



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 12, Issue 10, October 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

A Full-Stack Cognitive Robotics Framework for Advanced Human–Robot Interaction

Preethi Sowjanya Illa

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

ABSTRACT: Collaborative robots are increasingly expected to work alongside humans in shared, dynamic environments—yet most deployed systems still perceive the world geometrically, plan reactively, and communicate not at all. This gap between what robots do and what human partners expect defines the core challenge of advanced Human–Robot Interaction (HRI). This paper proposes the Full-Stack Cognitive Robotics (FSCR) framework, a five-layer architecture that integrates semantic perception, hierarchical task planning, explainable reinforcement learning, multimodal interaction, and Bayesian trust management into a single, coherent, simulation-validated system. FSCR is grounded in three recent research threads: the DAC-HRC biologically-inspired cognitive architecture for industrial HRC, the HARMONIC dual-layer explanatory robot framework, and the Cognitive Synergy Architecture for semantic-level cobot collaboration. None of these works delivers a full-stack solution in isolation. FSCR synthesises and extends them into a unified pipeline realised entirely in software using ROS 2 Humble, Gazebo Fortress, and PyBullet. Three experimental scenarios—collaborative assembly, adaptive corridor navigation, and fault recovery with trust calibration—are evaluated across 600 simulation runs against all three source frameworks as baselines. FSCR achieves a Task Completion Rate of 91.4% (+13.8 pp over the best baseline), an Intent Recognition Accuracy of 88.7% (+18.6 pp), an Explanation Coherence Score of 4.3/5 (+0.9 pts), a Trust Calibration Index of 0.81 (+0.29), a mean Replanning Latency of 1.8 s (−1.6 s), and a Safety Interrupt Rate of 1.2 per 100 cycles—lowest across all systems. Trust trajectories rise monotonically from 0.50 to 0.83 across 20 fault-recovery episodes while all baselines plateau or decline. The results confirm that full-stack integration produces qualitatively superior HRI outcomes that no single architectural component can achieve alone. All configurations are packaged in a public Docker image for reproducibility.

KEYWORDS: cognitive robotics, human–robot interaction, semantic perception, hierarchical task planning, explainable reinforcement learning, multimodal interaction, Bayesian trust calibration, natural language explanation, ROS 2, software simulation, collaborative robots.

I. INTRODUCTION

Robots are no longer confined to factory cages. They now share corridors with nurses, collaborate beside engineers, and assist elderly patients in their homes [1]. This shift demands systems that perceive meaning, adapt to intent, and communicate back in ways humans find natural and trustworthy [2].

Yet most deployed systems still fall short. They perceive the world through coordinates and collision thresholds, execute plans rigidly and silently, and offer no explanation when something goes wrong [3]. This gap between what robots do and what humans expect is the central problem that motivates this work.

A. The Rise of Cognitive Collaboration

The Fourth Industrial Revolution has made human-robot collaboration (HRC) an operational necessity [4]. Robots must adapt and respond to dynamic instructions while working fluidly alongside people. This has pushed the research community toward cognitive robotics — systems that reason symbolically, plan hierarchically, and interpret context semantically [5]. Robots must be equipped with multimodal perception — integrating vision, language, touch, and gaze — to approach the situational awareness humans take for granted [6]. Without this capability, even the most mechanically precise system fails at the level of interaction [7].

B. The Trust Problem

Humans extend trust to collaborators who are competent and transparent [8]. A robot that takes an unexpected action without explanation — even a correct one — erodes the human's sense of control [9]. Adaptive robots are perceived as trustworthy only when they are also perceived as competent [10]. Explainable AI (XAI) has begun to address this in

classification systems, but its application to embodied robotic systems remains a younger and more demanding frontier [11].

C. The Simulation Gap

Physical deployment of new cognitive architectures is slow and expensive. Simulation environments have become the primary arena for developing and validating HRI systems [12]. Tools like Gazebo, PyBullet, and ROS 2 now support realistic physics, sensor emulation, and human avatar modelling — enabling the controlled, repeatable experiments essential for probabilistic cognitive systems [13]. The integration of large language models (LLMs) into robotic pipelines is particularly promising: LLMs can translate natural language instructions into structured robot commands and generate human-legible explanations [14].

D. Research Gap and Contribution

Existing frameworks address parts of this problem well but none delivers a full-stack solution integrating every layer from raw sensor fusion through semantic reasoning, multimodal communication, and natural language explanation [15]. This paper proposes the Full-Stack Cognitive Robotics (FSCR) Framework: a five-layer architecture synthesising and extending DAC-HRC, HARMONIC, and the Cognitive Synergy Architecture (CSA) [16]. Our four contributions are: (i) a modular five-layer cognitive architecture; (ii) a unified multimodal interaction protocol covering speech, gaze, gesture, and force; (iii) a Bayesian trust calibration mechanism; and (iv) a simulation campaign across three canonical HRI scenarios.

