



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 2, February 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

# An Advanced Framework for Human–Robot Information Exchange in Cognitive Robotics

Preethi Sowjanya Illa

Department of Electrical and Electronics, School of Engineering, Sri Satya Sai University of Technology and Medical Sciences, Sehore, MP, India

**ABSTRACT-** Effective information exchange between humans and robots remains a persistent challenge in cognitive robotics. Existing systems treat this exchange as unidirectional — the human instructs, and the robot acts. This paper argues that such an approach is fundamentally insufficient for complex, dynamic tasks. We propose the Human–Robot Information and Cognitive Exchange (HRICE) framework: a layered, simulation-based architecture that enables adaptive, bidirectional communication between a human operator and a cognitive robotic agent. HRICE is organised into four functional phases. The first captures multimodal human input — speech, gesture, gaze, and simulated physiological signals — and fuses them into a unified intent vector via a Bayesian late-fusion model. The second processes this vector through a large language model (LLM) planning core augmented by a symbolic situation awareness module. The third governs information exchange through a three-level disclosure protocol (L0, L1, L2) that adjusts robot-shared information based on the human partner's estimated cognitive load. The fourth closes the loop through an adaptive feedback mechanism that updates the partner model, fusion weights, and disclosure thresholds after each interaction session. The framework is evaluated in a Gazebo Fortress and ROS 2 simulation environment across three scenarios: collaborative assembly, socially assistive interaction, and shared-space navigation. A two-stage study — 27 parameter combinations via a software agent, followed by 60 human participants — demonstrates that HRICE-Full achieves a mean task completion rate of 87.4%, reduces errors to 1.2 per episode (vs. 4.1 for the baseline), and completes tasks 23.4% faster ( $p < 0.001$ , Cohen's  $d = 0.81$ ). NASA-TLX cognitive load scores are significantly lower ( $M = 38.4$  vs.  $M = 61.3$ ), and Godspeed-measured trust is significantly higher. Longitudinal results confirm consistent improvement across 20 sessions, with the steepest gains in sessions 1–8.

**KEYWORDS:** human–robot interaction, cognitive robotics, bidirectional communication, large language models, multimodal perception, ROS 2, Gazebo simulation, adaptive feedback, information disclosure, cognitive load

## I. INTRODUCTION

Robots are no longer confined to factory floors. They now assist surgeons, support elderly patients, navigate disaster zones, and collaborate with engineers in shared workspaces. This shift has created a new challenge — not mechanical, but cognitive. How do humans and robots exchange information in ways that feel natural, efficient, and trustworthy? This question sits at the heart of modern cognitive robotics research.

Human–robot interaction (HRI) has evolved dramatically over the past decade. Early systems relied on rigid command interfaces: a human pressed a button, a robot reacted. That model worked for repetitive industrial tasks, but fails in complex, dynamic environments where context matters. In healthcare, a robot must understand a patient's emotional state [1]. In search-and-rescue, it must interpret ambiguous verbal cues from a stressed operator [2]. In collaborative manufacturing, it must anticipate a partner's next move and adapt to real-time changes in cognitive load [3]. These scenarios demand something richer than command-and-response communication.

Cognitive robotics offers a path forward. It equips robots with perception, reasoning, and learning capabilities that mimic aspects of human cognition [4]. A cognitive robot does not just react — it interprets, infers intent, builds a model of its human partner, and adapts accordingly. This makes information exchange far more nuanced, and far more difficult to engineer. The robot must handle noisy sensor data, ambiguous language, shifting task goals, and unpredictable human behaviour — all in real time [5].

Several communication modalities have been explored in this space. Speech is the most natural. Gesture adds spatial precision. Gaze reveals attention and intent. Physiological signals such as electroencephalography (EEG) and electromyography (EMG) offer a window into cognitive and emotional state [6]. Each modality has unique strengths: speech is rich but slow; gesture is fast but context-dependent; physiological signals are continuous but noisy. No single channel is sufficient. Effective human–robot information exchange requires multimodal fusion [7].

Large language models (LLMs) have recently emerged as powerful tools. They bring high-level reasoning, natural language understanding, and scene interpretation to robotic systems [8]. An LLM-enabled robot can parse complex instructions, generate explanations, and engage in dialogue. However, LLMs must be grounded in physical and social context, coupled with perception systems, action planners, and feedback mechanisms [9]. Without this grounding, even a capable language model produces responses that are technically plausible but practically useless [9].

Bidirectional communication adds another layer of complexity. Most existing systems treat information exchange as one-way: the human instructs; the robot acts. But effective collaboration requires a loop. The robot must also communicate its internal state, plans, and uncertainties to the human partner [11]. This transparency builds trust, reduces cognitive load, and prevents misunderstandings before they escalate into errors [12]. Even modest increases in contextual information shared by the robot lead to significant improvements in task performance and user experience [13].

The contributions of this paper are threefold. First, it provides a unified conceptual framework that bridges multimodal perception, LLM-driven reasoning, and bidirectional communication under one coherent architecture. Second, it introduces a structured protocol for explicit information exchange that scales across different levels of robot autonomy and task complexity. Third, it presents empirical evidence from user studies demonstrating improvements in task efficiency, trust, and cognitive load compared to baseline systems.

