

KnowTD Interface: An Interactive Neuro-Symbolic Workflow for Thermodynamics Problem Solving

Luisa Vollmer^{1,*}, Rébecca Loubet², Hans Hasse² and Heike Leitte¹

¹Visual Information Analysis Research Group, RPTU University Kaiserslautern-Landau, Germany

²Laboratory of Engineering Thermodynamics, RPTU University Kaiserslautern-Landau, Germany

Abstract

Large language models (LLMs) provide intuitive natural-language interfaces, but their outputs are unreliable when scientific problem solving depends on formal constraints, equation applicability, and numerical reasoning. We present a Streamlit-based interface for KnowTD, an ontology-grounded thermodynamics reasoning system, that combines natural-language interaction with symbolic validation and solving. Users can formulate thermodynamics problems in natural language, inspect and edit the LLM-generated YAML representation, validate it against the LinkML-based KnowTD ontology, and solve it using symbolic thermodynamic reasoning. In addition to text-based interaction, the interface supports structured form input, YAML file upload, interactive ontology exploration, curated benchmark examples, and scenario-based exercise generation.

A preliminary evaluation on 13 textbook-style thermodynamics problems and six LLM models shows that the best model correctly solved 12 of 13 problems through the complete pipeline. Remaining errors mainly concern sign conventions, variable selection, and symbol normalization, highlighting the importance of editable intermediate representations and schema-based validation. The demo illustrates a human-in-the-loop neuro-symbolic workflow in which LLMs lower the interaction barrier, while validation and thermodynamic reasoning remain explicit and ontology-grounded.

Keywords

Neuro-symbolic AI, Scientific problem solving, Thermodynamics education, Exercise generation, Knowledge graphs, Large language models, Explainable AI

1. Introduction

Large language models (LLMs) make it increasingly easy to formulate scientific tasks in natural language. This makes them a promising interface technology for educational exercise authoring and problem solving. However, for scientific and educational applications, fluent text is not sufficient: generated problem statements and exercise solutions must obey domain laws, respect units and variables, expose assumptions, and support inspection by humans. This is particularly important in thermodynamics, where small changes in process attributes determine which equations are applicable.

KnowTD is an ontology-grounded symbolic reasoning system for thermodynamics [1]. It represents thermodynamic concepts, variables, equations, attributes, and rules in an ontology and uses knowledge graphs and symbolic reasoning to compute explainable stepwise solutions [2]. While KnowTD provides transparent and scientifically grounded problem solving, its original workflow requires users to formulate problems in a structured representation compatible with the ontology.

In this demo, we present a web-based interface that combines LLM-based natural-language parsing with ontology-grounded validation and symbolic reasoning, as shown in Figure 1. An editable YAML representation connects the generative and symbolic components.

The central design principle is a separation of responsibilities between generative and symbolic components. LLMs are used exclusively for linguistic tasks, namely translating between natural-

2nd International Workshop on Symbolic and Generative AI for Science (SymGenAI4Sci), co-located with SEMANTiCS 2026: 22nd International Conference on Semantic Systems, September 15, 2026, Ghent, Belgium

*Corresponding author.

✉ lvollmer@rptu.de (L. Vollmer); rebecca.loubet@rptu.de (R. Loubet); hans.hasse@rptu.de (H. Hasse); leitte@rptu.de (H. Leitte)

ORCID 0009-0007-6698-9578 (L. Vollmer); 0009-0003-3633-0749 (R. Loubet); 0000-0003-4612-5995 (H. Hasse); 0000-0002-7112-2190 (H. Leitte)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

