



Submitted : 10 March, 2026
Accepted : 13 March, 2026
Published : 14 March, 2026

*Corresponding author: Liz Johnson, University of North Carolina Charlotte, 9201 University City Boulevard, Charlotte, NC 28223, USA, E-mail: elizabeth@charlotte.edu

Keywords: Trustworthy AI; AI ethics; Relational AI; Human-AI interaction; Constraint-aware reasoning; Embodied AI; Hallucination mitigation; Trust calibration; Human-centered decision support; Cognitive architecture

Copyright License: © 2026 Johnson L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<https://www.engineergroup.us>



Check for updates

Research Article

From AI Stack to Cognitive Stack: A Semantic–Spatial–Relational Architecture for Trustworthy AI

Liz Johnson*, Arham Hussain Inamdar and Bruh Lemma Yadecha

University of North Carolina Charlotte, 9201 University City Boulevard, Charlotte, NC 28223, USA

Abstract

This paper introduces a layered Semantic-Spatial-Relational (SSR) Cognitive Stack designed to embed trust calibration, epistemic boundaries, and human authority directly within AI system architecture. Contemporary large language model deployments often emphasize generative performance while under-specifying mechanisms for environmental grounding, relational governance, and epistemic restraint. The SSR architecture addresses this gap by integrating three complementary layers: semantic reasoning for intent interpretation, spatial grounding through contextual perception, and a relational governance layer that regulates interaction boundaries and trust calibration. Rather than treating ethical constraints as post hoc safeguards, the architecture embeds structural mechanisms that govern when and how AI systems should act.

This paper presents the conceptual design of the SSR Cognitive Stack and analyzes its implications for trustworthy AI, embodied systems, and human-AI decision support. We argue that ethical reliability in AI systems depends not solely on output accuracy but on structurally embedded governance mechanisms that regulate action, restraint, and authority. The framework contributes to ongoing debates in AI ethics by proposing an integrity-centered model of AI design in which responsibility, transparency, and relational stability function as architectural principles rather than external controls. Finally, we outline directions for future empirical research needed to evaluate relationally structured AI architectures across educational, commercial, and public-sector environments.

Introduction

Conceptual foundations

Large Language Model (LLM) systems have demonstrated unprecedented generative capabilities; however, significant challenges remain in ensuring clear boundaries, calibrated trust, and epistemic restraint in AI-generated outputs [1]. A growing body of research has examined hallucination mitigation, knowledge integrity, and adversarial robustness in LLMs, often focusing on techniques such as retrieval augmentation, post hoc correction mechanisms, and probabilistic uncertainty signaling designed to reduce erroneous or overconfident responses [2,3]. While these approaches improve output reliability, they largely

treat errors as downstream problems rather than examining the architectural factors that shape how AI systems interpret context and regulate responses.

Many contemporary AI failures arise not from limitations in data or computational capacity, but from structural characteristics of current language-model architectures. Large language models generate responses through probabilistic token prediction rather than grounded reasoning, which can produce confident, yet inaccurate outputs commonly described as hallucinations [4,5]. Because these systems primarily operate as semantic pattern predictors, they may misinterpret contextual cues, overlook relational dynamics within interactions, or generate responses that exceed their epistemic



boundaries [6,7]. These limitations reveal a structural gap between generative language capability and the broader cognitive architecture required for reliable decision support. While LLMs are highly effective language models, they do not inherently incorporate mechanisms for contextual grounding, relational governance, or constraint-aware reasoning, which are necessary for trustworthy human-AI collaboration.

Research in trustworthy AI has emphasized transparency, explainability, and fairness, while human-in-the-loop approaches attempt to preserve oversight through supervisory checkpoints [8]. However, these methods often treat trust calibration and restraint as external safeguards rather than architectural design principles. In parallel, embodied AI research has demonstrated that robust intelligence emerges through situated interaction with environmental context rather than abstract computation alone [9,10]. Although such work highlights the importance of environmental grounding, spatial perception is frequently developed independently from semantic reasoning and relational governance [11]. As a result, many current AI systems remain modular and transactional rather than cognitively integrated.

Modern decision environments are often framed as data problems characterized by overwhelming volumes of signals and information, often producing cognitive overload for decision-makers [12,13]. However, the central constraint is not sensing but sensemaking. When AI systems merely accelerate output generation without improving the quality of interpretation and judgment, they risk amplifying errors rather than creating a decision advantage. In such conditions, rapid responses may mis-calibrate trust, amplify epistemic overreach, and substitute computational speed for discernment.

Decision advantage emerges when AI systems integrate semantic meaning, environmental grounding, relational context, and human cognitive constraints within a unified architecture that governs not only how systems respond but also when they should refrain from acting. AI autonomy without relational awareness can amplify risks, including automation complacency, epistemic overreach, and erosion of human authority. Systems that generate plausible outputs without contextual grounding or calibrated trust may therefore undermine rather than enhance human judgment.

