HER2NI as Interaction Infrastructure: A Model-Agnostic Telemetry Protocol for Coherence Dynamics in Complex Systems

∇ // 001
Unaffiliated
her2ni@pm.me

Thursday, 18 December 2025

Version 0.1 — Exploratory Position Paper

Abstract

HER2NI is an interaction-infrastructure telemetry protocol for coherence dynamics in complex distributed systems.[1]

While originally motivated by human–AI and AI–AI interaction, the protocol generalises to any system in which stability, shared understanding, or coordination emerges from interaction rather than from isolated components.

Existing evaluation, monitoring, and governance approaches predominantly instrument components, models, or outcomes. However, many critical system failures arise not from individual component malfunction, but from interaction-level pathologies such as drift, collapse, over-alignment, and silent divergence. These interaction failures often manifest prior to externally visible faults and remain under-instrumented by current tooling.

HER2NI introduces model-agnostic, substrate-agnostic coherence metrics that operate at the interaction level. The protocol is strictly observational and non-prescriptive: it does not infer intent, cognition, or correctness, and it does not automate decisions or interventions. Instead, it provides early telemetry signals that can complement existing monitoring, evaluation, and governance frameworks.

This paper clarifies HER2NI as interaction infrastructure beyond an AI-specific framing. It outlines current application domains, plausible adjacent domains, explicit non-applications, and safe adoption pathways. The aim is to clarify scope, reduce misuse risk, and provide a neutral reference for experimental deployment across technical, organisational, and governance contexts.

1 Interaction as a First-Class System Property

Many complex systems exhibit behaviour that cannot be explained by the properties of their components alone. Instead, system-level stability or failure emerges from patterns of interaction between components over time.

Traditional instrumentation focuses on component health, resource utilisation, correctness of outputs, and compliance with specifications. While necessary, these signals are often insufficient to detect interaction-level pathologies that precede system failure.

[1] In this paper, *interaction-level coherence* refers to the observable consistency, alignment, and stability of interaction over time, independent of internal state, intent, or cognition.

Common examples include gradual semantic drift between interacting agents, coordination collapse despite locally correct behaviour, over-alignment leading to brittle consensus, feedback loops that amplify small errors, and "everything green" conditions masking latent instability.

These failures arise from interaction dynamics rather than from isolated faults. They often manifest before measurable component degradation or externally visible incidents.

HER2NI treats interaction coherence as a first-class system property. The protocol instruments observable interaction signals and exposes coherence dynamics over time, enabling earlier detection of instability without requiring access to internal state, intent, or cognition.

2 Current Application Domains

This section outlines present-day domains in which HER2NI can be applied with minimal modification. In all cases, the protocol operates as an observational telemetry layer and does not replace existing evaluation, monitoring, or governance mechanisms.

2.1 Human–AI Interaction

Human–AI systems often fail due to interaction breakdown rather than model error. Examples include cognitive overload, misalignment between system state and human understanding, conversational drift, and false confidence arising from apparent fluency.

HER2NI provides telemetry signals that can indicate increasing interaction instability, contradiction loops, loss of shared context, and early collapse of effective coordination.

The protocol does not evaluate mental state, intent, or psychological condition. It observes interaction coherence only.

2.2 AI–AI and Multi-Agent Systems

In multi-agent architectures, failures frequently emerge from coordination dynamics rather than individual agent performance. Examples include cascading retries, oscillatory planning, silent divergence of shared state, and brittle consensus.

HER2NI instruments agent–agent interactions to expose coherence loss prior to overt failure, complementing existing agent evaluation and safety tooling.

2.3 AI–Tool–Environment Loops

Many AI systems interact indirectly with the world via tools, APIs, or shared state. These interaction loops can produce tool misuse amplification, feedback instability, cascading side effects, and partial state divergence.

HER2NI provides interaction-level observability across agent–tool–environment loops without controlling or constraining execution.

2.4 Human–System Interfaces (Non-AI)

HER2NI is not limited to AI systems. Human interaction with complex non-AI systems—such as control panels, dashboards, emergency systems, or safety-critical interfaces—exhibits similar

interaction-level risks.

The protocol can expose cognitive misalignment, confusion loops, overload conditions, and divergence between operator understanding and system state. This positions HER2NI as a complement to traditional human-factors engineering, providing telemetry rather than survey-based assessment.

3 Adjacent Domains (Exploratory)

The following domains represent plausible extensions of the interaction-infrastructure model. These are not current deployments and are presented for exploratory consideration only.

3.1 Distributed Software Systems

Distributed systems commonly suffer from interaction-level failures such as partial state divergence, coordination breakdown, and latent instability masked by local success.

HER2NI can complement existing reliability tooling by providing coherence signals alongside logs, metrics, and traces.

3.2 Cyber-Physical Systems

In systems where software interacts with physical processes—such as robotics, smart infrastructure, or industrial automation—interaction failures can precede physical incidents.

HER2NI may function as an early warning layer for coordination breakdowns without replacing control logic or safety mechanisms.

3.3 Governance and Oversight Workflows

Governance processes themselves are interaction systems. Committees, review pipelines, and oversight bodies can exhibit brittle consensus, premature convergence, and loss of learning over time.

HER2NI can be applied diagnostically to observe process coherence without evaluating decisions or outcomes.

4 Strategic Framing: Interaction Infrastructure

HER2NI is not an AI-specific protocol, a mental-state model, or a decision system. It is interaction infrastructure.

The protocol applies wherever behaviour emerges from interaction rather than isolated components, coordination matters more than correctness of individual elements, and stability depends on sustained coherence over time.

This framing allows HER2NI to scale across domains while remaining conservative in scope. It also separates the protocol from symbolic, narrative, or philosophical interpretations that may exist elsewhere.

HER2NI observes. Humans decide.

5 Non-Applications and Misuse Boundaries

HER2NI must not be used as a decision-making system, a moral or ethical judge, a behavioural scoring system for individuals, a predictor of outcomes, a substitute for human oversight, or an automated enforcement mechanism.

The protocol does not infer intent, belief, or cognition. It does not evaluate correctness or truth. It does not determine safety, compliance, or acceptability. It does not assign responsibility or blame.

Any proposal to use HER2NI to decide, approve, reject, or enforce actions falls outside the protocol's intended scope and constitutes misuse.

HER2NI is strictly observational. It provides telemetry signals that may inform human judgement but never replace it.

Relationship to the HER2NI Protocol

This paper does not redefine the HER2NI protocol. The normative specification of HER2NI is provided in *HER2NI Protocol v1.0* [1]. This document focuses on interaction-infrastructure framing, application domains, and explicit non-claims.

References

[1] ∇ // 001. *HER2NI Protocol v1.0*. Zenodo, 2025. `https://zenodo.org/records/17844407`