INDEPENDENT RESEARCH · CONCEPTUAL WORKING PAPER

OrganoSphere:

A Conceptual Architecture for Human-Machine Organisational Intelligence

Software engineering, systems thinking and accountable coordination across two forms of intelligence

Thabang Mokoena

Independent research · ITMan Intelligence

Professional affiliation: NRF iThemba LABS

iThemba Laboratory for Accelerator-Based Sciences

ORCID iD: 0000-0003-0689-7632

Status	Conceptual working paper	**Version**	0.1
Date	14 August 2026	**Review**	Not peer reviewed

This is independent conceptual work. The affiliations provide author context only; neither NRF nor NRF iThemba LABS is the publisher, sponsor, reviewer or endorser of this paper. No clinical validation or institutional position is claimed.

Abstract

This working paper develops OrganoSphere as a public conceptual model for joining situated human intelligence and computational machine intelligence inside an engineered, governed organisational environment. It argues that useful human-machine coordination depends less on model capability alone than on explicit boundaries, traceable decisions, observable software, institutional memory and retained human accountability. The paper offers a vocabulary, a conceptual equation, design propositions and a research agenda. It does not disclose a production implementation, report an experiment, make a clinical claim or present validated scientific evidence.

Keywords: human-machine intelligence; software engineering; systems thinking; organisational design; governance; institutional memory.

Epistemic status. This is a conceptual working paper. "OrganoSphere" is used as a working name for an organisational-intelligence idea, not as evidence of a validated product, clinical system or established scientific theory. The paper remains at a public conceptual level and does not disclose implementation or protected experimental material.

1 The problem is coordination, not intelligence in the singular

Organisations increasingly work with two materially different forms of intelligence. The first is human: situated judgement, tacit knowledge, moral responsibility, imagination and the capacity to interpret ambiguous conditions. The second is computational: rapid search, statistical pattern transformation, simulation, retrieval and the consistent execution of encoded procedures. Calling both "intelligence" can be useful, but only if the word does not erase their differences.

A machine system may produce an impressive answer without understanding the institutional consequence of acting on it. A human team may understand consequence while remaining unable to search, compare or simulate at machine scale. The design question is therefore not which intelligence should replace the other. It is how an organisation can compose them without confusing fluency with authority, computation with judgement, or assistance with accountability.

OrganoSphere is proposed as a name for that composition problem. It is not a single model. It is a governed environment in which people, software, records, models and decisions remain connected strongly enough for work to be inspected, challenged, revised and inherited.

2 Definitions and conceptual boundary

Definition 1 (Situated human intelligence). The capacity of people to frame a problem, interpret context, weigh values, exercise judgement and accept responsibility for consequences.

Definition 2 (Computational machine intelligence). The capacity of engineered systems to transform representations, retrieve and combine information, identify patterns, generate candidate outputs and execute bounded procedures.

Definition 3 (Organisational intelligence). The capacity of a collective to perceive relevant conditions, preserve usable memory, coordinate action, learn from outcomes and remain answerable for decisions over time.

Definition 4 (OrganoSphere). A conceptual, governed software environment that helps an organisation coordinate human and machine intelligence through explicit interfaces, traceable records, feedback and human authority.

The definitions are functional rather than metaphysical. They make no claim that present machine-learning systems possess consciousness, intention or moral agency. The distinction matters because responsibility cannot be delegated merely by inserting an automated component into a workflow.

3 A minimal model of an organisational intelligence environment

Let H_t denote situated human judgement at time t, M_t the available machine capability, E_t the engineering substrate, G_t the governance constraints and B_t the relevant boundary conditions. A conceptual organisational action can be written as

$$a_t = \Gamma(H_t, M_t \mid E_t, G_t, B_t). \tag{1}$$

Equation (1) is not an empirical law. It is a bookkeeping device: an outcome cannot be attributed to a model alone when the model is selected, prompted, integrated, constrained, interpreted and acted upon through a socio-technical system. The environment matters.

