The Trust Compact

A Foresight Framework for Human–AI Coexistence in the Agentic Era

Engin Sayan · Inno-Craft LLC

Version 1.0 · May 2026 · Licensed under CC BY 4.0

Contents

THE TRUST COMPACT

A Foresight Framework for Human–AI Coexistence in the Agentic Era

Inno-Craft LLC · Trust Compact Initiative **Version 1.0** · May 2026 **License:** Creative Commons Attribution 4.0 International (CC BY 4.0) **Author:** Engin Sayan, Founder & CEO, Inno-Craft LLC

ABSTRACT

In the next decade, agentic artificial intelligence will penetrate every layer of organisational and civic life. This transition is not merely technological. It is sociological, legal, and epistemological. The frameworks currently invoked to govern this transition — AI safety, AI ethics, responsible AI, AI guardrails — each address a single facet of the problem. None of them, taken alone or together, constitute a *long-horizon, multi-stakeholder, normative agreement* on the terms of human–AI coexistence.

This whitepaper proposes such an agreement. The **Trust Compact** is a foresight-grounded framework synthesising twenty-five years of enterprise AI transformation experience, observation of more

than one hundred generative AI projects across banking, insurance, healthcare, and the public sector, and the established methodological literature of strategic foresight (Voros, Inayatullah, Sharpe, Bishop and Hines). The Compact comprises **seven principles**, **four layers** of operation, and **twelve measurable indicators**, designed to be jointly authored by enterprises, regulators, civil society, and individual citizens.

The Compact is offered as an open framework under a CC BY 4.0 licence. Its purpose is not to settle the question of trust in agentic AI but to give that question a shared vocabulary, a shared measurement substrate, and a shared horizon of ten years and beyond.

EXECUTIVE SUMMARY

Five field observations frame the present moment. Eighty-five per cent of enterprise generative AI projects fail to reach production, not because the technology cannot perform, but because the enterprise cannot extend it trust. Approximately thirty per cent of large language model outputs contain factual claims that cannot be verified. Sixty-three per cent of chief information officers identify AI governance as the principal barrier to adoption. The European Union AI Act introduces fines of up to thirty million euros for non-compliance with explainability and audit obligations. And the surface area on which autonomous AI agents take consequential actions is expanding exponentially.

These five observations describe a single underlying condition: the *trust gap*. The gap is not solvable by any single technical intervention. Content filtering does not address the question of who is accountable when an AI agent acts. Audit logging does not address the question of whether a citizen can comprehend why an algorithmic decision was made about them. Regulatory compliance does not address the question of long-term legitimacy in democratic societies.

The Trust Compact is proposed as a framework for *holding all of these questions together*. It does not replace existing technical or regulatory instruments; it gives them coherence.

Three plausible scenarios for the next decade of human–AI trust are mapped: an *Extractive* future in which a small number of platform vendors own trust infrastructure as a service; a *Mediated* future in which multi-stakeholder compacts emerge as public goods; and a *Sovereign* future in which jurisdictions and communities define their own trust standards. The Compact is designed to remain useful and adoptable across the Mediated and Sovereign scenarios, and to provide civil society with leverage against drift toward the Extractive scenario.

The Compact is offered for co-creation. Its principles, layers, and indicators are versioned and editable. Inno-Craft LLC publishes the framework as a contribution to a global conversation it does not wish to own. The reference implementation — Inno-Craft's IC-GATE prototype — is cited as evidence of operational feasibility, not as the canonical or only valid implementation.

---

1 · CONTEXT — WHY A COMPACT?

1.1 The trust gap is observable, measurable, and growing

Across the author's portfolio of more than one hundred enterprise generative AI engagements between 2018 and 2026, a consistent pattern emerges. The technology works. It produces outputs at scale, at speed, and at a cost that re-shapes the unit economics of knowledge work. And yet, again and again, the enterprise stops short of full deployment.