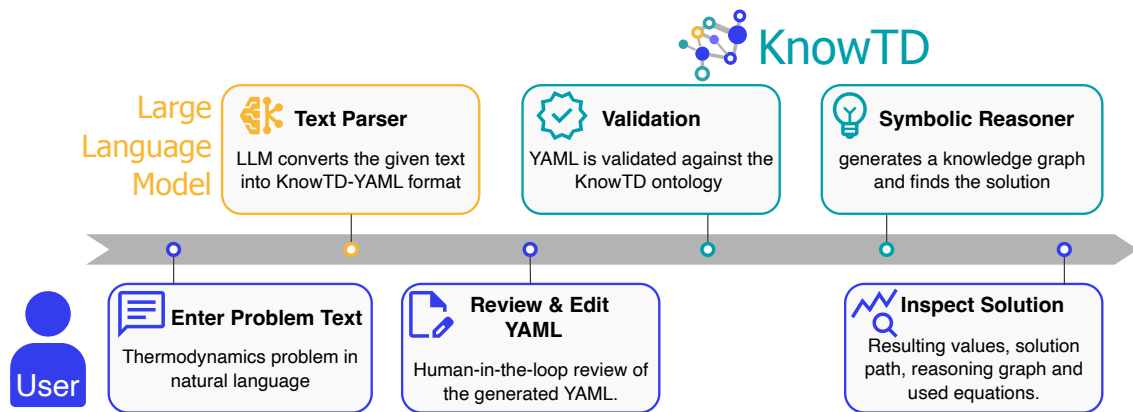


Figure 1: Overview of the KnowTD interface. The neuro-symbolic pipeline translates text into an inspectable YAML problem representation that is validated and solved by the symbolic KnowTD reasoner.

language and structured representations, while validation, equation selection, and thermodynamic reasoning remain fully symbolic and ontology-grounded. This human-in-the-loop workflow combines the accessibility of conversational interfaces with the transparency and reliability required for scientific problem solving.

The contributions are: (1) a unified interface with multiple input modalities for KnowTD; (2) an editable intermediate representation supporting human inspection and correction; and (3) integrated ontology exploration and exercise generation.

2. Background

KnowTD is a symbolic reasoning system for introductory engineering thermodynamics [1]. Its ontology represents thermodynamic concepts, variables, equations, process attributes, and their dependencies. The reasoner uses this knowledge to instantiate problem-specific knowledge graphs, identify applicable equations, and compute transparent stepwise solutions with explicit reasoning traces [2].

KnowTD currently focuses on textbook-style problems involving homogeneous closed ideal-gas systems. It covers changes of state, the first and second laws of thermodynamics, material properties, and process-specific constraints. However, its original workflow requires problems to be provided in an ontology-compatible YAML representation.

Recent work combines formal reasoning and LLMs to generate educational thermodynamics exercises [3]. Symbolic methods enumerate valid scenarios, determine solution paths, and assign consistent numerical values using KnowTD, while LLMs verbalize the structured specifications as textbook-style exercises.

The present interface supports the reverse direction by translating natural-language problem statements into editable KnowTD-compatible YAML specifications. These are validated against the ontology-derived schema and solved symbolically, thereby linking informal problem descriptions with formally grounded reasoning.

3. System Overview

The demo is implemented as a Streamlit-based interface¹ around KnowTD and provides natural-language access to ontology-grounded thermodynamic reasoning.

Figure 1 illustrates the overall workflow. Users can formulate thermodynamics problems in plain language, as structured forms, or as YAML files. Independent of the input modality, all workflows produce

¹Code available online: GitHub <https://github.com/LuisaVo/KnowTD-SymGenAI-Interface>
Live Demo: Streamlit Webapp <https://knowtd-interface.streamlit.app/>

a structured YAML representation that is validated against the KnowTD ontology and subsequently solved by the symbolic reasoner. The interface is organized into three functional areas: problem solving, exercise exploration and generation, and ontology exploration. The landing page provides an overview of the complete neuro-symbolic workflow, from problem formulation to symbolic reasoning, and offers direct access to the main functionalities.