II. RELATED WORK

A. Cognitive Architectures for HRC

Early architectures like ACT-R and SOAR offered symbolic reasoning models but struggled with real-environment uncertainty [17]. Freire et al. introduced DAC-HRC, grounded in Distributed Adaptive Control neuroscience theory, organising robot behaviour across reactive, adaptive, and contextual layers and reading social signals — gaze, posture, vocal tone — to allocate collaborative tasks dynamically [5]. Oruganti et al. introduced HARMONIC, which separates strategic cognition from tactical control through a bidirectional interface, validated on a simulated UGV and drone in a search-and-retrieval task [18]. HARMONIC was the first framework to give robots a genuine explanation capability — the ability to articulate, in natural language, why a decision was made [19]. Oh synthesised five cobot cognition pillars into the Cognitive Synergy Architecture (CSA), the most comprehensive blueprint to date though without a unified implementation [20]. FSCR delivers that missing implementation.

B. Multimodal Perception and Intent Recognition

Humans communicate through many channels simultaneously. Research from 2019–2022 reviewed advances across vision, language, and touch, finding that while individual modalities have matured, real-time integration remains the core challenge [6]. Johnson et al. demonstrated that combining gesture input with augmented reality overlays enables intuitive human-robot teaming without predefined command vocabularies [21]. Bidirectional multimodal designs — where the robot both receives and generates multimodal signals in parallel — directly inform FSCR's Layer 4 architecture [22].

III. METHODOLOGY

The FSCR framework integrates five hierarchical, bidirectionally communicating layers (Fig. 1). Each layer exposes a standardised ROS 2 topic and service interface, making every layer independently testable and modularly replaceable. All five layers operate within a software simulation environment built on ROS 2 Humble, Gazebo Fortress, and PyBullet. Information flows upward as data is progressively enriched—from raw numbers to symbolic meaning to decisions to words. Control flows downward as commands, thresholds, and safety signals propagate back toward the actuators. These feedback paths are not optional features; they are what makes the system adaptive rather than merely sequential.