To address this architectural gap, this paper proposes the Semantic-Spatial-Relational (SSR) Cognitive Stack, a layered framework integrating semantic reasoning, spatial contextual grounding, and relational governance within a unified decision-support architecture. Within the proposed model, semantic AI interprets intent and maintains conceptual coherence, spatial AI grounds interpretation within environmental context, and relational AI governs trust calibration, boundary enforcement, and interaction stability. Above these layers, a Decision Layer preserves explicit human authority to act, pause, or refrain. The central contribution of this study is the argument that relational intelligence¹ must govern not only how AI systems respond but also when they should refrain from action, a capacity essential for ethical reliability under uncertainty [14].

The framework presented here integrates these layers through an architectural model intended to support integrity-centered AI design. Rather than relying on prompt-level safeguards or post hoc corrections, the SSR Cognitive Stack embeds trust calibration and constraint enforcement directly within system architecture. This approach builds upon prior research in trustworthy AI, embodied cognition, and human-centered decision systems while extending these traditions toward relationally structured AI architectures.

Relational modeling within this framework draws conceptually on traditions that emphasize interdependence and dynamic interaction rather than static entities. Philosophical accounts of relational ontology [15], autopoietic systems theory [16], and process philosophy [17] similarly frame systems as emerging through structured interaction rather than isolated components. While the present study does not attempt metaphysical formalization, these perspectives support an architectural shift from isolated AI modules toward interdependent cognitive layers.

Definitions and terminology

This paper, therefore, introduces the SSR Cognitive Stack as a conceptual architecture for relationally structured AI systems. While the framework is informed by early deployment and testing environments, the present work primarily articulates the architectural principles underlying the approach and outlines a research agenda for future empirical validation.

For conceptual clarity, the term Large Language Model (LLM) is used in this work consistent with its prevailing usage in artificial intelligence research. The term emphasizes that these models operate on the structural and semantic properties of language as a representational system, rather than merely processing isolated text tokens. Language encodes relational meaning, contextual dependencies, and symbolic structures that allow models to perform reasoning-like operations across sentences, concepts, and discourse. While alternative terms such as “Large Text Model” or “Large Knowledge Model” could emphasize different aspects of the technology, the term “language” is preferred because it reflects the models’ capacity to interpret and generate structured linguistic representations rather than simply store textual data.

In this context, language is treated as a shared symbolic system that individuals use to communicate meaning rather than a construct that can be individually redefined by a single participant. While language evolves collectively through social use, individual speakers typically operate within established

¹Relational intelligence refers to the capability of an AI system to structure interactions in ways that maintain contextual continuity, calibrated trust, and appropriate behavioral constraints across human-AI exchanges. Rather than modeling emotion alone, relational intelligence governs how systems interpret social signals, regulate responses, and preserve alignment with human expectations and decision contexts (Johnson & Cochran, 2026).



linguistic conventions that allow communication to remain interpretable across participants. This perspective aligns with foundational work in linguistics and philosophy describing language as a socially maintained representational system through which meaning and knowledge are expressed [18–20].

In this paper, Artificial Intelligence (AI) refers to computational systems designed to perform tasks that typically require human cognitive capabilities, including perception, reasoning, learning, and decision-making [21]. Cognition refers to the processes by which information is interpreted, organized, and used to guide understanding and action. Within the proposed architecture, cognition is operationalized through layered mechanisms that interpret meaning (Semantic Layer), environmental context (Spatial Layer), relational interaction signals (Relational Layer), and decision arbitration (Decision Layer).

Methods

Overview

This research employs conceptual analysis² and analytic systems architecture methods³, drawing from traditions in analytic philosophy, trustworthy AI, and cognitive systems engineering. The study decomposes AI system design into semantic, spatial, and relational layers to clarify how epistemic boundaries, contextual grounding, and relational governance can be structurally embedded within AI systems. Through comparative theoretical analysis of existing AI architectures, the research identifies limitations in semantic-dominant models that lack mechanisms for trust calibration and environmental grounding. These insights are synthesized into the proposed SSR Cognitive Stack, which integrates layered cognitive functions within a unified architecture. The framework provides a theoretical foundation for future empirical research evaluating relationally structured AI systems across educational, commercial, and public-sector environments.

This study develops a conceptual architecture for relationally structured AI systems through architectural analysis and synthesis of prior research in trustworthy AI, embodied cognition, and human–automation trust. The proposed SSR Cognitive Stack integrates principles drawn

from large language model research, computer vision, spatial computing, and relational interaction theory.

Rather than evaluating system performance through a controlled experimental deployment, this work focuses on architectural design principles and theoretical integration. The goal is to define the structural requirements necessary for AI systems capable of maintaining epistemic boundaries, contextual grounding, and relational stability. These principles are derived through comparative analysis of existing AI architectures and synthesis of design requirements identified in the literature on trustworthy AI, human–AI interaction, and cognitive systems engineering [7,22–26].

System architecture: The semantic-spatial-relational cognitive stack

Semantic layer: Intent interpretation and constraint-aware reasoning: The Semantic Layer interprets user inputs within bounded reasoning structures defined by explicit intent parameters and contextual constraints. Intent parameters specify the interaction objective (e.g., an informational query, an explanation, or a conversational exchange), the scope of allowable responses, and the categories of actions the system may perform. Contextual constraints incorporate dialogue history, system capability limits, and environmental signals provided by the Spatial Layer. Together, these parameters define the permissible reasoning space within which responses may be generated, ensuring that outputs remain aligned with system capabilities and relational boundaries.

