Development of an AI-powered command-line agent to support methodological guidance selection in health economics analysis

Perera C[1], Hirst A[1], Heron L[1]
[1]Adelphi Values PROVE™, Bollington, Cheshire SK10 5JB, United Kingdom

ADELPHI VALUES | PROVE Expertise in Access and Value Evidence Outcomes

MSR71

Introduction

- > The widespread adoption of large language models (LLMs) across industries, including health economics and outcomes research (HEOR), has transformed how analysts access and utilize methodological guidance for evidence generation.[1]
- > One critical challenge that could benefit from artificial intelligence (AI) integration is the identification of appropriate methodological frameworks.
- > Traditional approaches to finding relevant guidance require analysts to manually navigate extensive documentation repositories, such as NICE Technical Support Documents (TSDs) which is time-consuming and may result in overlooking relevant resources.[2]
- > While TSDs provide gold-standard methodological guidance across diverse HEOR activities, identifying the most appropriate document for a specific project requires substantial domain expertise.
- > AI agents, which combine LLMs with specialized tools and structured workflows, offer potential to streamline this process by understanding natural language queries and matching them to relevant methodological guidance.
- > The aim of this research was to develop an AI-powered command-line agent that can recommend appropriate methodological guidance (i.e. NICE TSDs) based on user project descriptions, and to evaluate its accuracy, efficiency, and usability in supporting HEOR workflows.

Methods

AI Agent

- > An agent is a software program that can autonomously interact with its environment, reason, and take actions to achieve goals with minimal human intervention.[3]
- > An agent was developed using the LangGraph software development kit in Python 3.11 and powered by OpenAI o3-mini model.[4-5]
- > A zero-shot prompt was developed to enable the model to generate relevant responses without requiring specific task-related examples.
- > This approach allows the model to perform the task based on general understanding, leveraging the external data enhancing its ability to adapt to new or unseen data efficiently.
- > The agent architecture incorporated six specialised tools designed for interaction with the documents (**Table 1**).

Table 1. Tools used in the agent

Tool	Function
tsd_title	Retrieve all TSD titles
tsd_document	Extract URLs and metadata
download_document	Download TSD to local directory
identify_relevant_tsd	Match query to appropriate TSDs
read_pdf_content	Extract text from PDF
provide_tsd_guidance	Analyse content and provide guidance

Abbreviations: PDF, portable document format; TSD, technical support document; URL, uniform resource locator