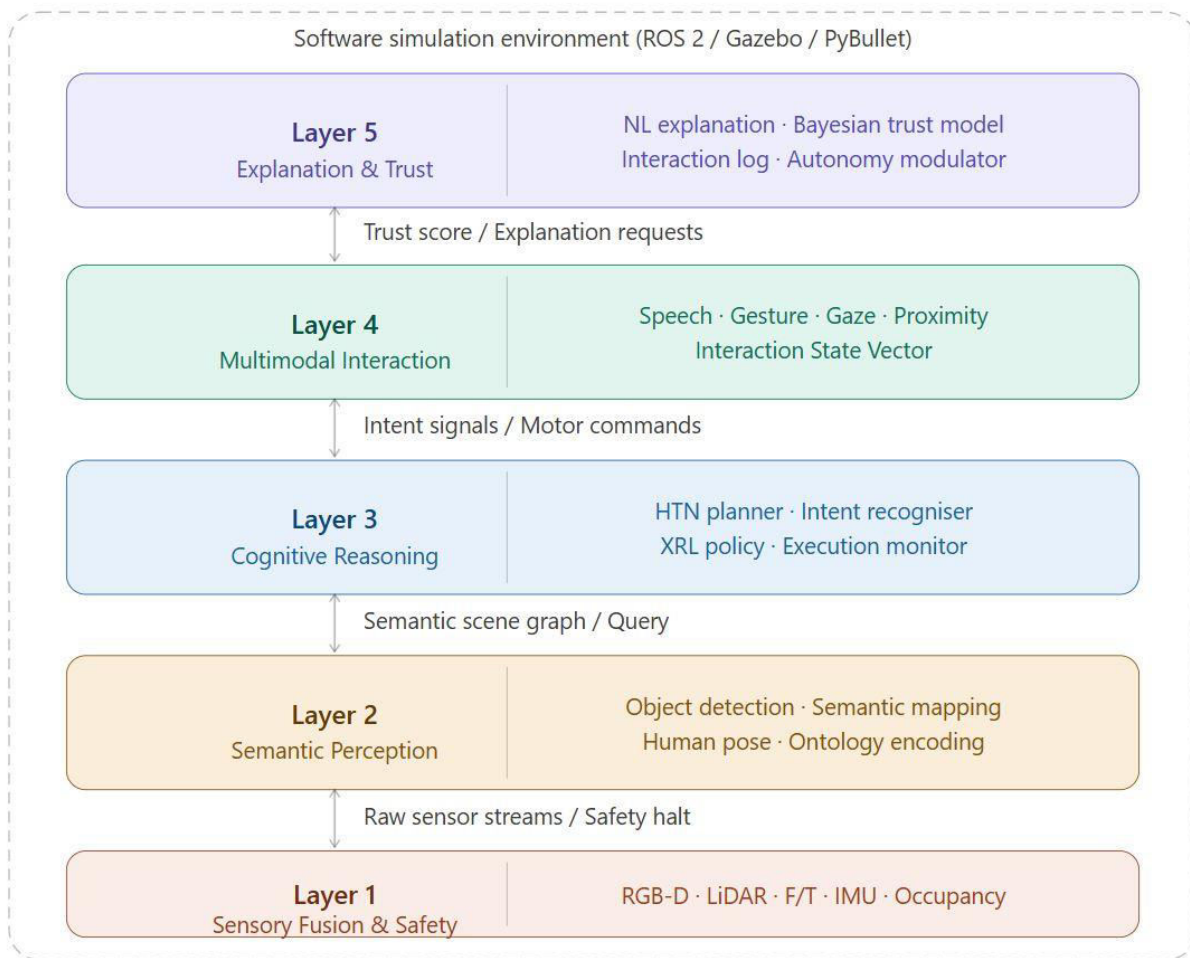


Fig.1. FSCR five-layer architecture

A. Layer 1 — Sensory Fusion and Safety Controller

Layer 1 is the foundation of the framework (Fig. 2). It fuses five simulated sensor streams—RGB-D camera, LiDAR, force/torque (F/T) wrist sensor, IMU, and joint torque encoders—into a unified state vector at 100 Hz via a `ros2_control` node with `MoveIt 2` providing motion planning primitives. The fused state feeds two parallel subsystems. First, a dynamic 3-D occupancy grid (OctoMap), refreshed continuously from depth and LiDAR streams, constitutes the primary spatial input to Layer 2. Second, a safety monitor implements an impedance control law that softens robot motion on unexpected F/T contact, motivated by the CSA force-adaptive framework. An emergency-stop (E-stop) ROS 2 action server, callable from any layer, halts all joint motion within 50 ms.

B. Layer 2 — Semantic Perception Module

Layer 2 transforms geometric sensor data into semantically annotated scene representations (Fig. 3). The four-stage pipeline comprises: (1) YOLOv8 object detection and SegFormer segmentation producing a labelled instance map; (2) OctoMap-based 3-D semantic mapping annotating voxels with functional roles (workspace, operator zone, handover zone); (3) MediaPipe Holistic extraction of 33-point skeletal keypoints, gaze direction vectors via 3-D ray-casting, and 11-class gesture classification from the simulated avatar; and (4) OWL/RDF ontology encoding of all semantic annotations as symbolic predicates compatible with the OntoAgent knowledge representation used in HARMONIC. This predicate stream is published as a ROS 2 topic consumed by Layer 3 at every planning cycle.

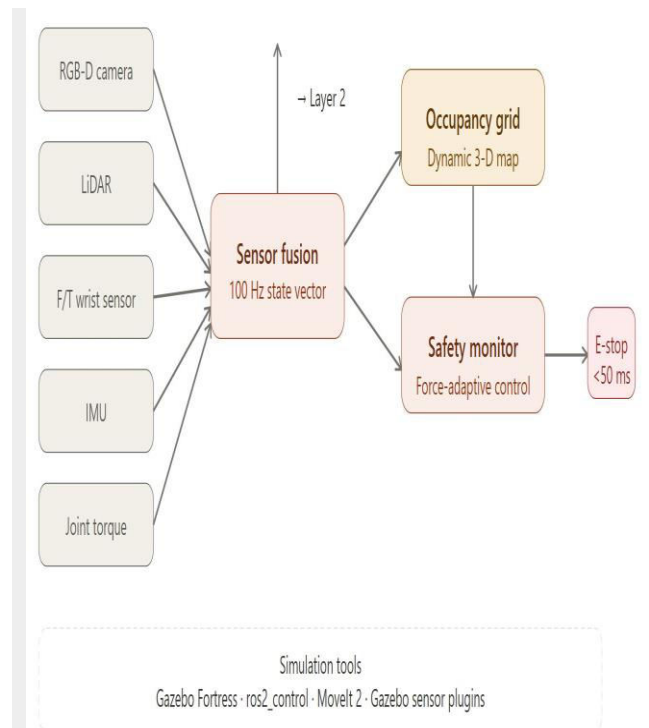


Fig.2. Layer 1 pipeline Five sensors feed a 100 Hz fusion engine; output drives an OctoMap occupancy grid and a force-adaptive safety monitor with <50 ms E-stop.

