From Trustworthy AI to Trusted, Operational Agentic Systems

A comparative capability and assurance framework analysis for China, the European Union and the United States

Spherity Research / Academic Paper

Dr. Carsten Stöcker
Spherity GmbH, Germany

Version 1.0 | 26 August 2026
Research evidence cut-off: 26 August 2026

Except where otherwise noted, this work is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). © 2026 Carsten Stöcker, Spherity GmbH.

Central Position

The decisive unit of competition in agentic AI is deployment capability. High-impact deployment requires cumulative competence in AI Governance, Trustworthy AI and Trusted AI.

Trusted AI supplies the operational trust layer: verifiable identity, legal authority, provenance, AI system BOM, AI Service Passport, current status, assurance monitoring, policy enforcement and auditable action evidence at the moment an agent acts.

Abstract

Agentic AI changes the object of assurance. Many established AI deployments primarily generate predictions, classifications or content; an agent can additionally select tools, acquire data, plan, decide and initiate actions with economic, legal or physical effects. This article develops a comparative taxonomy for the institutional and technical capacities that make such action governable and support justified reliance. It distinguishes three cumulative layers. AI Governance defines legitimate purposes, duties, oversight and remedies. Trustworthy AI defines the socio-technical qualities that systems should achieve across their lifecycle. Trusted AI, as defined in this article, comprises the machine-verifiable claims and enforcement mechanisms required for a specified actor, action, context and time. The analysis maps these layers across five control functions and applies a twelve-capability maturity rubric to China, the European Union and the United States. Within that rubric, China exhibits the most coherent emerging national stack for agent identity, credential lifecycle, delegation, tool access, traceability and content labelling. The United States exhibits the strongest combination of test, evaluation, verification and validation, cybersecurity, zero-trust engineering and sector-specific assurance. The European Union combines risk-based law, product and safety regulation, qualified trust services, trusted lists and authoritative registers with proposed European Business Wallet legislation and active pilot infrastructure. Evidence on legal-person credentials and agent powers of attorney supports a provisional, single-author rubric point estimate of 4 for legal-person identity, mandates and trust lists, while agent-specific runtime identity, provenance and action receipts remain priority integration areas. A conjunctive

readiness model and an ordinal assurance gradient identify potential adoption bottlenecks across consumer applications, connected products, government, cross-company processes, regulated industries, critical infrastructure, Industrial AI and Physical AI. The article argues that Europe can convert legal and industrial assets into a distinctive trusted-autonomy infrastructure if legislation, semantics, reference implementations, conformance testing and sector deployment advance through one coordinated programme.

Keywords. agentic AI; AI governance; trustworthy AI; trusted AI; eIDAS 2.0; agent identity; legal-person identity; European Business Wallet; verifiable credentials; verifiable evidence graph; AI Service Passport; authorisation; provenance; TEVV; Agentic Commerce; Industrial AI; Physical AI.

Research questions and contribution

The article addresses four research questions:

- RQ1. Which conceptual boundary separates AI Governance, Trustworthy AI and Trusted AI?
- RQ2. Which institutional, engineering and cryptographic capabilities support agentic AI deployment at increasing levels of autonomy and impact?
- RQ3. How mature are these capabilities in China, the European Union and the United States as of 26 August 2026?
- RQ4. Which capability combinations create a competitive advantage for safe deployment in industrial and regulated markets?

The contribution is a two-dimensional capability model, a transparent ordinal maturity assessment, a mathematically conservative prerequisite model, a comparative scenario analysis, and a normative European action programme. The framework treats jurisdictional maturity as a vector. This choice preserves bottlenecks that an aggregate index would conceal.

Contents

Research questions and contribution **2**

1 Introduction: the AI race is becoming a deployment race **5**

2 A three-layer taxonomy **6**
2.1 AI Governance . . . 6
2.2 Trustworthy AI . . . 6
2.3 Trusted AI . . . 6
2.4 Cumulative relationship and taxonomy boundary . . . 7

3 Methodology **7**
3.1 Comparative design and evidence classes . . . 7
3.2 Capability maturity rubric . . . 8
3.3 Prerequisite model and mathematical audit . . . 8
3.4 Limitations . . . 9

4 Two-dimensional capability model **9**
4.1 Technical trust building blocks . . . 9
4.2 Relationship to the Spherity Trusted AI thesis . . . 10

5 China: coordinated movement toward a national agent trust stack **10**
5.1 AI Governance taxonomy . . . 10
5.2 Trustworthy AI as a standardised property domain . . . 11
5.3 Agent identity, credentials and authorisation . . . 11
5.4 PKI, X.509, DID, VC, revocation and issuer governance . . . 12
5.5 Provenance, labels and runtime evidence . . . 12
5.6 Chinese taxonomy assessment . . . 12

6 European Union: legal trust infrastructure seeking an agent runtime layer **13**
6.1 Governance and Trustworthy AI . . . 13
6.2 eIDAS, trusted lists and the European Business Wallet . . . 13
6.3 WE BUILD, company registers and power-of-attorney semantics . . . 14
6.4 Why the EU score is 4 for legal-person identity, mandates and trust lists . . . 14
6.5 IPCEI-AI and the industrialisation pathway . . . 15
6.6 European integration gaps . . . 15
6.7 European taxonomy assessment . . . 15

7 United States: engineering assurance and enterprise identity **16**
7.1 AI Governance and Trustworthy AI . . . 16
7.2 Agent identity and authorisation . . . 16
7.3 US taxonomy assessment . . . 17

8 Comparative maturity heat map **17**
8.1 Comparative patterns . . . 18
8.2 Taxonomy clarity . . . 19

9 Assurance gradient and comparative scenario analysis **19**
9.1 Formal assurance-demand model . . . 21
9.2 From assurance demand to capability requirements . . . 22
9.3 Infrastructure scenarios and adoption stalls . . . 23
9.4 Scenario analysis by deployment domain . . . 23
9.5 Geography-specific bottlenecks . . . 24
9.6 The European regulatory pathway as adoption infrastructure . . . 25
9.7 Competitiveness and the regulated-industry adoption race . . . 26

10 Trusted infrastructure as a prerequisite for adoption **27**

11 Competitiveness in the global AI race **27**

12 A normative European programme for trusted autonomy **28**
12.1 Delivery sequence 29
12.2 Institutional coordination 29

13 Discussion, limitations and research agenda **29**

A China evidence map **30**

B European Union evidence map **30**

C United States evidence map **31**

D Score rationale and reproducibility record **31**

E Quality review protocol **32**

F Formal assurance model, scoring protocol and worked profiles **33**
F.1 Notation and level of measurement 33
F.2 Assurance-demand anchors 33
F.3 Scoring protocol 34
F.4 Worked assurance profiles 35
F.5 Calibration, uncertainty and validation 35

Declarations **36**

1 Introduction: the AI race is becoming a deployment race

AI policy has traditionally focused on models, datasets and decision support. Agentic systems expand this frame because they combine perception, memory, planning, interaction and execution. An agent can select an API, retrieve evidence, negotiate with another agent, initiate a transaction, modify software, control a connected product or influence an actuator [1, 2]. The relevant governance question therefore extends from output quality to whether a specific action was legitimate, authorised, sufficiently supported, safely executable and attributable.

This shift creates a deployment paradox. Model capability continues to advance, while institutions and firms require assurance before they permit autonomous action in material contexts. Consumer assistants can tolerate reversible inconvenience. Government benefits, financial transfers, medical workflows, industrial control, energy infrastructure and mobility systems require stronger grounds for reliance. This article treats the capacity to combine legal compliance, cybersecurity, functional safety, identity, authorisation, provenance, evaluation and operational evidence as a potentially binding adoption constraint [3, 4].

Independent enterprise evidence supports this claim as a research hypothesis and defines its present limits. In the 2022-23 OECD/BCG/INSEAD survey of 840 AI-adopting enterprises across G7 manufacturing and information and communication technology, 55% of manufacturers and 57% of ICT enterprises reported that privacy, data-protection or data-security concerns had limited AI use during the preceding year. Around 40% reported uncertainty about legal consequences and scarcity of cloud services that guarantee data security and regulatory compliance [5]. An OECD study of EU manufacturing also records legal, privacy, data-quality and system-compatibility concerns among reported reasons for AI non-adoption [6]. These surveys establish perceived barriers across sectors. Causal estimation of shared Trusted AI infrastructure and its relative contribution alongside skills, data quality and organisational capability remain open research tasks. The manuscript therefore treats infrastructure maturity as a testable potential constraint rather than an established general cause.

Existing scholarship offers rich accounts of AI ethics, responsible AI, trustworthy properties and organisational auditing [7–12]. International standards add management systems, risk management, impact assessment and assurance processes [13–18]. Digital identity standards provide reusable cryptographic building blocks [19–25]. These bodies of work remain partly separated. Agentic deployment requires them to operate as one assurance chain.

This article proposes a three-layer taxonomy. AI Governance answers which purposes, actors, duties, controls and remedies apply. Trustworthy AI answers which qualities the system should achieve and how those qualities are engineered and assessed. Trusted AI answers which claims a relying party can verify and enforce for the current action. The third layer is deliberately narrower than the second. It provides an operational trust substrate for evidence-based reliance. Substantive truth, fairness and safety arise from authoritative evidence, evaluation and governance.

The comparative analysis reveals three different development paths. China uses coordinated policy and national standardisation to assemble agent identity and interconnection infrastructure. The United States builds from voluntary risk management, cybersecurity, enterprise identity and TEVV. The European Union builds from fundamental rights, product regulation, conformity assessment and legally meaningful trust services. Each path creates strengths. Each path also leaves integration work at the boundary between legal authority, technical identity, runtime evidence and sector safety.

The strategic claim is normative. Europe should use its regulatory and industrial assets as components of a trusted-autonomy infrastructure. The European Business Wallet, authoritative

company registers, qualified trust services, trusted lists, digital powers of attorney, the AI Act, sector safety law, Data Spaces and Digital Product Passports can support a machine-verifiable chain from legal entity to agent action. Speed matters because technical interfaces and credential semantics become infrastructural lock-in points once global agent ecosystems scale.

Three questions for every material agent action

Governance: Is this action permitted, accountable and subject to effective oversight?

Trustworthiness: Does the system achieve the required qualities for this context and lifecycle state?

Trusted operation: Can the relying party verify the actor, authority, evidence, status, policy decision and action record now?

2 A three-layer taxonomy

2.1 AI Governance

AI Governance is the institutional and organisational system that allocates decision rights, duties, oversight and remedies across the AI lifecycle. Its instruments include legislation, administrative rules, standards mandates, procurement conditions, corporate policies, risk classification, incident reporting, market surveillance and judicial or administrative redress. Governance determines legitimate use and accountable actors. It also assigns authority to approve, suspend or terminate deployment.

A governance taxonomy should separate substantive rules from implementation mechanisms. A legal prohibition, a risk-management duty, a conformity assessment, an internal approval gate and a runtime access policy operate at different levels. Their coordination creates effectiveness. Their conflation creates semantic ambiguity and weakens auditability.

2.2 Trustworthy AI

Trustworthy AI is the broad socio-technical condition under which an AI system earns context-appropriate confidence. Common properties include validity, reliability, safety, security, resilience, privacy, transparency, explainability, accountability, fairness, human oversight and sustainability [7, 10, 13–18, 26, 27]. These properties require lifecycle engineering, organisational governance and empirical evaluation.

Trustworthiness is multi-dimensional and context-dependent. A highly accurate system can remain unsafe under distribution shift. A transparent system can remain insecure. A robust model can operate under invalid authority. The properties therefore require explicit trade-offs, measurable acceptance criteria and a safety or assurance case appropriate to the sector.

2.3 Trusted AI

Trusted AI is the verifiable operational subdomain required for material decisions and actions. It binds machine-verifiable claims about legal and technical actors, delegated authority, system and model version, data and evidence provenance, evaluation state, runtime policy, lifecycle status and executed action. A relying party evaluates these claims at the point of action and produces one of four outcomes: permit, restrict, escalate or deny.

Trusted AI is technology-neutral at the architectural level. A deployment can use public key infrastructure, X.509 certificates, qualified trust services, federated identity, OAuth, workload identities, decentralised identifiers, verifiable credentials, trusted registries, hardware attestation, transparency logs and signed receipts. The governing requirement is evidentiary coherence: the trust anchor, issuer authority, semantic meaning, status mechanism, subject binding and policy effect should align.

Cryptographic verification proves properties such as issuer control, signature integrity, key possession and status under a defined trust framework. Substantive claims such as factual accuracy, fairness, safety and legal sufficiency require authoritative issuance, evaluation, governance and sector-specific evidence. Trusted AI therefore operationalises justified reliance while preserving the independent work of Trustworthy AI.

2.4 Cumulative relationship and taxonomy boundary

Table 1: Proposed taxonomy for AI Governance, Trustworthy AI and Trusted AI.

Layer	Primary question	Representative evidence	Decision function
AI Governance	Legitimate purpose, duties, accountable actors, oversight, remedies	Law, policy, institutional responsibility, enforcement record	Permitted and accountable deployment
Trustworthy AI	Required system qualities and lifecycle performance	Risk assessment, QMS, TEVV, safety case, monitoring evidence	Context-appropriate confidence in system behaviour
Trusted AI	Current identity, authority, evidence, status and action	Credentials, trust lists, provenance, policy decision, runtime attestation, receipt	Verifiable and enforceable reliance at the point of action

The layers are cumulative and non-substitutable as a normative requirement for high-impact deployment. Strong cryptographic identity should never substitute for safety engineering. Strong TEVV should never substitute for lawful authority. Comprehensive regulation should produce executable controls and current evidence if autonomous systems are expected to act across organisational boundaries.