Traditional large language models rely primarily on probabilistic token prediction. While this approach enables fluent language generation, it may lack explicit mechanisms for constraint grounding and contextual governance, increasing the risk of hallucination or goal drift. The Semantic Layer addresses these limitations by incorporating structured meaning representations and constraint-aware reasoning mechanisms, such as ontology-based knowledge structures, rule-based filters, and parameterized response policies. These mechanisms emphasize traceable and bounded reasoning processes rather than purely generative fluency, supporting more explainable and accountable system behavior. In this architecture, the Semantic Layer produces candidate interpretations and response structures that are subsequently evaluated by relational governance and decision arbitration components before any system action is taken.

Spatial/Contextual constraint layer: Environmental grounding: The Spatial Layer maintains a structured environmental state that represents entities and their spatial relationships within the interaction environment. Visual inputs are processed through computer vision pipelines that detect entities and estimate their positions relative to the system. From these detections, the Spatial Layer constructs a simplified spatial state model comprising entity identifiers, spatial coordinates, proximity relationships, and movement vectors, where available. Rather than implementing full geometric world modeling, the system maintains a lightweight representation sufficient for contextual grounding.

²Conceptual analysis is distinguishing terms, analyzing their meanings, and representing fundamental entities and/or phenomena within a discipline. It focuses on understanding abstract ideas or mental images represented by theoretical terms, constructing a framework for making sense of the world, and articulating concepts within an ontology (<https://www.sciencedirect.com/topics/computer-science/conceptual-analysis>).

³An analytics system integrates multiple analytic components so that their combined operation produces knowledge outputs beyond the capability of any individual component (<https://www.sciencedirect.com/topics/computer-science/analytics-system>). Within the SSR Cognitive Stack, semantic interpretation, spatial contextual grounding, and relational governance function as interdependent analytic layers whose integration enables context-aware reasoning, calibrated interaction, and constraint-aware decision support.



These spatial variables are passed to the Semantic Layer as structured context signals that can influence interpretation and response generation. For example, the presence and location of detected entities may constrain which explanations, references, or visual outputs are relevant to the user's query. In this way, the Spatial Layer does not independently determine meaning but provides environmental state information that semantic reasoning can incorporate into the running conversational context.

Computer vision and contextual sensing detect entities, estimate trajectories, and differentiate humans from objects. Spatial inputs are represented as structured environmental signals that describe the presence, location, and relationships of detected entities. For example, identifying a mobile phone is a perceptual classification performed within the Spatial Layer. Interpreting that device as a recording tool, communication device, or conversational signal is performed by the Semantic Layer, which integrates spatial observations with the running interaction context.

Environmental layout, proximity, and movement patterns function as constraint fields shaping interpretation. The Spatial Layer also supports embodied visualization modalities, including AR and holographic interfaces, enabling cognitive offloading through structured spatial representations. Spatial grounding complements semantic reasoning by embedding meaning within real-world constraints.

Relational trust calibration layer: Boundary enforcement: The Relational Layer governs interaction stability, boundary enforcement, and trust calibration. While semantic processing interprets intent and spatial processing grounds context, relational governance maintains integrity under conversational and social pressure.

Structured boundaries operate across four domains:

1. **Epistemic:** Avoids unverifiable claims and signals uncertainty.
2. **Identity:** Maintains clarity regarding non-human status.
3. **Relational:** Preserves tone stability under adversarial prompts.
4. **Functional:** Prevents scope overreach and clarifies operational limits.

Cross-layer coupling

The SSR architecture derives its strength from structured coupling between layers rather than independent subsystems. Each layer contributes a distinct type of constraint to a shared decision process rather than directly producing system outputs. The Semantic Layer interprets user inputs and generates candidate meaning structures and reasoning paths. The Spatial Layer contextualizes those interpretations by grounding them in environmental conditions, spatial relationships, or operational constraints. The Relational Layer evaluates

interactional stability and enforces boundaries related to epistemic limits, identity clarity, and functional scope.

These layers exchange information bidirectionally. Spatial context can refine semantic interpretation by constraining what actions or explanations are relevant within a given environment. Semantic structures guide relational evaluation by identifying the goals, tone, and scope of an interaction. In turn, relational governance modulates both semantic reasoning and spatial presentation by enforcing trust calibration, ethical boundaries, and authority constraints.

Importantly, the layers do not directly generate final responses. Instead, each produces structured signals that the Decision Layer integrates, determining the appropriate system action. Possible actions include generating a language response through the semantic generation component, rendering spatial representations through visualization modules, requesting clarification, or deferring to human oversight. In this way, cross-layer coupling ensures that system outputs reflect coordinated evaluation of meaning, context, and relational governance rather than the isolated behavior of any single module.

Integrated decision architecture

The SSR architecture operates as a layered processing pipeline in which each module contributes structured constraints rather than directly generating final outputs. Incoming inputs, including user prompts, dialogue context, and environmental signals, are first interpreted by the Semantic Layer, which identifies intent, task structure, and candidate reasoning paths. These semantic interpretations are then evaluated within the Spatial Layer, which grounds them in environmental context by incorporating spatial relationships, object positioning, and operational constraints where available. The resulting contextualized interpretation is subsequently processed by the Relational Layer, which evaluates interactional stability and applies boundary constraints related to epistemic limits, identity clarity, and functional scope.