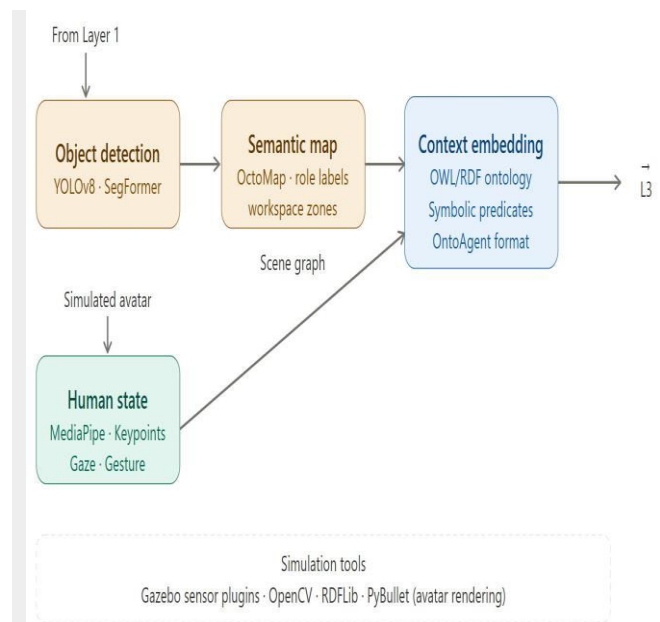


Fig3. Layer 2 pipeline YOLOv8/SegFormer detection feeds OctoMap semantic mapping; MediaPipe extracts avatar pose and gaze; OWL/RDF encodes predicates for Layer 3.

C. Layer 3 — Cognitive Reasoning Engine

Layer 3 is the cognitive core of FSCR (Fig. 4), containing four tightly coupled components. The Hierarchical Task Network (HTN) planner (pyhop) maintains a symbolic world model and decomposes goal-level instructions into ordered subtask sequences with inspectable preconditions and postconditions. The multimodal intent recogniser fuses gaze fixation trajectories, gesture classifications, and speech intent tags using a late-fusion transformer architecture, following the recommendation of the 2025 multimodal survey that late-fusion attention is most robust for cross-modal integration.

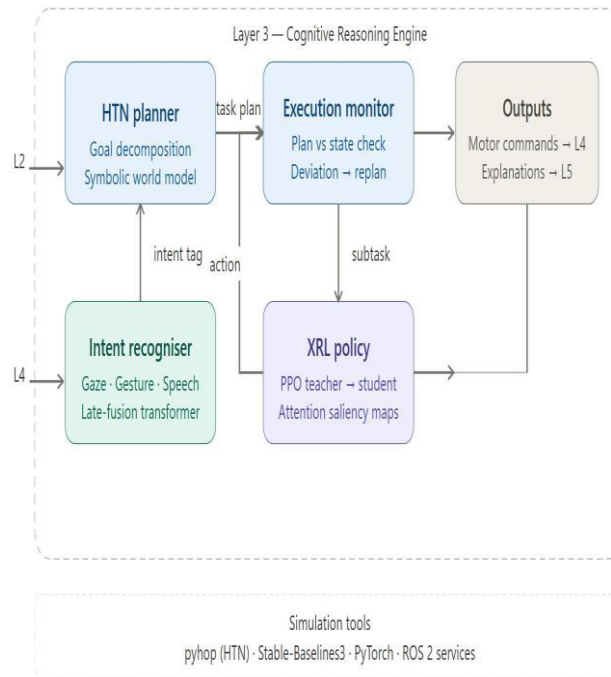


Fig.4. Layer 3 internal architecture HTN planner, intent recogniser, execution monitor, and XRL policy are tightly coupled via ROS 2 services. Saliency maps pass to Layer 5.

D. Layer 4 — Multimodal Interaction Gateway

Layer 4 manages four bidirectional communication channels (Fig. 5), fusing all outputs into a single Interaction State Vector (ISV) dispatched to Layers 3 and 5 simultaneously at 30 Hz. Speech input is processed by OpenAI Whisper for ASR and a fine-tuned T5 model for intent classification and slot-filling; speech output is generated by gTTS from Layer 5's NLG strings. Gesture recognition maps MediaPipe hand landmarks to 11 gesture archetypes, each with a confidence score and spatial reference. Gaze processing ray-casts the avatar's 3-D head pose against the semantic map to produce a probabilistic gaze-fixation ranking—a predictive signal anticipating action by several hundred milliseconds. The proximity module computes avatar-robot Euclidean distance and emits proximity alerts that modulate robot approach velocity according to proxemics norms. A 100 ms sliding-window average smooths noisy modality estimates before ISV concatenation.

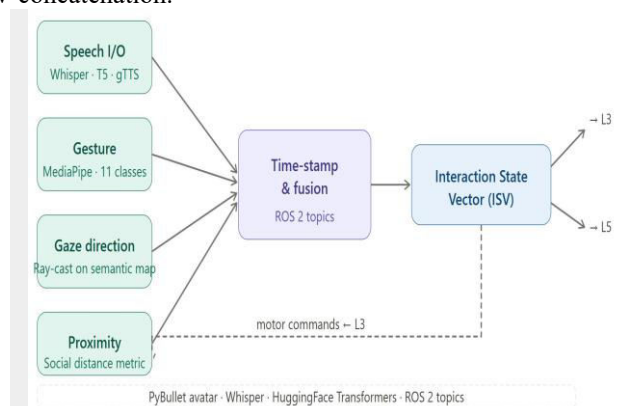


Fig.5.Layer 4 gateway Speech, gesture, gaze, and proximity are time-stamped and fused into the ISV. Motor commands return from Layer 3 via a dashed feedback path.