3 Methodology

3.1 Comparative design and evidence classes

The study uses structured comparative analysis across China, the European Union and the United States. The evidence cut-off is 26 August 2026. Five evidence classes are separated to preserve legal and empirical status:

- E1. Enacted law and binding administrative measures.
- E2. Official national or international standards, including their mandatory, recommended, guiding or draft status.
- E3. Official strategies, frameworks, programme documents and legislative proposals.
- E4. Public technical specifications, research and institutional implementation evidence.
- E5. Participant-supplied field evidence from pilots and pre-standardisation work.

The study gives priority to primary sources. Academic literature supports conceptual framing and methodological choices. Public claims about Spherity pilots receive E5 treatment unless an independent public source confirms the same deployment fact. This distinction supports reproducibility and protects the boundary between programme ambition and operational maturity.

3.2 Capability maturity rubric

Table 2: Ordinal capability maturity rubric.

Score	Operational definition
1 Initial	Principles, isolated projects or exploratory discussions.
2 Emerging	Official direction and partial implementation across selected actors or sectors.
3 Established	Coherent framework and repeatable sector practice with recognised technical building blocks.
4 Advanced	Broad legal or technical foundations plus operational pilots and reusable infrastructure.
5 Integrated and assured	Cross-sector and cross-organisational lifecycle operation with conformance, current status and runtime evidence.

Scores are ordinal expert judgements. A one-point difference signals an ordered maturity difference under the rubric and has ordinal meaning. Descriptive layer means serve solely as visual profiles. Readiness decisions preserve the complete capability vector. Each score in this edition is a provisional, single-author point estimate derived from the published rubric and evidence classes. It is therefore a structured comparative hypothesis rather than a measured population parameter or expert-consensus estimate. Transparent construction, alternative specifications and sensitivity testing follow established guidance for cross-country indicators [28].

Confirmatory validation requires independent raters with legal, technical and sector expertise. The protocol should define inclusion rules, evidence weights, conflict resolution and consensus thresholds before scoring. Raters should work independently during the first round, report weighted inter-rater agreement and publish the full distribution of scores. A Delphi extension should state its consensus criterion in advance and preserve minority positions, following established reporting guidance [29].

3.3 Prerequisite model and mathematical audit

The formal model separates assurance demand from capability supply. The assurance-demand vector characterises the inherent action context before controls receive credit. A requirements mapping combines that context with applicable law and sector assurance regimes. Supply is represented at three levels: jurisdictional infrastructure, sector or ecosystem implementation, and organisational capability. Deployment readiness is conjunctive: every relevant capability must reach its required ordinal threshold. Chapter 9 presents Equations (1)-(9); Appendix F defines the notation, anchors and scoring protocol.

The model uses the ordered maturity set $\{1, 2, 3, 4, 5\}$. Requirement vectors additionally use 0 to denote a capability that is irrelevant to a defined use case. Maxima, minima and order comparisons preserve ordinal information. Arithmetic operations across maturity levels require interval-scale assumptions that the evidence does not support. Figure 2 therefore labels arithmetic means as descriptive visual summaries and excludes them from readiness decisions.

The scalar assurance tier is a conservative communication device rather than a risk estimate. Operational approval retains the complete demand vector, the capability-threshold vector and the

bottleneck set. This distinction preserves heterogeneous legal, safety and security requirements that an average score could conceal.

3.4 Limitations

The analysis captures published and supplied evidence at a fast-moving point in time. Chinese standards contain translation and access constraints. EU legislative status remains dynamic. US implementation varies by agency, state and sector. Public documentation reveals programme design more readily than field performance. The score set should therefore be read as a transparent comparative hypothesis, supported by an evidence register and suitable for periodic revision. The current results express one documented application of the rubric. Independent replication, inter-rater testing and sector-level validation remain required before the scores can support statistical inference or claims of expert consensus.

4 Two-dimensional capability model

The capability model uses two dimensions. The vertical dimension contains the three cumulative layers. The horizontal dimension contains five control functions that recur across institutional, engineering and technical systems. The resulting fifteen cells define a complete deployment architecture while preserving conceptual boundaries.

Table 3: Two-dimensional AI governance and trust capability model.

Layer	Rules and accountability	Identity and authority	Evidence, provenance and semantics	Runtime control and enforcement	Assurance, lifecycle and status
AI Governance	Legal applicability, risk tiers and accountable institutions	Responsible actors, roles and decision rights	Documentation, records and data duties	Human oversight, incidents and enforceable controls	Conformity, supervision and remedies
Trustworthy AI	Context properties, values and trade-offs	Accountable owner, operator and affected parties	Model, data and system lineage; impact evidence	Robustness, security, fairness and safety controls	QMS, TEVV, monitoring and safety case
Trusted AI	Machine-readable policy and reliance conditions	Legal person, agent identity, roles and mandate	Signed claims, provenance and semantic schemas	Dynamic authorisation, tool control and action receipts	Issuer trust, status, revocation and runtime attestation

4.1 Technical trust building blocks

Table 4: Technology-neutral building blocks for Trusted AI.

Building block	Core function	Assurance conditions
PKI and X.509	Binds public keys to identified subjects; supports signatures, TLS, device and workload certificates.	Certificate policy, subject semantics, key protection, path validation, expiry and revocation.
Revocation and status	Expresses current validity for certificates, credentials, mandates and system states.	Low-latency status, reason codes, privacy, archived evidence and fail-safe policy.

Table 4 continued

Building block	Core function	Assurance conditions
Trust lists and registries	Identifies recognised issuers, trust service providers, schemas and sector roles.	Governance authority, machine readability, update cadence and cross-border mapping.
OAuth and OIDC	Supports delegated API access and authentication assertions.	Audience, scope, proof of possession, token exchange, sender constraint and agent delegation semantics.
DIDs and VCs	Supports portable identifiers and signed, selectively disclosed claims.	Authoritative issuance, DID method governance, schema interoperability, subject binding and status.
VDR and evidence registries	Publishes identifiers, schemas, issuer metadata, status and governance artefacts.	Integrity, governance, privacy, availability, legal effect and long-term validation.
Semantic schemas	Makes roles, mandates, AI service properties, provenance and receipts machine interpretable.	Controlled vocabularies, versioning, profile governance and policy-test vectors.
Runtime attestation and receipts	Binds execution state, policy decision, evidence and action into auditable records.	Freshness, secure time, workload binding, selective disclosure, retention and independent verification.

These building blocks become Trusted AI infrastructure when governance supplies authoritative issuers, stable semantics, status policy, relying-party rules and conformance tests. A certificate or credential format alone provides syntax. A governed trust framework provides reliance.

4.2 Relationship to the Spherity Trusted AI thesis

Spherity Research separates a legal control plane from an evidence data plane [30–32]. The control plane connects legal-person identity, representation, a bounded mandate, agent or workload identity, current status, policy decision and action receipt. The data plane connects source provenance, validation, freshness, transformations, uncertainty and issuer-bound evidence. The two planes answer different assurance questions. Authority establishes who may act. Evidence establishes what supports the action.

The AI Service Passport is a proposed bridge between these planes. It can bind the service and operator, purpose, version, risk and legal status, evaluation evidence, conformity information, incidents, monitoring state and credential status. At execution time, the relying party can combine an AI Service Passport with a legal-person credential, an agent power of attorney, sector roles, product evidence and runtime context.

This article generalises that thesis into a comparative institutional taxonomy. The maturity model asks whether each geography can issue authoritative claims, express their semantics, verify current status, enforce policy and preserve an auditable action record. The scenario model then identifies the deployment frontier created by each capability combination.

5 China: coordinated movement toward a national agent trust stack

5.1 AI Governance taxonomy

China has a clear policy taxonomy for AI governance at the level of lifecycle responsibility, safety, ethics, content governance, security assessment and sector deployment. The 2021 New Generation Artificial Intelligence Ethics Code requires human well-being, fairness, privacy, controllability,

accountability and improved ethical literacy across management, research, supply and use [33]. The 2023 Global AI Governance Initiative adds explainability, predictability, human control, reviewability, monitorability, traceability and risk-tiered testing [34]. The AI Safety Governance Framework 2.0 organises risks, risk levels and governance measures around a safe, trustworthy and controllable ecosystem [35].

The May 2026 Implementation Opinions on the Standardised Application and Innovative Development of Intelligent Agents create a dedicated governance layer for autonomous systems [36]. They distinguish decisions reserved for the user, decisions made under user authorisation, and decisions assigned to agent autonomy. They also call for behavioural boundaries, permission management, risk classification, lifecycle security, third-party evaluation and verifiable, traceable behaviour in important applications.

This policy family provides a coherent governance direction, although the term Trusted AI has yet to become a formally separated national top-level category. China instead distributes operational trust capabilities across agent interconnection standards, cryptographic guidance, PKI, distributed identity, content labelling, logging and planned audit standards.

5.2 Trustworthy AI as a standardised property domain

GB/T 47507-2026, Artificial Intelligence: Trustworthiness General Rules, was published on 30 April 2026 and entered into effect on 1 August 2026 [37]. Its publication is taxonomically significant. It establishes trustworthiness as an explicit national standardisation domain and connects earlier ethical and safety language to technical rules. The broader programme contains standards for large models, datasets, terminals, privacy, risk, safety evaluation and sector applications. Trustworthiness therefore has institutional recognition, a standards pipeline and a pathway toward testing.

5.3 Agent identity, credentials and authorisation

The GB/Z 185-2026 series creates a seven-part national guiding framework for agent interconnection: overall architecture, identity code, identity management, agent description, discovery, interaction and tool invocation [38–40]. SAMR describes the series as a closed chain from identity through capability description and discovery to interaction and tool invocation. Identity becomes the trust foundation for cross-domain cooperation.

GB/Z 185.3 covers registration and verification, identity accounts, credential management and authentication [39]. Registration evidence includes function, supported tasks, version, system composition, delegated authority and behavioural boundaries. The credential lifecycle covers issuance, update, lock, unlock, expiry and revocation. Authentication can verify a delegation chain from the entrusting party to the agent, after which a relying party combines the authentication assertion with its own authorisation policy.

The architecture distinguishes authentication from authorisation and extends identity controls to agent-to-agent and agent-to-tool interaction. This separation supports attribute-based and policy-based access control. SAMR has announced future work on agent auditing and transactions, together with possible conversion of the identity-code specification into a mandatory national standard [40].

5.4 PKI, X.509, DID, VC, revocation and issuer governance

China has an established domestic electronic-certification and PKI substrate. GB/T 20518-2018 specifies the digital-certificate format, GB/T 19713-2025 specifies an online certificate-status protocol, and the 2024 e-government electronic-certification rules establish supervisory requirements for service providers [41–43]. Agent credentials can anchor in certificate chains. This foundation supports lifecycle controls such as issuance, validity, key management, status checking and revocation.

GB/T 47702-2026 specifies technical requirements for decentralised identity systems. It was published on 25 May 2026 and is scheduled to enter into effect on 1 September 2026 [44]. A June 2026 CAC consultation draft extends distributed digital identity to individuals, public bodies, legal persons, organisations, industrial devices and products [45]. It describes identifiers, verifiable credentials, presentations, authoritative identity institutions, credential status, issuer lists, use-case lists and lifecycle controls. For legal persons, the draft links credential issuance to the unified social credit code authority. It also provides for data-authorisation and business-delegation credentials.

The draft therefore connects DID and VC syntax to authoritative legal-person and industrial identity sources. Its consultation status requires careful interpretation. It signals architecture and policy direction; it carries prospective rather than binding force as of the evidence cut-off.

5.5 Provenance, labels and runtime evidence

China already operates a binding content-provenance layer. The Measures for Labelling AI-Generated and Synthesised Content entered into force on 1 September 2025 and require explicit and implicit labels [46]. Metadata can identify AI-generated status, the service provider and the content. Propagation services add platform and propagation identifiers. GB 45438-2025 supplies the mandatory technical standard [47].

The 2026 Cryptography Application Guide for Generative AI Systems, issued by the Chinese Cryptology Society, provides a broader technical design [48]. It proposes unique agent identifiers and certificates, mutual authentication, dynamic and revocable authorisation tokens, signed plans, signed sensitive instructions, protected tool calls, provenance for retrieval sources, signed runtime and tool logs, secure timestamps and human-signed confirmation for highly privileged actions. Its status is industry guidance. Its scope nevertheless shows a sophisticated operational conception of Trusted AI.

China therefore has two provenance tracks. The first records the origin and propagation of generated content. The second records agent behaviour, delegation, tool access and runtime events. The announced audit and transaction standards could connect these tracks into a more complete evidence graph.

5.6 Chinese taxonomy assessment

China in one sentence

China has explicit taxonomies for AI Governance and AI trustworthiness, plus an emerging operational Trusted AI stack whose components are more mature than its umbrella terminology.

China's comparative strength is coordinated infrastructure formation. Policy, national standards, pilots, open protocols, conformance tooling and prospective mandatory identity requirements

reinforce one another. The principal strategic questions concern legal-person authority across organisations, interoperability beyond domestic trust anchors, independent assurance, privacy-preserving status, and the evidentiary sufficiency of runtime records in regulated and cross-border disputes.

Relative to the Spherity thesis, China already converges on principal-to-agent delegation, credential lifecycle, authentication, tool control and traceability. Spherity places greater emphasis on the preceding legal-person and representation chain, machine-readable powers of attorney, AI Service Passports, evidence graphs and signed action receipts. China's planned audit and transaction standards could narrow this difference quickly [30, 31, 40].

6 European Union: legal trust infrastructure seeking an agent runtime layer

6.1 Governance and Trustworthy AI

The European Union offers the clearest legal distinction between unacceptable, high-risk, transparency-related and general-purpose AI obligations through Regulation (EU) 2024/1689 [49,50]. The governance system combines the European AI Office, national competent authorities, notified bodies, market surveillance and judicial protection. Product safety, cybersecurity, data protection, fundamental-rights and sector law add parallel obligations.

