# LANGUAGE MODELS IN HUMAN-ROBOT COLLABORATION: PAST, PRESENT AND FUTURE


VYTAUTAS MAGNUS UNIVERSITY
Faculty of Informatics

**AUTHORS:**

**Eimantas Zaranka**
eimantas.zaranka@vdu.lt

**Roel Pieters**
roel.pieters@tuni.fi

**Tomas Krilavičius**
tomas.krilavicius@vdu.lt


**Survey the field for me:**

1. **Classify** how language models (LMs) and vision-language-action (VLA) models have been integrated into HRC.
2. **Identify** bottlenecks for language models application in Human-Robot Collaboration.
3. **Outline** future research directions.

H

✓✓ Read

R

> survey --peer-reviewed --years 2020–2026

**48** **sources**
LMs and VLA models usage in HRC, excluding HRI without LMs integration

SURVEYED VENUES

CoRL · RSS · ICML · NeurIPS · ICLR · AAAI · ICRA · Robotics journals · arXiv

## Why does this matter now?

H

R

**4.6 million+** active units in the global fleet

Robotics is shifting toward the **Industry 5.0** paradigm, emphasizing sustainability and human-centricity. Demand for seamless HRC has grown rapidly.

Nevertheless, the **safe integration** of robots into shared human environments remains the **critical bottleneck**. A primary challenge in deploying LMs and VLAs lies in combining high-level semantic reasoning with timely and precise physical execution.

Leveraging the models enables robots to understand natural language, anticipate human intent, and adapt safely and dynamically to complex and changing environments.

## How did we get here?

H

R

- **2020** — **Symbolic Grounding** (Grammars)
- **2021** — **CLIPort** Semantic-Spatial Bridging
- **2022** — **SayCan & Inner Monologue** Closed-loop planning & feedback
- **2023–2024** — **RT-2 & OpenVLA** Generalist VLAs (Open X-Embodiment)
- **2025–2026** — **Recurrent SLMs & Compact VLAs**

## Where are LMs used?

H

R

**Planning & Decomposition**
- SayCan
- LLM+P
- TAMP
- Voyager

**Intent & Anticipation**
- GG-LLM
- HOI4ABOT
- LIT

**Grounding & Communication**
- Inner monologue
- VoxPoser
- CoPA

**Shared Autonomy & HRI**
- VoicePilot
- Assistron
- ChatGPT-Rob

**Reward & Distillation**
- Eureka
- ADAPT
- TraKDis

## Trained on what?

H

R

**Manipulation**
Large-scale cross-embodiment corpora for generalist manipulation policies.
Open X-Embodiment

**Conditioned Task**
Procedural task suites with language annotations for controlled benchmarking.
CALVIN · LIBERO · ManiSkill2 · RoboCasa · BEHAVIOR-1K

**Interaction-centric**
Human-in-the-loop, multi-agent scenarios for collaborative behaviour.
Habitat 3.0 · TEACh · Watch-And-Help · PARTNR

## How's the coverage?

H

R

**7 of 15 cells** are knowledge gaps: no signal at all.

| | Manipulation datasets | Conditioned task suites | Interaction-centric suites |
|---|---|---|---|
| Physical grounding & spatial precision | HC | HC | MC |
| Real-time latency & edge inference | LC | LC | KG |
| Adversarial safety & hallucination defense | LC | KG | KG |
| Continual adaptation & memory persistence | KG | LC | MC |
| Standardized certification frameworks | KG | KG | KG |

HC – High Coverage · MC – Medium Coverage · LC – Low Coverage · KG – Knowledge Gap

## What's not getting through?

H

R

- **Physical grounding** — lost in translation
- **Real-time latency** — message timed out
- **Hallucinations in safety-critical contexts** — unverified command
- **Catastrophic forgetting** — memory overwritten
- **Evaluation & verification gaps** — safety check failed

## What's next?

H

- **Neuro-symbolic HRC** with runtime assurance
- **Calibrated uncertainty** as the foundation for intent alignment and help-seeking
- **Parameter-efficient continual personalization** for edge deployment
- **Interaction-centric world models** for anticipatory collaboration
- **Evaluation science** for LM-mediated Human-Robot Collaboration

## So, in short?

H

This review provides a structured summary of the **current state**, **limitations** and **possible future directions** of safe, robust, and collaborative robotic systems in human-centric environments.

✓✓ Delivered

## Thanks! See ya!

H



Robot is typing…

Ask us a follow-up…

SustAIn Liv Work

ENS Tampere University