The primary entry point is the **Text Input** view, which follows the five-step workflow shown in Figure 1. First, users enter a thermodynamics problem in natural language and select an LLM backend for parsing. Second, the selected LLM translates the problem into a KnowTD-compatible YAML specification. The current interface supports only single-step problems, although KnowTD also models sequential and cyclic problems. Therefore, the exercise type is currently fixed to the single-step problem class rather than inferred automatically. The corresponding YAML template is retrieved by querying the ontology for the concepts, attributes, and variables required to represent this problem class, and the resulting template is incorporated into the LLM prompt. Third, the generated YAML representation is displayed in an editable editor, allowing users to inspect and correct extracted variables, values, symbols, or assumptions before proceeding. This human-in-the-loop step enables users to identify and resolve common LLM errors before solving. Fourth, the edited representation is then validated against the LinkML-based KnowTD schema. When validation fails, the LinkML error message is provided to the user for manual corrections. Finally, if validation succeeds, the YAML specification is passed to the symbolic KnowTD reasoner, which computes the requested quantities and returns solution values, selected equations, a solution path, and a reasoning graph.

Besides natural-language interaction, the interface supports two additional solving modes. The **Form Input** view provides a guided authoring environment in which users manually specify the problem class, process attributes, given quantities, and required variables. The **File Input** view enables users to upload existing YAML problem-specifications and solve them directly through the same validation and reasoning pipeline. Since all input modes ultimately produce the same structured representation, users can switch between interaction styles while relying on a common symbolic backend.

To facilitate exploration and reproducibility, the demo includes a **Try an Example** page containing the 13 benchmark exercises used in the evaluation. Users can browse these curated examples and directly load them into the text-input interface. They are also included in the form input. In addition, the **Generate Exercises** page integrates the scenario-based exercise generation framework [3]. Users can inspect valid thermodynamic scenarios, generate new textbook-style exercises, and review the corresponding symbolic solutions computed by KnowTD.

Finally, the **Explore Ontology** view exposes the underlying knowledge model. Users can browse thermodynamic concepts, variables, equations, and their relationships, making the reasoning process transparent rather than hidden behind the solver. Search and filtering facilities support inspection of the ontology and help users understand which concepts and equations are involved in solving a particular problem.

4. Preliminary Evaluation

We evaluated the natural-language workflow on the 13 benchmark exercises from the original KnowTD publication. The same exercises were previously used to assess the ability of LLMs to solve thermodynamics problems directly. While older models such as GPT-4 [4] did not answer these problems reliably, the reasoning model o3 [5] solved all examples as well as more challenging exam problems. These results motivated us to investigate whether such models are also suitable for translating natural-language exercises into KnowTD-compatible representations. The evaluation focuses on the interface layer, specifically whether the generated YAML representation can be processed by KnowTD to produce the expected reference solution.

For each problem, an ontology-derived YAML template was included in the prompt. The generated YAML was validated against the LinkML-based KnowTD schema and subsequently solved using the symbolic KnowTD reasoner. Correctness was assessed by comparing the resulting solution with expert-

annotated reference solutions. Each model was prompted once per problem without prompt variation or few-shot examples. The results therefore indicate preliminary feasibility rather than a statistically robust model ranking.

`gpt-5.5-2026-04-23` [6] achieved the best performance, correctly solving 12 of 13 benchmark problems. The three evaluated Mistral models `mistral-large-2512` [7], `mistral-medium-3-5` [8], and `mistral-small-2603` [9] solved 10, 9 and 8 of 13 problems respectively, while `gpt-5.4-2026-03-05` [10] and `gpt-5.4-mini-2026-03-17` [11] solved 6 and 4 problems, respectively. The errors were primarily caused by incorrect sign conventions, inappropriate selection of required variables, and symbol normalization errors (e.g., confusing c_v and c_{vm}). The output of the LLMs as well as a detailed analysis can be found in the code repository. For each of the failed cases, the generated YAML was manually corrected by adjusting the affected signs, required variables, or variable references. After these small corrections, the solver obtained the corresponding reference solution in all cases. The interface therefore allows users to identify and correct such errors in the YAML view and highlights common pitfalls during the *Review & Edit YAML* step.