The European Trustworthy AI tradition is explicit. The High-Level Expert Group framed trustworthy AI as lawful, ethical and robust, supported by seven requirements [26]. The AI Act translates parts of this tradition into risk management, data governance, technical documentation, record-keeping, transparency, human oversight, accuracy, robustness and cybersecurity. ISO and CEN-CENELEC work add management systems, risk processes, impact assessment and conformity pathways [13–18, 51].

Europe's strongest capability lies where AI meets regulated products, industrial systems and legal accountability. Functional safety traditions in machinery, automotive, rail, energy and medical devices provide mature methods for hazard analysis, safety integrity, validation and post-market responsibility [52–54]. The integration of probabilistic AI and adaptive agents into these safety cases remains a priority research and standardisation task.

6.2 eIDAS, trusted lists and the European Business Wallet

Regulation (EU) 2024/1183, commonly called eIDAS 2.0, provides a legally recognised trust-services framework for electronic signatures, seals, timestamps, registered delivery, certificates and wallet-based identity [55]. EU trusted lists identify supervised qualified trust service providers and the qualified services they offer. Their status has constitutive legal effect and the lists are available in machine-processable form [56]. In July 2026, the Commission reported an expanding eIDAS Dashboard that includes trusted entities such as PID providers, wallet providers, registrars and registers [57]. This infrastructure gives Europe a distinctive capacity to connect cryptographic verification with regulated issuer status and legal effect.

The Commission's proposal COM(2025) 838 would establish European Business Wallets for companies and public bodies [58]. The Council adopted its negotiating position on 9 June 2026 and stated an objective of political agreement by the end of 2026 [59]. The European Parliament procedure remained at committee stage on 26 August 2026 [60]. This paper therefore treats Q4 2026 adoption as a forward-looking scenario aligned with the Council objective and supplied by

the author. The score reflects current legal foundations and live implementation work, rather than enacted EBW legislation.

The proposed wallet functions include legal-person identity, trusted documents, signatures, seals, timestamps, secure communications and delegation of legal authority. These functions create the components of a legal control plane for agents. A company can establish its identity, issue or receive role and mandate credentials, bind a technical agent to a defined power, and allow a relying party to check issuer trust and current status before execution.

6.3 WE BUILD, company registers and power-of-attorney semantics

WE BUILD is a European large-scale pilot with more than 200 organisations across 28 countries and thirteen use cases [61]. Its company-representation use case explicitly targets secure, verifiable digital powers of attorney and representation rights. The work includes rulebooks and attestation definitions that link natural persons and legal persons to permitted activities [62].

A February 2026 WE BUILD participant non-paper proposes alignment between wallet infrastructure and trusted identities for AI agents [63]. The author group includes representatives from the consortium coordination, Bosch, Bundesanzeiger Verlag, the Netherlands Chamber of Commerce, Ericsson, Google, Spherity and Visa. The document is an expert contribution to programme development. It demonstrates ecosystem convergence around legal-person identity, agent identity and delegated authority.

Bundesanzeiger Verlag publicly describes company identity, trade-register credentials, legal representatives and powers of attorney as components of its European Business Wallet work [64]. Its position as operator of German corporate disclosure and register services supplies an authoritative-data pathway. Spherity reports participant-supplied field evidence in which legal-person credentials from the German company-register context and scoped agent authorisation credentials, described as powers of attorney for agents, are issued for testing [65]. The same evidence shows verification against a verifiable data registry and eIDAS-oriented trust anchors. These field claims are material for maturity assessment and remain separated from independently verified production deployment.

6.4 Why the EU score is 4 for legal-person identity, mandates and trust lists

Table 5: Evidence supporting the EU maturity score of 4.

Evidence status	Capability contribution	Class
Enacted foundation	eIDAS 2.0, qualified trust services, electronic seals and EU trusted lists	E1
Pending legislation	EBW proposal; Council general approach; political agreement targeted by end-2026	E3
Public pilot	WE BUILD legal representation use case, rulebooks and attestation definitions	E4
Authoritative issuer path	Bundesanzeiger company-register data and planned business credentials	E4
Field validation	Reported legal-person credentials, agent PoA credentials, VDR and trust-anchor verification	E5
Industrial programme	IPCEI-AI initiative and a Spherity proposal for Trusted AI as a transversal foundation	E3 / E5

A score of 4 signifies advanced legal foundations, trusted-list infrastructure, authoritative issuers, reusable pilot semantics and live field validation. The score also recognises that ecosystem-wide interoperability, final EBW legislation, harmonised agent-mandate profiles and cross-sector

conformance remain active work. This interpretation follows the rubric and preserves the distinction between advanced maturity and integrated assurance at level 5.

The score is a provisional rubric point estimate. A conservative sensitivity case that excludes participant-supplied field evidence and gives limited credit to pending legislation yields a score of 3. The documented base case retains enacted eIDAS foundations, operational trusted lists, authoritative-issuer pathways and public pilot semantics and yields 4. The resulting plausible range is therefore 3-4. This range expresses model and evidence sensitivity rather than a statistical confidence interval. Independent multi-rater assessment remains the required next validation step.

6.5 IPCEI-AI and the industrialisation pathway

Germany's IPCEI-AI initiative seeks a European industrial AI ecosystem oriented toward industrial needs and deployment [66]. Its current programme status supports strategic mobilisation. It should be distinguished from an approved cross-border IPCEI with settled work packages and state-aid decisions.

Spherity has proposed Trusted AI as a transversal foundation for such a programme [65]. The proposed architecture connects governance and compliance, operational trust infrastructure, functional safety and TEVV, interoperability, open reference implementations and sector industrialisation. This design is participant-supplied programme evidence. Its relevance lies in the systems logic: industrial AI requires identity and authority, evidence provenance, policy enforcement, cybersecurity and functional safety to mature together.

An IPCEI-scale implementation could fund shared infrastructure whose benefits spread across energy, manufacturing, mobility, pharma, finance, defence and government. Candidate common assets include legal-person and agent credential profiles, semantic schemas, trust registries, conformance suites, evidence-graph services, runtime receipts, reference policy engines and sector safety profiles.

6.6 European integration gaps

Europe's legal control plane is ahead of its agent runtime plane. Priority work includes a standard agent or workload identity profile, binding between a wallet-held mandate and an ephemeral workload, status and revocation suitable for machine-speed decisions, common semantic schemas for powers of attorney and AI Service Passports, signed policy-decision evidence, tool-call receipts, data and model provenance, and conformity assessment for the complete chain. The opportunity is structural because these gaps can be filled by extending existing trust and product infrastructures.

6.7 European taxonomy assessment

European Union in one sentence

The European Union has the clearest normative taxonomy for lawful, ethical and robust AI and the strongest legal trust-service foundation, while its operational Trusted AI layer remains distributed across eIDAS, the proposed European Business Wallet, cybersecurity, product and sector-assurance frameworks.

Europe's comparative strength is its capacity to connect normative duties to legally recognised digital evidence. The AI Act and the Trustworthy AI tradition define governance and system-

quality requirements [26, 49, 50]. eIDAS supplies supervised trust services and trusted lists [55]. The proposed European Business Wallet, WE BUILD and authoritative-register work extend this foundation toward legal-person identity, representation and machine-readable mandates [58–64]. This combination creates a credible institutional bridge from compliance requirements to verifiable agent authority.

The remaining taxonomy task is operational integration. A distinct European Trusted AI subdomain should connect legal-person and workload identity, mandate semantics, issuer and credential status, AI Service Passport evidence, runtime attestation, data and model provenance, policy-decision records and signed action receipts. Common profiles and conformance tests should define how these components support reliance at the point of action. This integration would translate Europe's legal and industrial assets into reusable deployment infrastructure for regulated, industrial and cyber-physical AI [51, 65, 66].

7 United States: engineering assurance and enterprise identity

7.1 AI Governance and Trustworthy AI

The United States uses a decentralised governance model. Federal agency policy, sector regulators, state law, procurement rules, voluntary frameworks and private standards allocate responsibility. NIST's AI Risk Management Framework provides the most influential common taxonomy through Govern, Map, Measure and Manage, together with characteristics of trustworthy AI [27].

OMB Memorandum M-25-21 establishes minimum practices for high-impact AI in federal agencies, including pre-deployment testing, impact assessment, ongoing monitoring, risk mitigation and discontinuation duties when risk remains unacceptable [67]. Sector institutions add safety, health, financial, transport, employment and consumer-protection controls. This architecture creates strong implementation guidance while preserving jurisdictional fragmentation.

The US technical advantage is TEVV. NIST's August 2026 TEVV-Athlon draft offers a four-stage, extensible method that explicitly includes agentic systems [68]. NIST is also developing a trustworthy-AI profile for critical infrastructure [4]. Combined with mature cybersecurity, zero-trust and software-supply-chain practice, these assets support high scores for lifecycle assurance, resilience and critical-infrastructure engineering.

7.2 Agent identity and authorisation

NIST launched the AI Agent Standards Initiative in February 2026 to advance industry-led standards, community protocols, agent authentication, identity infrastructure and security evaluations [1]. The NCCoE concept paper on software and AI agent identity and authorisation converts this agenda into an implementation programme [2].

The concept paper brings together OAuth 2.0 and emerging OAuth 2.1 profiles, OpenID Connect, SPIFFE and SPIRE workload identity, SCIM lifecycle management, NGAC policy graphs and NIST zero-trust guidance. It addresses dynamic policy, least privilege, on-behalf-of delegation, human binding, tamper-evident logs, non-repudiation, prompt and data provenance, and controlled tool access. Its initial scope emphasises enterprise environments where organisations retain visibility and control.

This modular approach can evolve rapidly because it builds on deployed cloud and identity infrastructure. It also has a distinct boundary. A workload identity proves which technical process is acting inside a trust domain. It provides limited evidence about the legal person

represented, statutory authority, company-register status, regulated market role or cross-company power of attorney.

7.3 US taxonomy assessment

United States in one sentence

The United States has the strongest engineering vocabulary for trustworthy performance and agent security, while its cross-organisational legal-authority layer remains institutionally fragmented.

The US pathway is well suited to enterprise agents, cloud-native workloads and sector-specific assurance. Scaling into legally consequential cross-company agent commerce requires interoperable organisational identity, authoritative mandates, issuer governance, status and portable evidence semantics. Market-led networks can supply parts of this function. Nationwide legal equivalence and common trust lists would require further institutional coordination.

Adoption policy increasingly recognises the infrastructure constraint. NIST's 2026 analysis of agent-security responses records novel threats and security concerns as barriers to adoption [3]. CISA's careful-adoption guidance translates this concern into staged enterprise controls [69]. America's AI Action Plan identifies slow adoption in large organisations and regulated sectors as a competitiveness problem and links progress to clearer governance, standards and evaluation [70].

8 Comparative maturity heat map

Table 6 and Figure 1 apply the twelve-capability rubric. Scores are provisional, single-author ordinal point estimates based on the evidence available at the cut-off date. They describe jurisdictional systems at a high level and support hypothesis formation and standards-roadmap analysis. Their evidentiary role is limited to comparative hypothesis formation and standards-roadmap analysis. Population measurement, expert consensus and use-case conformity require separate validation. Use-case-specific conformity and safety assessment remains a mandatory deployment gate. Appendix D records score rationales; Appendix F specifies replication, sensitivity and validation requirements.

Table 6: Comparative maturity scores. Scale: 1 initial, 2 emerging, 3 established, 4 advanced, 5 integrated and assured.

Layer	Capability	China	EU	US
AI Governance	Risk-based legal and policy baseline	4	4	2
AI Governance	Institutions, accountability and enforcement	4	3	3
AI Governance	AI-specific standards and conformity pathways	3	2	3
Trustworthy AI	Explicit property taxonomy	3	3	3
Trustworthy AI	Lifecycle governance, QMS and TEVV	3	3	4
Trustworthy AI	AI cybersecurity and resilience	3	3	4
Trustworthy AI	Functional safety, Physical AI and critical infrastructure	3	4	4
Trusted AI	Agent and workload identity	3	1	2
Trusted AI	Legal-person identity, mandates and trust lists	2	4	1

Table 6 continued

Layer	Capability	China	EU	US
Trusted AI	Dynamic authorisation, delegation and status	3	2	3
Trusted AI	Provenance, labels and semantic evidence	3	2	2
Trusted AI	Signed runtime evidence and action receipts	3	2	2

