuration-bound credentials reduce acceptance of valid but inapplicable evidence.
- **P3:** explicit issuer role and mandate reduce acceptance of claims made outside authority.
- **P4:** linked runtime drift and credential status reduce the time between loss of an assurance assumption and restriction of operational reliance. The evaluation protocol can reject these propositions if performance does not exceed the document baseline or if graph complexity creates offsetting errors.

Standardisation should focus on an open semantic core for component identity, configuration, safety function, assumption, test evidence, assurance scope and status. Controlled trust frameworks can govern issuers, access, confidentiality and credential profiles. Relevant forums include ISO/TC 299, IEC functional-safety committees, CEN-CENELEC JTC 24 for DPP, W3C credential and provenance work, IETF attestation work and OMG assurance modelling.

18 Limitations and threats to validity

This is a conceptual design-science contribution. Public standard summaries support scope statements, while licensed normative texts and deployment-specific legal assessment remain necessary for implementation. The seven-axis model excludes coupled whole-body and bimanual dynamics. Numerical examples are synthetic and establish no universal contact, PL/SIL or timing limits.

Graph completeness depends on governance and incentives. A signed false statement remains false. Selective disclosure can prevent full independent review if policy is weak. Status services can become availability dependencies. Hardware identity cannot by itself prove mechanical continuity. Behavioural drift detection has imperfect sensitivity and specificity. These limitations are represented as assumptions and evaluation targets rather than hidden by a binary trust label.

Empirical validity remains open: hardware fault injection, independent safety review, representative fleet data and comparative user studies are required to establish deployment performance and assurance utility. The Edge Wallet is a proposed architectural profile; interoperability, certification requirements and final European Business Wallet legislation require further work. These boundaries should govern the paper's submission as a conceptual contribution and any later claims of demonstrated safety or economic benefit.

18.1 Data, ethics and submission declarations

Data and code availability: the companion material contains synthetic example data, exact-byte digests, demonstration credentials, an offline verification script and the editorial traceability record. It contains no experimental human or robot dataset. Ethics: this conceptual study and synthetic specimen involved no recruited participants or patient testing. Future human-interaction studies require prospective review. Author contributions, funding, competing interests and any required AI-assistance disclosure must be completed and confirmed by the submitting authors under the selected journal's policy.

19 Conclusion, collaboration needs and research agenda

Modular humanoid robot arms require evidence that survives organisational boundaries and remains bound to the current configuration, application and operating state. The proposed VSAEG combines semantic structure, cryptographic provenance, engineering metadata, assurance arguments and operational status. It supports unknown future use by separating reusable source evidence from scoped application credentials.

The immediate research agenda is to implement the complete profile, validate independent issuer/holder/verifier interoperability and exercise the demonstrator against the falsification cases. Subsequent paired studies should measure evidence work and decision quality against conventional document management. Industrial, collaborative and care profiles should then be reviewed with qualified manufacturers, integrators, operators and assessors. Standardisation should address portable evidence, mandate and status semantics while preserving the responsible assessment process.

No single organisation controls the complete evidence chain. Specialised component suppliers create source evidence for sensors, actuators, controllers, secure elements and software. Joint and arm OEMs own the system architecture, integration assumptions and robot-level safety case. Application integrators own the task, tool, payload, workplace, safeguards and final application validation. AI, simulation and external-perception providers contribute model, compute, scenario and runtime-monitoring evidence. Maintenance and fleet-service providers create replacement, calibration, update and operational-health records. Operators and care providers supply use-context, human-factor and incident evidence. Independent laboratories and conformity-assessment bodies contribute competent assessment within their legal scope.

Trust-infrastructure companies provide organisational identity, wallet, credential, seal, timestamp, access-control and status services. Their role is essential because cross-company evidence must remain attributable, current and selectively shareable. Their role is also bounded: they provide the means to issue and verify claims; they do not decide whether a motor, safety function or robot application is safe. eIDAS trust-service providers connect digital evidence to legal persons and regulated trust services. Standards bodies and sector consortia govern shared semantics, profiles and issuer rules.