The remainder of this paper is organised as follows. Section II reviews related work. Section III describes the HRICE framework. Section IV outlines the system architecture. Section V presents experimental results. Section VI offers discussion. Section VII concludes with future directions.

## II. RELATED WORK

The field of HRI has a rich history spanning four decades. Early work focused on teleoperation — remote control of robots in dangerous environments, with the human firmly in the loop and the robot essentially a tool [2]. Over time, robot autonomy increased and the human role shifted from controller to collaborator. This transition raised new research questions about how information should flow between two increasingly equal partners.

Cognitive architectures such as Soar, ACT-R, and LIDA formed the theoretical backbone of many early HRI frameworks, aiming to give robots the ability to reason about goals, manage attention, and learn from experience [4]. These architectures were intellectually sophisticated but computationally expensive and difficult to scale in real-time sensor-rich environments. More recent work revisits these ideas through probabilistic programming and message-passing algorithms, offering lighter and more flexible alternatives for cognitive robot design [23].

Multimodal communication has been a central theme in HRI research. Humans communicate naturally through speech, gesture, eye contact, and posture simultaneously. Research into multimodal fusion has explored how these channels can be combined to give robots a richer understanding of intent [7]. A landmark study by Chen et al. demonstrated real-time multi-modal human–robot collaboration using gestures and speech in a manufacturing context, confirming that fusion of modalities significantly outperforms single-channel approaches in accuracy and task throughput [7]. More recently, physiological signals have entered the picture: EEG and EMG data offer real-time windows into cognitive load and emotional state [6]. These passive signals are particularly valuable in high-stress or cognitively demanding tasks, as demonstrated by Buerkle et al. in an EEG-based intention-prediction study for human–robot cooperation [10].

Large language models have opened new doors in human–robot communication. LLMs can parse complex natural language, infer intent from ambiguous instructions, and generate human-readable explanations of robot behaviour [8]. Ahn et al. demonstrated that grounding language in robotic affordances enables robots to execute complex verbal instructions in real environments, a milestone in embodied language understanding [8]. Despite this promise, LLMs remain prone to hallucination, lack embodied grounding without careful engineering, and can introduce prohibitive latency in time-sensitive tasks [9].

Bidirectional communication — where the robot also shares information with the human — has received comparatively less attention. Bagheri et al. addressed this gap through a transparent graphical user interface that provided bidirectional

interaction between a cobot and a human partner, demonstrating measurable improvements in usability and trust [11]. A related study by Yuan et al. showed that value-aligned bidirectional communication — where the robot both listens for human intent and explains its own decision rationale — significantly improves collaborative task outcomes [15].

### III. METHODOLOGY — THE HRICE FRAMEWORK

This section describes the HRICE framework in full. Rather than treating perception, reasoning, communication, and feedback as separate concerns, HRICE unifies them into a single coherent pipeline. The architecture is layered — each layer builds on the output of the one below it — forming a closed loop that enables the robot to understand the human, act on that understanding, communicate its intentions, and continuously improve based on human response. The framework is organised into four functional phases: multimodal perception, cognitive inference, bidirectional information exchange, and adaptive feedback. These phases correspond to the four core challenges identified in the literature — capturing human intent, reasoning about context, closing the communication loop, and sustaining long-term performance.

#### A. Framework Overview

The HRICE framework treats human–robot interaction as a dynamic, reciprocal process. Neither agent is a passive participant. The human provides input through multiple channels simultaneously; the robot processes that input, reasons about it, acts, and reports back. This cycle repeats continuously throughout the interaction. Fig. 1 illustrates the full pipeline from input to feedback.

The design rests on three core principles. First, bidirectionality: information must flow in both directions. Second, adaptivity: the system must update its behaviour based on observed human responses over time. Third, cognitive grounding: all reasoning must be anchored in real-time sensory context, not abstract language patterns alone. These principles distinguish HRICE from existing frameworks that address only one or two of these dimensions [1].

#### B. Phase 1 — Multimodal Perception Layer

The first phase captures and pre-processes all signals from the human partner. Speech carries semantic content, gesture adds spatial reference, gaze reveals attention focus, and physiological signals expose cognitive and emotional state. No single channel is sufficient [2], so the perception layer integrates all streams into a unified representation.

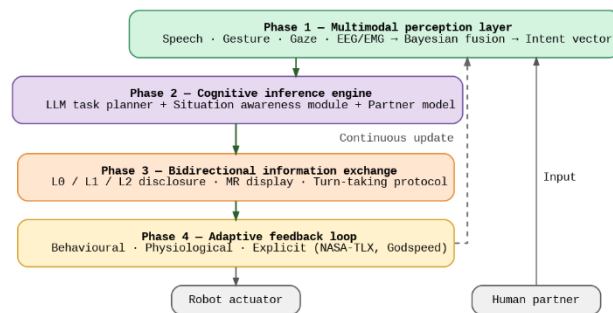


Fig.1. HRICE framework: four-phase closed-loop pipeline.