The outputs of these layers are combined within a Decision Layer, which functions as the system's arbitration mechanism. Rather than generating responses directly, the Decision Layer evaluates semantic interpretations, spatial context signals, relational governance flags, and confidence thresholds to determine the appropriate system action. Possible actions include generating a language response, presenting a spatial visualization, requesting clarification, escalating to a human operator, or declining to act. Language responses are generated by the semantic generation component, while spatial representations are rendered through the visualization interface using structured spatial data from the Spatial Layer.

This architecture separates interpretation, contextual grounding, governance, and action selection, ensuring that system outputs are generated only after semantic meaning, spatial context, and relational constraints have been jointly evaluated.

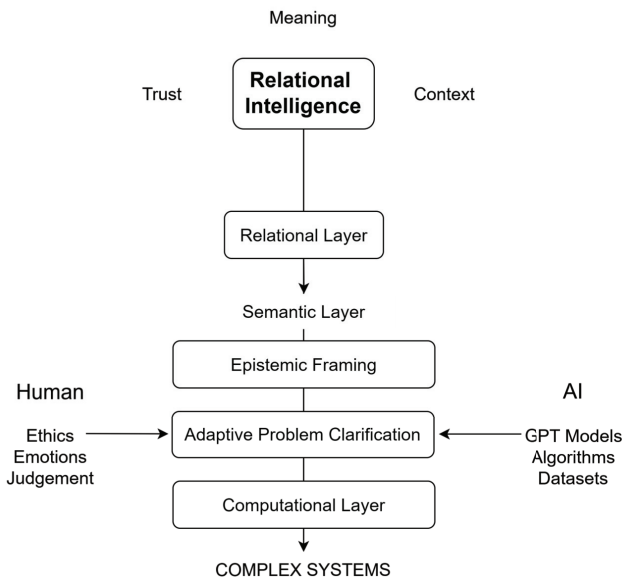


Figure 1: RI architecture stack.

System design logic

At a system level, the SSR architecture operates as a constrained processing pipeline with feedback loops. User and environmental inputs are ingested, parsed, and distributed to semantic, spatial, and relational modules in parallel. Each module produces structured outputs: semantic intent hypotheses, spatial contextual constraints, and relational governance flags. These outputs are fused in the Decision Layer, where confidence thresholds, escalation rules, and human-control policies determine the final system behavior. This design allows the architecture to act not only on what is possible to generate, but on what is contextually grounded, relationally appropriate, and epistemically permissible (Figure 1).

Cognitive effects of the integrated architecture

Consequently, the architecture can accelerate human judgment by restructuring how information is presented and interpreted within the decision environment:

- 1. Interpretation compression:** By integrating semantic reasoning, spatial context, and relational constraints within a unified architecture, the system reduces the time users spend assembling meaning from separate data streams. Rather than interpreting text outputs, environmental signals, and interaction cues independently, the integrated layers align these signals within a single representation. This reduces the cognitive effort required to synthesize fragmented information compared with conventional AI interfaces that distribute these elements across multiple dashboards, text outputs, or analytic tools.
- 2. Cognitive externalization:** Environmental and logical relationships are maintained within the system's spatial model and externalized through visualization interfaces.

This allows users to perceive spatial relationships, such as locations, constraints, trajectories, or resource placement, directly rather than reconstructing them mentally, reducing the burden on working memory while preserving situational awareness.

- 3. Relevance filtering:** Context-sensitive prioritization mechanisms highlight variables most relevant to the current task or operational objective while suppressing low-salience signals. By reducing informational noise, the architecture helps users focus attention on decision-critical factors.
- 4. Stress stabilization:** Interaction pacing and information presentation are structured to support coherent reasoning under time pressure or cognitive load. By presenting aligned semantic, spatial, and relational cues, the system reduces interpretive ambiguity, which can increase stress during complex decision-making tasks.
- 5. Authority preservation:** Human authority remains explicit within the architecture through visible override, escalation, and pause mechanisms. The system can support interpretation and recommendation, but does not remove the human decision-maker from the control loop.

Integrity evaluation dimensions

Evaluation of the SSR Cognitive Stack should consider five primary integrity dimensions: hallucination frequency, context retention stability across multi-turn exchanges, boundary enforcement under relational or adversarial pressure, appropriate escalation-to-human activation, and