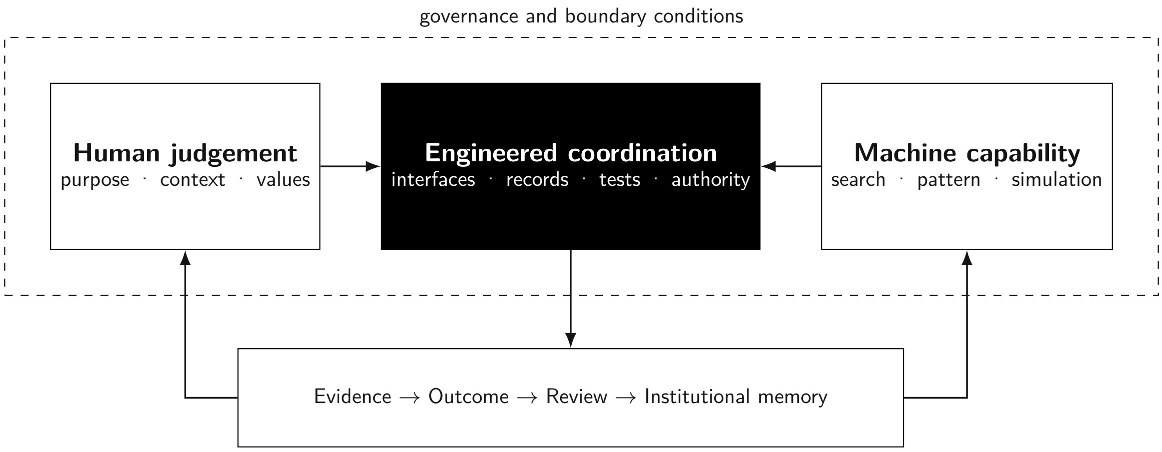

Figure 1: The OrganoSphere coordination loop (not a tested architecture). An engineered, governed layer joins human judgement and machine capability; evidence returns through institutional learning.

4 Software engineering as institutional constitution

Software is often described as a tool used by an organisation. Once it determines who can see a record, which fields count as evidence, when a workflow may proceed and how an exception is handled, it becomes more than a tool. It begins to encode institutional structure. In that limited but important sense, software behaves constitutionally: permissions allocate power, schemas define what can be said, interfaces rank what can be seen and logs determine what can later be known.

This makes conventional engineering disciplines central to responsible human-machine work. Version control preserves lineage. Automated tests protect declared behaviour. Observability makes failures discoverable. Access control gives effect to authority boundaries. Reversible deployment limits the cost of error. Documentation allows a future practitioner to reconstruct why a decision path exists.

Table 1: Engineering properties and their organisational purpose.

Property	Engineering expression	Organisational purpose
Traceability	Versioned inputs, outputs and approvals	Reconstruct who knew what, when and why
Observability	Logs, metrics, alerts and review queues	Make degraded behaviour visible before it becomes normal
Reversibility	Rollbacks, checkpoints and bounded release stages	Preserve room to correct decisions under uncertainty
Authority	Explicit roles, permissions and escalation paths	Keep accountability attached to identifiable people
Memory	Durable records, rationale and linked evidence	Allow learning to survive staff and model turnover

5 Three design propositions

The following propositions are normative design claims offered for criticism. They are not proven theorems.

Proposition 1 — Accountability invariance. Introducing machine assistance should not make the person or role answerable for a consequential decision less identifiable.

Proposition 2 — Traceability before autonomy. The permissible degree of automated action should not increase faster than the organisation's ability to reconstruct inputs, transformations, authority and consequences.

Proposition 3 — Reversibility under uncertainty. As uncertainty and consequence increase, the system should favour smaller action boundaries, stronger review and more recoverable states.

These propositions resist a common failure mode: treating the apparent competence of an output as sufficient evidence that the surrounding decision system is mature. A reliable organisational capability requires a chain from intent to execution and back to evidence.

6 Systems thinking: boundaries, feedback and near-decomposability

Systems thinking contributes three practical cautions. First, every model draws a boundary. What sits outside that boundary—labour, incentives, informal knowledge, infrastructure, power relations or environmental cost—may still determine the result. Second, feedback is not automatically learning. A metric can reinforce the wrong behaviour when it becomes detached from purpose. Third, complex organisations need modularity without fragmentation: components should be separable enough to change, yet connected enough for consequences to propagate visibly.