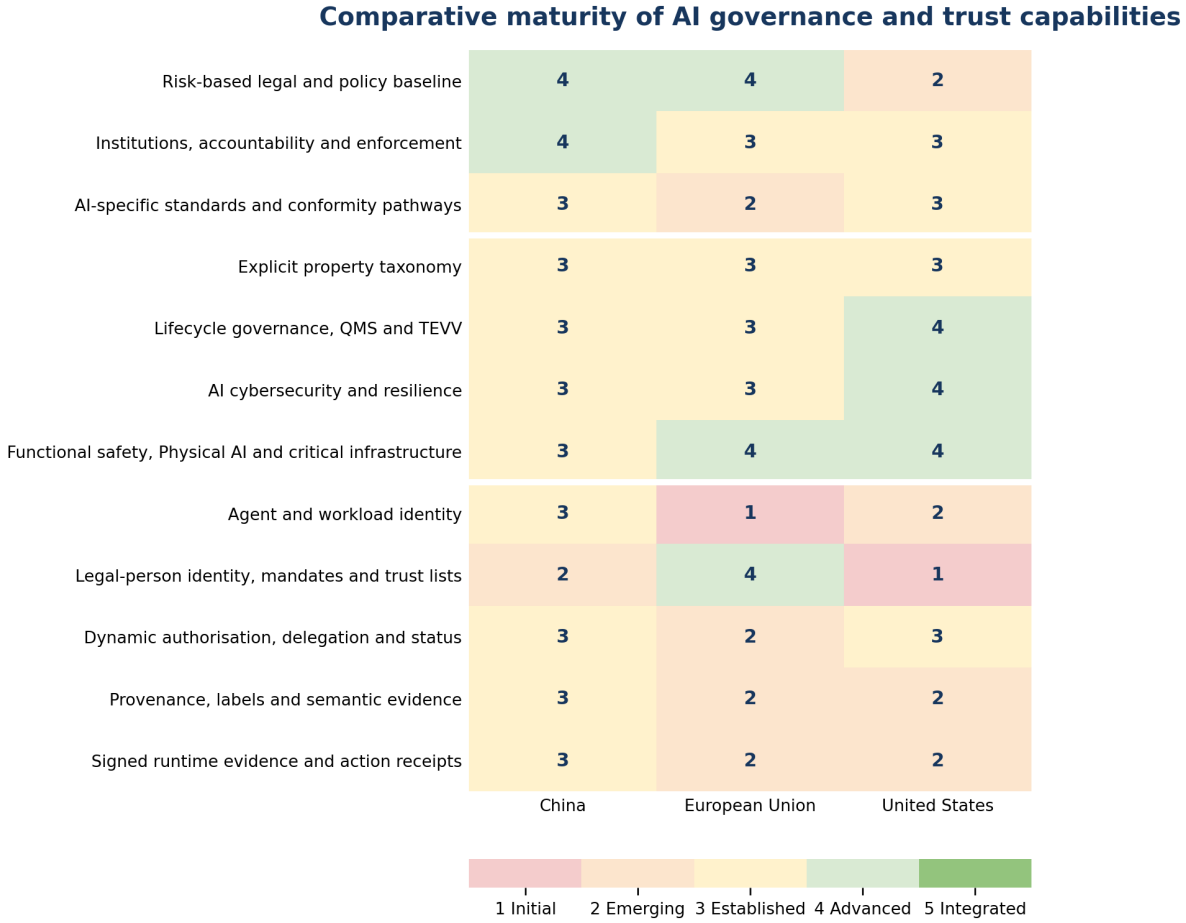

Figure 1: Comparative maturity heat map across twelve capabilities.

Source note. Single-author, rubric-based assessment using the evidence register in Appendices A-D. Scores are provisional ordinal point estimates; the EU legal-person score has a sensitivity range of 3-4.

8.1 Comparative patterns

Table 7: Comparative patterns derived from the capability vectors.

Geography	Comparative strength	Priority integration area
China	Coordinated governance and national standards; agent identity and interconnection; content labels; emerging DID and credential infrastructure.	Legal-person authority and cross-border reliance; independent runtime assurance; mature audit and transaction standards.
European Union	Risk-based law, legal-person identity, qualified trust services, trusted lists, sector safety and cross-border mandate pilots.	Agent workload identity; runtime authorisation binding; semantic evidence profiles; action receipts and continuous assurance.

Table 7 continued

Geography	Comparative strength	Priority integration area
United States	TEVV, cybersecurity, zero trust, cloud workload identity, modular authorisation standards and critical-infrastructure practice.	Portable legal-person identity; authoritative cross-company mandates; common issuer governance and trust-list semantics.

8.2 Taxonomy clarity

China has the clearest emerging institutional separation between general trustworthiness standards and agent-specific operational standards, even though a formal Trusted AI umbrella category remains emergent. The EU has the clearest normative taxonomy for trustworthy, lawful and robust AI, together with the strongest legal trust-service vocabulary. Its operational Trusted AI category remains dispersed across eIDAS, EBW, cybersecurity, product and pilot work. The United States has the clearest engineering taxonomy through the AI RMF, cybersecurity and TEVV, together with an emerging agent-identity programme. Its legal-person trust taxonomy remains sectoral and network-specific.

9 Assurance gradient and comparative scenario analysis

The comparison becomes operational when a deployment team can translate a use case into assurance demand and compare that demand with available governance, engineering and trust infrastructure. Figure 3 orders deployment contexts from A1, informational sandbox use, to A5, action affecting industry, essential services or critical infrastructure. The gradient ranks required assurance. Model intelligence, economic value and novelty sit outside this ranking.

The colour transition expresses an ordinal increase. Each tier is a cumulative control envelope. A use case enters a higher tier when any of seven dimensions reaches that level. A low-autonomy system can therefore require A5 assurance through rights or safety impact. Human approval can reduce autonomy exposure, while identity, evidence, cybersecurity, due-process and safety duties continue to follow the context. The full vector remains the primary analytical object. The tier is a communication device.

Table 8: Assurance tiers and deployment contexts.

Tier	Deployment class	Representative contexts	Minimum assurance emphasis
A1	Information and sandbox	Recommendations, drafting, experimentation	Basic access control, disclosure and evaluation
A2	Low-risk consumer assistance	Reversible bookings, shopping support, personal organisation	User identity, consent, bounded tools, logs and recovery
A3	Enterprise and connected products	Internal workflows, maintenance, home and vehicle functions	Workload identity, least privilege, cybersecurity, product evidence and monitoring
A4	Regulated and cross-organisational action	Government, finance, health, B2B commitments and sensitive processes	Legal-person identity, mandate, current status, policy evidence, TEVV and auditability
A5	Industrial, Physical AI and critical infrastructure	Industrial control, energy, transport, robotics and safety-critical autonomy	Independent assurance, functional safety case, runtime attestation, provenance, fail-safe control and incident evidence

Table 8 translates the gradient into representative deployment classes and minimum areas of emphasis. The listed controls are cumulative. An A4 system inherits the access control, recovery, workload identity and monitoring disciplines associated with lower tiers, then adds legal-person

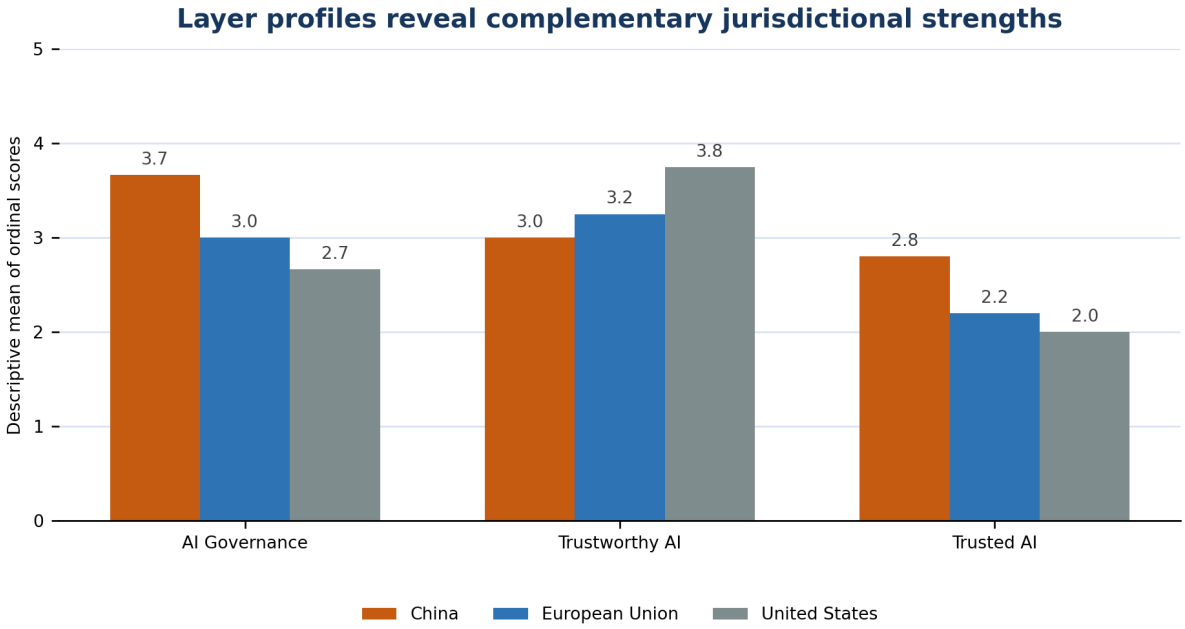

Figure 2: Descriptive layer profiles.

Source note. Means summarise ordinal scores solely for visual comparison. Readiness decisions use the complete capability vector.

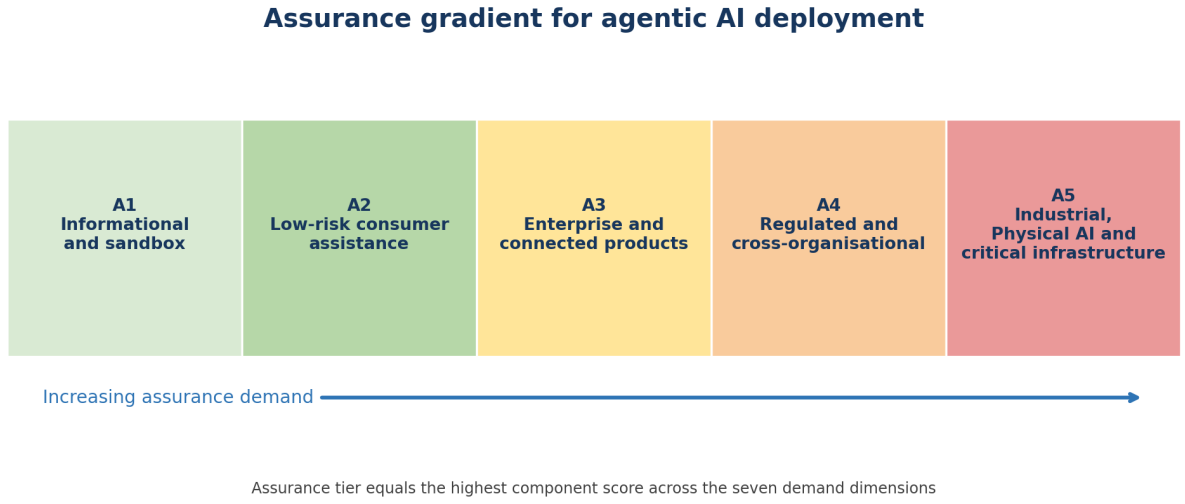

Figure 3: Assurance gradient for agentic AI deployment.

authority, current mandates, policy evidence and regulated auditability. An A5 system adds independent assurance, functional-safety reasoning, verified runtime state, provenance and fail-safe behaviour. The examples guide initial classification. A sector-specific hazard analysis and the applicable legal framework determine the final requirement profile.

9.1 Formal assurance-demand model

The assurance-demand model begins with the action envelope: the agent purpose, users, counterparties, tools, data, assets, value limits, physical interfaces, escalation conditions and operating period. Scores describe the inherent deployment context before controls receive credit. This sequencing prevents a strong control in one domain from suppressing the requirement created by another domain.

For each use case u, define the assurance-demand vector a_u on the common ordered scale $\mathcal{L} = \{1,\ 2,\ 3,\ 4,\ 5\}$:

$$a_u = (\alpha_u, \theta_u, \iota_u, \varphi_u, \delta_u, \rho_u, \sigma_u) \in \mathcal{L}^7, \qquad \mathcal{L} = \{1, 2, 3, 4, 5\} \tag{1}$$

The components represent autonomy (alpha), authority (theta), irreversibility (iota), physical impact (phi), data sensitivity (delta), rights impact (rho) and systemic reach (sigma). Each component uses anchored ordinal judgements calibrated to the same five assurance classes. The anchors express order and carry no assumption of equal numerical distance between adjacent levels. The scalar communication tier is the lattice join, or upper envelope, of the seven components:

$$A_{\text{required}}(u) = \max\{\alpha_u, \theta_u, \iota_u, \varphi_u, \delta_u, \rho_u, \sigma_u\} \tag{2}$$

Appendix F specifies operational questions and anchors for Levels 1, 3 and 5. Levels 2 and 4 represent reasoned intermediate states. The assessor records a factual rationale for every score and preserves the full vector alongside the tier.

Equation (2) is deliberately non-compensatory. A safety-critical physical effect remains binding despite extensive logging. Legal authority remains binding despite strong cybersecurity. One high-impact dimension therefore retains its force alongside several low-scoring dimensions. This property reflects legal and safety settings in which one applicable duty or hazard can determine the required control class. The maximum also respects ordinal measurement. For every strictly increasing recoding h of the scale, $h(max a_{ui}) = max\ h(a_{ui})$. The construction is order-equivariant rather than numerically invariant. A weighted sum would require defensible interval-scale assumptions, commensurability across heterogeneous harms and empirically grounded weights. The current evidence supports ordered anchors and threshold decisions; cardinal aggregation requires a stronger empirical foundation.

The upper envelope is conservative and transparent. It also compresses information. Two use cases can both receive A5 while presenting different engineering problems. A high-impact public-benefits system may reach A5 through rights impact. An industrial control agent may reach A5 through physical impact and systemic reach. Their capability requirements therefore differ. Decision-makers should publish the vector together with the tier, record the evidence supporting each component and conduct domain-specific interaction analysis. Correlated dimensions deserve particular attention. High autonomy combined with broad systemic reach can accelerate propagation. High authority combined with weak reversibility can increase legal and financial exposure. Physical impact combined with stale state evidence can invalidate a previously sound safety case.