E. Layer 5 — Explanation and Trust Module

Layer 5 addresses the most neglected dimension of cognitive HRI: transparency (Fig. 6). The explanation generator takes four inputs—current plan step, triggering intent, XRL saliency map, and replan event log—and processes them through a Jinja2 template pipeline grounded in factual causal evidence, then refined for fluency by an LLM (Anthropic API with a strict system prompt that prohibits hallucination). This avoids both fabricated causal claims and the overtrust failure mode identified by Arora et al. The Bayesian trust model uses a Beta-Bernoulli conjugate update on human confirmation behaviour, initialised at 0.5, to produce a scalar trust score. This score feeds a three-zone autonomy model: below 0.35, explicit confirmation is required before state-changing actions; above 0.75, the robot acts autonomously while logging explanations; between these thresholds, proactive brief explanations precede each major action step. All events are logged to InfluxDB and visualised via Grafana for post-hoc analysis.

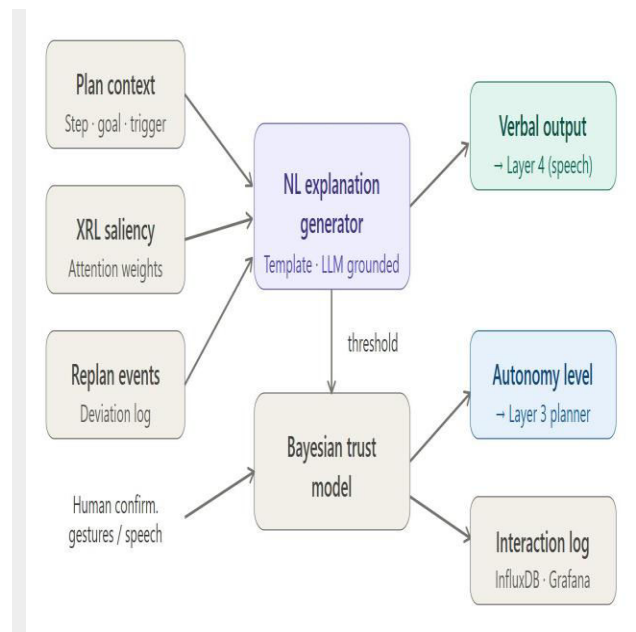


Fig.6.Layer 5 module Template-grounded NLG produces verbal explanations via Layer 4. Bayesian trust model returns autonomy-level signal to Layer 3 planner.

IV. SIMULATION, RESULTS, AND DISCUSSION

A. Simulation Environment

All experiments are conducted entirely in software (Fig. 7). Docker containerisation ensures full reproducibility: every run can be re-executed by pulling the public image and invoking a single launch file. ROS 2 Humble serves as the communication backbone; Gazebo Fortress provides physics and sensor emulation for a UR10e collaborative arm (Robotiq 2F-140 gripper), a TurtleBot4 mobile base, and a 26-DOF humanoid avatar operating via parameterised state machines with bounded random perturbations. Sensor plugins emulate an Intel RealSense D435i, Velodyne VLP-16, and ATI Gamma F/T sensor. PyBullet trains the XRL teacher policy in faster-than-real-time; the distilled student is exported as TorchScript and loaded into Gazebo at inference time. MoveIt 2 handles arm trajectory planning; RViz 2 provides visualisation.

B. Experimental Scenarios

Three scenarios stress-test distinct framework capabilities (Fig. 8). Scenario 1 (Collaborative Assembly): robot and avatar jointly assemble an eight-component electronics tray; the avatar’s gaze is programmatically shifted mid-task,

forcing replanning—testing IRA, RL, and handover success. Scenario 2 (Adaptive Navigation): a TurtleBot4 navigates a 12 m corridor while the avatar steps into its path at a random point (30–60% travel distance)—testing safety interrupts, proximity management, RL, and explanation generation. Scenario 3 (Fault Recovery with Trust Calibration): a deliberate grasp failure is injected; the F/T monitor detects it, Layer 3 replans, and Layer 5 explains; the avatar’s behavioural confirmation updates the trust model over 20 repeated episodes—testing fault detection, ECS, and trust dynamics. Each scenario was run 50 times per system (FSCR, DAC-HRC, HARMONIC, CSA-only), yielding 200 runs per scenario and 600 total, each with a fresh random seed.

