

New directions in integrating Large Language Models for knowledge elicitation

Vlad Luca Sebastian Spataru | v.l.s.spataru-1@student.tudelft.nl | EEMCS Faculty of Delft University of Technology

Responsible Professor: Ujwal Gadiraju | Supervisor: Shreyan Biswas

June 28, 2024

1 Introduction

Motivation

- AI Race potentially leads to entities cutting corners for competitive advantages [1]. Higher priority than following research & development best practices. This can lead to knowledge gaps.
- Knowledge elicitation is a domain within the field of Large Language Model (LLM) applications.
- Examples already exist, but where to seek next?

Research Question

- What new directions can be explored in integrating Large Language Models for knowledge elicitation?

Sub-questions

- Subquestion to clarify domain of search:
 - What directions has the research community and industry identified that have yet to be thoroughly researched or productized?
- Subquestions to guide analysis:
 - What are the possible strengths and benefits of exploring and integrating the identified directions?
 - What possible concerns exist with the integration of the identified directions?

2 Background

Key concepts

- Knowledge Elicitation
- Large Language Models (LLMs)
- Observation: knowledge elicitation and knowledge extraction used often interchangeably in literature of LLMs

Usecases

- Some usecases already exist in products available:
 - Chatbots (Chat GPT), Sentiment Analysis tools, Code suggestion tools (GitHub Copilot), Games with a purpose (GWAPs), AI-generative art etc.

3 Methodology

Literature Review

- PRISMA - screening process visualizer [2]

Tools

- Journal Databases (Google Scholar, Arxiv, Science Direct)
- Search Engines (Google, Bing)

Queries

- First query included 14 parameters and 13 logic operators. Output of 13,000 entries, but unsupported on some databases.

Final Query: ("knowledge elicitation" OR "knowledge extraction") AND ("Large Language Models" OR "LLMs") AND ("integration" OR "applications") AND ("unexplored" OR "new direction")

4 Findings

Screening Process

- Total entries 1,756 results
- Inclusion: publication date newer than 2023
- After all exclusions 17 papers included for in-depth review.

New Directions

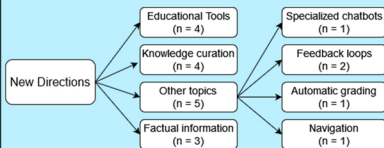


Figure 2: New directions overview

5 Analysis

DIRECTION	BENEFIT
Educational tools	Automate teaching process: speed, user-friendliness, logical functionality, more accessible, always available
Knowledge curation	Help researchers in processing large quantity of data to enable diverse fields of research
Factual information	Improve reliability of LLMs sharing facts and sources
Specialized chatbots	Assist domains that are burdened by staff shortages (ex. nursing)
Feedback loops	Improve capability, reliability, and effectiveness of LLMs based on user feedback.
Automatic grading	Less effort spent on tedious grading of open-answer questions
Navigation	Multimodal processing, advancements for autonomous driving

Table 1: Benefits analysis synthetization

- Common concerns identified on all topics: reliability (hallucinations), safety, security, legality, privacy, transparency
- Concerns on educational tools also state a preference for human teaching

6 Limitations

Limitations & Potential Bias

- Non-exhaustive search query - more variations can be identified
- Limited time frame
- Selection bias - non-exhaustive tools selection
- Publication bias - potentially mostly literature with positive results was available
- Language bias - only papers in English have been selected

7 Conclusion

- Performed a systematic literature review on new directions in integrating Large Language Models for knowledge elicitation.
- Identified 17 insightful research papers and synthesized them into 7 directions (educational tools, knowledge curation, factual information, specialized chatbots, feedback loops, automatic grading, and navigation) for future research
 - Highlighted key benefits for each direction of researching further and integrating them into tools.
 - Raised awareness on the main concerns identified in literature for all directions
 - Highlighted the main limitations of the research and encouraged future research into the directions identified and other other systematic analyses in other domains.

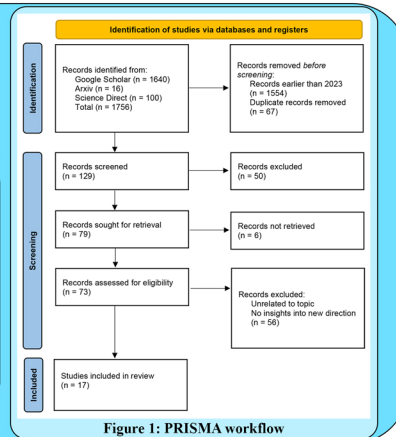


Figure 1: PRISMA workflow

[1] Cave, Stephen, and Seán S. ÓhÉigeartaigh. "An AI race for strategic advantage: rhetoric and risks." Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. 2018.

[2] Moher, David, et al. "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement." International journal of surgery 8.5 (2010): 336-341.