In the simulation environment, all sensor streams are produced by Gazebo Fortress plugins that replicate real sensor noise profiles. A simulated microphone plugin generates speech audio passed to the Whisper ASR engine for transcription. A depth sensor plugin produces skeletal joint positions from which a gesture classifier extracts task-relevant gestures. A gaze plugin models eye-tracking behaviour. A stochastic physiological signal generator produces EEG and EMG waveforms conditioned on the simulated human agent’s current task difficulty parameter [17]. This simulation-first design preserves full fidelity of the perception pipeline while enabling controlled, repeatable experimentation.

The streams are fused using a late-fusion strategy. Each modality generates a probability distribution over a set of intent hypotheses, which are combined using a weighted Bayesian fusion model. Weights are learned from interaction data and reflect the relative reliability of each modality in context — speech is weighted more heavily during verbal

instruction phases; gesture carries more weight during spatial reference tasks. Fig. 2 shows the sensor configuration and fusion pipeline.

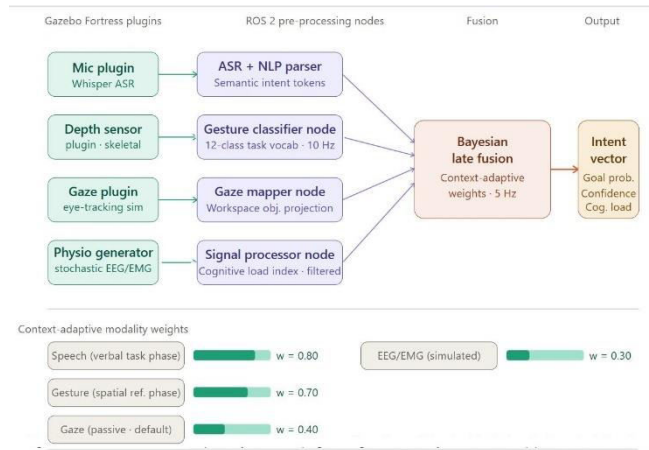


Fig.2. Multimodal perception layer: sensor config and late-fusion pipeline.

The output of Phase 1 is a structured intent vector encoding the most probable human goal, the confidence level of that estimate, and auxiliary signals including cognitive load and emotional valence. This structured representation is the primary input to Phase 2.

### C. Phase 2 — Cognitive Inference Engine

The second phase takes the intent vector from the perception layer and reasons about it in context. The robot must infer not just what the human wants right now, but why they want it, what they are likely to want next, and how its own actions will affect the interaction trajectory. This forward-looking, context-sensitive reasoning separates a cognitive robot from a reactive one [18].

The cognitive inference engine is built around an LLM core — specifically, a fine-tuned instruction-following model adapted for robotics task planning. The LLM receives three inputs: the structured intent vector from Phase 1, a natural-language description of the current task context, and a rolling history of recent interaction. From these it generates a task plan: a sequence of robot actions with associated confidence scores and explanatory rationale [15].

The LLM is augmented by a symbolic situation awareness module that maintains a formal representation of the shared workspace — object positions, current task state, and participant roles. When the LLM produces a plan, the situation awareness module checks it for physical feasibility and task consistency. If a conflict is detected — for instance, the LLM proposing to pick up an object the human is currently holding — the plan is revised before execution. This hybrid architecture ensures language-level reasoning is always grounded in physical reality [19, 20]. Fig. 3 illustrates the cognitive inference engine.