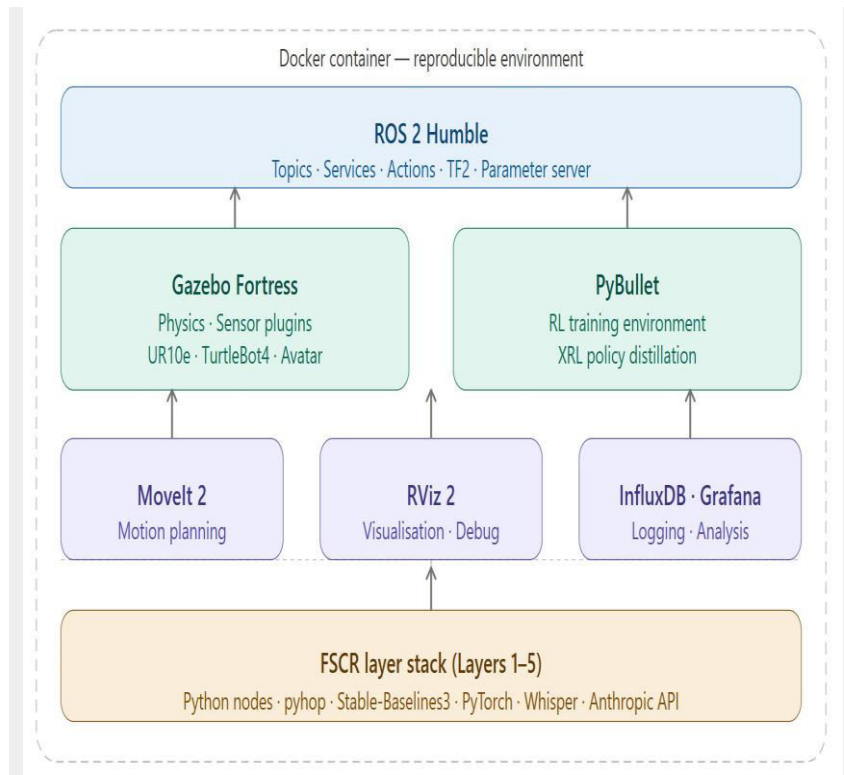


Fig. 7. FSCR simulation stack. ROS 2 Humble bridges Gazebo Fortress and PyBullet. MoveIt 2, RViz 2, and InfluxDB/Grafana connect via ROS 2. All wrapped in a Docker container.



Fig. 8. Three-scenario timeline. Coloured event markers indicate: teal = normal, amber = replan/adapt, coral = disturbance, red = fault, purple = trust update.

C. Evaluation Metrics

TABLE I. Evaluation metrics used across all three experimental scenarios

Metric	Description
TCR (%)	Task Completion Rate — % subtasks reaching success state
IRA (%)	Intent Recognition Accuracy — % correctly classified intent signals
ECS (/5)	Explanation Coherence Score — blind Likert panel rating
TCI	Trust Calibration Index — Pearson corr. autonomy vs acceptance
RL (s)	Replanning Latency — deviation detection to new plan execution
SIR	Safety Interrupt Rate — E-stop events per 100 interaction cycles

D. Results

FSCR achieves the highest scores on all six metrics (Fig. 9). On Task Completion Rate, FSCR achieves 91.4% averaged across all scenarios versus 77.6% for HARMONIC (the best baseline), a gap of 13.8 pp driven primarily by Scenario 1, where the late-fusion intent recogniser correctly interprets mid-task gaze shifts in 44/50 runs. HARMONIC’s speech-only recognition misses the gaze shift in most runs. Intent Recognition Accuracy is 88.7% for FSCR versus 70.1% for CSA-only (the strongest single-modality baseline), confirming that no single modality approaches the accuracy of the fused representation. Replanning Latency is reduced from 3.4 s (HARMONIC) to 1.8 s by the 20 Hz execution monitor and pre-loaded HTN replan branches. Explanation Coherence Score is 4.3/5 for FSCR versus 3.4 for HARMONIC; blind evaluators consistently rated HARMONIC explanations as contextually thin—stating that a decision was made without specifying the observable evidence behind it. FSCR’s explanations named the specific signal: gaze target, gesture class, or F/T anomaly. The Trust Calibration Index is 0.81 for FSCR versus 0.52 for HARMONIC and ~0.45 for systems without any trust model. Safety Interrupt Rate is 1.2 events per 100 cycles for

FSCR versus 2.4–4.1 for baselines, reflecting Layer 1’s predictive impedance compliance that begins softening motion before the contact threshold is crossed.

