In Situ Analytics: Instrumentation as the Unit of Enterprise Intelligence

Akbar Akbari Esfahani
Manager, Enterprise Intelligence & Automation, Central California Alliance for Health
Adjunct Faculty, Carnegie Mellon University

Draft — August 2026

Abstract

That intelligence belongs inside the work rather than in a report is widely asserted and rarely explained. This paper supplies the account. Artificial intelligence can generate plausible records from existing information; it cannot recover ground truth about an event the organization never instrumented. It is also how the instrument that would observe that event gets built, so the asymmetry and its remedy come from the same technology on two different clocks. That makes observational capacity — not any corpus it produces — the durable organizational asset. An instrument is defined here by three necessary properties, giving a scoreable test of whether a system qualifies; it is also a delegation system, moving observational and decision authority to the point of work while centralizing accountability in its architecture; and the choice between building one and buying inference turns on marginal cost per unit of work rather than on capability, since instrumentation amortizes and inference recurs. The argument is cybernetic in Beer's managerial sense rather than technological. This is a conceptual paper; one instrumented function illustrates the constructs and is not offered as a test of them.

Keywords: operational instrumentation, proprietary observation, work-as-done, management cybernetics, requisite variety, enterprise AI

Disclosure. The author leads the enterprise intelligence function at the organization whose operational unit appears in Section 7 and helped build the system described there. That illustration is practitioner-reported and should be weighed accordingly.

Acknowledgment. The author's application of motivational interviewing to operational elicitation was

formed through work alongside Allan Zuckoff, and the evaluation discipline of Section 6 through work under Anita Cattrell. Neither bears responsibility for the argument made here.

---

1. What is established, and what is not

Enterprise AI fails at rates that model quality and data readiness do not explain, for a reason long established elsewhere. Every operational function has a documented process and an actual one, and competent work happens in the gap between them.^1,2^ Classification renders some work visible and the rest invisible, and the invisible portion disproportionately holds the system together.^3^ Prevailing deployment models treat that gap as a specification problem and optimize for a process nobody performs.

The remedy is equally well established, and this paper does not claim it. Zuboff distinguished technology that *automates* from technology that *informates* — generating information about the work as a byproduct of the work — and observed that informating serves surveillance as readily as insight.^9^ Process mining has argued for two decades that event logs recover the actual process better than documentation,^10^ and the learning health system literature has held since 2007 that knowledge should be captured as a byproduct of delivery.^11^ One payer's chief analytics officer marks as a milestone the point at which a prediction reached staff natively in the workflow rather than as an emailed spreadsheet.^12^

The ground is occupied. What the field asserts, it has not explained. Why is identical analytical content worth more inside the work than outside it? What makes the resulting asset durable rather than replicable? What decides between building an instrument and buying inference? And what must be true of a system before it is an instrument at all?

The questions are cybernetic ones, which locates the argument in management theory rather than in technology. Beer's management cybernetics asked how an enterprise acquires enough internal variety to regulate the variety it faces, treating that capacity as the object of design.^16^ Conant and Ashby's theorem states what a regulator must therefore contain; second-order cybernetics adds the condition this account does not try to escape, that the observer belongs to the system observed.^8^ What follows is not a new cybernetics but a missing unit for an old one — the instrument as the place where requisite variety is acquired.

On lineage. The framework descends from work on inference under incomplete observation — uncertainty as information deficiency rather than noise, limited measurement admitting many models rather than one, an event's origin located from the structure of sparse traces rather than their number^4^ — and from instrument design, where a population becomes visible to policy only through apparatus built to make it so.^6^ The second line ran through *artificial adaptive systems*, which are defined by what they refuse to do: they elicit the underlying model from the data rather than impose a model on it as the means of determining what relates to what.^5^ That construct returns in Sections 3 and 4 as the thing an instrument is once it is running, and in Section 5 as the discipline governing how one is built. This paper carries the line into the interior of an operating organization.

2. The asymmetry

Proprietary data was long assumed to be the durable advantage. Models now infer missing values, reconstruct populations, and close from public sources much of the gap a proprietary corpus once represented, which makes a dataset a depreciating asset.