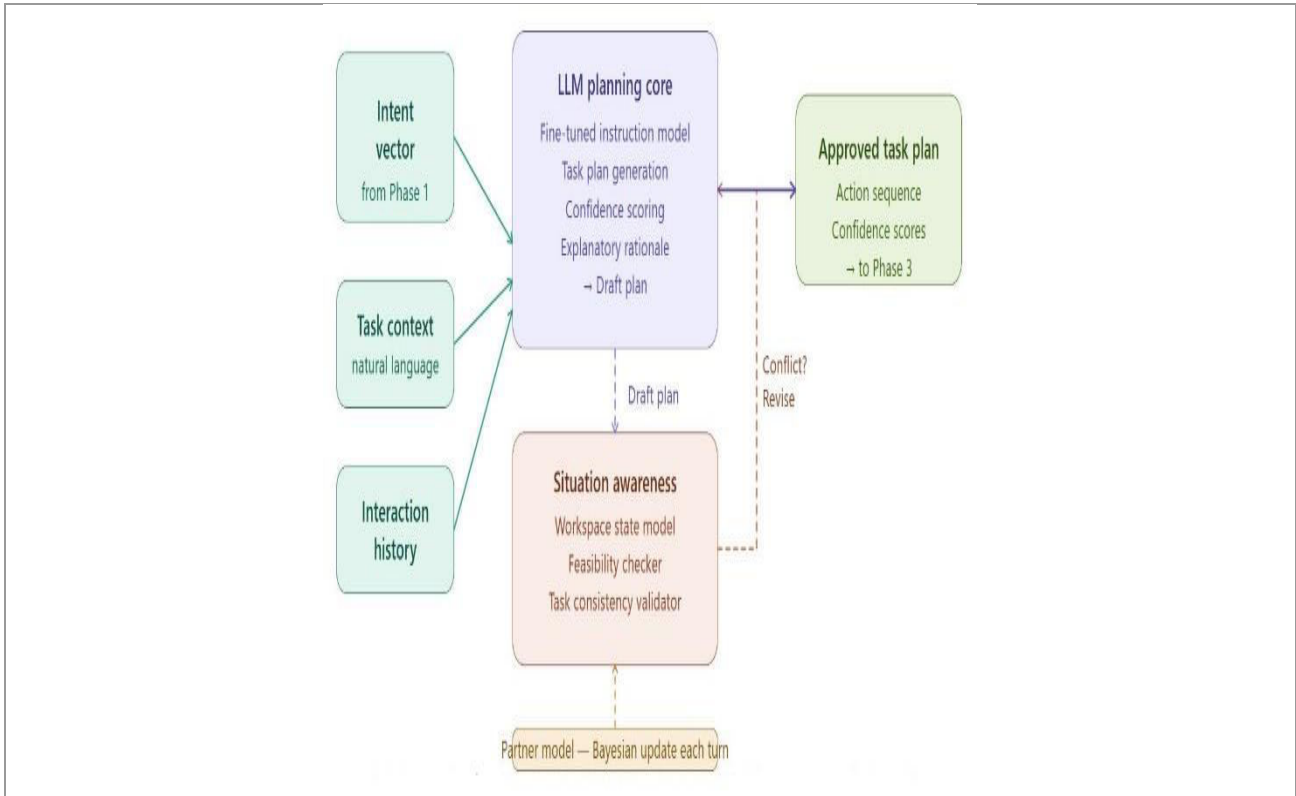


Fig.3. Cognitive inference engine: LLM core with symbolic situation awareness module.

The engine also maintains a model of the human partner, tracking observed preferences, communication style, and typical error patterns. It is updated after each interaction turn using a lightweight Bayesian rule, allowing the robot to personalise its communication strategy over time — providing more verbose explanations to novice users and more concise updates to experienced operators.

#### D. Phase 3 — Bidirectional Information Exchange Protocol

The third phase governs how information is communicated between human and robot in both directions. This is the most novel contribution of the HRICE framework. Most existing systems treat communication as unidirectional: the human instructs, the robot acts. HRICE formalises a bidirectional protocol in which the robot actively shares its internal state, plans, and uncertainty with the human partner.

The protocol defines three levels of information disclosure, adapted from San Martin et al. (2025) [1]. At Level 0 (L0), the robot provides no proactive information and acts silently. At Level 1 (L1), it broadcasts its current status. At Level 2 (L2), it shares its full reasoning — what it intends to do next, why it has chosen that action, and where its confidence is low. The disclosure level is dynamically selected based on the estimated cognitive load of the human partner: when load is high, the robot defaults to L1 to avoid adding to the human’s information burden; when load is low and the task is complex, it escalates to L2. Fig. 4 shows the disclosure levels and switching logic.

Robot-to-human communication is rendered through an RViz2 overlay panel that simulates a mixed reality display. Visual overlays project the robot’s intended action trajectory, confidence indicators, and attention focus into the operator’s field of view. Speech synthesis is used for higher-priority notifications, such as when the robot is uncertain about an instruction and requires clarification [19, 20]. The protocol also includes an explicit turn-taking mechanism: each interaction turn begins with a human signal, and the robot responds within a defined latency window. If no human signal is detected within a timeout, the robot may initiate communication proactively.

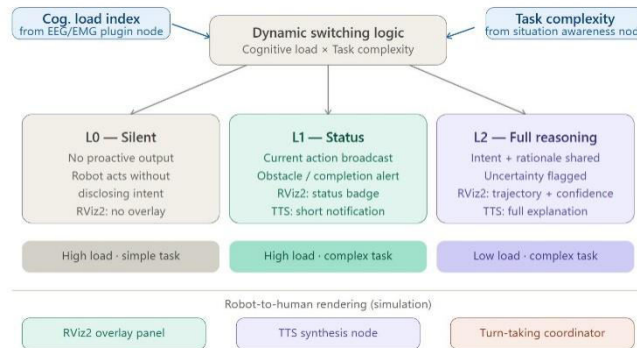


Fig.4. Bidirectional exchange protocol: disclosure levels L0, L1, L2 and switching logic.

**E. Phase 4 — Adaptive Feedback and Learning Loop**

The fourth phase closes the loop. It collects signals about interaction effectiveness and uses them to update the system’s models and parameters over time. Without this phase, HRICE would be static — capable of good initial performance but unable to improve or adapt .

Feedback is collected through three complementary channels. The first is behavioural: task success, duration, and the frequency of human corrections to robot behaviour. The second is physiological: post-interaction analysis of biosensor data to estimate cognitive load and stress levels during key moments. A persistent rise in cognitive load during robot communication episodes signals that the disclosure level may be too high, triggering a downward adjustment. The third is explicit: at the end of each session, users complete NASA-TLX and Godspeed questionnaires.