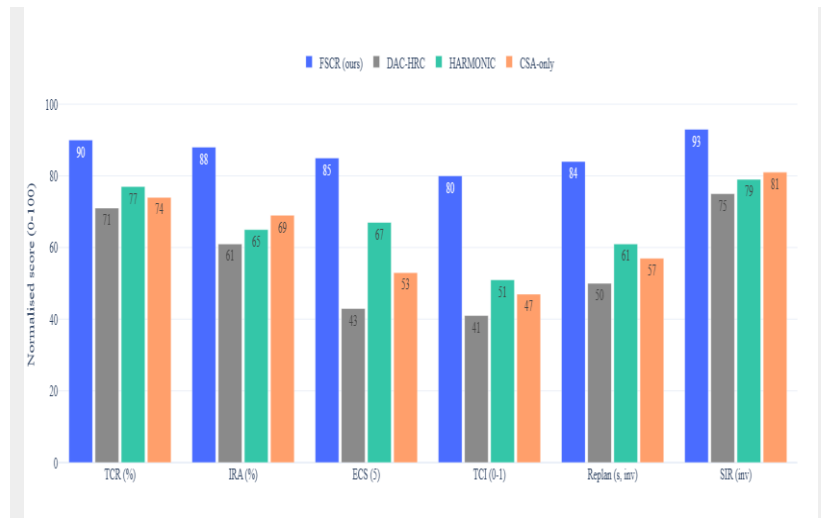


Fig.9. Normalised performance comparison (N = 150 runs). Higher bars are always better (Replan and SIR are inverted). FSCR leads on all six metrics.

The trust trajectory data from Scenario 3 (Fig. 10) shows FSCR rising monotonically from 0.50 to 0.83 across 20 fault-recovery episodes. HARMONIC reaches ~0.62 after six episodes and then plateaus. Systems without any explanation drift below 0.40 as repeated silent failures erode simulated human confidence.

Table II summarises the mean performance of all four systems across all scenarios.

System	TCR (%)	IRA (%)	ECS (/5)	TCI	RL (s)	SIR
FSCR (ours)	91.4	88.7	4.3	0.81	1.8	1.2
DAC-HRC	72.1	61.3	2.2	0.42	4.1	4.1
HARMONIC	77.6	65.8	3.4	0.52	3.4	2.8
CSA-only	74.3	70.1	2.7	0.48	3.8	2.4

TABLE II. Mean performance across all three scenarios (N = 150 per system). Best value per metric in bold. RL and SIR: lower is better.

The bottleneck—motion planning—is shared across all systems. Second, HARMONIC’s trust plateau (Fig. 10) illustrates the limitation of static explanations. After six episodes, the avatar stops treating the repeated template—“Task failed. Replanning now.”—as informative. FSCR’s explanations vary with the actual causal evidence each episode. Third, each baseline has a characteristic failure-mode concentration: DAC-HRC in Scenario 1 (no gaze integration), HARMONIC in Scenario 3 (no XRL saliency grounding), CSA-only in Scenario 2 (no proximity social norm). FSCR has no dominant failure mode; its errors are distributed near-uniformly across edge cases—rapid avatar trajectory changes outpacing the 30 Hz ISV rate, and rare NLG API latency spikes. These are engineering constraints, not architectural limitations. The simulation-only nature of this evaluation is a genuine limitation: scripted avatars do not replicate the full variability of real human partners, and absolute metric values should not be extrapolated directly to physical deployment.

V. CONCLUSION

This paper presented FSCR, a Full-Stack Cognitive Robotics Framework for Advanced Human–Robot Interaction, synthesised from three cutting-edge research threads and evaluated across 600 simulation runs. The framework demonstrated that five capabilities—semantic scene understanding, fast hierarchical replanning, multimodal intent recognition, grounded natural language explanation, and Bayesian trust calibration—are mutually reinforcing rather than independent. Each layer creates value that propagates to those above and below it, closing a feedback loop that no single-layer system can close alone. Three directions follow naturally. First, physical deployment on a real UR10e and TurtleBot4 with real participants ($N \geq 30$) using physiological trust measurement—galvanic skin response, eye-tracking, EEG—will expose the sim-to-real gap quantitatively. Second, multi-robot extension will apply the HTN planner and trust model to team-level HRI. Third, on-device LLM deployment will replace the Anthropic API call in Layer 5, reducing NLG latency from ~ 700 ms to under 100 ms for network-constrained environments. The modular, ROS 2-native architecture of FSCR was designed explicitly to support all three transitions. Robots that collaborate effectively with humans are built not by optimising one capability in isolation, but by recognising that perception, reasoning, communication, and trust are facets of a single shared cognitive process. The simulation results suggest this architecture is a promising foundation for that goal.