The Edge Wallet makes these partnerships operational. OEMs attest installed hardware and approved firmware; operators issue task mandates; service providers maintain bounded update and diagnostic rights; trust-infrastructure providers supply interoperable identity and status; qualified providers deliver regulated signatures, seals and timestamps where required. Contracts should allocate key custody, data-controller roles, delegated authority, incident notification, support duration and exit procedures. An operator's mandate and a supplier's warranty or safety assumption must remain separately reviewable.

Effective deployment therefore requires a partnership model with explicit responsibility, liability, confidentiality and change-control interfaces. A minimum consortium for a demonstrator should contain a component or joint supplier, an arm OEM, an application integrator, an operator, an AI or perception provider, a test laboratory, an independent assessor and a trust-infrastructure provider. Joint governance should publish the semantic core, retain competitive evidence behind controlled access, test portability between vendors and define a credible exit path from any infrastructure provider. Collaboration is not an organisational supplement to the architecture; it is a condition for assembling the distributed assurance argument.

A Example JSON-LD Verifiable Credential evidence object

This complete unsecured JSON-LD VC payload describes a synthetic stationary brake proof-load test. All custom terms use an explicit example namespace; the standard VC context supplies the core credential and status terms. The credential subject is the test result, with separate links to the tested brake, joint and arm. Values are illustrative. The measured load, uncertainty, dwell and angular-displacement criterion support the stated proof-load decision. The full specimen and secured envelope accompany the manuscript.


```
{
  "@context": [
    "https://www.w3.org/ns/credentials/v2",
    {
      "ex": "https://example.org/vsaeg#"
    }
  ],
  "id": "urn:example:evidence:brake-test:001",
  "type": [
    "VerifiableCredential",
    "ex:BrakeHoldingTestCredential"
  ],
  "issuer": "https://lab.example",
  "validFrom": "2026-09-07T10:00:00Z",
  "credentialSubject": {
    "id": "urn:example:test-result:001",
    "ex:testedComponent": {
      "id": "urn:example:brake:B08"
    },
    "ex:installedJoint": {
      "id": "urn:example:joint:J4:8F31"
    },
    "ex:arm": {
      "id": "urn:example:arm:A17"
    },
    "ex:installationPosition": "J4",
    "ex:configurationSha256": "bf9eef67784d14d6ca6f809221c5f1d410a6ba7770a7e52326339581ad0fa1df",
    "ex:testMethod": {
      "id": "urn:example:method:brake-proof-load:v1"
    },
    "ex:testSoftware": {
      "id": "urn:example:software:fixture-generator:v1"
    },
    "ex:firmware": {
      "id": "urn:example:firmware:demo-controller-1.0"
    },
    "ex:calibration": {
      "id": "urn:example:calibration:torque-rig:2026-09-01"
    },
    "ex:stimulus": {
      "ex:targetTorqueNm": 91.4,
      "ex:dwellSeconds": 60
    },
    "ex:measurement": {
      "ex:deliveredTorqueNm": 91.4,
      "ex:expandedUncertaintyNm": 1.1,
      "ex:coverageFactor": 2,
      "ex:maxAngularDisplacementDeg": 0.08,
      "ex:expandedAngularUncertaintyDeg": 0.01
    },
    "ex:acceptanceCriterion": {
      "ex:minimumVerifiedLoadNm": 88.0,
      "ex:maxDisplacementDeg": 0.1,
      "ex:decisionRule": "torque lower bound >= minimum load; displacement upper bound <= limit"
```

```
    },
    "ex:result": "pass",
    "ex:trace": {
      "id": "https://lab.example/traces/001",
      "ex:sha256": "d6c88751dbd2a7b2418b319931a8bb08dc5c1c340f72efe495b9b895e7809b09"
    },
    "ex:executor": {
      "id": "urn:example:rig:03"
    },
    "ex:reviewer": {
      "id": "urn:example:reviewer:01"
    },
    "ex:mandate": {
      "id": "urn:example:mandate:lab-rig-03"
    },
    "ex:policyDecision": {
      "id": "urn:example:decision:test-001"
    },
    "ex:observationTime": "2026-09-07T09:59:00Z",
    "ex:timeUncertaintyMs": 5,
    "ex:environment": {
      "ex:ambientTemperatureC": 23
    },
    "ex:assumptionOfUse": "Stationary test fixture; only the stated configuration and proof load.",
    "ex:claimScope": "Static brake proof-load result; application validation is separate.",
    "ex:synthetic": true
  },
  "credentialStatus": {
    "id": "https://status.example/lists/17#481",
    "type": "BitstringStatusListEntry",
    "statusPurpose": "revocation",
    "statusListIndex": "481",
    "statusListCredential": "https://status.example/lists/17"
  }
}
```