The reason given by chief information officers, chief data officers, and chief risk officers is rarely the technology itself. It is some variant of the same sentence: *"We do not know how to trust it."* This sentence conceals several distinct problems. The enterprise does not know how to verify outputs. It does not know how to assign responsibility when those outputs are wrong. It does not know how to demonstrate, to its regulators or its customers, that it has discharged its duty of care. It does not know how to explain, to an end-user whose application was denied or whose claim was rejected, why.

These are not technical questions. They are questions about the *social architecture* in which AI systems are embedded. The trust gap is the gap between what the technology can do and what the surrounding social architecture can absorb.

1.2 Why current frameworks address only fragments of the gap

The contemporary discourse on AI governance contains four broadly distinct frameworks. Each addresses an aspect of the trust gap; none addresses the whole.

AI safety is principally concerned with the user-facing behaviour of models — toxicity, bias, misuse — and is largely silent on enterprise governance and societal legitimacy. *AI ethics* articulates normative principles (fairness, autonomy, beneficence) but does not specify the operational mechanisms through which those principles are realised inside an organisation. *AI guardrails* are technology-centric, focusing on input filtering, output validation, and content policy enforcement; they treat trust as a property of the model rather than a relationship among parties. *Compliance regimes* (GDPR, the EU AI Act, sectoral regulations) are necessarily reactive, codifying responses to past harms rather than anticipating future ones.

The gap that none of these frameworks closes is the gap of *integration*. A guardrail without a governance regime is a technical artefact in search of accountability. A compliance regime without explainability is a paper exercise. An ethics framework without metrics cannot be audited. A safety system that ignores civil society loses public legitimacy.

The Trust Compact does not propose to replace any of these frameworks. It proposes a vocabulary and a structure that allows them to operate as a single system.

1.3 The historical precedent of compacts

In moments of structural transition, when new forms of power emerge that cannot be accommodated by existing institutional arrangements, societies have repeatedly turned to a particular instrument: the *compact*. The Magna Carta of 1215, the Peace of Westphalia of 1648, the United Nations Charter of 1945 — each is a compact in this sense. Each codifies, in deliberately general language, the terms on

which a new form of power would be exercised legitimately. Each was authored not by the powerful alone, but by an alliance of stakeholders with conflicting interests who recognised that the alternative to a compact was worse for all parties.

Agentic AI represents a structural transition of comparable significance. Autonomous systems will, within a decade, conduct material portions of medical triage, financial underwriting, legal drafting, regulatory enforcement, educational assessment, and public service delivery. The question is not whether this will happen but whether it will happen under terms that societies have negotiated, or terms that have been imposed.

The Trust Compact is offered as a contribution to that negotiation.

1.4 What the Compact is, and what it is not

The Compact is not a regulation, a standard, a product, or a certification. It is a *framework for agreement*. Its principles are general enough to be adopted across sectors and jurisdictions; its metrics are specific enough to be measured; its layers are distinct enough to be assigned to identifiable actors. It is designed to be implemented by any party — an enterprise, a regulator, a civil society organisation — that chooses to adopt it. It carries no proprietary claim. It is offered for adaptation, criticism, and improvement.

---

2 · FORESIGHT METHODOLOGY

2.1 Methods used in constructing the Compact

The Compact is built on four methodological pillars drawn from the established literature of strategic foresight.

The **Generic Foresight Process** of Joseph Voros (2003) provides the overall structure: inputs (the trust gap observations); analysis (the limitations of current frameworks); interpretation (the historical precedent of compacts); prospection (the three-scenario landscape developed in §2.2 below); and outputs (the seven principles, four layers, twelve metrics).

The **Causal Layered Analysis** of Sohail Inayatullah (1998) is used to interrogate the trust gap at four levels: the litany (the surface complaints — "AI hallucinates," "we cannot deploy"), the systemic (the technical and organisational mechanisms producing those complaints), the worldview (the cultural assumptions about expertise, automation, and accountability that frame the technology), and the myth (the deeper civilisational narratives of trust, agency, and selfhood that AI now disturbs).