One asymmetry survives. **Artificial intelligence can generate plausible records from existing information. It cannot recover ground truth about an event the organization never instrumented.** Synthetic records demonstrate this rather than contradict it: generated from the distribution of what was recorded, they are indistinguishable, from inside the model, from records of events that occurred. A model queried about something never observed does not fail loudly; the absence never appears in the output.

This is the inverse problem at organizational scale. Uncertainty from imprecise measurement is a different object from uncertainty from absent measurement: the first is narrowable by better inference, the second only by better instrumentation, because there is nothing for inference to operate on. Synthesis is now abundant; instrumentation remains scarce.

Operational systems record the claim, the authorization, the disposition — the residue of work already finished. What no system records is the work itself: the second screen open because one system holds eligibility and another authorization, the undocumented rule for which cases escalate. That interior exists only if observed while happening.

Two constructs must be separated. *An observation* is a trace: a machine-recorded account of a unit of work at the moment it occurred. *Observational capacity* is the standing ability to produce such traces as a byproduct of work. Read as traces the asymmetry is near-trivial; read as capacity it is the

claim that matters, since traces are reproducible by anyone who installs the same capacity while the capacity itself cannot be bought as a finished generic artifact — however developed, it must be instantiated inside the work it observes. Capacity generates traces; traces do not generate capacity.

Capacity also outlasts its outputs. Every representation of work is a map and the territory keeps moving beneath it,^7^ so a single act of observation begins ageing immediately; what resists decay is the standing capacity. And it is proprietary only to the degree the workflow it observes is one the organization controls — where a workflow is standardized across an industry, the instrument is replicable.

3. The instrument

Conventional architecture makes the report the destination, and the moment a worker leaves the work to read one, the architecture has already failed. This is specimen-based analytics: examining dead tissue under a microscope and calling it intelligence. The alternative takes *instrument* in its scientific sense — it does not observe passively but structures how observation occurs and alters the system observed.

An instrument holds three properties at once. It *captures a decision-relevant event as part of the work* rather than through a separate analytical episode. It *conditions action while the decision remains open*. And it *returns what was learned from prior action to condition what follows*. A system holding two of the three is a report, a monitoring tool, or a dashboard — each may be worth having, none is an instrument.

Placement also changes the distribution of authority. An instrument centralizes accountability in its architecture — the permitted actions, evidence standards, escalation thresholds, security boundaries, and outcome definition — while distributing observational and decision authority to the point of work. The worker no longer petitions an analytical or technical centre to reconstruct the case, interpret its state, or alter the workflow. Those powers become locally exercisable within a centrally governed boundary. An instrument is therefore not only a measurement system but a delegation system: it converts organizational policy into bounded operational authority.

That answers the first of the questions above. Outside the work, analytical content informs the person who already controls the information; inside it, the same content relocates the authority to use it. **Adoption is the diffusion of that authority, not the utilization of the interface.** Logins and page views establish exposure. Adoption occurs when eligible work passes through the instrument,

decisions are completed locally within its governed boundaries, and the resulting observations alter subsequent action. A technically deployed system can therefore remain unadopted, while an adopted instrument changes who can know, decide, and act.

The properties are structural and deliberately separate from whether the instrument works. **Instrument status is a question of form; instrument effectiveness is a question of evidence.** A badly designed instrument remains an instrument and can be shown to be a poor one; a well-designed report does not become an instrument by being useful. That keeps the category assessable without outcome data, and stops the definition from guaranteeing its own success.

The third property carries the most weight. A system that improves the next action is one whose representation is revised by what it observes. An effective regulator must embody a model of the system it regulates,^8^ and a model that cannot update cannot regulate a system that moves.

Two neighbours differ from this. Process mining shares the premise that the trace is the object of analysis but is log-first; this account operates upstream, on how a log worth mining comes to exist and on what terms with the people whose work produces it. ITIL Version 5 integrates product, service, experience, transformation, and governance across a common lifecycle, and treats value as co-created rather than delivered.^17^ It supplies a management architecture for distributed, continually improving digital work. It does not explain how an organization constructs observational capacity where the relevant work is absent from existing systems, nor how authority actually moves from a technical centre to the point of work. The instrument is the mechanism through which those aspirations become operational and observable.