The JSON shown above is the payload, rather than a credential with an embedded proof. The companion brake_test_vc.jws secures its claims as a JWS using the VC JOSE/COSE profile (protected typ = vc+jwt and cty = vc) [56]. The demonstration public key is supplied for offline verification. Context bytes and a local signed status-list fixture are bundled; the example URLs are not operational services. The revocation entry includes its required identifier. A suspension profile would add a separate entry with statusPurpose = suspension [67].

The source trace is synthetic, and its 60 s dwell is a method parameter rather than a universal brake-test requirement. A pass verifies the illustrated load/displacement criterion only. Expanded uncertainty and coverage factor are explicit; interpreting a coverage probability additionally requires the measurement model and distribution assumptions [68]. Production issuance would bind complete method, firmware, calibration, executor, mandate and review records with governed status and access.

A qualified organisational seal, when required, is obtained through the applicable eIDAS service over the credential or a manifest digest chain as described in Section 9.7. The demonstration JWS supplies no qualified seal, qualified timestamp, legal-person identity verification or independent safety certification.

B Example scoped arm-validation credential fields

- Subject and authority: decision ID; responsible principal; robot/workload ID and key; mandate chain; requester; action and resource; policy and profile versions.
- Assurance scope: credential and assessment IDs; configuration digest; use case; ODD; jurisdiction; assumptions; applicable evidence snapshot.
- Verification outcome: each prerequisite and its true/false/unknown/conflict value; signature/issuer/status checks; freshness; missing or contradictory evidence.
- Enforcement: permitted action subset; limits; expiry and maximum offline age; sequence counter; required local checks; degraded or stopping response.
- Audit: observation and decision times; clock uncertainty; evidence references; decision signature; human override or review; superseding decision.

C Integration evidence checklist

For each figure, preserve the distinction between physical signal flow, evidence dependence, delegated authority and credential status. Arrow labels and endpoints shall name the relationship shown. The nine figure-specific amendment instructions in the editorial review identify the Edge Wallet extensions; the present revision preserves the supplied artwork. Final figure export should verify legibility at journal column width, text and arrow clearance, monochrome distinctions, caption accuracy and accessible descriptions.

D Demonstrator deliverables

A demonstrator package should include a controlled BOM/configuration, hazard and threat models, applicability rationale, assurance argument, ontology/context and validation shapes, issuer/mandate policies, representative evidence and status objects, raw-trace storage, safety-controller baseline, test procedures, fault scenarios and reproducibility instructions. Report measured data separately from synthetic fixtures and planned experiments. The initial companion specimen provides a subset of these artefacts for inspection.

References

References

[1] European Parliament and Council. Regulation (EU) 2024/2847 on horizontal cybersecurity requirements for products with digital elements (Cyber Resilience Act). Official Journal of the European Union, 2024. https://eur-lex.europa.eu/eli/reg/2024/2847/oj/eng

[2] European Parliament and Council. Regulation (EU) 2023/1230 on machinery. Official Journal of the European Union, 2023. https://eur-lex.europa.eu/eli/reg/2023/1230/oj/eng

[3] European Parliament and Council. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union, 2024. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

[4] Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. Journal of Management Information Systems. 2007;24(3):45–77. https://doi.org/10.2753/MIS0742-1222240302

[5] ISO. ISO 10218-1:2025, Robotics — Safety requirements — Part 1: Industrial robots. Geneva: ISO, 2025. https://www.iso.org/standard/73933.html

[6] ISO. ISO 10218-2:2025, Robotics — Safety requirements — Part 2: Industrial robot applications and robot cells. Geneva: ISO, 2025. https://www.iso.org/standard/73934.html

[7] ISO. ISO 12100:2010, Safety of machinery — General principles for design — Risk assessment and risk reduction. Geneva: ISO, 2010.