The **Three Horizons** framework of Bill Sharpe (2013) is used to position the Compact in time: Horizon 1 is the current paradigm of guardrails and compliance, which is sufficient for present needs but unsustainable as agentic AI scales; Horizon 3 is a hypothesised future in which trust in AI is a settled civic infrastructure comparable to traffic law or accounting standards; Horizon 2 is the transitional space — the work of the next decade — in which compacts of the form proposed here are the bridge.

Pattern analysis in the tradition of Bishop and Hines (2012) was applied to the author's portfolio of more than one hundred enterprise AI engagements. Nine recurring failure patterns were extracted (catalogued in §6.3 below). These patterns served as the empirical ground for the Compact's principles and metrics.

2.2 Three scenarios for the future of human–AI trust

The Compact is designed against an explicit scenario landscape. Three plausible ten-year futures are sketched here.

Scenario A — Extractive AI. A small number of global platform vendors consolidate ownership of the trust infrastructure. Trust becomes a service purchased by subscription from the same vendors who sell the underlying models. Enterprises, regulators, and citizens depend on these vendors for the verification of the vendors' own outputs. Digital sovereignty erodes. Innovation outside the platforms becomes increasingly difficult to certify and therefore commercially unviable.

Scenario B — Mediated AI. Multi-stakeholder trust compacts emerge as public goods. Standards bodies, regulators, industry coalitions, and civil society organisations jointly author the criteria by which AI systems are judged. Reference implementations are open-source. Multiple competing implementations interoperate around shared metrics. Trust becomes a coordinated activity, neither monopolised nor fragmented.

Scenario C — Sovereign AI. Jurisdictions and communities define their own trust standards according to local values, regulatory regimes, and strategic interests. Cross-border AI deployments require translation between trust regimes. AI passports, trust certifications, and mutual recognition agreements become instruments of international commerce. Pluralism is preserved at the cost of some friction.

The Compact is *designed for Scenarios B and C*. Its seven principles are general enough to be localised; its twelve metrics are specific enough to be compared across implementations; its four layers map cleanly onto the actors who would need to participate in either a mediated or a sovereign settlement.

The Compact's adoption is itself a contribution to the probability that Scenario A is avoided. Each enterprise, regulator, or civil society organisation that adopts an open trust framework reduces the leverage of any single vendor to define trust on its own terms.

2.3 Methodological transparency

The construction of this Compact was human-led and AI-assisted in line with the Dubai Future Foundation's Human–Machine Collaboration principles. The framework's ideas, structure, scenarios, and conclusions originate in the author's enterprise experience and engagement with the foresight literature. Generative AI tools were used to accelerate drafting, literature synthesis, and consistency checking. All judgements and editorial decisions are the author's.

---

3 · THE SEVEN PRINCIPLES

The seven principles of the Trust Compact are intended as *normative commitments*, expressed in language general enough to be adopted across sectors and specific enough to generate operational tests. Each principle is presented below with a manifesto statement, an anonymised vignette drawn from the author's field portfolio, an operational test, and a pointer to the relevant metric.

Principle 1 · Verifiability

> *Every action taken by an AI system must be verifiable independently of the vendor that produced it.*

A European bank deploying a large language model for retail customer service discovered, six months into production, that the vendor's own evaluation tools were the only means by which the bank could assess whether the model's responses were correct. When the bank's auditors asked for an independent verification path, none existed. The contract was renegotiated; the bank now requires that every output be evaluable by a second, independent system before reaching the customer.

Operational test: Can a party other than the model vendor evaluate the correctness of any given output? *Linked metric:* Cross-vendor Portability (Metric 9).