One measure follows. If the asset is continuous reconciliation, the quantity worth managing is the lag in it: *adaptation latency*, the time between a meaningful divergence in operational reality and a governed change in the organization's executable response. Conventional programmes measure delivery against a specification frozen at the outset, which is fidelity to a map rather than contact with a territory. Adaptation latency is the regulator's time constant, and only an instrumented organization can report it.

The primary deliverable of an analytics function is therefore neither a model nor a dashboard but the instrument. Absent one, an organization observes late, and a framework certifying that rules were followed cannot report that the rules stopped mapping to reality. **The failure is not detected. It is survived.**

What adapts is not the software. The instrument, the unit whose understanding is revised by what

the instrument returns, and the workflow altered by both together constitute one artificial adaptive system, its structure elicited from what passes through it rather than imposed in advance.^5^ Adaptation latency is then not only a management measure but that system's learning rate: the same quantity read from outside and from within.

4. The economics

What decides between building an instrument and buying inference is not capability but marginal cost per unit of work. The discriminator is often put as instrument versus inference, and put that way it is wrong — the illustration in Section 7 carries a model inside the instrument. The distinction that holds is between what capitalizes and what does not: **inference that capitalizes into structure amortizes; inference that stays in the loop recurs.**

Artificial intelligence therefore appears on two clocks. On the construction clock it is a fabrication technology — compressing elicitation, generating code, encoding a boundary the unit could not articulate — and that is fixed cost, paid once, which is exactly the F_I the crossing point below already assumes. On the operating clock what runs is an adaptive system at low marginal cost per unit of work. The objection that adaptive systems carry their own recurring cost of adaptation is correct and not damaging: that cost is the learning, and the learning is what is being bought.

Both carry fixed and variable components, and the comparison is only honest if both are written out. Over *n* units of work:

$$C_I(n) = F_I + n \cdot V_I$$

$$C_G(n) = F_G + n \cdot (V_G + R_G + H_G)$$

where F_I is instrument construction and integration, V_I its execution and maintenance per unit; F_G is inference-system integration, V_G model and retrieval cost per unit, R_G retry and context reconstruction, and H_G the incremental human review a consequential determination still requires. The crossing is

$$n^* = (F_I - F_G) / (V_G + R_G + H_G - V_I)$$

and exists only where the denominator is positive. Where per-unit inference cost net of retry and review falls below the instrument's, there is no crossing and inference dominates at every volume —

the honest form of the claim, and the form that makes it testable.

The relation inverts wherever the crossing is not reached — at low volume, where nothing amortizes; at high case variance, where there is no stable boundary to encode; on unstructured input; and under bounded exposure, where errors are cheaply caught. There, general inference is what is being purchased, and purchasing it is correct.

The two clocks require two roles, and conflating them is where the argument would fail. As *scaffolding*, artificial intelligence accelerates construction and comes out at handover, leaving no trace in the running system; its cost is fixed. As a *component* it stays and becomes part of the adaptive system's transfer function — the confidence verb of Section 7 is a component, not scaffolding. A component is a recurring cost by construction, and earns its place only where what it contributes cannot be capitalized into structure. What the argument excludes is neither role but the request-and-return pattern, where a worker leaves the task to consult a system and comes back, whatever implements it.

5. Building one

Instruments are built inside the unit whose work they will observe, across six stages.

0. Permission. Staff do not name friction freely, and their reasons are sound: a workaround admits the official process does not work, and hours lost to a system sound like a complaint about whoever chose it. The pairing states what will and will not be recorded, and that what is named will not be held against the person naming it. Without this the remaining stages yield work-as-imagined, because that is what people describe when description carries risk.

1. The walk-through. Staff describe their process under a motivational-interviewing stance — open questions, reflective listening, no correction — with adherence monitored rather than assumed.^13^ Two framings do most of the work: ask about the *last* case, since the typical one is work-as-imagined; and ask what made it harder than it needed to be, rather than what the problems are.

2–5. Wishlist, minimum viable ask, co-owned iteration, architectural addition. A specification freezes the process as currently understood, so what the unit produces is a wishlist — evidence of the need, not scope, and always an understatement of it. The refusal is the same one that defines an adaptive system, applied to people rather than to data: elicit the model, do not impose it. The smallest coherent item is built to production, because a working artifact establishes that reported

problems produce fixes rather than ticket numbers. Iteration is co-owned through two commitments — the instrument adopts the unit's language and the unit's own measures — not through consultation. The final stage proposes what the system should become; an addition the unit does not adopt is not shipped.