[8] ISO. ISO 13849-1:2023. Safety of machinery — Safety-related parts of control systems — Part 1: General principles for design. 2023. https://www.iso.org/standard/73481.html

[9] IEC. IEC 62061:2021+AMD1:2024+AMD2:2026 CSV, Safety of machinery — Functional safety of safety-related control systems. Geneva: IEC, 2026. https://webstore.iec.ch/en/publication/112847

[10] IEC. IEC 61800-5-2:2016. Adjustable speed electrical power drive systems — Part 5-2: Safety requirements — Functional. 2016. https://webstore.iec.ch/en/publication/24555

[11] ISO. ISO 13849-2:2012. Safety of machinery — Safety-related parts of control systems — Part 2: Validation. 2012. https://www.iso.org/standard/53640.html

[12] ISO. ISO 13850:2015, Safety of machinery — Emergency stop function — Principles for design. Geneva: ISO, 2015.

[13] ISO. ISO 14119:2024, Safety of machinery — Interlocking devices associated with guards. Geneva: ISO, 2024.

[14] IEC. IEC 61496-1:2020. Safety of machinery — Electro-sensitive protective equipment — Part 1: General requirements and tests. 2020.

[15] IEC. IEC 60204-1:2016+AMD1:2021, Safety of machinery — Electrical equipment of machines — Part 1. Geneva: IEC.

[16] ISO. ISO/TS 15066:2016, Robots and robotic devices — Collaborative robots. Geneva: ISO, 2016. https://www.iso.org/standard/62996.html

[17] ISO. ISO/AWI 15066-1. Collaborative Safety – Physical contact with robots — Part 1: Biomechanical thresholds and data. Project under development. https://www.iso.org/standard/91522.html

[18] ISO. ISO 13482:2014. Robots and robotic devices — Safety requirements for personal care robots. 2014. https://www.iso.org/standard/53820.html

[19] ISO. ISO/FDIS 13482. Robotics — Safety requirements for service robots. Draft successor to ISO 13482:2014. https://www.iso.org/standard/83498.html

[20] ISO. ISO 22166-1:2021. Robotics — Modularity for service robots — Part 1: General requirements. 2021. https://www.iso.org/standard/72715.html

[21] ISO. ISO 22166-201:2024, Robotics — Modularity for service robots — Part 201: Common information model. Geneva: ISO, 2024. https://www.iso.org/standard/82334.html

[22] ISO. ISO 22166-202:2025, Robotics — Modularity for service robots — Part 202: Information model for software modules. Geneva: ISO, 2025. https://www.iso.org/standard/84589.html

[23] European Parliament and Council. Regulation (EU) 2017/745 on medical devices. 2017. https://eur-lex.europa.eu/eli/reg/2017/745/oj/eng

[24] IEC. IEC 80601-2-78:2019+AMD1:2024 CSV. Medical electrical equipment — Part 2-78: Particular requirements for basic safety and essential performance of medical robots for rehabilitation, assessment, compensation or alleviation. 2024. https://webstore.iec.ch/en/publication/93014

[25] ISO. ISO 14971:2019, Medical devices — Application of risk management to medical devices. Geneva: ISO, 2019.

[26] IEC. IEC 62304:2006+AMD1:2015, Medical device software — Software life cycle processes. Geneva: IEC.

[27] IEC. IEC 62366-1:2015+AMD1:2020, Medical devices — Application of usability engineering. Geneva: IEC.

[28] IEC. IEC 81001-5-1:2021, Health software and health IT systems safety, effectiveness and security — Security activities in the product life cycle. Geneva: IEC, 2021.

[29] ISO/IEC. ISO/IEC TR 5469:2024, Artificial intelligence — Functional safety and AI systems. Geneva: ISO/IEC, 2024. https://www.iso.org/standard/81283.html

[30] EUR-Lex. Cyber Resilience Act summary and application dates. Updated 22 May 2026. https://eur-lex.europa.eu/EN/legal-content/summary/horizontal-cybersecurity-requirements-for-products-with-digital-elements-cyber-resilience-act.html

[31] IEC. IEC 62443-4-1:2018. Security for industrial automation and control systems — Part 4-1: Secure product development lifecycle requirements. 2018.