9.2 From assurance demand to capability requirements

Assurance demand and infrastructure maturity occupy different analytical spaces. The seven-dimensional vector describes the use case. The twelve-capability maturity vector in Chapter 8 describes jurisdictional supply. A policy mapping connects these spaces. Let l_u denote the applicable legal and regulatory regime and s_u the applicable sector, cybersecurity and functional-safety regime. The required-capability vector belongs to $\mathcal{L}_0^{12}$, where $\mathcal{L}_0 = \{0,\ 1,\ 2,\ 3,\ 4,\ 5\}$ and 0 denotes an irrelevant capability:

$$r_u = g(a_u, l_u, s_u) \in \mathcal{L}_0^{12} \tag{3}$$

The mapping g translates the action context and applicable regimes into a threshold for each of the twelve capabilities. Inputs can include the EU AI Act classification and provider or deployer duties, eIDAS reliance conditions, the proposed EBW framework, data protection, product and cybersecurity law, professional obligations, procurement conditions and sector safety rules. A material change in law, intended purpose, authority, tools, data, deployment environment or hazard analysis therefore requires a fresh mapping. The function is a documented policy mapping rather than an estimated statistical relationship.

For capability k, three independently assessed sources can establish the threshold: requirements arising from the agent context, from applicable law and from the sector assurance regime. The strongest applicable requirement governs:

$$r_{uk} = \max\{r_{uk}^{\text{agent}}, r_{uk}^{\text{law}}, r_{uk}^{\text{sector}}\} \tag{4}$$

Capability supply is equally layered. Let c_{jk} denote jurisdictional maturity, x_{dk} the maturity of the relevant sector or deployment ecosystem d, and o_{ok} the implemented capability of organisation o. The effective capability is the weakest of these three complements:

$$e_{jdok} = \min\{c_{jk}, x_{dk}, o_{ok}\} \tag{5}$$

This weakest-link construction expresses a specific institutional claim. Enacted law and national trust services create limited deployment value when a sector lacks recognised semantic profiles or relying-party rules. A mature sector framework creates limited value when an organisation lacks implementation, governance or operational evidence. Organisational excellence remains bounded when authoritative issuers, status services or legal recognition are unavailable.

Readiness is a componentwise threshold relation, and the bottleneck set contains every capability deficit:

$$\text{Ready}(j, d, o, u) \iff e_{jdok} \geq r_{uk} \quad \text{for every } k \text{ with } r_{uk} > 0 \tag{6}$$

$$B(j, d, o, u) = \{k \in \{1, \ldots, 12\} : r_{uk} > 0 \wedge e_{jdok} < r_{uk}\} \tag{7}$$

Equations (4)-(7) are non-compensatory and use ordinally admissible operations. Excellent model performance cannot compensate for an expired mandate, an unavailable status service, an exceeded tool scope, missing provenance or an unverifiable runtime state. The bottleneck set also supports policy prioritisation: public investment should target deficits shared across many high-value use cases, while organisations should target deficits within their implementation boundary. Appendix F provides the operational anchors and a reproducible scoring protocol.

9.3 Infrastructure scenarios and adoption stalls

Table 9: Comparative infrastructure scenarios and adoption frontiers.

Scenario	Trust infrastructure condition	Adoption frontier	Next binding constraint
S1 Local trust	Local accounts, API keys, platform logs and bilateral policy	A1 and selected A2 uses	Cross-company delegation, regulated reliance and portable accountability
S2 Enterprise trust	Workload identity, OAuth/OIDC, lifecycle IAM, policy engine and tamper-evident logs	A2-A3 within controlled organisations	Legal-person authority, cross-border acceptance and external issuer status
S3 Legal cross-organisational trust	Authoritative legal-person credentials, digital mandates, trust lists, status, semantic profiles and signed decisions	A3-A4 across firms and public bodies	Independent runtime assurance, physical-state evidence and sector safety integration
S4 Assured cyber-physical trust	S3 plus runtime attestation, evidence graphs, continuous TEVV, safety cases, fail-safe controls and signed action receipts	A4-A5 regulated, industrial and critical uses	Residual sector hazards, systemic concentration and geopolitical trust boundaries

Table 9 is a supply-side ladder. Its S1-S4 scenarios describe combinations of trust infrastructure, while A1-A5 describe demand. The labels therefore have independent meanings. S3 and A3 serve independent analytical roles. S3 can support many A4 uses when its legal-person credentials, mandates, trust lists, status services, semantic profiles and signed policy decisions satisfy the relevant thresholds. Selected A5 uses may still require S4 capabilities because cyber-physical assurance depends on verified operational state, independent safety evidence, continuous evaluation and fail-safe execution.

The adoption frontier is the highest class of use cases for which all mandatory capabilities are available in a defined jurisdiction, organisation and sector. It is a frontier rather than a forecast. Commercial demand, usability, liability allocation, insurance and organisational change also shape adoption. The model isolates a narrower question: whether a relying party can justify permission for the agent to act. Adoption stalls at the first binding capability deficit. S1 supports experimentation because bilateral controls can contain exposure. S2 supports enterprise productivity because workload identity and policy remain within an organisational trust domain. S3 enables portable legal reliance across organisations. S4 extends that reliance into systems whose physical or systemic consequences require independent and continuously refreshed assurance.

9.4 Scenario analysis by deployment domain

Table 10: Where adoption stalls under incomplete Trusted AI infrastructure.

Domain	Tier	Binding prerequisite set	Adoption stall
Consumer applications	A2	Identity, consent, disclosure, bounded tools, transaction limits, recovery	Adoption concentrates in reversible assistance; material purchasing and financial delegation remains human-approved.
Connected products	A3	Product identity, software state, cybersecurity, user authority, update status, telemetry provenance	Autonomous control remains bounded when current device state and product safety evidence are weak.
Government	A4	Legal authority, due process, identity, mandate, records, explanation, appeal and audit	Material decisions remain advisory or require official approval when authority and evidence chains are incomplete.

Table 10 continued

Domain	Tier	Binding prerequisite set	Adoption stall
Cross-company B2B	A4	Legal-person identity, market role, agent mandate, scope, value limit, status and receipt	Negotiation can scale while legally binding commitment remains constrained by representation uncertainty.
Sensitive and regulated industries	A4	Sector licence, professional role, privacy, TEVV, policy evidence, provenance and supervision	Deployment remains bounded to decision support when reliance evidence lacks regulated acceptance.
Critical infrastructure	A5	Zero trust, functional safety, independent assurance, runtime attestation, fail-safe control and incident forensics	Autonomous actuation remains restricted to tightly bounded envelopes until a complete safety and security case exists.
Industrial AI and Physical AI	A5	Machine and component identity, digital product evidence, sensor provenance, authority, safety state and action receipt	Pilot value appears early; cross-site and cross-company autonomy waits for interoperable evidence and liability allocation.

Table 10 applies the demand and supply models to seven deployment domains. The tier column provides a representative baseline. Each concrete deployment still receives its own vector. Consumer applications often remain at A2 when actions are reversible and transaction limits are low. A consumer agent can move to A4 or A5 when it controls substantial funds, handles health decisions or affects a safety-relevant connected product. Government services often require A4 controls for authenticated, reviewable administrative workflows. A system that materially determines benefits, liberty or access to essential services can reach A5 through the rights-impact dimension even when physical impact remains at level 1.

The final column describes a recurrent stall mechanism. Model capability generates pilot value before reliance infrastructure reaches the required threshold. A procurement agent can discover suppliers, draft terms and recommend a transaction under S2. Binding acceptance across firms requires an authoritative legal-person identity, a valid agent mandate, scope and value limits, current status and a durable receipt. An industrial agent can optimise a simulation under S2 or S3. Direct actuation requires machine identity, trustworthy sensor state, safety constraints, verified software state, fail-safe controls and incident evidence. The gap between recommendation and authorised execution therefore marks the central adoption boundary for agentic AI.

9.5 Geography-specific bottlenecks

China can move rapidly from S1 to S2 and toward S3 because national agent identity, interconnection, credentials, labels and pilots share a coordinated policy direction [36, 38–40, 44–48]. Cross-organisational legal-person delegation and independent audit form the main bottleneck set for regulated and cross-border reliance.

The European Union already holds an advanced S3 foundation for legal identity, issuer trust and mandates [30, 51, 55, 58–64, 66]. Agent workload identity, machine-speed status, policy receipts and cyber-physical evidence form its main bottleneck set. This pattern creates a short path to high-assurance deployment if technical integration accelerates.

The United States operates strongly at S2 and in selected S4 sectors because cloud identity, cybersecurity, TEVV and sector assurance are mature [1–4, 27, 67–70]. Portable legal-person identity, authoritative delegation and common cross-network issuer governance form the main bottleneck set for broad agent commerce.

9.6 The European regulatory pathway as adoption infrastructure

Europe can use regulation as an adoption architecture when legal duties are converted into reusable, machine-verifiable capabilities. The AI Act supplies a risk-based governance frame and, for high-risk systems, requirements concerning risk management, data governance, technical documentation, logging, human oversight, accuracy, robustness, cybersecurity, quality management, conformity assessment, registration and post-market monitoring [49, 50]. These duties define parts of the Trustworthy AI evidence set. eIDAS and the European Digital Identity Framework supply supervised trust services, qualified certificates and attestations, electronic signatures, seals and time stamps, together with Member State trusted lists whose entries carry constitutive legal significance [55–57]. These capabilities supply part of the Trusted AI validation layer.

The proposed European Business Wallet framework adds a route for legal persons to hold and present authoritative organisational evidence and to execute digital business processes [58–60]. WE BUILD develops company-representative and power-of-attorney semantics in a large-scale pilot [61–63]. Bundesanzeiger Verlag has begun authoritative business-identity and representation work for field testing [64]. Spherity field evidence reports legal-person credentials and agent-authorisation credentials issued in pilot settings [30, 32, 65]. IPCEI-AI can connect these legal and technical capabilities to shared industrial infrastructure [66]. Their combined maturity supports the score of 4 for legal-person identity, mandates and trust lists. Legislative proposals and pilots retain their stated evidence classes, while enacted eIDAS trust services provide the current legal foundation.

The viable European path is therefore cumulative. The AI Act can state which controls and records a deployment requires. eIDAS can establish supervised issuers, legally recognised signatures, seals and time evidence. EBW can provide an organisational presentation and transaction layer. Company registers can issue authoritative legal-person and representation attributes. Agent power-of-attorney credentials can express purpose, scope, value, duration, geography, delegation depth and approval conditions. Trust lists and status services can establish current issuer and credential validity. Policy engines can verify these claims at the point of action. Signed decision and execution receipts can preserve the resulting assurance evidence.

This path becomes economically significant when it replaces repeated bilateral due diligence with interoperable verification. Let all terms below be measured as expected discounted monetary values over a common decision horizon. Let τ_u denote the expected time to authorised deployment of use case u. A simplified adoption surplus is:

$$V_{\text{adopt}}(u, \tau_u) = \mathbb{E}[PV(B_u \mid \tau_u)] - PV(C_{\text{AI},u} + C_{\text{assurance},u} + C_{\text{integration},u}) - \mathbb{E}[PV(L_{\text{residual},u} \mid \tau_u)] \tag{8}$$

Expected present-value benefit $E[PV(B_u \mid \tau_u)]$ includes productivity, quality, speed and new service value after deployment timing is incorporated. $C_{\text{AI,u}}$ covers the technical system; $C_{\text{assurance,u}}$ covers assessment, evidence production and verification; and $C_{\text{integration,u}}$ covers connection to identity, data, tools and sector systems. $E[PV(L_{\text{residual,u}} \mid \tau_u)]$ captures expected safety, security, legal and operational loss after controls. Deployment delay enters the timing of benefits and residual losses through τ_u; a separate delay-cost term would risk double counting unless it represented an independently measured adjustment.

The formal deployment rule combines the ordinal prerequisite with the monetary surplus:

$$\text{Adopt}(j, d, o, u) = 1 \iff \text{Ready}(j, d, o, u) \wedge V_{\text{adopt}}(u, \tau_u) > 0 \tag{9}$$

Equation (9) is a normative decision rule for full deployment. It preserves legal, security and safety thresholds as hard constraints. A positive economic surplus cannot override a failed threshold. When the readiness condition fails, the operational decision can still be to restrict the action envelope, require human approval, escalate the decision or defer deployment. The monetary expression complements the ordinal readiness test and requires empirical calibration before investment appraisal.

The independent survey evidence identifies where empirical calibration should begin. Reported concerns about privacy, security, legal consequences and compliant cloud services justify their inclusion in assurance expenditure, integration expenditure, residual loss and deployment-time estimates [5, 6]. The surveys provide prevalence measures rather than causal parameters. A future evaluation should compare otherwise similar deployments that use mature shared identity, status, mandate and evidence services with deployments that rely on bilateral or fragmented controls, while controlling for sector, firm size, data maturity, skills, system complexity and regulatory intensity. The central empirical hypothesis is that reusable trust infrastructure reduces time to authorised deployment and recurring verification cost after accounting for its fixed implementation cost. The present paper specifies that hypothesis and its measurement structure and leaves the treatment effect for future estimation.

Regulation can initially increase fixed compliance expenditure because organisations must build controls and evidence. Harmonised infrastructure can then lower recurring assurance and integration costs across transactions, counterparties and Member States. Current status, common semantics and recognised trust anchors can reduce uncertainty and residual loss. Conformance profiles and reusable evidence can shorten deployment cycles. The economic effect depends on implementation quality. Fragmented national profiles, manual evidence, slow status services and incompatible schemas preserve high transaction costs. Common machine-readable profiles, open test suites and mutual recognition create scale effects.

9.7 Competitiveness and the regulated-industry adoption race

The adoption race concerns the conversion of model capability into authorised economic action. A1 and A2 consumer software can scale within platform controls. Regulated industry, government, connected products and critical infrastructure require evidence of authority, compliance, security, safe operational state and attributable execution. Under the model, trusted infrastructure can therefore shape how rapidly model capability enters high-value production. The magnitude of this effect remains an empirical question.

China can reduce domestic coordination costs through national agent standards and identity infrastructure. The United States can combine frontier models with mature cloud identity, cybersecurity and TEVV inside enterprise and sector trust domains. Europe has a potential advantage in cross-organisational legal reliance and industrial assurance. Company registers, eIDAS trust services, trusted lists, wallets, product conformity and functional-safety expertise can form an integrated fabric across firms and borders.