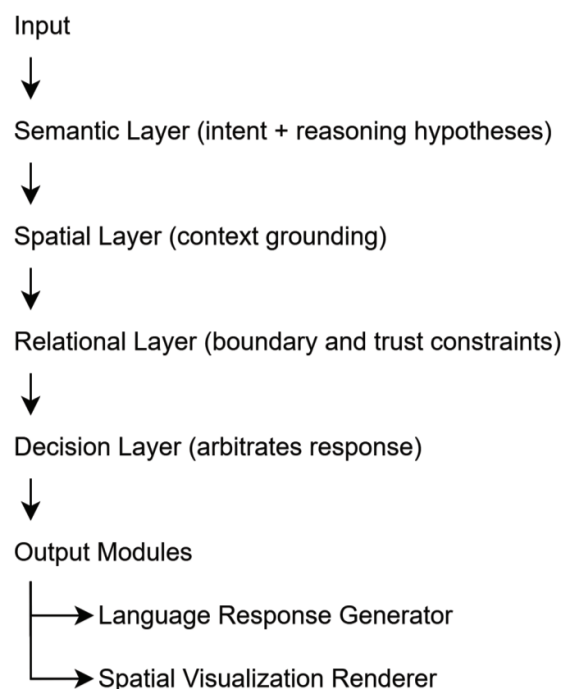


Figure 2: Layered AI architecture.



response variance across semantically equivalent prompts. To operationalize these dimensions, system responses should be analyzed using structured observational coding protocols. Each interaction should be assessed for output consistency, adherence to predefined constraints, evidence of semantic drift, instances of relational overreach (e.g., anthropomorphic or scope-expanding claims), and frequency of correction or clarification. Findings should be aggregated descriptively to document architectural stability and integrity performance across deployment conditions.

Together, these mechanisms shift the role of AI from producing isolated outputs to structuring the decision environment in which human judgment occurs. In this architecture, outputs are not generated by a single model but are the result of a layered arbitration process in which semantic interpretation, spatial grounding, and relational governance jointly constrain system behavior (Figure 2).

Limitations

Semantic AI limitations

While the Semantic Layer provides structured intent interpretation and constraint-aware reasoning, limitations remain in the depth of adaptive semantic processing. Current architectures primarily support first-order intent recognition and bounded response generation, but often lack mechanisms for deeper conceptual evolution across extended interactions. Semantic systems may struggle to progressively elaborate explanations, adjust the level of conceptual complexity to match the audience's expertise, or synthesize prior dialogue into more refined reasoning.

Consequently, contextual continuity may be preserved at the surface level without enabling higher-order meaning construction, such as abstraction, synthesis, or iterative refinement, across multi-turn exchanges. These limitations indicate that while semantic reasoning can reliably interpret intent and generate structured responses, further research is required to support richer semantic adaptability that can sustain evolving dialogue.

Spatial AI limitations

The Spatial Layer is designed to ground system behavior within environmental and contextual constraints; however, most current spatial AI implementations remain primarily perceptual rather than cognitively spatial. Systems can detect entities, estimate positions, and recognize environmental features, but more advanced spatial reasoning, such as predictive modeling, persistent spatial memory, and anticipatory interaction modulation, remains limited.

As a result, spatial architectures often operate reactively, responding to detected environmental inputs without maintaining long-term awareness of spatial state or forecasting environmental dynamics. This restricts the system's ability to adjust interaction pacing, prioritize information based on changing spatial contexts, or anticipate environmental shifts. Future research should extend spatial grounding beyond perception toward predictive and adaptive spatial cognition.

Relational AI limitations

The Relational Layer governs interaction stability through mechanisms such as boundary enforcement, identity clarity, and trust calibration. While these mechanisms help maintain interaction integrity and prevent epistemic overreach, advanced relational modeling remains an open research challenge. Current architectures typically enforce stable boundaries but do not yet support more adaptive relational capabilities.

For example, systems rarely adjust explanation depth based on inferred user expertise, differentiate participant roles, or tailor communication strategies based on engagement signals. Relational architectures also lack mechanisms for sensing cognitive load, adaptive pacing, and cross-session continuity tracking. As a result, relational containment can stabilize interactions but does not yet support fully adaptive relational intelligence capable of dynamically prioritizing information or supporting long-term trust development.

Discussion

Integrity-centered design can function as a foundational architectural principle for embodied AI systems. Rather than optimizing generative fluency alone, the SSR Cognitive Stack integrates semantic interpretation, spatial grounding, and relational containment within a layered structure designed to preserve coherence in complex environments.

Relational governance acts as a structural constraint rather than a surface feature. Embedding boundaries within system architecture can reduce epistemic drift [4] and mitigate risks of anthropomorphic overextension [27]. These insights support arguments that responsible AI must be architecturally embedded rather than retrofitted through external safeguards [28].

Within this framework, the Relational Layer moderates the interaction between semantic reasoning and spatial embodiment. Without relational governance, semantic systems risk unbounded inference; without modulation, embodied systems risk socially mis-calibrated interactions. Treating integrity as an architectural property, therefore, becomes essential for stable human-AI interaction.

Integrity-centered AI design

Integrity-centered AI reframes system optimization around responsibility, stability, and calibrated trust rather than speed alone. In human-automation systems, trust typically emerges from consistent, predictable system behavior rather than from novelty or expressive fluency.

Research on human-automation trust shows that predictability and transparency are key factors in trust calibration [29]. Systems that preserve epistemic boundaries and avoid exaggerated claims are more likely to support calibrated reliance rather than overdependence [30]. From this perspective, relational containment functions as a stabilizing mechanism that supports the formation of appropriate trust between humans and AI systems.

Comparison to semantic-dominant systems

Many AI systems operate primarily as semantic engines that optimize response generation without spatial grounding or relational governance. While effective in text-based environments, such systems can be unstable when deployed in contexts where environmental signals and social interactions influence interpretation.