[32] IEC. IEC 62443-4-2:2019. Security for industrial automation and control systems — Part 4-2: Technical security requirements for IACS components. 2019. https://webstore.iec.ch/en/publication/34421

[33] European Parliament and Council. Regulation (EU) 2024/1781 establishing a framework for ecodesign requirements for sustainable products. Official Journal of the European Union, 2024. https://eur-lex.europa.eu/eli/reg/2024/1781/oj/eng

[34] European Parliament and Council. Regulation (EU) 2024/1183 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework. Official Journal of the European Union, 2024. https://eur-lex.europa.eu/eli/reg/2024/1183/oj/eng

[35] European Parliament and Council. Regulation (EU) No 910/2014 on electronic identification and trust services for electronic transactions in the internal market, consolidated 18 October 2024. Articles 25, 35 and 41. https://eur-lex.europa.eu/eli/reg/2014/910/2024-10-18/eng

[36] European Commission. COM(2025) 838 final, Proposal for a Regulation on the establishment of European Business Wallets. Brussels, 19 November 2025. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025PC0838

[37] European Parliament. European Business Wallet: legislative train, procedure 2025/0358(COD). Status: tabled; record updated 1 August 2026. https://www.europarl.europa.eu/legislative-train/theme-a-new-plan-for-europe-s-sustainable-prosperity-and-competitiveness/file-european-business-wallet

[38] Object Management Group. Structured Assurance Case Metamodel (SACM), version 2.3. October 2023. https://www.omg.org/spec/SACM/2.3/About-SACM

[39] Adisorn T, Tholen L, Götz T. Towards a digital product passport fit for contributing to a circular economy. Energies. 2021;14(8):2289. https://doi.org/10.3390/en14082289

[40] Jansen M, Meisen T, Plociennik C, Berg H, Pomp A, Windholz W. Stop guessing in the dark: identified requirements for digital product passport systems. Systems. 2023;11(3):123. https://doi.org/10.3390/systems11030123

[41] Koppelaar RHEM, Pamidi S, Hajósi E, et al. A digital product passport for critical raw materials reuse and recycling. Sustainability. 2023;15:1405. https://doi.org/10.3390/su15021405

[42] Denney E, Pai G, Habli I. Dynamic safety cases for through-life safety assurance. In: Proceedings of the 37th IEEE/ACM International Conference on Software Engineering, volume 2. 2015:587–590. https://doi.org/10.1109/ICSE.2015.199

[43] Kelly T, McDermid J. Safety case construction and reuse using patterns. In: Safety of Computer Control Systems 1997 (SAFECOMP '97). Springer; 1997:55–69. https://doi.org/10.1007/978-1-4471-0997-6_5

[44] Lasota PA, Fong T, Shah JA. A survey of methods for safe human–robot interaction. Foundations and Trends in Robotics. 2017;5(4):261–349. https://doi.org/10.1561/2300000052

[45] Merckaert K, Convens B, Nicotra MM, Vanderborght B. Real-time constraint-based planning and control of robotic manipulators for safe human–robot collaboration. Robotics and Computer-Integrated Manufacturing. 2024;87:102711. https://doi.org/10.1016/j.rcim.2023.102711

[46] Abdi J, Al-Hindawi A, Ng T, Vizcaychipi MP. Scoping review on the use of socially assistive robot technology in elderly care. BMJ Open. 2018;8:e018815. https://doi.org/10.1136/bmjopen-2017-018815

[47] German Aerospace Center (DLR). Rollin' Justin: robot system description. https://www.dlr.de/en/rm/research/robotic-systems/humanoids/rollin-justin

[48] Asfour T, Kaul L, Wächter M, et al. ARMAR-6: a collaborative humanoid robot for industrial environments. In: 2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids). 2018:447–454. Author manuscript: https://h2t.iar.kit.edu/pdf/Asfour2018a.pdf

[49] Unitree Robotics. G1: mechanical dimensions, arm/wrist options and configuration notes. Manufacturer product information. https://www.unitree.com/g1/

[50] W3C. RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation, 2014. https://www.w3.org/TR/rdf11-concepts/

[51] W3C. JSON-LD 1.1. W3C Recommendation, 2020. https://www.w3.org/TR/json-ld11/

[52] W3C. CBOR-LD 1.0. Working Draft, 2 September 2026. https://www.w3.org/TR/2026/WD-cbor-ld-10-20260902/