Recursion. Each completed cycle updates three objects: the unit's understanding of its desired outcome, the local instrument through which that outcome is pursued, and the shared architecture from which later instruments are built. The output is therefore not only an operational result but new evidence about the work and new capacity for producing the next result. An instrument compounds only when all three are allowed to change; iteration confined to the interface is maintenance, not recursion.

Stages 0–4 are protocol, with entry conditions, an artifact, and an exit test; Stage 5 is craft. An engagement ends when the unit can observe its own measures, exercise bounded decision authority through the instrument, and identify the next required adaptation without technical mediation. The scaffolding is out by then; what remains is the instrument, whatever components it carries, and the unit.

6. The boundary, and where the account fails

Observation of work becomes observation of workers unless prevented by construction — the more so where the instrument also carries authority. The object is the digital interaction, not the employee: the record is keyed to the transaction, the person is not its key, and reporting is by function. Person-level attribution exists only in the audit logging regulated environments require, governed separately and never joined to the observation store. The boundary is enforced through access control rather than undertaking, since good intentions are not a structural guarantee.^14^ **An organization that asks staff to expose the seams in their own work and then measures them on what they exposed will not be told anything true a second time.**

A second risk is behavioural. A measure participants understand can become a target rather than a description.^15^ Once staff know how an allocation rule uses a measure, behaviour can adapt to the measure rather than to the work — which bears on the account's most distinctive prediction, that an instrumented function keeps improving rather than plateauing. Sustained improvement is equally consistent with participants learning what the instrument rewards.

A single-level instrument cannot separate the two. A recursive one can, which is where the

second-order condition becomes a design requirement rather than a defeat. An operational layer returns to each worker the measures of their own work; a supervisory layer observes those same measures alongside the outcome they are supposed to stand for. Adaptation to the measure has a signature that exists only at the second level: the local measure improves while the outcome it proxies does not. Continuous deployment makes that signature actionable, since a drifting measure is redefined at the latency of a release rather than of a planning cycle. This is Beer's recursion applied to the instrument,^16^ and the learning rate decides whether it works — the loop is adequate when detection plus correction runs shorter than the time an adapted measure needs to entrench. One limit remains: a second level cannot certify an outcome the same function also instruments, so at least one must sit outside the instrument's own reporting surface.

The account has boundaries. It requires a digital surface; volume, since returns come from friction repeated many times; a modifiable workflow; architectural authority; and a unit not under threat, because where reduction in force is live, permission cannot be established honestly.

What would falsify it: if instrumented functions show no improvement in the conversion of work time to service against their own prior baseline; if improvement appears at deployment and decays rather than compounding; or if improvements are attributable to volume, staffing, or case-mix change rather than to the instrument.

7. Illustration

This section shows what the constructs look like in operation. It is not an evaluation and offers no causal estimate.

A provider dispute resolution function at a Medi-Cal managed care plan: twelve staff, two case types, the case as unit of work. A two-person pairing built over three months, covering claim inquiries. Model output reached the function as a spreadsheet in November 2025; the applications launched 15 December. What was built: a unified case environment; a model whose output was a **confidence verb** — Very Confident through Uncertain — rather than a probability, preserving agency where the model is least certain, and a component rather than scaffolding in the sense of Section 4; automated allocation; and a management application carrying the unit's outcome indicators beside the staff-level measures. Determinations that had required a technical or analytical intermediary became locally completable within the boundary the architecture sets, which is what the delegation claim looks like in operation; the nested pair is Section 6's recursion, both layers deploying continuously from one codebase.

Being one function at one site, described by the team that built it, the case shows that the constructs have referents and establishes nothing further.

A failure, by the same test. The first proof of concept in this portfolio was a call-centre knowledge base, built to specification and delivered as specified. It failed because a knowledge base is a *destination*: a representative on a call still had to leave the work to search it and return. Against the three properties it satisfied at most one — it could inform an open decision, but the informing was a separate episode, and nothing learned from prior calls conditioned the next. Such failure is invisible to conventional deployment measures: on schedule, matched to specification, demonstrable. What it did not do was change the work.