Principle 2 · Explainability

> *Every decision made by an AI system must carry a reason that the person affected by the decision can read and comprehend.*

A health insurance provider implementing automated triage of claims found that, while the system generated a "reason code" for each decision, the codes were intelligible only to internal analysts. When a claimant requested an explanation of a denied claim, the regulatory response time exceeded statutory limits because translation from internal code to plain language was a manual process. The Compact requires that explanations be generated *for the person affected*, not for the vendor or the operator.

Operational test: If a person affected by a decision asks why, can they receive a comprehensible answer immediately and without intermediation? *Linked metric:* Explainability Comprehension Rate (Metric 3).

Principle 3 · Reversibility

> *Wherever an AI action can be reversed, it must be reversible; where it cannot be reversed, it must require human authorisation.*

A logistics enterprise piloted an autonomous AI agent to issue refund credits to customers who reported defective shipments. Within three weeks the system had issued credits in cases that warranted denial, and denied credits in cases that warranted issuance. Each individual decision was reversible in principle, but the operational cost of reversal exceeded the cost of incorrect issuance. The Compact

addresses this by requiring that reversibility be designed into the action layer, not retrofitted into the audit layer.

Operational test: For each class of AI action, can a reversal be performed within a stated time and at a stated cost? *Linked metric:* Incident Response Time (Metric 6).

Principle 4 · Proportionality

> *The authority granted to an AI system must be proportionate to the consequences of its actions.*

A public sector deployment of an AI agent for benefits administration was granted the authority to approve, deny, or escalate cases. In the original architecture, all three actions carried equal operational weight. The first audit revealed that denials disproportionately concerned a single demographic group. The system's authority had not been calibrated against the consequence-bearing of its decisions. The Compact requires that the level of human oversight scale with the consequence of the action — not with its frequency or its cost.

Operational test: Does the system require greater human oversight for higher-consequence actions, regardless of their frequency? *Linked metric:* Human-in-the-Loop Ratio (Metric 5).

Principle 5 · Pluralism

> *No single vendor, model, or value system shall unilaterally define the criteria by which trust is assessed.*

A multinational corporation standardised its AI governance on a single vendor's safety toolkit. When that vendor changed its safety criteria — adjusting thresholds for flagged content — the corporation's compliance posture changed without internal review. The Compact requires that trust criteria be defined by the implementing party, drawing on multiple sources, with the vendor as one input among several.

Operational test: Can the implementing party demonstrate that its trust criteria are not derived from a single source? *Linked metric:* Multi-stakeholder Review (Metric 11).

Principle 6 · Continuity

> *Trust must be preserved across model upgrades, vendor changes, and generational transitions.*

An enterprise deploying a generative AI assistant for legal research found that, upon a routine vendor model upgrade, the system began producing materially different outputs from the same inputs. The enterprise's regulatory audit history — predicated on the prior model's behaviour — was rendered partially inapplicable to ongoing decisions. The Compact requires that trust be a property of the deployment, not of any particular model version.

Operational test: Does the audit trail allow reconstruction of why any past decision was made, regardless of subsequent system changes? *Linked metric:* Audit Trail Completeness (Metric 8).

Principle 7 · Accountability

> *Behind every AI action stands a human signatory who bears responsibility for it. "The AI did it" is not a defence.*