Europe requires rapid integration from legal person to representative, mandate, agent, policy decision and action receipt, joined to evidence from product, model, data and sensor state. EBW, WE BUILD, IPCEI-AI, Data Spaces, Digital Product Passports, AI Factories and sector test facilities can supply the components. Common schemas, low-latency status and revocation, open implementations and procurement-backed conformance can create scale. The resulting permission to automate can accelerate manufacturing, energy, mobility, pharmaceuticals, finance, public administration and connected products. Timely delivery can make legal certainty, safety engineering and cross-border trust a competitive capability for Europe's industrial core.

10 Trusted infrastructure as a prerequisite for adoption

Adoption depends on justified reliance. A consumer requires confidence that an agent respects consent and financial limits. A company requires evidence that a counterparty agent represents a valid organisation and holds a current mandate. A regulator requires auditable compliance. A safety engineer requires verified configuration, hazard controls and operational state. An insurer requires attributable events and bounded exposure. A court requires durable evidence and recognised authority.

These stakeholders create a cumulative prerequisite. Compliance defines legitimate operation. Security protects identities, credentials, data, tools and execution paths. Functional safety controls hazardous effects and supports fail-safe behaviour. Trusted AI connects these domains through current, machine-verifiable evidence. Once the assurance tier rises, human review alone becomes too slow and too costly to support agent-speed interaction. Automated verification becomes an economic infrastructure requirement.

The model generates a testable adoption-stall sequence. At S1, pilots demonstrate model capability. At S2, firms gain internal productivity. At S3, organisations can delegate material cross-company and public-sector actions. At S4, autonomous systems can enter industrial and cyber-physical environments under independent assurance. Investment in foundation models expands the supply of intelligence. Investment in trusted infrastructure can expand the set of contexts in which institutions permit its use. The relative contribution of each investment remains sector- and use-case-specific.

Normative deployment principle

Every material agent action should carry sufficient evidence to answer five questions: which legal person is accountable; which agent instance acted; which current mandate applied; which evidence and policy supported the decision; and which signed record proves the executed result.

Safety-critical actions should add verified system state, functional-safety constraints, independent assurance and fail-safe execution evidence.

11 Competitiveness in the global AI race

The global AI race has at least four production factors: models, compute, data and deployment permission. The fourth factor is increasingly decisive in sectors where an agent can bind an organisation, move money, affect rights, operate machinery or change infrastructure. Jurisdictions that make high-assurance deployment cheaper, faster and legally predictable can convert technical capability into productivity sooner.

China's advantage is coordination. National policy, standards, pilots, open technical infrastructure and sector deployment can move in a common direction. This structure can reduce integration cost and produce rapid domestic scale. It can also export protocol and semantic choices through products, platforms and standards participation.

The United States gains from frontier models, cloud platforms, deep cybersecurity capability, enterprise identity infrastructure and a strong evaluation ecosystem. Its market-led modularity encourages rapid experimentation. Cross-company legal reliance remains a federation problem, which large platforms and sector networks may solve through proprietary trust domains.

Europe's position is frequently described through regulation and model-market weakness. Independent statistics make the scale of its industrial base more precise. Eurostat reports that the

EU business economy generated EUR 10.5 trillion in value added in 2023; industry accounted for 29% of that value added and 21% of employment [71]. In 2024, extra-EU goods exports reached EUR 2.583 trillion. Machinery and vehicles contributed EUR 1.013 trillion, or 39.2%, and chemicals contributed EUR 560 billion, or 21.7% [72]. These figures establish the scale of physical, regulated and cross-border deployment contexts. They measure economic scale rather than AI readiness or future productivity. Europe's manufacturing, energy, mobility, chemical, pharmaceutical, financial and public-sector systems nevertheless create a large domain in which lawful authority, product conformity, safety engineering, trusted data and accountable cross-company processes shape deployment.

Trusted autonomy can become a European unique selling proposition. The product is broader than a compliant AI model. It is an interoperable system in which legal identity, mandate, licence, product state, data provenance, AI service evidence, authorisation and action receipt can be verified across organisational borders. This proposition aligns with Europe's institutional structure and industrial core.

12 A normative European programme for trusted autonomy

Europe should pursue a coordinated programme with eight workstreams. The programme should connect legislative implementation, standards, open infrastructure, pilots, conformity assessment and industrial procurement.

1. Agent identity profile. Define interoperable identities for persistent agents, ephemeral instances and workloads. Bind software, model, configuration, operator and execution environment. Align PKI/X.509, SPIFFE, DIDs, VCs and wallet protocols.

2. Legal-person and mandate semantics. Finalise EBW profiles for legal-person identity, representation, powers of attorney, role, purpose, value limit, geography, duration, delegation depth and human approval. Publish machine-readable policy test vectors.

3. European trust and status fabric. Extend trusted-list concepts to recognised credential issuers, schema authorities, status services, agent-service issuers and sector assurance bodies. Support rapid revocation, archived validation and privacy-preserving queries.

4. AI Service Passport and evidence graph. Standardise service identity, version, purpose, risk status, evaluation evidence, model and data provenance, incidents, monitoring state and issuer authority. Connect AI evidence with Digital Product Passports and Data Spaces.

5. Point-of-action enforcement. Define a policy-decision interface that verifies identity, mandate, purpose, audience, context, status and risk. Produce signed permit, restrict, escalate or deny evidence and bind it to the action.

6. Runtime and physical assurance. Integrate remote attestation, secure time, machine identity, sensor provenance, functional-safety constraints, human override, fail-safe behaviour and signed actuation receipts.

7. Conformance and sector profiles. Create open reference implementations, adversarial test suites and certification profiles for government, finance, health, energy, mobility, manufacturing and connected products.

8. Industrial scale and international mapping. Use WE BUILD, IPCEI-AI, AI Factories, Data Spaces and public procurement to create demand. Map European credentials and assurance levels to Chinese and US trust domains through explicit gateway policy.

12.1 Delivery sequence

Table 11: Indicative European delivery sequence.

Horizon	Priority outputs	Deployment outcome
0-12 months	Common vocabulary, agent identity profile, EBW mandate schema, trust-list extension design, reference receipt, three sector test beds	Interoperable proof of legal entity, agent and scoped authority
12-24 months	Conformance suites, qualified issuer and status profiles, AISP and evidence graph, procurement clauses, industrial safety integration	Repeatable A4 deployment across B2B, government and regulated sectors
24-48 months	Cross-border production, independent assurance network, cyber-physical profiles, international trust mapping and PQC migration	Scalable A5 deployment in industrial and critical systems

12.2 Institutional coordination

The Commission, Member States, ENISA, the European AI Office, eIDAS cooperation bodies, CEN-CENELEC JTC 21, ETSI, sector regulators, notified bodies, company registers, trust service providers, industrial firms and research organisations should maintain one shared architecture map. WE BUILD can validate business identity and mandate semantics. IPCEI-AI can finance shared industrial infrastructure. AI Factories and Testing and Experimentation Facilities can integrate TEVV. Data Spaces and Digital Product Passports can supply provenance. Public procurement can create early demand and open conformance requirements.

13 Discussion, limitations and research agenda

The comparison supports a central pattern: governance traditions shape technical trust architecture. China translates coordinated policy into national agent standards. The United States translates cybersecurity and enterprise practice into modular identity and TEVV. Europe translates legal effect and regulated-market institutions into wallet and trust-service infrastructure. International convergence is therefore likely at the protocol layer, while trust anchors, issuer governance and legal semantics remain jurisdiction-specific.

Trusted AI should develop as an interoperability discipline between law, identity, security, safety and AI assurance. Its objects are claims, credentials, policies, status, evidence paths and receipts. Its unit of analysis is the action in context. Its assurance claim is bounded by issuer authority, semantic scope, lifecycle state and independent evaluation.

Future research should validate the score matrix with a multi-expert Delphi process, publish machine-readable evidence records, test inter-rater reliability and develop sector requirement vectors. Empirical studies should measure whether Trusted AI infrastructure reduces integration time, compliance cost, incident severity and human approval load. Safety research should examine how evidence freshness, degraded operation and agent composition affect cyber-physical assurance.

A second research stream should compare trust portability. Chinese PKI and DID infrastructures, European qualified trust services and wallet credentials, and US cloud-native workload identities can interoperate technically while retaining different legal meanings. Gateway frameworks should express equivalence, limitation and residual risk explicitly.

A third stream should address semantic governance. Powers of attorney, market roles, AI Service Passports, provenance claims, evaluation results and action receipts require controlled

vocabularies, schema authorities, version rules and conformance examples. Semantic precision is as important as signature validity because automated authorisation depends on machine-interpretable meaning.

The strategic policy test is practical: a European manufacturer, hospital, bank or public authority should be able to verify a foreign or domestic agent's accountable organisation, permitted action, current system state, evidence basis and execution record before granting consequential access. Achieving this capability would convert Europe's regulatory and industrial structure into deployable infrastructure for the next phase of AI competition.

A China evidence map

Table 12: China evidence map.

Capability	Primary source	Class	Interpretation
AI ethics and governance	New Generation AI Ethics Code; Global AI Governance Initiative; AI Safety Governance Framework 2.0	E3	Lifecycle ethics, controllability, monitoring, traceability and risk governance
Trustworthiness taxonomy	GB/T 47507-2026	E2	National recommended standard; effective 1 August 2026
Agent policy	Implementation Opinions on Intelligent Agents, 8 May 2026	E3	Identity, registration, authority boundaries, permission and traceability
Agent interconnection	GB/Z 185.1-185.7	E2	Seven national guiding technical documents; identity-to-tool chain
Identity lifecycle	GB/Z 185.3	E2	Registration, account, credentials, authentication, delegation and status
DID and VC	GB/T 47702-2026; CAC DID consultation draft	E2 / E3	DID system requirements, scheduled for 1 September 2026, plus proposed legal-person, device and delegation credentials
PKI and X.509	National certificate and cryptography standards	E2	Certificate chains, electronic certification, lifecycle and revocation substrate
Content provenance	AI-generated content labelling measures; GB 45438-2025	E1 / E2	Explicit and implicit labels, provider and content metadata
Runtime cryptography	Cryptography Application Guide for Generative AI Systems v1.0	E4	Mutual authentication, dynamic authorisation, signed plans and logs, provenance
Deployment evidence	SAMR press conference and pilots	E3 / E4	More than 50 firms reported in pilot application; audit and transaction standards planned

B European Union evidence map

Table 13: European Union evidence map.

Capability	Primary source	Class	Interpretation
Risk-based AI law	Regulation (EU) 2024/1689 and implementation material	E1	Harmonised risk duties, governance and enforcement
Trustworthy AI taxonomy	HLEG Ethics Guidelines; ISO and CEN-CENELEC work	E2 / E3	Lawful, ethical and robust framing; lifecycle standards
Identity and trust services	Regulation (EU) 2024/1183; EU trusted lists	E1	Qualified trust services and legally recognised trust anchors
European Business Wallet	COM(2025) 838; Council general approach; EP procedure	E3	Pending legislation; political agreement targeted by end-2026

Table 13 continued

Capability	Primary source	Class	Interpretation
Legal representation semantics	WE BUILD company-representative use case	E4	Digital PoA, representation rights, rulebooks and attestation definitions
Agent identity alignment	WE BUILD participant non-paper	E4	Expert proposal aligning agents, wallets and payments
Authoritative company credentials	Bundesanzeiger EBW programme	E4	Company identity, register credentials, representatives and PoA
Field issuance and agent PoA	Spherity Trusted AI field evidence	E5	Reported test issuance of legal-person and agent-authorisation credentials
Industrial programme	IPCEI-AI initiative; Spherity transversal architecture proposal	E3 / E5	Industrial AI ecosystem and proposed Trusted AI foundation
Functional safety	EU product and sector law; ISO/IEC safety standards	E1 / E2	Mature hazard, conformity and market-surveillance traditions

C United States evidence map

Table 14: United States evidence map.

Capability	Primary source	Class	Interpretation
Trustworthy AI taxonomy	NIST AI RMF 1.0	E3	Govern, Map, Measure, Manage and trustworthiness characteristics
Federal high-impact AI	OMB M-25-21	E3	Testing, impact assessment, monitoring and risk mitigation
Agent standards	NIST AI Agent Standards Initiative	E3	Industry-led standards, protocols, identity and evaluation research
Agent identity	NCCoE agent identity and authorisation concept paper	E3	Enterprise reference implementation direction
Authorisation stack	OAuth, OIDC, SPIFFE/SPIRE, SCIM, NGAC and zero trust	E2 / E4	Dynamic access, workload identity, lifecycle and policy graphs
TEVV	NIST AI 200-2 initial public draft	E3	Four-stage extensible evaluation framework including agents
Critical infrastructure	NIST AI RMF critical-infrastructure profile concept	E3	AI assurance across IT, OT and ICS contexts
Security adoption	NIST RFI analysis; CISA agentic AI guidance	E3	Threat mitigation and careful enterprise adoption
Legal-person trust	Sector and network-specific identity frameworks	E1 / E4	Strong domain systems; fragmented portable mandate layer

D Score rationale and reproducibility record

Table 15: Selected score rationales, including the EU legal-person score.

Geography	Score focus	Rationale
China	Agent/workload identity = 3	National guiding standards, lifecycle credentials, authentication and pilots; further audit integration under development.
China	Legal person/mandates/trust lists = 2	Authoritative legal-person and delegation direction exists in a consultation draft; agent-principal and PKI roots exist; cross-organisational legal semantics remain emerging.
China	Runtime evidence = 3	Traceability requirements, signed-log guidance and planned audit standards create an established direction with pilot implementation.

