
Learning in Agentic Cognitive Robots: LLM-Supported Lexicon and Ontology Acquisition

Kwanlada Srijomkwan
Sergei Nirenburg

SRIJOK@RPI.EDU
NIRENS@RPI.EDU

Department of Cognitive Science, Rensselaer Polytechnic Institute, Troy, NY 12180 USA

Overview

Trustworthy knowledge-based AI needs extensive amounts of explicit knowledge. Manual acquisition of such knowledge is often considered too expensive to be practical at scale. We present a learning approach that uses large language models (LLMs) to drastically lower the cost of knowledge acquisition for knowledge-based AI. The pipeline outsources cognitively costly steps to LLMs while reserving the final validation of new knowledge for human experts. Our empirical results show that this learning approach produces quality knowledge at practical scales, and suggest a trend of quality improvement over time.

1. The LEIA OntoAgentic Environment

LEIAs (Language-Endowed Intelligent Agents) are agentic systems built on the OntoAgent cognitive architecture (McShane, Nirenburg, & English, 2024). They understand language, reason, and act using explicit knowledge bases, which allows them to explain their reasoning—a crucial component of trustworthy AI. This work targets two of these knowledge bases. The *ontology* is the agent’s world model, a hierarchy of concepts defined by their properties. The *lexicon* is a collection of word senses that link the syntactic structure (syn-struc) of input text to ontological concepts in their semantic structure (sem-struc). Unlike LLMs, LEIAs learn explicitly: each new piece of knowledge is a readable entry that can be validated and corrected, and our system automates part of this process.

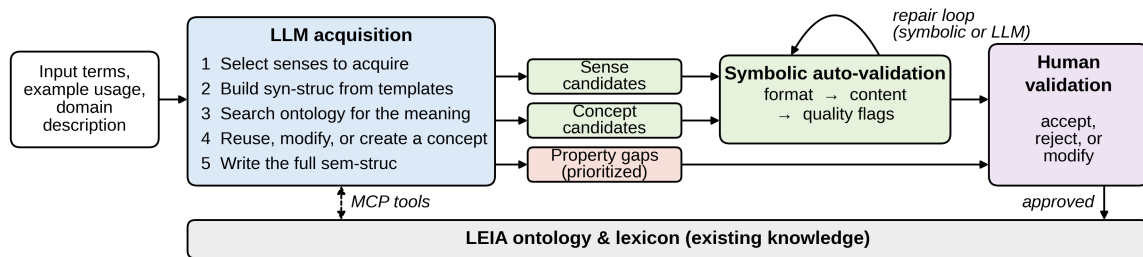


Figure 1. The acquisition and validation pipeline. The input is a list of terms with example usage and a brief domain description. Outputs are proposed senses and concepts, which go through symbolic auto-validation with an LLM-based repair loop, and a prioritized list of properties needing human acquisition.

2. LLM-Supported Acquisition and Validation

The acquisition pipeline (Fig. 1) is a series of LLM prompts, run as headless Claude Code calls, that extract knowledge latent in LLM weights into symbolic form. The LLM accesses the existing ontology and lexicon through a custom model context protocol (MCP) toolset and, to discourage hallucination, is instructed to use only concepts and properties that already exist there. All entries, even those passing automatic validation, then go to human validation, where knowledge engineers accept, reject, or modify them.

3. Experiments and Results

Table 1. Results of both experiments. All LLM calls used Claude Opus 5. Pass rates are from the first round of automatic validation; Exp. 2 also shows failures remaining after each repair round.

Exp. 1: welding & robotics (968 terms)	Senses	Concepts
Proposed	935	819
Passed auto-validation	740 (79.1%)	698 (85.2%)
Passed after adding 97 LLM-proposed properties	912 (97.5%)	803 (98.0%)
Exp. 2: shipboard engineering (identical input)	Run 1	Run 2
Senses / concepts proposed	317 / 139	344 / 327
Failing validation, repair rounds 1 → 2 → 3	19 (4.2%) → 3 → 0	19 (2.8%) → 0 → 0

Experiment 1 acquired ~1,000 welding and robotics terms extracted from introductory textbooks (Bowditch, Bowditch, & Bowditch, 2017; Craig, 2005). Of 968 filtered terms, 38 were flagged as unacquirable because the ontology cannot yet express their meanings—a desirable behavior, since the LLM should flag these rather than hallucinate. Nearly all of the 116 concepts needing human review were blocked by the same ~97 missing properties, which the LLM had proposed on its own following LEIA property templates rather than hallucinating them. If these properties are validated, pass rates rise to about 98% (Table 1), so human review becomes a much smaller property acquisition task.

Experiment 2 acquired ~350 shipboard engineering terms (Prabhu, 2021) twice, on top of the same base lexicon and ontology, to observe variation due to LLM stochasticity. The repair loop fixed all failures within three rounds in both runs. The large difference in concepts is because the first run tends to reuse existing concepts with modifications, whereas the second tends to create new ones. This is not necessarily harmful, since the world can be described in multiple equally-correct ways: “brother” could be an existing concept SIBLING who is male, or a new concept BROTHER.

4. Conclusion and Future Work

The LLM-supported pipeline converts knowledge hidden in LLM weights into explicit lexical and ontological knowledge at practical scale. Future work will reduce the cognitive load of human validation, quantify costs in human time and computation, and develop a quality evaluation method that admits the nuances of ontological work.

Acknowledgements

We are grateful to Dr. Marjorie McShane, Dr. Sanjay Oruganti, and Dr. Jesse English for their mentorship and guidance, and to Sophia March for her help with software implementation.

References

- Bowditch, W. A., Bowditch, K. E., & Bowditch, M. A. (2017). *Welding fundamentals* (5th ed.). Tinley Park, IL: Goodheart-Willcox.
- Craig, J. J. (2005). *Introduction to robotics: Mechanics and control* (3rd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- McShane, M. J., Nirenburg, S., & English, J. (2024). *Agents in the long game of AI: Computational cognitive modeling for trustworthy, hybrid AI*. Cambridge, MA: MIT Press.
- Prabhu, T. L. (2021). *Marine engineering: An introduction to marine engineering*. Nestfame Creations.