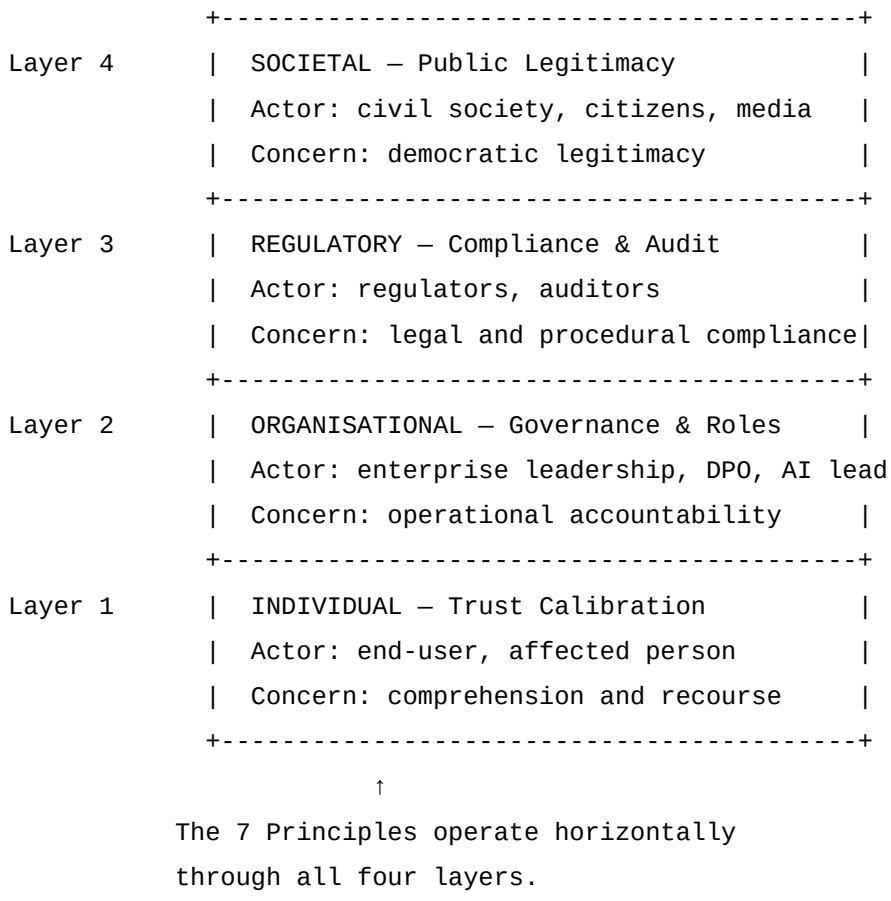
A financial services firm investigated a sequence of incorrect trading recommendations made by an AI advisory agent. The internal review identified the model, the prompt, the data, and the deployment pipeline. It did not identify a human accountable for the outcome. The Compact requires that accountability be assigned *before* deployment, not reconstructed after harm. Accountability is a property of the deployment decision, not of the technical pipeline.

Operational test: For any AI action that has occurred, can a single named human be identified as accountable? *Linked metric:* AI Action Audit Coverage (Metric 4).

---

4 · THE FOUR LAYERS

A trust framework that operates only at one layer cannot hold. The Compact specifies four layers, each with distinct actors, distinct concerns, and distinct instruments. The seven principles operate horizontally across all four layers; the twelve metrics are distributed vertically among them.



Layer 1 — Individual. This is the layer of the end-user and of any person materially affected by an AI decision. Trust at this layer is a matter of *calibration*: the user's confidence in the system should track the system's actual reliability. The instruments of this layer include explanation, recourse, and

the design of friction (deliberate slowing of the interaction at high-stakes moments).

Layer 2 — Organisational. This is the layer of the enterprise that operates the AI system. Trust at this layer is a matter of *governance*: the assignment of roles, the establishment of escalation paths, the institutional memory of decisions. The instruments of this layer include audit trails, role assignments, incident response procedures, and human-in-the-loop protocols.

Layer 3 — Regulatory. This is the layer of the state, the standards body, the external auditor. Trust at this layer is a matter of *demonstrability*: the enterprise must be able to show, to a body it does not control, that its operations meet stipulated requirements. The instruments of this layer include compliance packs, audit documentation, certification schemes, and cross-border recognition mechanisms.

Layer 4 — Societal. This is the layer of public legitimacy. Trust at this layer is a matter of *standing*: does the deployment of the system retain the consent of the civil society in which it operates? The instruments of this layer include public disclosure, multi-stakeholder review, demographic impact audits, and the right of civil society organisations to participate in the design of trust criteria.