Table 15 continued

Geography	Score focus	Rationale
EU	Agent/workload identity = 1	Strong general identity technologies exist; a harmonised agent-specific profile and deployment fabric remain initial.
EU	Legal person/mandates/trust lists = 4; sensitivity range 3-4	The base case combines enacted eIDAS trust services and trusted lists with EBW work, WE BUILD semantics, authoritative issuer pathways and reported field credentials. Excluding participant evidence and reducing credit for pending legislation yields 3.
EU	Runtime evidence = 2	Research, pilots and adjacent cyber/product evidence exist; common agent action-receipt and runtime assurance profiles remain emerging.
US	Agent/workload identity = 2	Strong component standards and an active NCCoE implementation programme; agent-specific cross-domain practice remains emerging.
US	Legal person/mandates/trust lists = 1	Sector and platform identity systems exist; a portable national legal-person and mandate trust fabric remains initial.
US	Lifecycle and TEVV = 4	AI RMF, federal practices, security evaluation and TEVV-Athlon create advanced reusable assurance capability.

E Quality review protocol

Table 16: Manuscript quality review process.

Review gate	Acceptance criterion	Status
1. Source provenance	Prioritise official law, regulators, standards bodies and institutional documents; label public research and participant evidence separately.	Completed
2. Temporal validity	Verify status and dates against sources available by 26 August 2026; mark legislative proposals, drafts and future assumptions explicitly.	Completed
3. Claim traceability	Map every comparative score to an evidence class and concise rationale; preserve evidence gaps as research questions.	Completed
4. Mathematical audit	Use ordinal comparisons for readiness; use maxima only for conservative assurance communication; restrict arithmetic means to descriptive figures; test equation order and notation.	Completed
5. Conceptual validity	Test layer boundaries, non-substitutability and technology neutrality against academic and standards literature.	Completed
6. Language review	Apply British spelling, short sentences, neutral academic tone, normative precision and an em-dash exclusion check.	Completed
7. Citation review	Check first-occurrence numbering, titles, dates, status labels, URLs and primary-source proximity.	Completed
8. Visual and accessibility review	Inspect every rendered page, table, caption and figure; confirm readable colour contrast and meaningful figure descriptions.	Completed
9. Conflict-of-interest review	Declare author position and distinguish Spherity field evidence from independent public evidence.	Completed
10. Replication pathway	Publish the rubric, score vector, equations, evidence register and update date for future review.	Completed
11. Causal-claim discipline	Distinguish reported adoption barriers, model hypotheses and estimated causal effects; define an empirical calibration pathway.	Completed
12. Independent score validation	Pre-register evidence rules and consensus criteria; use independent raters, sensitivity analysis and published inter-rater agreement.	Protocol specified

Recommended external review. Before journal submission, the manuscript should receive independent legal review for Chinese and EU source interpretation, a standards review by PKI and identity experts, a functional-safety review, and blind academic peer review. A confirmatory empirical edition should apply the pre-registered multi-rater protocol and publish score distributions, sensitivity results and inter-rater agreement.

F Formal assurance model, scoring protocol and worked profiles

F.1 Notation and level of measurement

Table 17: Formal notation used in Chapter 9.

Symbol	Meaning	Measurement role
u	Defined agentic AI use case, including purpose, users, tools, environment and action envelope	Unit of analysis
a_u	Seven-dimensional assurance-demand vector	Ordinal vector
$A_{\text{required}}(u)$	Maximum component of the assurance-demand vector	Ordinal communication tier
l_u	Applicable legal and regulatory regime	Structured policy input
s_u	Applicable sector, cybersecurity and functional-safety regime	Structured assurance input
r_u	Required maturity across the twelve capabilities	Ordinal threshold vector
c_j	Published capability maturity for jurisdiction j	Ordinal evidence-based vector
x_d	Implemented maturity in sector or deployment ecosystem d	Ordinal assessed vector
o_o	Implemented capability maturity for organisation o	Ordinal assessed vector
e_{jdo}	Effective capability, componentwise minimum of jurisdictional, ecosystem and organisational maturity	Ordinal availability vector
$B(j,\ d,\ o,\ u)$	Capabilities whose effective maturity falls below the requirement	Bottleneck set
τ_u	Expected time to authorised deployment of use case u	Time parameter
$V_{\text{adopt}}(u,\ \tau_u)$	Expected discounted adoption surplus	Monetary decision quantity

The model uses ordinal levels because the evidence supports ordered states and threshold comparisons. Level 4 indicates greater assurance demand or maturity than level 3, while the numerical distance between levels remains unspecified. Maxima, minima and order comparisons preserve meaning under every strictly increasing recoding. Addition, subtraction and averaging require stronger measurement assumptions. Figure 2 uses arithmetic means solely as descriptive profiles, as stated in Chapter 8. Equations (1)-(7) therefore remain within ordinal operations. Equation (8) uses expected discounted monetary values and occupies a separate cardinal measurement domain. Equation (9) joins the two domains through a conjunction, rather than through arithmetic aggregation.

Equations (3) and (4) separate the requirement-setting process from measurement of capability supply. The policy mapping should be specified through traceable legal and technical rules. It should receive expert validation and version control. Equations (5)-(7) then preserve non-substitutability across jurisdictional, sector or ecosystem, and organisational supply. A validator and status policy remain necessary for a machine-readable trust list to create organisational value. A capable enterprise policy engine remains bounded by the maturity of authoritative issuers and mandate semantics. Every readiness conclusion remains conditional on correct requirements, valid evidence and continued monitoring.

F.2 Assurance-demand anchors

Table 18: Operational anchors for the seven assurance-demand dimensions. Levels 2 and 4 represent reasoned intermediate states.

Dimension	Operational question	Level 1 anchor	Level 3 anchor	Level 5 anchor
Autonomy, α	How independently can the agent select goals, plans, tools and timing?	Advisory output or sandbox action	Bounded execution with defined escalation	Self-initiated, multi-step and multi-system execution with sparse intervention
Authority, θ	Which legal, financial, professional or public powers can the action exercise?	Informational support	Bounded internal or low-value transactional authority	Power to bind a legal person, exercise a regulated role, transfer major value or invoke public authority
Irreversibility, ι	How completely and quickly can an adverse action be reversed?	Prompt and low-cost recovery	Material remediation within an operational window	Irreversible effect or recovery outside the applicable harm tolerance
Physical impact, ϕ	Can the action change a physical system or safety state?	Digital effect within an isolated environment	Connected-product or bounded machinery effect	Safety-critical actuation, operational technology or critical-infrastructure effect
Data sensitivity, δ	Which confidentiality, privacy or security interests attach to the data?	Public or low-sensitivity data	Confidential business or ordinary personal data	Special-category, medical, security-sensitive or critical operational data
Rights impact, ρ	How strongly can the action affect rights, access, due process or essential services?	Minor convenience effect	Material eligibility, employment, credit or service effect	Fundamental-rights, liberty, health, due-process or essential-service effect
Systemic reach, σ	How far can errors, compromise or correlated behaviour propagate?	Single user and isolated task	Multi-party workflow or bounded product fleet	Population-scale, market-wide, supply-network or critical-infrastructure propagation

F.3 Scoring protocol

1. Define the action envelope. Record the agent purpose, deployment environment, users, counterparties, tools, data, assets, value limits, physical interfaces, escalation rules and expected operating period.

2. Score the inherent context. Assign levels to the seven dimensions before crediting controls. Cite a short factual rationale and the applicable Level 1, 3 or 5 anchor. Use Levels 2 and 4 for documented intermediate states.

3. Record interactions. Identify combinations that can amplify harm or propagation, including autonomy with systemic reach, authority with irreversibility, and physical impact with uncertain state.

4. Map applicable regimes. Identify AI, data, cybersecurity, product, public-law, professional, sector and functional-safety requirements. Resolve conflicts through the stricter applicable threshold.

5. Derive capability thresholds. Translate the context and applicable regimes into r_u across the twelve capabilities. Record the source and rationale for every positive threshold.

6. Assess effective capability. Evaluate jurisdictional infrastructure, sector or ecosystem implementation, and organisational implementation separately. Use their componentwise minimum as effective capability and preserve evidence dates.

7. Decide and monitor. Apply Equations (6)-(9). Approve full deployment only when readiness and discounted value are positive; otherwise restrict, escalate or defer. Recalculate after material changes to tools, authority, data, software, model, environment, law or safety assumptions.

F.4 Worked assurance profiles

Table 19: Illustrative assurance profiles. Scores support method demonstration and require validation for each actual deployment.

Illustrative use case	Vector (α, θ, ι, ϕ, δ, ρ, σ)	Tier	Dominant assurance implications
Reversible consumer scheduling assistant	(2, 1, 2, 1, 2, 1, 1)	A2	Consent, bounded calendar access, recovery, logs and clear user control
Cross-company procurement agent with a defined value limit	(4, 4, 3, 1, 3, 3, 4)	A4	Legal-person identity, current mandate, counterparty role, scope, value limit, policy decision and signed transaction receipt
Public-benefits decision agent affecting entitlement	(3, 4, 4, 1, 4, 5, 4)	A5	Legal authority, due process, high-quality records, explanation, human review, appeal, equality testing and durable audit evidence
Industrial energy agent controlling distributed assets	(4, 4, 4, 5, 4, 4, 5)	A5	Machine identity, authorised operating envelope, sensor and software state, functional safety, runtime attestation, fail-safe control and incident evidence

The procurement profile illustrates the A4 cross-organisational boundary. Its physical-impact score remains low, while autonomy, authority and systemic reach create a requirement for portable legal reliance. The public-benefits profile reaches A5 through rights impact. This demonstrates why domain labels provide guidance rather than a ceiling. The industrial-energy profile reaches A5 through physical impact and systemic reach, with authority and irreversibility reinforcing the safety case. These profiles produce distinct r_u vectors even when two profiles share the same communication tier.

F.5 Calibration, uncertainty and validation

A journal-grade empirical application should use at least two independent raters with legal, security and sector expertise. Raters should score the seven dimensions and the twelve capability thresholds independently, then resolve differences through documented adjudication. Weighted agreement statistics can describe inter-rater consistency because the categories are ordered. Sensitivity analysis should test every plausible adjacent score. A readiness conclusion is robust when all plausible profiles produce the same bottleneck set or the same deployment decision. This protocol follows the principles of transparent indicator construction and sensitivity analysis [28].

A larger validation study should pre-register the evidence cut-off, source hierarchy, decision rules, rater selection, number of rounds, stopping rule and consensus threshold. A Delphi panel can support structured expert convergence when direct measurement remains unavailable. Consensus

should be defined before the first round and reported together with the score distribution and dissenting judgments [29]. Under this protocol, the current EU legal-person point estimate of 4 becomes a hypothesis for replication. The evidence sensitivity range is 3-4: the lower value follows when participant evidence and pending legislation receive minimal weight, while the base value follows from the full documented evidence set.

Evidence freshness forms part of validity. Credentials, mandates, issuer status, software state, monitoring results and legal rules can change during operation. A mature implementation should attach an evidence time, validity interval and refresh policy to each readiness input. Revocation and incident events should trigger immediate reassessment where policy defines them as material. Historical verification should preserve the status evidence and policy version that applied at the moment of action.

Equation (8) invites empirical calibration through observed assurance expenditure, integration expenditure, incident loss, deployment timing and realised benefit. Researchers can compare bilateral or fragmented verification with shared trust infrastructure across otherwise similar transactions. Deployment timing should enter the present-value calculation through observed or estimated cash-flow dates. A separate delay penalty should be added only when it captures an independently defined welfare or option-value effect. Causal inference requires careful control for sector, firm size, data maturity, skills, system complexity and regulatory intensity. The OECD enterprise surveys identify relevant reported barriers and covariates [5, 6]. Equation (8) therefore provides a transparent economic decomposition and a design for future estimation rather than a numerical forecast.

Declarations

Funding. This independent research paper was conducted independently of a dedicated external research grant.

Competing interests. The author is founder and Chief Executive Officer of Spherity GmbH. Spherity develops organisational identity, verifiable credential and Trusted AI infrastructure. Section 6 labels Spherity field evidence as participant-supplied evidence and separates it from enacted law, official standards and public programme documentation.

Data and materials. The appendices provide a source-level evidence register. All public sources were accessed by 26 August 2026. Jurisdictional scores are provisional, single-author rubric assessments designed for comparative analysis and future replication. The EU legal-person score of 4 has an evidence sensitivity range of 3-4. It should receive independent multi-rater validation before use as a consensus maturity estimate.

References

[1] NIST (2026). AI Agent Standards Initiative. Available from: https://www.nist.gov/artificial-intelligence/ai-agent-standards-initiative (accessed 26 August 2026).

[2] NIST NCCoE (2026). Accelerating the Adoption of Software and AI Agent Identity and Authorization: Draft Concept Paper. Available from: https://www.nccoe.nist.gov/sites/default/files/2026-02/accelerating-the-adoption-of-software-and-ai-agent-identity-and-authorization-concept-paper.pdf (accessed 26 August 2026).

[3] NIST (2026). Summary and analysis of responses regarding security considerations for AI agents. Available from: https://www.nist.gov/publications/summary-analysis-responses-request-information-regarding-security-considerations-ai (accessed 26 August 2026).

[4] NIST (2026). Concept Note: AI RMF Profile on Trustworthy AI in Critical Infrastructure. Available from: https://www.nist.gov/programs-projects/concept-note-ai-rmf-profile-trustworthy-ai-critical-infrastructure (accessed 26 August 2026).

[5] OECD, Boston Consulting Group and INSEAD (2025). The Adoption of Artificial Intelligence in Firms: New Evidence for Policymaking. OECD Publishing, Paris. Available from: https://doi.org/10.1787/f9ef33c3-en (accessed 26 August 2026).