[53] W3C. OWL 2 Web Ontology Language Primer, Second Edition. W3C Recommendation, 2012. https://www.w3.org/TR/owl2-primer/

[54] W3C. Shapes Constraint Language (SHACL). W3C Recommendation, 2017. https://www.w3.org/TR/shacl/

[55] W3C. Verifiable Credential Data Integrity 1.0. W3C Recommendation, 2025. https://www.w3.org/TR/vc-data-integrity/

[56] W3C. Securing Verifiable Credentials using JOSE and COSE. W3C Recommendation, 2025. https://www.w3.org/TR/vc-jose-cose/

[57] W3C. RDF Dataset Canonicalization. Recommendation, 21 May 2024. https://www.w3.org/TR/rdf-canon/

[58] W3C. PROV-O: The PROV Ontology. W3C Recommendation, 2013. https://www.w3.org/TR/prov-o/

[59] W3C. Verifiable Credentials Data Model v2.0. W3C Recommendation, 2025. https://www.w3.org/TR/vc-data-model-2.0/

[60] IETF. RFC 9334. Remote ATtestation procedureS (RATS) Architecture. January 2023. https://www.rfc-editor.org/rfc/rfc9334

[61] IETF. RFC 9711. The Entity Attestation Token (EAT). April 2025. https://www.rfc-editor.org/rfc/rfc9711

[62] Trusted Computing Group. TPM 2.0 Library Specification. Specification resource and revision information. https://trustedcomputinggroup.org/resource/tpm-library-specification/

[63] European Parliament and Council. Regulation (EU) 2016/679 (General Data Protection Regulation). 2016. https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng

[64] NIST. SP 800-207, Zero Trust Architecture. Gaithersburg, MD: National Institute of Standards and Technology, 2020. https://doi.org/10.6028/NIST.SP.800-207

[65] IETF. RFC 8693. OAuth 2.0 Token Exchange. January 2020. https://www.rfc-editor.org/rfc/rfc8693

[66] IETF. RFC 9449. OAuth 2.0 Demonstrating Proof of Possession (DPoP). September 2023. https://www.rfc-editor.org/rfc/rfc9449

[67] W3C. Bitstring Status List v1.0. W3C Recommendation, 2025. https://www.w3.org/TR/vc-bitstring-status-list/

[68] Joint Committee for Guides in Metrology. JCGM 106:2012. Evaluation of measurement data — The role of measurement uncertainty in conformity assessment. BIPM, 2012. https://doi.org/10.59161/JCGM106-2012

[69] NVIDIA. Inside NVIDIA Halos for Robotics: a full-stack functional safety system for physical AI. Developer Blog, 22 June 2026. https://developer.nvidia.com/blog/inside-nvidia-halos-for-robotics-a-full-stack-functional-safety-system-for-physical-ai/

[70] NVIDIA. Halos Outside-In Safety Blueprint. Official source repository. https://github.com/NVIDIA/halos-outside-in-safety

[71] NVIDIA. IGX Thor industrial edge AI platform and IGX software resources. https://www.nvidia.com/en-gb/edge-computing/products/igx/; https://developer.nvidia.com/igx

[72] Infineon Technologies. AURIX microcontrollers, OPTIGA security solutions and robotics functional-safety application information. Accessed 31 August 2026. https://www.infineon.com/aurix; https://www.infineon.com/optiga

[73] NXP Semiconductors. Functional safety and robotics solutions; EdgeLock secure elements. Accessed 31 August 2026. https://www.nxp.com/applications/technologies/functional-safety:FUNCTIONAL-SAFETY; https://www.nxp.com/products/security-and-authentication:SECURITY-AND-AUTHENTICATION

[74] Texas Instruments. Industrial Functional Safety for C2000 Real-Time Microcontrollers. Application brief SWAB013. https://www.ti.com/lit/swab013

[75] STMicroelectronics. X-CUBE-STL functional-safety package and STSAFE secure elements. Accessed 31 August 2026. https://www.st.com/en/embedded-software/x-cube-stl.html; https://www.st.com/en/secure-mcus/stsafe.html

[76] Renesas Electronics. Robotics functional-safety solutions. Accessed 31 August 2026. https://www.renesas.com/en/key-technologies/motor-control-robotics/robot-functional-safety