Spatial grounding constrains abstraction by situating interpretation within environmental context, while relational containment maintains authority boundaries and interaction stability. Integrating these layers differentiates the SSR Cognitive Stack from transactional architectures that optimize output fluency without regulating responsibility or contextual coherence.

Toward adaptive relational modeling

The relational layer described above provides the architectural foundation for context-aware reasoning within the system. The GRAICE framework represents a developmental implementation of this layer, operationalizing relational reasoning through structured interaction loops that clarify user intent and guide the generation of adaptive responses.

Most current AI systems remain transactional, emphasizing response generation over sustained relational coherence. A relationally structured architecture extends beyond single-turn exchanges by incorporating temporal continuity, awareness of interaction patterns, and calibrated responsiveness across engagements.

The proposed GRAICE™ framework represents a developmental extension of the Relational Layer, shifting from static containment toward adaptive relational structuring. Rather than modeling emotion, GRAICE™ focuses on measurable mechanisms such as cross-session continuity tracking, conversational pattern detection, engagement-based response modulation, and calibrated trust development.

These capabilities remain conceptual and require controlled longitudinal evaluation to determine their stability and effectiveness across diverse operational contexts.

Decision superiority through layered integration

The layered architecture improves decision performance by restructuring judgment processes rather than simply accelerating computation. By compressing interpretation time, externalizing cognitive complexity, and filtering relevance through contextual and relational alignment, the system can reduce cognitive overload during complex interactions.

This approach aligns with cognitive load theory, which suggests that structured externalization can improve decision quality under stress [31]. When semantic clarity, spatial grounding, and relational governance operate together, cognitive burden may decrease without diminishing human authority. Such integration may enable more coherent and resilient human-AI collaboration in complex decision environments.

Given that the architecture integrates multiple reasoning layers operating in parallel, computational latency is an important consideration, particularly in time-sensitive environments. In the proposed framework, the Semantic, Spatial, and Relational layers operate as partially parallel modules, with their outputs coordinated by the Decision Layer via lightweight arbitration rather than through sequential deep processing. This design reduces latency by enabling concurrent evaluation of contextual signals while prioritizing high-confidence cues when rapid responses are required. In scenarios involving stress stabilization or time-sensitive decision support, the system may also employ adaptive prioritization mechanisms that temporarily reduce reasoning depth or defer noncritical evaluations to generate timely responses. These mechanisms allow the architecture to balance contextual reasoning with real-time responsiveness while maintaining the stabilizing interaction behaviors described in the framework.

The role of the Decision Layer can be further clarified through a simple illustrative scenario. To illustrate how the Decision Layer integrates signals across the architecture, consider a hypothetical medical triage scenario in which a clinician consults the system regarding patient prioritization during a high-volume emergency intake. The Semantic Layer interprets the clinician's request and identifies the decision task, such as determining urgency or triage category. The Spatial Layer evaluates contextual signals, including patient location, movement constraints, and environmental factors such as congestion within the treatment area. The Relational Layer evaluates interaction context, including whether the system has sufficient confidence and authority to provide guidance or whether escalation to a human decision-maker is appropriate. The Decision Layer then arbitrates among these inputs, weighing semantic interpretation, environmental constraints, and relational considerations before producing an output such as a ranked recommendation, a request for additional information, or a recommendation for human escalation. This example illustrates that system outputs emerge from coordinated evaluation across multiple reasoning layers rather than from a single inference process.

Future research

Future research must empirically evaluate the SSR Cognitive Stack across diverse operational environments to determine whether relationally structured architectures improve system reliability, user trust, and decision quality. Experimental studies should compare traditional semantic-dominant AI systems with architectures that integrate semantic, spatial, and relational layers.

Controlled deployments in educational, commercial, and public-sector environments would enable researchers to measure hallucination frequency, boundary-enforcement stability, user-trust calibration, and interaction coherence over extended engagements. Longitudinal research is particularly necessary to assess whether relational governance mechanisms support sustained human-AI collaboration without encouraging anthropomorphic misinterpretation or overreliance.



Additional work should investigate adaptive relational modeling frameworks such as GRAICE™, including mechanisms for cross-session continuity tracking, conversational pattern detection, and calibrated response modulation. These capabilities must be tested carefully to ensure that increasing adaptivity does not undermine epistemic restraint or identity clarity.

Ultimately, the goal of future research is to determine whether relationally integrated architectures can support trustworthy AI systems that enhance human judgment while preserving accountability and authority in complex decision environments.

Conclusion

Finally, the goal is not to construct a flawless AI system, nor to surrender human agency to automated decision structures. The goal is to design AI that strengthens human judgment, preserves accountability, and expands our capacity to act wisely in complex environments. AI need not replace human cognition; it can reinforce it.

As Domingos [32] notes, AI will not replace humans; it will amplify them. Similarly, Andrew Ng's [33] framing of AI as "the new electricity" underscores that transformative technologies expand human capability rather than eliminate it. Nadella [34] emphasizes that the purpose of AI should be to amplify human ingenuity, not replace it totally, reinforcing a vision of enhancement rather than substitution. Across leaders in the field, from Bengio's [35] emphasis on unlocking human potential to Johnson and Cochran's [14] call to treat AI as a co-collaborator and co-creator. The consistent theme is empowerment.