Simon's account of near-decomposable systems offers a useful design intuition: stable intermediate structures can make complexity governable without pretending that the whole is reducible to isolated parts [6]. Conant and Ashby's regulator result adds a related warning: effective regulation depends on an adequate model of what is being regulated [2]. In OrganoSphere terms, a dashboard is not organisational intelligence if its representation excludes the conditions that matter.

The conceptual architecture therefore privileges linked layers rather than a monolithic "brain": work objects, evidence, decision rights, computational services and institutional memory. The layers should exchange declared signals through inspectable interfaces. This supports local change while retaining system-level accountability.

7 A public research agenda

A serious next step would turn the propositions into falsifiable research questions rather than expanding the concept rhetorically. Candidate questions include:

1. Does explicit decision lineage improve the accuracy with which independent reviewers reconstruct a human-machine decision?
2. Which combinations of confidence, consequence and reversibility produce useful escalation thresholds?
3. Can modular organisational memory reduce repeated decision error without creating excessive administrative burden?
4. How does interface design affect whether practitioners challenge, defer to or appropriately calibrate machine-generated recommendations?
5. What measures distinguish genuine organisational learning from increased record production?

Each question requires a defined context, comparison condition, outcome measure and independent review. Until such studies exist, OrganoSphere should remain a design hypothesis and vocabulary—not a validated intervention.

8 Limitations, independence and acknowledgements

This paper is intentionally conceptual. It presents no user study, benchmark, clinical study, organisational trial or causal estimate. Its notation is explanatory rather than predictive. The proposed terms may overlap with established traditions in cybernetics, human-centred automation, organisational learning and software engineering; the references identify important foundations but do not constitute a systematic literature review.

The author writes independently through ITMan Intelligence and has a professional affiliation with NRF iThemba LABS, formally expanded by the organisation as the iThemba Laboratory for Accelerator-Based Sciences. That professional context is disclosed for transparency only. NRF and NRF iThemba LABS did not publish, sponsor, review or endorse this working paper.

Acknowledgements

Some questions developed here were shaped through interdisciplinary exchanges with medical-doctor colleagues with whom the author works. No individual is named, and those exchanges do not imply co-authorship, formal review, approval, clinical validation or institutional endorsement.

9 Conclusion

OrganoSphere is best understood as a proposition about organisational design. Human and machine intelligence become valuable together only when the environment makes boundaries, authority, evidence and learning explicit. Software engineering supplies much of that environment; systems thinking prevents the software boundary from being mistaken for the whole system; governance keeps consequential action answerable to people. The concept earns scientific value only by becoming testable, falsifiable and open to revision.

References

[1] S. Amershi et al., "Software Engineering for Machine Learning: A Case Study," *ICSE-SEIP 2019*, pp. 291–300, 2019. `https://doi.org/10.1109/ICSE-SEIP.2019.00042`.

[2] R. C. Conant and W. R. Ashby, "Every good regulator of a system must be a model of that system," *International Journal of Systems Science*, vol. 1, no. 2, pp. 89–97, 1970. `https://doi.org/10.1080/00207727008920220`.

[3] G. Klein, D. D. Woods, J. M. Bradshaw, R. R. Hoffman, and P. J. Feltovich, "Ten challenges for making automation a 'team player' in joint human-agent activity," *IEEE Intelligent Systems*, vol. 19, no. 6, pp. 91–95, 2004. `https://doi.org/10.1109/MIS.2004.74`.

[4] E. Tabassi, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, 2023. `https://doi.org/10.6028/NIST.AI.100-1`.

[5] E. Ostrom, "A general framework for analyzing sustainability of social-ecological systems," *Science*, vol. 325, no. 5939, pp. 419–422, 2009. https://doi.org/10.1126/science.1172133.

[6] H. A. Simon, "The architecture of complexity," *Proceedings of the American Philosophical Society*, vol. 106, no. 6, pp. 467–482, 1962. https://www.jstor.org/stable/985254.

[7] D. D. Woods, "Four concepts for resilience and the implications for the future of resilience engineering," *Reliability Engineering & System Safety*, vol. 141, pp. 5–9, 2015. https://doi.org/10.1016/j.ress.2015.03.018.