[6] OECD (2026). Progress in Implementing the European Union Coordinated Plan on Artificial Intelligence, Volume 2: Uptake in High-Impact Sectors. OECD Publishing, Paris. Available from: https://doi.org/10.1787/3ac96d41-en (accessed 26 August 2026).

[7] Jobin, A., Ienca, M. and Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389-399. Available from: https://doi.org/10.1038/s42256-019-0088-2 (accessed 26 August 2026).

[8] Floridi, L. et al. (2018). AI4People: An ethical framework for a good AI society. Minds and Machines, 28, 689-707. Available from: https://doi.org/10.1007/s11023-018-9482-5 (accessed 26 August 2026).

[9] Jacovi, A., Marasović, A., Miller, T. and Goldberg, Y. (2021). Formalizing trust in artificial intelligence. Proceedings of FAccT 2021. Available from: https://doi.org/10.1145/3442188.3445923 (accessed 26 August 2026).

[10] Li, B. et al. (2023). Trustworthy AI: From principles to practices. ACM Computing Surveys, 55(9). Available from: https://doi.org/10.1145/3546872 (accessed 26 August 2026).

[11] Raji, I. D. et al. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of FAccT 2020. Available from: https://doi.org/10.1145/3351095.3372873 (accessed 26 August 2026).

[12] Mitchell, M. et al. (2019). Model cards for model reporting. Proceedings of FAT* 2019. Available from: https://doi.org/10.1145/3287560.3287596 (accessed 26 August 2026).

[13] ISO/IEC (2023). ISO/IEC 42001:2023, Artificial intelligence management system. Available from: https://www.iso.org/standard/42001 (accessed 26 August 2026).

[14] ISO/IEC (2023). ISO/IEC 23894:2023, Artificial intelligence guidance on risk management. Available from: https://www.iso.org/standard/77304.html (accessed 26 August 2026).

[15] ISO/IEC (2020). ISO/IEC TR 24028:2020, Overview of trustworthiness in artificial intelligence. Available from: https://www.iso.org/standard/77608.html (accessed 26 August 2026).

[16] ISO/IEC (2022). ISO/IEC 38507:2022, Governance implications of the use of artificial intelligence by organisations. Available from: https://www.iso.org/standard/56641.html (accessed 26 August 2026).

[17] ISO/IEC (2025). ISO/IEC 42005:2025, AI system impact assessment. Available from: https://www.iso.org/standard/42005 (accessed 26 August 2026).

[18] ISO/IEC (2025). ISO/IEC 42006:2025, Requirements for bodies providing audit and certification of AI management systems. Available from: https://www.iso.org/standard/42006 (accessed 26 August 2026).

[19] W3C (2025). Verifiable Credentials Data Model v2.0. Available from: https://www.w3.org/TR/vc-data-model-2.0/ (accessed 26 August 2026).

[20] W3C (2022). Decentralized Identifiers (DIDs) v1.0. Available from: https://www.w3.org/TR/did-core/ (accessed 26 August 2026).

[21] W3C (2025). Bitstring Status List v1.0. Available from: https://www.w3.org/TR/vc-bitstring-status-list/ (accessed 26 August 2026).

[22] IETF (2012). RFC 6749, The OAuth 2.0 Authorization Framework. Available from: https://www.rfc-editor.org/rfc/rfc6749 (accessed 26 August 2026).

[23] IETF (2025). RFC 9700, Best Current Practice for OAuth 2.0 Security. Available from: https://www.rfc-editor.org/rfc/rfc9700 (accessed 26 August 2026).

[24] OpenID Foundation (2014). OpenID Connect Core 1.0. Available from: https://openid.net/specs/openid-connect-core-1_0.html (accessed 26 August 2026).

[25] SPIFFE Project (2026). SPIFFE overview and specifications. Available from: https://spiffe.io/docs/latest/spiffe-about/overview/ (accessed 26 August 2026).

[26] European Commission High-Level Expert Group on AI (2019). Ethics Guidelines for Trustworthy AI. Available from: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed 26 August 2026).

[27] NIST (2023). Artificial Intelligence Risk Management Framework 1.0. Available from: https://www.nist.gov/itl/ai-risk-management-framework (accessed 26 August 2026).

[28] OECD, European Union and European Commission Joint Research Centre (2008). Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD Publishing, Paris. Available from: https://doi.org/10.1787/9789264043466-en (accessed 26 August 2026).

[29] Diamond, I. R. et al. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. Journal of Clinical Epidemiology, 67(4), 401-409. Available from: https://doi.org/10.1016/j.jclinepi.2013.12.002 (accessed 26 August 2026).

[30] Stöcker, C. (2026). European Business Wallets as the Legal Control Plane for Zero Trust AI Agents. Spherity Research. Available from: https://spherity.github.io/spherity-research/ebw-zero-trust-ai-agents.html (accessed 26 August 2026).

[31] Stöcker, C. (2026). Evidence Graphs for Industrial AI. Spherity Research. Available from: https://spherity.github.io/spherity-research/evidence-graphs-industrial-ai-data-plane.html (accessed 26 August 2026).

[32] Spherity Research (2026). Identity, Infrastructure, Resilience. Available from: https://spherity.github.io/spherity-research/ (accessed 26 August 2026).

[33] Ministry of Science and Technology of China (2021). New Generation Artificial Intelligence Ethics Code. Available from: https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html (accessed 26 August 2026).

[34] Cyberspace Administration of China (2023). Global AI Governance Initiative. Available from: https://www.cac.gov.cn/2023-10/18/c_1699291032884978.htm (accessed 26 August 2026).

[35] National Cybersecurity Standardization Technical Committee (2025). Artificial Intelligence Safety Governance Framework 2.0. Available from: https://www.cac.gov.cn/2025-09/15/c_1759653448369123.htm (accessed 26 August 2026).

[36] CAC, NDRC and MIIT (2026). Implementation Opinions on the Standardised Application and Innovative Development of Intelligent Agents. Available from: https://www.cac.gov.cn/2026-05/08/c_1779979789523320.htm (accessed 26 August 2026).

[37] SAMR and SAC (2026). GB/T 47507-2026, Artificial Intelligence: Trustworthiness General Rules. Available from: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=FE401CB127E80B2CAAFEF5EB4EC512D8 (accessed 26 August 2026).

[38] SAMR and SAC (2026). GB/Z 185.1-185.7, Artificial Intelligence: Agent Interconnection. Available from: https://openstd.samr.gov.cn/bzgk/std/std_list?p.p1=0&p.p2=GB%2FZ+185 (accessed 26 August 2026).

[39] SAMR and SAC (2026). GB/Z 185.3-2026, Artificial Intelligence: Agent Interconnection, Part 3, Agent Identity Management. Available from: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=BFCCE65447B30927953D90059AB89E23 (accessed 26 August 2026).

[40] State Administration for Market Regulation (2026). Press conference transcript on agent interconnection standardisation, 26 June 2026. Available from: https://www.samr.gov.cn/hd/zxft/art/2026/art_9bc3d59fd48b4be8af5dcd136aa6f66b.html (accessed 26 August 2026).

[41] SAMR and SAC (2018). GB/T 20518-2018, Information Security Technology: Public Key Infrastructure: Digital Certificate Format. Available from: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=26BD056FD628817775239AEDDB039C3A (accessed 26 August 2026).

[42] SAMR and SAC (2025). GB/T 19713-2025, Information Technology Security Techniques: Online Certificate Status Protocol. Available from: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=3F1BFB2155685C71A71A3CF477D55370 (accessed 26 August 2026).

[43] State Cryptography Administration (2024). Measures for the Administration of Electronic Certification Services for Electronic Government Affairs. Available from: https://www.sca.gov.cn/sca/xxgk/2024-11/04/content_1061187.shtml (accessed 26 August 2026).

[44] SAMR and SAC (2026). GB/T 47702-2026, Blockchain and Distributed Ledger Technology: General Service: Technical Requirements of Decentralized Identity System. Available from: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=8C8DF07885D1082D5B9F78E7F79A7D58 (accessed 26 August 2026).

[45] Cyberspace Administration of China (2026). Draft Provisions on Promoting Interoperability and Mutual Recognition of Distributed Digital Identity. Available from: https://www.cac.gov.cn/2026-06/18/c_1783525605384124.htm (accessed 26 August 2026).

[46] Cyberspace Administration of China et al. (2025). Measures for Labelling AI-Generated and Synthesised Content. Available from: https://www.cac.gov.cn/2025-03/14/c_1743654684782215.htm (accessed 26 August 2026).

[47] SAMR and SAC (2025). GB 45438-2025, Cybersecurity technology: Labelling method for content generated by artificial intelligence. Available from: https://openstd.samr.gov.cn/bzgk/std/newGbInfo?hcno=F32EA2A561F1886CD8D606513512D547 (accessed 26 August 2026).

[48] Chinese Association for Cryptologic Research (2026). Cryptography Application Guide for Generative Artificial Intelligence Systems, version 1.0. Available from: https://cmsfiles.zhongkefu.com.cn/cmsmima/backend_upload/file/20260807/1786096412171088.pdf (accessed 26 August 2026).

[49] European Union (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Available from: https://eur-lex.europa.eu/eli/reg/2024/1689/oj (accessed 26 August 2026).

[50] European Commission (2026). Regulatory framework for artificial intelligence. Available from: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (accessed 26 August 2026).

[51] CEN-CENELEC JTC 21 (2026). Artificial Intelligence working groups and standards programme. Available from: https://jtc21.eu/working-groups/ (accessed 26 August 2026).

[52] ISO (2018). ISO 26262, Road vehicles: Functional safety. Available from: https://www.iso.org/standard/68383.html (accessed 26 August 2026).

[53] IEC (2010). IEC 61508, Functional safety of electrical, electronic and programmable electronic safety-related systems. Available from: https://www.iec.ch/functionalsafety (accessed 26 August 2026).

[54] ISO (2022). ISO 21448:2022, Road vehicles: Safety of the intended functionality. Available from: https://www.iso.org/standard/77490.html (accessed 26 August 2026).

[55] European Union (2024). Regulation (EU) 2024/1183 amending Regulation (EU) No 910/2014 as regards the European Digital Identity Framework. Available from: https://eur-lex.europa.eu/eli/reg/2024/1183/oj (accessed 26 August 2026).

[56] European Commission (2026). List of Qualified Trust Service Providers in the European Union. Available from: https://digital-strategy.ec.europa.eu/en/policies/eu-trusted-lists (accessed 26 August 2026).

[57] European Commission (2026). eIDAS Dashboard Expands Its Role in Europe's Digital Identity Trust Framework. Available from: https://ec.europa.eu/digital-building-blocks/sites/spaces/DIGITAL/pages/973932461/eIDAS%2BDashboard%2Bexpands%2Bits%2Brole%2Bin%2BEurope%2Bs%2BDigital%2BIdentity%2BTrust%2BFramework (accessed 26 August 2026).

[58] European Commission (2025). COM(2025) 838, Proposal for a Regulation establishing European Business Wallets. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52025PC0838 (accessed 26 August 2026).

[59] Council of the European Union (2026). European business wallets: Council adopts negotiating position. Available from: https://www.consilium.europa.eu/en/press/press-releases/2026/06/09/european-business-wallets-council-adopts-negotiating-position/ (accessed 26 August 2026).

[60] European Parliament (2026). Procedure 2025/0358(COD), Establishment of European Business Wallets. Available from: https://oeil.europarl.europa.eu/oeil/en/procedure-file?reference=2025/0358(COD) (accessed 26 August 2026).

[61] WE BUILD Consortium (2026). WE BUILD: European Digital Identity Wallet large-scale pilot. Available from: https://www.webuildconsortium.eu/ (accessed 26 August 2026).

[62] WE BUILD Consortium (2026). Company representative acting on behalf of a company. Available from: https://www.webuildconsortium.eu/use-cases/company-representative-acting-on-behalf-of-a-company (accessed 26 August 2026).

[63] WE BUILD Consortium participants (2026). Trusted Identities for AI Agents: An Opportunity for Europe. Available from: https://www.webuildconsortium.eu/news/non-paper-trusted-identities-for-ai-agents-an-opportunity-for-europe-8396 (accessed 26 August 2026).

[64] Bundesanzeiger Verlag (2026). European Business Wallet: Verified Business Identity. Available from: https://www.bundesanzeiger-verlag.de/evidenzwesen/eubw/ (accessed 26 August 2026).

[65] Spherity GmbH (2025). Trusted AI, version 13, 12 September 2025. Participant-supplied field and programme-design evidence.

[66] German Federal Ministry for Economic Affairs and Energy (2026). IPCEI Artificial Intelligence. Available from: https://www.bundeswirtschaftsministerium.de/Redaktion/EN/Dossier/ipcei-ai.html (accessed 26 August 2026).

[67] US Office of Management and Budget (2025). Memorandum M-25-21, Accelerating Federal Use of AI through Innovation, Governance, and Public Trust. Available from: https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf (accessed 26 August 2026).

[68] NIST (2026). NIST AI 200-2 initial public draft, TEVV-Athlon Framework for Evaluating AI Systems. Available from: https://doi.org/10.6028/NIST.AI.200-2.ipd (accessed 26 August 2026).

[69] CISA (2026). Careful Adoption of Agentic AI Services. Available from: https://www.cisa.gov/resources-tools/resources/careful-adoption-agentic-ai-services (accessed 26 August 2026).

[70] The White House (2025). America's AI Action Plan. Available from: https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf (accessed 26 August 2026).

[71] Eurostat (2025). EU businesses: EUR 10.5 trillion in value added in 2023. Available from: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251013-2 (accessed 26 August 2026).

[72] Eurostat (2025). Exports of machinery and vehicles reached EUR 1 013 billion. Available from: https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20250908-1 (accessed 26 August 2026).