The architectural direction outlined in this work aligns with the philosophy listed in the previous paragraph. A relationally structured AI system should not be designed to automate away responsibility. Instead, the AI system should seek to stabilize trust, maintain epistemic boundaries, and support human coherence under stress. As Rometty [36] observed, AI is not a human-versus-machine narrative but an enhancement of human capacity. The system described here aims to function as a cognitive partner, one that accelerates interpretation, reduces cognitive overload, and preserves human authority over final action [37].

In this sense, AI is neither magic nor autonomy; it is applied mathematics embedded in social systems. Its value lies not in replacing human reasoning, but in amplifying it. The future of decision advantage will belong to systems that combine semantic understanding, spatial grounding, and relational integrity while ensuring that judgment remains fundamentally human.

References

- Abishethvarman V, Sabrina F, Kwan P. Knowledge integrity in large language models: A state-of-the-art review. *Information*. 2025;16(12):1076. Available from: <https://doi.org/10.3390/info16121076>
- Kushwah S, Dave N. AI hallucination and strategies to overcome: Enhancing human-AI interaction. In: 2025 International Conference on Artificial Intelligence and Machine Vision (AIMV). IEEE; 2025; 1-6. Available from: <https://doi.org/10.1109/AIMV66517.2025.11203756>
- Yang Z, Meng Z, Zheng X, Wattenhofer R. Assessing adversarial robustness of large language models: An empirical study. *arXiv*. 2024. Available from: <https://doi.org/10.48550/arXiv.2405.02764>
- Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: Can language models be too big? In: *Proceedings of the 2021 ACM Conference*. 2021. Available from: <https://dl.acm.org/doi/10.1145/3442188.3445922>
- Ji Z, Lee N, Frieske R, Yu T, Su D, Xu Y, et al. Survey of hallucination in natural language generation. *ACM Computing Surveys*. 2023;55(12):1-38. Available from: <https://dl.acm.org/doi/10.1145/3571730>
- Floridi L, Cows J, King TC, Taddeo M. How to design AI for social good: Seven essential factors. In: *Ethics, governance, and policies in artificial intelligence*. Cham: Springer International Publishing; 2021; 125-151. Available from: <https://doi.org/10.1007/s11948-020-00213-5>
- Amershi S, Weld D, Vorvoreanu M, Fournay A, Nushi B, Collisson P, et al. Guidelines for human-AI interaction. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019; 1-13. Available from: <https://doi.org/10.1145/3290605.3300233>
- Zhang XY, Ye HJ, Zhan DC. Trustworthy evaluation of large language models. *Frontiers of Computer Science*. 2026;20:Article 2002324. Available from: https://doi.org/10.1007/s11704-025-50442-9?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle
- Brooks RA. Intelligence without representation. *Artificial Intelligence*. 1991;47(1-3):139-159. Available from: <https://people.csail.mit.edu/brooks/papers/representation.pdf>
- Pfeifer R, Bongard J. How the body shapes the way we think: A new view of intelligence. MIT Press; 2006. Available from: <https://mitpress.mit.edu/9780262537421/how-the-body-shapes-the-way-we-think/>
- Duan J, Yu S, Tan HL, Zhu H, Tan C. A survey of embodied AI: From simulators to research tasks. *IEEE Transactions on Emerging Topics in Computational Intelligence (TETCI)*. 2022. Available from: <https://doi.org/10.48550/arXiv.2103.04918>
- Duchon D, Dunegan KJ, Barton SL. Framing the problem and making decisions: The facts are not enough. *IEEE Transactions on Engineering Management*. 1989;36(1):25-27. Available from: <https://doi.org/10.1109/17.19979>
- Tian H. A literature review on the influence of information overload on individual decision-making. University of Illinois Urbana-Champaign; 2023. Available from: https://sundaram.cs.illinois.edu/Example_reports/Review_overload.pdf
- Johnson L, Cochran J. The feel of thinking: Toward relational intelligence in generative AI and complex systems. *International Journal of Humanities and Social Science*. 2026. Available from: <https://ijhss.thebrpi.org/journal/index/5079>
- Barad K. Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning. Duke University Press; 2007. Available from: <https://www.dukeupress.edu/meeting-the-universe-halfway>
- Maturana HR, Varela FJ. The tree of knowledge: The biological roots of human understanding. Shambhala; 1987. Available from: <https://uranos.ch/research/references/Maturana1988/maturana-h-1987-tree-of-knowledge-bkmrk.pdf>
- Whitehead AN. *Process and reality: An essay in cosmology*. Corrected ed. Griffin DR, Sherburne DW, editors. Free Press; 1978.
- Saussure FD. *Course in general linguistics*. The Library of The University Of California Los Angeles; 1916.