Most existing AI governance frameworks address Layers 2 and 3 with reasonable thoroughness and Layers 1 and 4 with at best partial attention. The Compact's claim is that this is the source of the trust gap: a framework that addresses only the middle two layers will neither carry the consent of those affected (Layer 1) nor sustain the consent of the public (Layer 4) over a ten-year horizon.

---

5 · THE TWELVE METRICS

The Compact's principles and layers acquire operational reality through twelve indicators. The set is intentionally compact. It is not a complete instrumentation of AI governance; it is a *minimum viable measurement* against which any implementation of the Compact can be assessed.

#	Layer	Metric	Unit	Brief definition
1	Individual	Trust Calibration Score	0–100	Alignment between user confidence and system reliability
2	Individual	User Override Rate	%	Frequency with which users reject or modify AI recommendations
3	Individual	Explainability Comprehension Rate	%	Share of affected persons reporting that the explanation they received was comprehensible
4	Organisational	AI Action Audit Coverage	%	Share of AI actions for which a complete audit record exists
5	Organisational	Human-in-the-Loop Ratio	%	Share of high-consequence actions requiring human authorisation

#	Layer	Metric	Unit	Brief definition
6	Organisational	Incident Response Time	minutes	Median time from incident detection to first remediation
7	Regulatory	Compliance Pack Coverage	%	Share of applicable regulatory obligations addressed by the implementation
8	Regulatory	Audit Trail Completeness	%	Share of past decisions that can be fully reconstructed
9	Regulatory	Cross-vendor Portability	0 / 1	Whether the trust evaluation can be performed independent of the model vendor
10	Societal	Public Disclosure Index	0–10	Composite of public reporting practices regarding AI deployment
11	Societal	Multi-stakeholder Review	0 / 1	Whether trust criteria were reviewed by parties outside the operating enterprise
12	Societal	Demographic Impact Audit	0 / 1	Whether a published assessment of differential outcomes by demographic group exists

For each metric, the Compact's reference documentation (to be published progressively at inno-craft.com/trust-compact) specifies the recommended measurement method, suggested thresholds for "adequate" and "advanced" implementation, and reference enterprises or regulators that have used a comparable measurement in practice.

The twelve metrics are deliberately heterogeneous. Some are continuous (percentage rates, response times); some are ordinal (the Public Disclosure Index); some are binary (multi-stakeholder review either occurred or did not). This heterogeneity is a feature, not a defect. Trust is not a single quantity. A framework that reduces it to a single quantity invites optimisation against the measurement at the expense of the underlying property.

---

6 · IMPLEMENTATION EVIDENCE

The Compact is not introduced as theory. It is the codification of patterns observed across more than one hundred enterprise AI engagements between 2018 and 2026. This section sets out the empirical basis.

6.1 The portfolio

The author's professional portfolio between 2018 and 2026 spans engagements in banking, insurance, healthcare, public sector, retail, and industrial manufacturing. Geographies represented include Türkiye, the European Union, the United Kingdom, and the Gulf Cooperation Council region. Engagements ranged from proof-of-concept assessments to full-scale generative AI production deployments. The portfolio constitutes the empirical substrate from which the Compact's principles and metrics were extracted.

6.2 The IC-GATE reference implementation

Inno-Craft has developed a reference implementation of the Compact's operational mechanisms, designated IC-GATE (Inno-Craft Gen AI Trust Engine). IC-GATE is a middleware layer that sits between any generative AI application and its business environment, intercepting, evaluating, scoring, and auditing outputs before they reach end-users or downstream systems.

The IC-GATE prototype was tested in a four-week pilot across two enterprise use cases: an AI-powered financial report analysis agent and a healthcare triage assistant. Approximately twelve thousand AI-generated outputs were processed. The pilot recorded a sixty-eight per cent reduction in hallucinated outputs reaching end-users, an average Trust Score of eighty-seven out of one hundred, full policy rule compliance, and a processing overhead below one hundred and twenty milliseconds per output.