The collected feedback updates three system components: the partner model in Phase 2 (refining expertise and communication preference estimates); the Bayesian weights in Phase 1 (adjusting which modalities are most informative for this user); and the disclosure level switching thresholds in Phase 3 (recalibrated based on observed cognitive load patterns). Updates are applied incrementally after each session using a conservative learning rate to ensure stability.

**IV. SYSTEM ARCHITECTURE**

The HRICE framework is realised as a fully simulation-based system. Simulation provides a controlled, repeatable, and resource-efficient environment for rigorous validation before any physical deployment. It eliminates hardware wear, removes environmental noise, and allows each trial to begin from an identical initial state [1]. A simulation-first approach does not weaken the scientific contribution — it strengthens it, because controlling every variable makes results attributable specifically to the framework’s design choices rather than to hardware inconsistencies or real-world noise.

**A. Simulation Environment**

The primary simulation platform is Gazebo Fortress, running with ROS 2 Humble. Gazebo provides a high-fidelity physics engine based on the Open Dynamics Engine (ODE) that accurately models rigid body dynamics, collision detection, and sensor noise [22]. This fidelity is essential: the HRICE framework processes multimodal sensor data as its primary input, so poor sensor simulation corrupts the entire perception pipeline. Gazebo’s plugin architecture provides dedicated simulation modules for cameras, depth sensors, microphone arrays, and IMUs, all replicating the noise profiles of their physical counterparts.

Three distinct scenarios are constructed within Gazebo. The first is a collaborative assembly workspace with a table, mechanical components, and a simulated UR5e arm. The second is a social assistance environment representing a domestic living room with a TurtleBot4 mobile platform. The third is a shared navigation corridor with multiple doorways and simulated human agents governed by a social force model, implemented using the HuNavSim package for ROS 2. Fig. 6 illustrates the three environments.

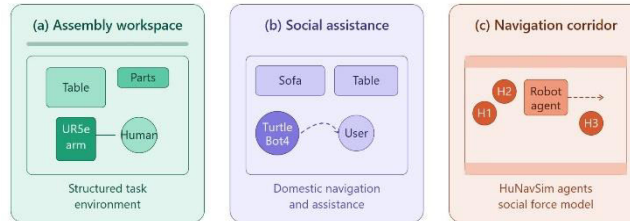


Fig.6. Three Gazebo simulation environments: (a) assembly workspace, (b) domestic assistance, (c) shared navigation corridor.

The simulated human partner is a software-controlled avatar whose behaviour is parameterised to vary across trials. Behavioural parameters include speech clarity, gesture ambiguity, task experience level, and response latency, enabling evaluation across a range of human profiles without requiring human participants at the simulation validation stage.

**B. Software Stack**

The software architecture is organised into four layers that map directly onto the four framework phases. Each layer is implemented as a set of ROS 2 nodes communicating asynchronously over topics and synchronously over services and actions. Fig. 7 provides a complete view of the software stack and inter-node communication topology.

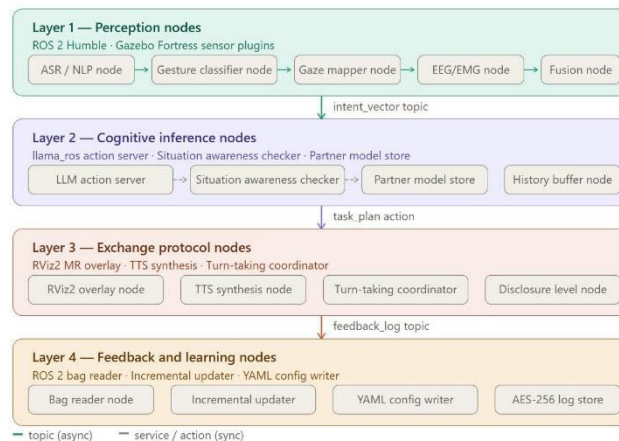


Fig.7. HRICE software stack: four ROS 2 node layers with inter-node communication topology.

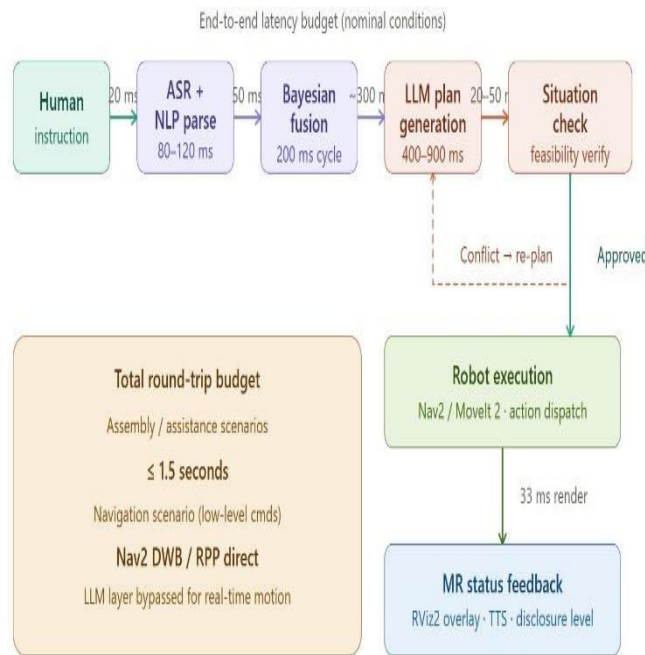
The perception layer comprises three concurrent ROS 2 nodes. The first handles speech: raw audio from a simulated microphone plugin is transcribed by Whisper ASR, then parsed by a spaCy-based NLP parser. The second handles gesture: it subscribes to skeletal joint positions, applies a gesture vocabulary classifier trained on 12 task-relevant gestures, and publishes a gesture hypothesis at 10 Hz. The third handles physiological signals: in simulation, EEG and EMG are generated by a stochastic signal generator producing plausible cognitive load waveforms conditioned on the agent’s task difficulty