The results demonstrate that LLMs can effectively lower the input barrier to symbolic reasoning, while schema validation and user inspection remain essential components of reliable scientific workflows.

5. Conclusion

We presented a human-in-the-loop interface that enables natural-language access to KnowTD, a symbolic reasoning framework for thermodynamics. Unlike previous KnowTD workflows requiring manually authored structured specifications, the interface integrates text-based, form-based, and file-based problem entry with ontology exploration and exercise generation within a unified environment.

A central feature of the demo is the editable intermediate YAML representation that separates LLM-based language interpretation from symbolic validation and thermodynamic reasoning. This human-in-the-loop design preserves transparency and allows users to inspect, validate, and correct extracted problem representations before solving.

In future work, the class of the provided problem could be selected by the LLM in a preceding classification step by selecting from the ontology-defined classes and their descriptions. The system would then retrieve the corresponding YAML template and use it for translation. Future work will extend the benchmark and investigate automatic repair strategies based on validation feedback. A user study will evaluate whether students and instructors find correcting generated YAML easier than entering equivalent specifications through the form or file interface.

Acknowledgments

We gratefully acknowledge the funding by Deutsche Forschungsgemeinschaft DFG in the context of the Priority Program 2331 "Machine Learning in Chemical Engineering" and Research Unit 5359 "KI-FOR: Deep Learning on Sparse Chemical Process Data".

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT-5.5 in order to: Grammar and spelling check, Improve writing style. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

- [1] L. Vollmer, S. Fellenz, F. Jirasek, H. Leitte, H. Hasse, KnowTD—an actionable knowledge representation system for thermodynamics, *Journal of Chemical Information and Modeling* 64 (2024) 5878–5887. doi:10.1021/acs.jcim.4c00647.
- [2] L. Vollmer, R. Loubet, F. Jirasek, S. Fellenz, H. Hasse, H. Leitte, Enabling transparent problem solving in thermodynamics with ontologies and knowledge graphs, in: *Joint Proceedings of the ESWC 2025 Workshops and Tutorials*, volume 3977 of *CEUR Workshop Proceedings*, CEUR-WS.org, Aachen, 2025, pp. 1–12. URL: <https://ceur-ws.org/Vol-3977/XAIKG-3.pdf>.
- [3] L. Vollmer, R. Loubet, D. Neider, H. Hasse, H. Leitte, Combining formal reasoning and llms for scenario-based educational stem exercises, *Human-Centric Intelligent Systems* (2026). doi:10.1007/s44230-026-00174-1.
- [4] OpenAI, GPT-4 Model, <https://developers.openai.com/api/docs/models/gpt-4>, 2023. Accessed: 2026-07-31.
- [5] OpenAI, GPT-o3 Model, <https://developers.openai.com/api/docs/models/o3>, 2025. Accessed: 2026-07-31.
- [6] OpenAI, GPT-5.5 Model, <https://developers.openai.com/api/docs/models/gpt-5.5>, 2026. Accessed: 2026-07-31.
- [7] Mistral AI, Mistral Large 3, <https://docs.mistral.ai/models/model-cards/mistral-large-3-25-12>, 2025. Accessed: 2026-07-31.
- [8] Mistral AI, Mistral Medium 3.5, <https://docs.mistral.ai/models/model-cards/mistral-medium-3-5-26-04>, 2026. Accessed: 2026-07-31.
- [9] Mistral AI, Mistral Small 4, <https://docs.mistral.ai/models/model-cards/mistral-small-4-0-26-03>, 2026. Accessed: 2026-07-31.
- [10] OpenAI, GPT-5.4 Model, <https://developers.openai.com/api/docs/models/gpt-5.4>, 2026. Accessed: 2026-07-31.
- [11] OpenAI, GPT-5.4 mini Model, <https://developers.openai.com/api/docs/models/gpt-5.4-mini>, 2026. Accessed: 2026-07-31.