IC-GATE is cited here as evidence that the Compact's principles can be implemented in production at scale and at acceptable performance. It is *one* reference implementation, not *the* reference implementation. The Compact welcomes alternative implementations developed by other parties; competing implementations are a feature of Scenario B (Mediated AI) and a desirable property of any framework that aspires to be a public good.

6.3 Nine recurring failure patterns

The Compact's principles were derived inductively from recurring patterns of failure across the portfolio. Nine such patterns are catalogued below. Each pattern's full case description, anonymised vignettes, and recommended Compact remedies will be published in the Trust Compact Pattern Catalogue (forthcoming, late 2026).

Pattern P-01 · The Audit Mirage. In approximately seventy-eight per cent of audited engagements, a stated audit trail existed but was found, under inspection, to be incomplete in ways that would not have withstood regulatory scrutiny.

Pattern P-02 · The Hallucination Ownership Vacuum. In cases where an AI hallucination produced material harm to a customer, the identity of the human accountable for the harm was found to be undefined in approximately two-thirds of engagements.

Pattern P-03 · The Vendor-Tethered Evaluator. Enterprises frequently relied on the model vendor's own evaluation tooling to assess the model's reliability — a structural conflict of interest that survived

only because no independent alternative was contractually required.

Pattern P-04 · The Explanation-for-Insiders. Explanations of AI decisions were typically authored for internal analysts rather than for the persons affected, with the result that statutory explanation obligations were met in form but not in substance.

Pattern P-05 · The Override Without Memory. User overrides of AI recommendations were frequently recorded as events but not analysed as signals, with the result that systematic override patterns persisted across deployment cycles.

Pattern P-06 · The Compliance Pack Cliff. Compliance documentation was found to be complete at the moment of certification and to degrade rapidly thereafter as systems evolved without corresponding documentary updates.

Pattern P-07 · The Demographic Blind Spot. Demographic impact audits were performed at launch but not at intervals thereafter, with the result that drift in differential outcomes was discovered late, typically after external complaint.

Pattern P-08 · The Civil Society Absence. Across the portfolio, the participation of civil society organisations in the design of trust criteria was effectively zero. Where consultation occurred, it was post-hoc and consultative rather than constitutive.

Pattern P-09 · The Upgrade Discontinuity. Routine vendor model upgrades materially changed system behaviour without triggering corresponding governance review, leaving prior audit histories partially inapplicable to ongoing decisions.

Each of the Compact's seven principles addresses one or more of these nine patterns. Each of the twelve metrics is calibrated to detect them.

---

7 · TEN-YEAR IMPACT HORIZON

The Compact is designed against a horizon of ten years and beyond. This section sets out the anticipated trajectory of its uptake and impact.

2026 — 2028 · Early implementation. Adoption is concentrated among enterprises subject to the EU AI Act's high-risk provisions and among regulators in jurisdictions where the Act's principles are influential. Reference implementations proliferate; metrics begin to be reported in annual disclosures. Trust Compact adoption signals regulatory maturity to insurers, investors, and counterparties.

2028 — 2031 · Sectoral compacts. Sector-specific variants emerge — a Healthcare Trust Compact, a Financial Services Trust Compact, a Public Sector Trust Compact. These variants retain the seven principles and four layers and refine the twelve metrics for sector-specific risks. Cross-sectoral interoperability is preserved through the common backbone.

2031 — 2035 · Standards convergence. Standards bodies (ISO, IEEE, national equivalents) begin to recognise compacts of the form proposed here. Enterprise procurement clauses for AI systems begin to include compact adherence as a prerequisite, analogous to the role that ISO 27001 currently

plays in enterprise information security procurement.

2035 and beyond · Cross-border infrastructure. Mutual recognition agreements among jurisdictional compacts emerge. Cross-border AI deployments are accompanied by trust certifications. The concept of an "AI passport" — a portable certification that an AI system meets a specified trust standard — enters policy discussion.