8. Conclusion

An enterprise cannot govern what it observes too late, cannot improve what it only reconstructs, and cannot trust intelligence built on observation never structured to be trustworthy. That instruments belong inside the work is not a new claim; what is offered here is why it holds — an asymmetry making observational capacity durable where its outputs are not, a definition making "instrument" a testable category rather than a metaphor, a delegation structure explaining what adoption actually is, and a cost structure deciding when to build one. That is not merely a technology argument. It is an epistemological and organizational one: the placement of observation determines what can be known, the placement of the instrument determines who can act, and the architecture surrounding it determines who remains accountable. Instrumentation centralizes the integrity of the outcome while diffusing the power to produce it.

References

1. Suchman, L. A. (1987). *Plans and Situated Actions.* Cambridge University Press.
2. Hollnagel, E. (2015). Why is Work-as-Imagined Different from Work-as-Done? In *Resilient Health Care, Vol. 2.* Ashgate.
3. Bowker, G. C., & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences.* MIT Press.
4. Minsley, B., Deszcz-Pan, M., Esfahani, A., Bedrosian, P., Abraham, J., & Finn, C. (2013). Quantifying Uncertainty in Airborne Electromagnetic Data Using Bayesian MCMC Methods.

Symposium on the Application of Geophysics to Engineering and Environmental Problems (SAGEEP). See also Friedel, M. J., Esfahani, A., & Iwashita, F. (2016). Toward Real-Time Three-Dimensional Mapping of Surficial Aquifers Using a Hybrid Modeling Approach. *Hydrogeology Journal,* 24(1), 211–229.

5. Bronstein, A. C., Buscema, M., Esfahani, A., Lodwick, W. A., & Grossi, E. (2013). Locating the Source of Public Health Events Using Intelligent Adaptive Systems: 2011 United States Listeriosis Outbreak Linked to Whole Cantaloupes. *Clinical Toxicology,* 51(7), 625–626. For the methodological statement of artificial adaptive systems, see Buscema, P. M., Massini, G., Breda, M., Lodwick, W. A., Newman, F., & Asadi-Zeydabadi, M. (2018). *Artificial Adaptive Systems Using Auto Contractive Maps.* Springer International Publishing.
6. Esfahani, A. A., Jans, M., Ponce, N. A., & Ganz, C. (2015). How to Code School Names More Efficiently: Common Sense, Scripting, and a Novel SAS Application. *American Association for Public Opinion Research, 70th Annual Conference,* Hollywood, Florida.
7. Korzybski, A. (1933). *Science and Sanity.* International Non-Aristotelian Library.
8. Conant, R. C., & Ashby, W. R. (1970). Every Good Regulator of a System Must Be a Model of That System. *International Journal of Systems Science,* 1(2), 89–97. See also Ashby, W. R. (1956). *An Introduction to Cybernetics.* Chapman & Hall; von Foerster, H. (2003). *Understanding Understanding: Essays on Cybernetics and Cognition.* Springer.
9. Zuboff, S. (1988). *In the Age of the Smart Machine.* Basic Books.
10. van der Aalst, W. M. P. (2016). *Process Mining: Data Science in Action* (2nd ed.). Springer; Kubrak, K., et al. (2022). Prescriptive Process Monitoring: Quo Vadis? *PeerJ CS,* 8, e1097.
11. Institute of Medicine (2007). *The Learning Healthcare System.* National Academies Press.
12. Clarke, R. (2024). Interview, *Agile Giants,* Episode 67, with S. Ammirati.
13. Miller, W. R., & Rollnick, S. (2013). *Motivational Interviewing* (3rd ed.). Guilford Press.
14. Suchman, L. A. (1994). Do Categories Have Politics? *Computer Supported Cooperative Work,* 2(3), 177–190.
15. Ridgway, V. F. (1956). Dysfunctional Consequences of Performance Measurements. *Administrative Science Quarterly,* 1(2), 240–247; Simons, R. (1995). *Levers of Control.* Harvard Business School Press.
16. Beer, S. (1979). *The Heart of Enterprise.* Wiley; Beer, S. (1972). *Brain of the Firm.* Allen Lane.
17. PeopleCert (2025). *ITIL 5,* including the Transformation module. PeopleCert International.