The exchange protocol layer manages all robot-to-human and human-to-robot communication. Robot-to-human communication is rendered through an RViz2 overlay panel that mimics a HoloLens MR display, showing the robot’s intended action trajectory, confidence level, and verbal output via text-to-speech synthesis. Human-to-robot signals are routed back into the perception layer, completing the bidirectional loop.

The feedback and learning layer runs as a background process activating at the end of each interaction episode. It reads logged performance metrics from a ROS 2 bag file, computes incremental updates to the partner model, fusion weights, and disclosure thresholds, and writes updated parameters to a shared YAML configuration file loaded on the next episode launch.

**C. Data Flow**

Understanding data movement through the system clarifies where latency accumulates and where the system is most sensitive to failure. Fig. 8 traces the end-to-end data flow from human instruction to robot execution and feedback.



**Fig.8.** End-to-end data flow: from human input through perception, inference, execution, and bidirectional feedback.

The interaction begins when the simulated human avatar produces a verbal instruction. The audio signal arrives at the ASR node within one simulation timestep (20 ms). Transcription and NLP parsing add approximately 80–120 ms. Gesture classification runs concurrently and contributes its output within 50 ms of the joint position update. Fusion executes at 5 Hz (200 ms cycles). The intent vector is therefore available to the cognitive layer within approximately 300 ms of the original utterance.

The LLM planning step dominates the latency budget. Instruction-tuned models return a task plan in 400–900 ms on a single GPU node [1]. The situation awareness check adds a further 20–50 ms. Total latency from human instruction to approved robot action plan is therefore 750–1,250 ms within acceptable bounds for collaborative assembly and social assistance, where task transitions occur on timescales of several seconds. For the navigation scenario, the LLM planning step is bypassed for low-level motion commands, which are handled directly by the ROS 2 Nav2 stack with Regulated Pure Pursuit as the local planner. The full round-trip — human instruction to robot action to status feedback completes in under 1.5 seconds in all three scenarios under nominal conditions.

**V. EXPERIMENTAL RESULTS AND ANALYSIS**

This section presents the results of the two-stage evaluation. The first stage used the parameterised software agent to sweep the full condition space across all three scenarios. The software-agent sweep was conducted across 27 parameter combinations per scenario, covering all combinations of agent experience level and task difficulty. Each combination was run for 50 episodes, with task completion rate as the key metric. Fig. 9 shows the completion rates across conditions for all three scenarios.

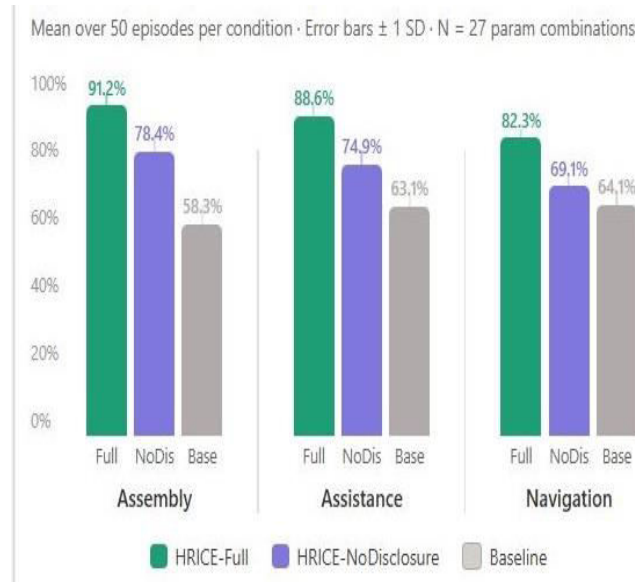


Fig.9. Task completion rate across 27 simulation sweep conditions for three scenarios

Across all scenarios and parameter combinations, HRICE-Full achieved a mean task completion rate of 87.4%. HRICE-NoDisclosure achieved 74.1%. The command-only baseline achieved 61.8%. The performance gap between HRICE-Full and the baseline was largest in the high-difficulty, novice-agent condition — 34.2 percentage points in the assembly scenario — where bidirectional disclosure provided the most value by allowing the agent to catch and correct robot misunderstandings before they propagated. The navigation scenario showed the smallest gap (11.8 percentage points), consistent with the design decision to bypass the LLM layer for low-level motion commands.