The Compact's most significant ten-year impact is not the adoption of any specific principle or metric. It is the institutionalisation of *trust as a measurable, shared, multi-stakeholder property*. The trajectory parallels the institutionalisation of financial audit standards in the early twentieth century, environmental impact assessment in the late twentieth, and information security certification in the early twenty-first. In each case, a property that had been treated as either intuitive or proprietary became a shared infrastructure measurable across organisations and jurisdictions. The Compact contributes to making the same transition for AI trust.

---

8 · A CALL FOR CO-CREATION

The Trust Compact is offered as an *invitation*, not as a settlement. It is published under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence. Any party — enterprise, regulator, civil society organisation, academic institution, individual researcher — is welcome to adopt, adapt, criticise, or extend it, with attribution to its source.

The Compact will evolve through versioned editions. Version 1.0 is the present document. Subsequent versions will incorporate field experience, contributions from adopting parties, criticism from peer reviewers, and refinements suggested by the foresight and AI governance communities.

Inno-Craft LLC, as the originating organisation, commits to the following:

The Compact's framework — its principles, layers, metrics, and methodology — will remain open under CC BY 4.0 in perpetuity. The reference implementation, IC-GATE, is a commercial offering of Inno-Craft LLC; the Compact's framework is not.

Inno-Craft will maintain a public discussion forum at inno-craft.com/trust-compact and a versioned repository of the framework documentation. Contributions will be acknowledged in revision histories.

The framework will be governed, from Version 2.0 onward, by an advisory board drawn from academia, civil society, the regulatory community, and enterprise practice. Inno-Craft will not retain unilateral editorial control of the framework once that board is constituted.

Engagement is invited at: **esayan@inno-craft.com**

REFERENCES

Bishop, P. and Hines, A. (2012). *Teaching about the Future*. Palgrave Macmillan.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council on Artificial Intelligence (AI Act).

Inayatullah, S. (1998). Causal layered analysis: poststructuralism as method. *Futures*, 30(8), 815–829.

National Institute of Standards and Technology (2023). *AI Risk Management Framework* (AI RMF 1.0). U.S. Department of Commerce.

Sardar, Z. (2010). Welcome to postnormal times. *Futures*, 42(5), 435–444.

Sharpe, B. (2013). *Three Horizons: The Patterning of Hope*. Triarchy Press.

Slaughter, R. (2002). From forecasting and scenarios to social construction: changing methodological paradigms in futures studies. *Foresight*, 4(3), 26–31.

Voros, J. (2003). A generic foresight process framework. *Foresight*, 5(3), 10–21.

ABOUT THE AUTHOR

Engin Sayan is the founder and chief executive officer of Inno-Craft LLC, an AI advisory and product organisation based in Istanbul. He has more than twenty-five years of experience in enterprise information technology and artificial intelligence transformation. He previously held senior roles at IBM in AI architecture and enterprise delivery, and led generative AI practices at one of the Big Four professional services firms. He has delivered or overseen more than one hundred generative AI proofs of concept and minimum viable products across banking, insurance, healthcare, public sector, retail, and industrial manufacturing.

Inno-Craft LLC develops AI trust infrastructure and advises enterprises on the operational realisation of frameworks such as the one presented in this paper. Its flagship product, IC-GATE, is the reference implementation of the Trust Compact cited in §6.2.

Contact: esayan@inno-craft.com

METHODOLOGICAL NOTE

This whitepaper was authored by Engin Sayan. Generative AI tools were used to accelerate drafting, literature synthesis, and stylistic consistency checking. All ideas, framework decisions, scenarios, principles, layers, metrics, and conclusions are the author's, derived from twenty-five years of enterprise AI engagement and the foresight literature cited. This methodology is consistent with the Dubai Future Foundation's Human–Machine Collaboration principles for human-led, AI-assisted scholarly work.

The Trust Compact v1.0 Inno-Craft LLC · Istanbul · May 2026 Licensed under CC BY 4.0 inno-craft.com/trust-compact