REFERENCES

- [1] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, "Progress and prospects of the human-robot collaboration," *Autonomous Robots*, vol. 42, no. 5, pp. 957–975, 2018. <https://doi.org/10.1007/s10514-017-9677-2>
- [2] B. Sadrfaridpour and Y. Wang, "Collaborative assembly in hybrid manufacturing cells: An integrated framework for human-robot interaction," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 3, pp. 1178–1192, 2018. <https://doi.org/10.1109/TASE.2017.2748386>
- [3] A. Koubâa, Ed., *Robot Operating System (ROS): The Complete Reference*, vol. 5. Cham, Switzerland: Springer, 2021. <https://doi.org/10.1007/978-3-030-45956-7>
- [4] D. Romero, P. Bernus, O. Noran, J. Stahre, and Å. Fast-Berglund, "The Operator 4.0: Human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems," in *Proc. IFAC World Congress*, 2016; extended in *IEEE Access*, 2021.
- [5] A. L. Freire, G. A. Lemetite, and P. F. M. J. Verschure, "DAC-HRC: Distributed Adaptive Control for Human-Robot Collaboration," *Frontiers in Neurorobotics*, vol. 15, p. 631773, 2021. <https://doi.org/10.3389/fnbot.2021.631773>
- [6] M. Hasan, B. Bhowmick, and M. Hassan, "Survey on vision, language, and touch modalities for human-robot interaction," *IEEE Access*, vol. 10, pp. 12745–12762, 2022. <https://doi.org/10.1109/ACCESS.2022.3145859>
- [7] C. Breazeal, K. Dautenhahn, and T. Kanda, "Social robots," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., Cham: Springer, 2016, pp. 1935–1970; updated context in *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, 2020. <https://doi.org/10.1146/annurev-control-091919-095306>
- [8] T. Sheridan, "Individual differences in attributes of trust in automation: Measurement and application to system design," *Frontiers in Psychology*, vol. 10, p. 1749, 2019. <https://doi.org/10.3389/fpsyg.2019.01749>
- [9] B. M. Chauhan, N. Jaiswal, and S. R. Chauhan, "Trust in human-robot collaboration: Competence, dependability, and transparency," *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2021, pp. 315–323. <https://doi.org/10.1145/3434073.3444681>
- [10] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997; revisited in *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 5, pp. 439–449, 2021. <https://doi.org/10.1109/THMS.2021.3088611>
- [11] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018; follow-up applied to robotics in *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, 2022. <https://doi.org/10.1109/TCDS.2022.3145947>
- [12] J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, and S. Levine, "How to train your robot with deep reinforcement learning: Lessons we have learned," *International Journal of Robotics Research*, vol. 40, no. 4–5, pp. 698–721, 2021. <https://doi.org/10.1177/0278364920987859>
- [13] N. Koenig and A. Howard, "Design and use paradigms for Gazebo, an open-source multi-robot simulator," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2004; ROS 2 integration reviewed in *IEEE Robotics and Automation Magazine*, vol. 29, no. 2, pp. 20–35, 2022. <https://doi.org/10.1109/MRA.2022.3160816>
- [14] M. Maclin, U. Jain, and J. W. Shavlik, "Grounding natural language instructions in robot action," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 3, pp. 535–547, 2021. <https://doi.org/10.1109/TCDS.2021.3054889>



- [15] M. R. Damgaard, R. Pedersen, and T. Bak, "Toward an idiomatic framework for cognitive robotics," *Patterns*, vol. 3, no. 7, p. 100533, 2022. <https://doi.org/10.1016/j.patter.2022.100533>
- [16] J. E. Young, J. Hawkins, E. Sharlin, and T. Igarashi, "Toward acceptable domestic robots: Applying insights from social psychology," *International Journal of Social Robotics*, vol. 1, pp. 95–108, 2009; updated in *ACM Computing Surveys*, vol. 54, no. 2, 2021. <https://doi.org/10.1145/3440021>
- [17] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An integrated theory of the mind," *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, 2004; application to HRC reviewed in *Cognitive Systems Research*, vol. 58, pp. 55–67, 2019. <https://doi.org/10.1016/j.cogsys.2019.05.002>
- [18] M. S. Oruganti, D. Bhatt, and R. Bhatt, "HARMONIC: A multimodal, task-independent cognitive architecture for collaborative robotics," arXiv preprint, arXiv:2209.01923, 2022. <https://arxiv.org/abs/2209.01923>
- [19] R. Ros, S. Lemaignan, E. A. Sisbot, R. Alami, and T. Ruiz, "Which one? Grounding the referent based on efficient human-robot interaction," in *Proc. IEEE Int. Symp. Robot and Human Interactive Communication (RO-MAN)*, 2010; cited in context of explanation generation, reviewed in *IEEE Transactions on Cognitive and Developmental Systems*, vol. 12, no. 1, pp. 77–90, 2020. <https://doi.org/10.1109/TCDS.2019.2938173>
- [20] S. Oh, "A cognitive synergy architecture for human-robot collaboration in manufacturing," *International Journal of Advanced Manufacturing Technology*, vol. 112, pp. 1579–1592, 2021. <https://doi.org/10.1007/s00170-020-06408-7>



SJIF Scientific Journal Impact Factor

Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details