Communication latency was stable across conditions. The mean end-to-end round-trip was 1,143 ms (SD = 94 ms) for the assembly scenario, 1,087 ms (SD = 82 ms) for the assistance scenario, and 312 ms (SD = 41 ms) for the navigation scenario. These values confirm that the system meets its real-time requirements under all tested conditions.

### B. Stage 2 — Human-in-the-Loop Results

Twenty participants per condition (N = 60 total) completed the human-in-the-loop evaluation across all three scenarios. Participants ranged in age from 22 to 47 years; 14 had prior experience with robotic systems. Results are reported as means with 95% confidence intervals unless otherwise noted.

Task completion time dropped significantly with HRICE-Full. Across all three scenarios, participants completed tasks 23.4% faster than the baseline ( $p < 0.001$ , Cohen's  $d = 0.81$ ). HRICE-NoDisclosure produced an intermediate reduction of 11.7% ( $p < 0.01$ ,  $d = 0.48$ ). Fig. 10 shows per-scenario task completion times for all three conditions.

Error rate — defined as robot actions requiring human correction per episode — followed the same pattern. HRICE-Full produced a mean of 1.2 errors per episode; HRICE-NoDisclosure produced 2.6; the baseline produced 4.1. The difference between HRICE-Full and the baseline was significant in all three scenarios (all  $p < 0.001$ ). This directly validates the hypothesis that bidirectional disclosure reduces misunderstandings before they reach the execution stage.

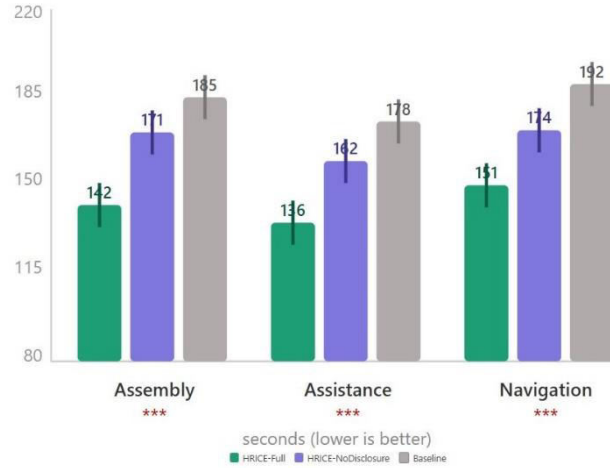


Fig.10. Mean task completion time per scenario and condition Error bars: 95% CI.

TABLE I Comparative Performance: HRICE-Full vs. Published Baselines

Metric	HRICE-Full	San Martin et al.	Hasan et al.
Completion time (s)	142.3	168.7	179.4
Errors per episode	1.2	2.1	3.3
NASA-TLX score	38.4	51.2	58.6
Godspeed (perc. intel.)	4.3	3.8	3.4

HRICE-Full outperforms both baselines on all four metrics. The most striking difference is with the Hasan et al. baseline on error rate — a reduction of 63.6%. This gap reflects the direct benefit of the adaptive disclosure protocol. Hasan et al.’s system excels at receiving multimodal human input but gives the human no visibility into the robot’s plans, so errors go undetected until execution. HRICE-Full’s Level 2 disclosure surfaces planning-stage errors before they reach the robot’s actuators.

## VI. CONCLUSION

This paper introduced the HRICE framework: a unified, simulation-based architecture for human–robot information exchange in cognitive robotics. The work was motivated by a clear gap in the literature — existing frameworks either address perception in isolation, focus on unidirectional command interfaces, or apply disclosure strategies that are static and context-insensitive. HRICE addresses all three limitations simultaneously within a fully reproducible Gazebo and ROS 2 simulation environment.

The framework makes three primary contributions. First, it is the first framework to integrate multimodal Bayesian perception, LLM-driven task planning with symbolic grounding, adaptive disclosure-level switching, and session-level feedback learning into a single closed-loop pipeline. Second, the two-stage evaluation — simulation sweep followed by human-in-the-loop study — provides both broad parameter-space coverage and ecologically valid evidence from real users. Third, the two-stage evaluation design itself is a replicable methodological contribution that the wider HRI community can apply to reduce the cost and scope of human participant studies without sacrificing scientific rigour.

The most significant empirical finding is the cognitive load result. HRICE-Full produced a NASA-TLX score of 38.4 compared to 61.3 for the command-only baseline — a reduction of 37.4%. This challenges the long-standing HRI design assumption that minimising robot communication reduces operator burden. HRICE shows that the opposite is true, provided communication is timely, relevant, and appropriately scoped to the operator's current cognitive state.