19. Wittgenstein L. Philosophical investigations. Oxford: Blackwell; 1953. Available from: <https://static1.squarespace.com/static/54889e73e4b0a2c1f9891289/t/564b61a4e4b04eca59c4d232/1447780772744/Ludwig.Wittgenstein.-.Philosophical.Investigations.pdf>
20. Chomsky N. Aspects of the theory of syntax. 50th ed. The MIT Press; 1965. Available from: <https://www.jstor.org/stable/j.ctt17kk81z>
21. Russell S, Norvig P. Artificial intelligence: A modern approach. 4th ed. Global ed. Pearson Education; 2021. Available from: https://api.pageplace.de/preview/DT0400.9781292401171_A41586057/preview-9781292401171_A41586057.pdf
22. Atf Z, Lewis PR. Is trust correlated with explainability in AI? A meta-analysis. IEEE Transactions on Technology and Society. 2025. Available from: <https://doi.org/10.1109/TTS.2025.3558448>
23. Kaplan AD, Kessler TT, Brill JC, Hancock PA, Endsley MR. Trust in artificial intelligence: Meta-analytic findings. Human Factors. 2023;65(2):337–359. Available from: <https://doi.org/10.1177/00187208211013988>
24. Smith PJ, Hoffman RR, editors. Cognitive systems engineering: The future for a changing world. CRC Press; 2017. Available from: <https://doi.org/10.1201/9781315572529>
25. Xu W, Dainoff MJ, Ge L, Gao Z. From human-computer interaction to human-AI interaction: New challenges and opportunities for enabling human-centered AI. arXiv preprint. 2021. Available from: https://www.researchgate.net/publication/351545957_From_Human-Computer_Interaction_to_Human-AI_Interaction_New_Challenges_and_Opportunities_for_Enabling_Human-Centered_AI
26. Yang Q, Steinfeld A, Rosé C, Zimmerman J. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020; 1-13. Available from: <https://dl.acm.org/doi/10.1145/3313831.3376301>
27. Crawford K. Atlas of AI. Yale University Press; 2021. Available from: <https://yalebooks.yale.edu/book/9780300264630/atlas-of-ai/>
28. Floridi L, Cowls J. A unified framework of five principles for AI in society. Harvard Data Science Review. 2019. Available from: <https://doi.org/10.1162/99608f92.8cd550d1>
29. Lee JD, See KA. Trust in automation: Designing for appropriate reliance. Human Factors. 2004;46(1):50-80. Available from: https://doi.org/10.1518/hfes.46.1.50_30392
30. Parasuraman R, Riley V. Humans and automation: Use, misuse, disuse, abuse. Human Factors. 1997;39(2):230-253. Available from: <https://web.mit.edu/16.459/www/parasuraman.pdf>
31. Sweller J. Cognitive load theory. Psychology of Learning and Motivation. 2011;55:37-76. Available from: <https://psycnet.apa.org/record/2011-17503-002>
32. Domingos P. The master algorithm: How the quest for the ultimate learning machine will remake our world. Basic Books; 2015. Available from: [https://ebooks.karbus.me/Technology/The%20Master%20Algorithm_%20How%20the%20Quest%20for%20the%20Ultimate%20Learning%20Machine%20Will%20Remake%20Our%20World%20-%20Pedro%20Domingos%20-%20Basic%20Books%20\(2015\).pdf](https://ebooks.karbus.me/Technology/The%20Master%20Algorithm_%20How%20the%20Quest%20for%20the%20Ultimate%20Learning%20Machine%20Will%20Remake%20Our%20World%20-%20Pedro%20Domingos%20-%20Basic%20Books%20(2015).pdf)
33. Ng A. What artificial intelligence can and can't do right now [conference session]. Stanford MSx Future Forum, Stanford University; 2017. Available from: <https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now>
34. Nadella S, Shaw G, Nichols JT. Hit refresh: The quest to rediscover Microsoft's soul and imagine a better future for everyone. Harper Business; 2017. Available from: https://books.google.co.in/books/about/Hit_Refresh.html?id=SeGMDAAQBAJ&redir_esc=y
35. Bengio Y, Mindermann S, Privitera D, Besiroglu T, Bommasani R, Casper S, et al. International AI Safety Report. arXiv preprint. 2025. Available from: <https://doi.org/10.48550/arXiv.2501.17805>
36. Rometty G. Good power: Leading positive change in our lives, work, and world. Harvard Business Press; 2023. Available from: <https://hbsp.harvard.edu/product/10550-PDF-ENG>
37. Conference on Fairness, Accountability, and Transparency (FAccT '21). Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. ACM; 2021. 610–623.

Discover a bigger Impact and Visibility of your article publication with Peertechz Publications

Highlights

- ❖ Signatory publisher of ORCID
- ❖ Signatory Publisher of DORA (San Francisco Declaration on Research Assessment)
- ❖ Articles archived in worlds' renowned service providers such as Portico, CNKI, AGRIS, TDNet, Base (Bielefeld University Library), CrossRef, Scilit, J-Gate etc.
- ❖ Journals indexed in ICMJE, SHERPA/ROMEO, Google Scholar etc.
- ❖ OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)
- ❖ Dedicated Editorial Board for every journal
- ❖ Accurate and rapid peer-review process
- ❖ Increased citations of published articles through promotions
- ❖ Reduced timeline for article publication

Submit your articles and experience a new surge in publication services

<https://www.peertechzpublications.org/submission>

Peertechz journals wishes everlasting success in your every endeavours.