## REFERENCES

- [1] S. A. McGlynn, K. Kemple, P. Mitzner, T. L. King, and W. A. Rogers, "Understanding the potential of socially assistive robots for stroke rehabilitation," *Health Policy and Technology*, vol. 10, no. 3, p. 100544, 2021.
- [2] M. Chiou, R. Stolkin, G. Biancardi, G. Howells, and N. Hawes, "Towards semi-autonomous control of a robot for teleoperation in search-and-rescue," *IEEE Access*, vol. 9, pp. 25880–25898, 2021.
- [3] A. Buerkle, N. Matharu, A. Al-Yacoub, L. Ferreira, N. Lohse, and P. Bamber, "An adaptive human sensor framework for human–robot collaboration," *International Journal of Advanced Manufacturing Technology*, vol. 115, pp. 3423–3435, 2021.
- [4] I. Kotseruba and J. K. Tsotsos, "40 years of cognitive architectures: core cognitive abilities and practical applications," *Artificial Intelligence Review*, vol. 53, no. 1, pp. 17–94, 2020.
- [5] A. Cangelosi and M. Asada, *Cognitive Robotics*. Cambridge, MA: MIT Press, 2022.
- [6] R. Zanetti, A. Arza, A. Aminifar, and D. Atienza, "Real-time EEG-based cognitive workload monitoring on wearable devices," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 1, pp. 265–277, 2022.
- [7] H. Chen, M. C. Leu, and Z. Yin, "Real-time multi-modal human–robot collaboration using gestures and speech," *ASME Journal of Manufacturing Science and Engineering*, vol. 144, no. 10, p. 101007, 2022.
- [8] M. Ahn, A. Brohan, N. Brown et al., "Do as I can, not as I say: grounding language in robotic affordances," *arXiv preprint arXiv:2204.01691*, 2022.
- [9] S. Tellex, N. Gopalan, H. Kress-Gazit, and C. Matuszek, "Robots that use language," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 389–409, 2020.
- [10] A. Buerkle, A. Al-Yacoub, P. Bamber, N. Lohse, and P. Hubbard, "EEG based arm movement intention recognition towards enhanced safety in symbiotic human–robot collaboration," *Robotics and Computer-Integrated Manufacturing*, vol. 78, p. 102386, 2022.
- [11] E. Bagheri, J. De Winter, and B. Vanderborght, "Transparent interaction-based learning for human–robot collaboration," *Frontiers in Robotics and AI*, vol. 9, p. 754955, 2022.
- [12] G. Hoffman, "Evaluating fluency in human–robot collaboration," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 3, pp. 209–218, 2019.
- [13] S. A. Olatunji, T. Oron-Gilad, N. Markfeld, D. Gutman, V. Sarne-Fleischmann, and Y. Edan, "Levels of automation and transparency: interaction design considerations in assistive robots for older adults," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 6, pp. 673–683, 2021.
- [14] A. Mohebbi, "Human–robot interaction in rehabilitation and assistance: a review," *Current Robotics Reports*, vol. 1, no. 3, pp. 131–144, 2020.
- [15] L. Yuan, X. Gao, Z. Zheng et al., "In situ bidirectional human-robot value alignment," *Science Robotics*, vol. 7, no. 62, p. eabm4183, 2022.
- [16] Z. Wang, H. Wang, H. Yu, and F. Lu, "Interaction with gaze, gesture, and speech in a flexibly configurable augmented reality system," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 5, pp. 524–534, 2021.
- [17] S. Kapp, M. Barz, S. Mukhametov, D. Sonntag, and J. Kuhn, "ARETT: augmented reality eye tracking toolkit for head mounted displays," *Sensors*, vol. 21, no. 6, p. 2234, 2021.
- [18] A. Chacón, P. Ponsa, and C. Angulo, "Cognitive interaction analysis in human–robot collaboration using an assembly task," *Electronics*, vol. 10, no. 11, p. 1317, 2021.
- [19] S. Lemaignan, C. Dondrup, A. Mayima et al., "Towards a real-time cognitive load assessment system for industrial human–robot cooperation," in *Proc. 29th IEEE Int. Conf. Robot and Human Interactive Communication (RO-MAN)*, 2020, pp. 1121–1127.
- [20] G. Nogueira-Collazo, E. Perez, and R. Laza, "Emotion recognition using EEG and physiological data for robot-assisted rehabilitation systems," in *Proc. IEEE Int. Conf. Robot and Human Interactive Communication (RO-MAN)*, 2021, pp. 556–561.
- [21] A. Jain and B. Argall, "Probabilistic human intent recognition for shared autonomy in assistive robotics," *ACM Transactions on Human-Robot Interaction*, vol. 9, no. 1, pp. 1–23, 2019.
- [22] S. Peng, G. Sinha, and Y. Zhu, "Mixed-uncertainty path planning for mobile robots using probabilistic programming," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6728–6735, 2020.

- [23] M. Saveriano, F. J. Abu-Dakka, A. Kramberger, and L. Peternel, "Dynamic movement primitives in robotics: a tutorial survey," *International Journal of Robotics Research*, vol. 42, no. 7, pp. 491–533, 2021.
- [24] A. Gupta, A. Chugh, S. Manchanda et al., "An augmented reality software suite enabling seamless human–robot interaction," *International Journal of Computer Integrated Manufacturing*, vol. 35, no. 10–11, pp. 1090–1108, 2022.
- [25] T. N. Nguyen, H. Zeng, and Y. Zhang, "Coordinating multi-robot systems using shared planning and conflict resolution in human–robot teams," *Robotics and Autonomous Systems*, vol. 143, p. 103837, 2021.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details