Agent Tools and Functionality

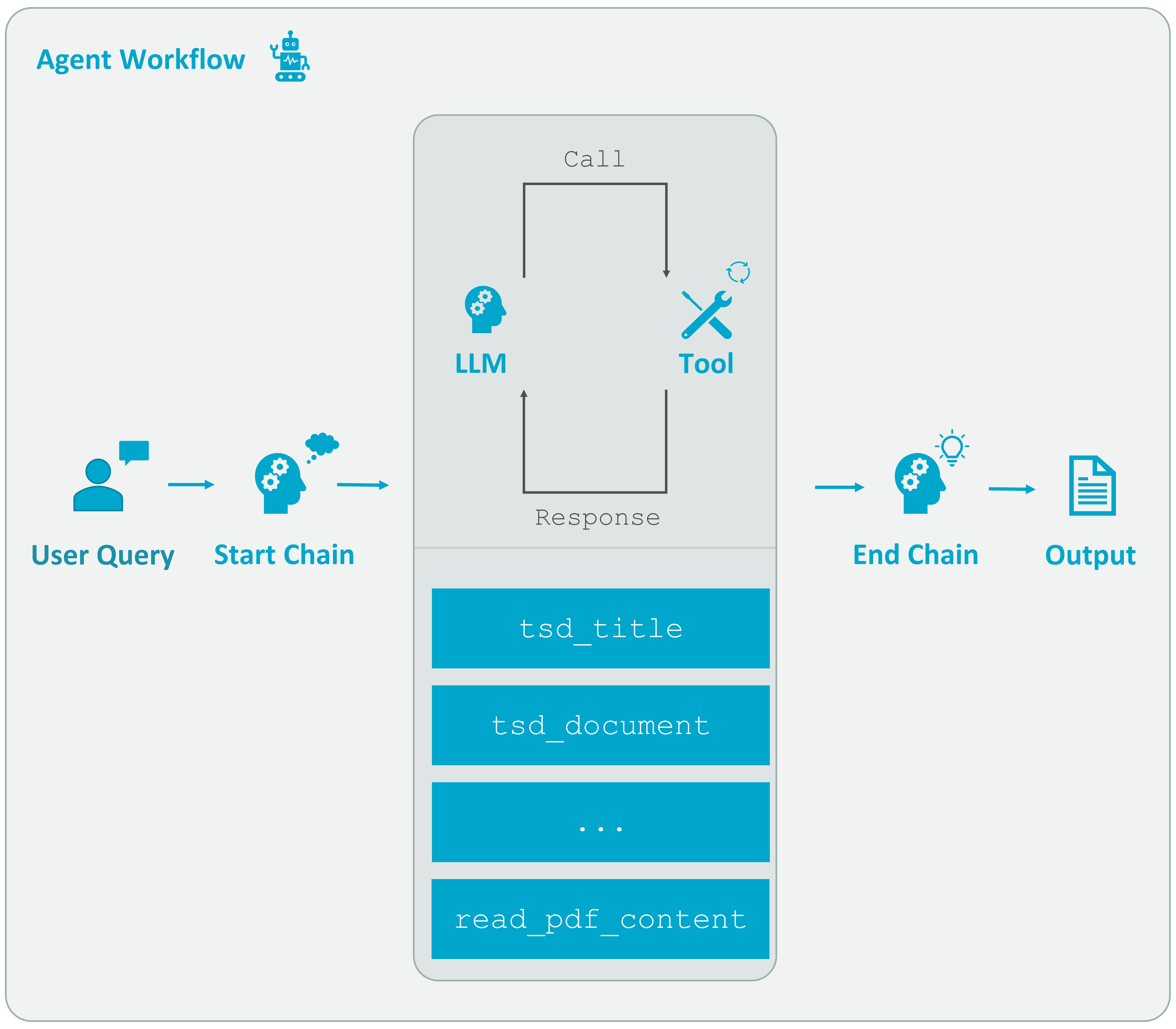
- > The agent workflow followed a structured process where the users describe their project in natural language, the agent processes the query through keyword mapping and contextual understanding (**Figure 1**).
- > The workflow was aligned to a human workflow to ensure quality of the output was maintained.
- > The agent could also download documents and provide specific methodological guidance based on user input.
- > The agent was implemented as a command-line interface with real-time web scraping of the NICE DSU TSD files and state management through LangGraph's StateGraph.[5]
- > The core logic utilized a conditional decision loop where the LLM determines whether to invoke tools or conclude the interaction based on user needs.[6]

Evaluation Framework

- > To assess the accuracy of the agent generated content, outputs were evaluated across various metrics:
  - – **Accuracy:** The accuracy was tested by comparing agent-generated TSD recommendations against expert human selection across diverse HEOR activities.
  - – **Efficiency:** Measured in time savings relative to manual methodological guidance documentation search.
  - – **Usability:** Qualitative user testing with HEOR analysts assessed intuitiveness and workflow integration.
  - – **Validity:** Manual validation to confirm agent comprehension and appropriateness of the recommendations provided.

Results

Figure 1. Overview of agentic framework



Precision

- > The AI agent demonstrated high accuracy in TSD recommendations that was comparable to human expert selection across diverse HEOR activities.
- > Manual validation by domain experts confirmed the agent's comprehension ability, with recommendations consistently aligned with expected methodological guidance and no major errors or inappropriate suggestions identified.

Efficiency

- > The efficiency analysis revealed significant time savings compared to traditional manual search approaches.
- > The agent provided instant access to TSD metadata and download capability, with automated extraction and synthesis od relevant content eliminating the need to manually navigate the repository.
- > Time to identify appropriate guidance was reduced significantly (95%-time savings compared to manual approach).

Usability

- > The agent successfully provided contextual explanations of TSD content and demonstrated understanding of how specific methods should be applied to user project.
- > User testing yielded positive qualitative feedback, with the CLI deemed intuitive to use, however complicated for non-technical users.
- > Users suggested that development of a graphical chat-bot interface could enhance accessibility for less technical analysts

Key functionalities demonstrated:

- ✓ Accurate interpretation of user queries.
- ✓ Context-aware recommendations based on project descriptions.
- ✓ Methodological guidance based on natural language processing and content analysis.

Conclusions

- > The findings from this study showcase a potential use case in which AI agents have been leveraged for the purpose of identification and synthesis of appropriate methodological guidance for HEOR.
- > The agent achieved accuracy comparable to human expert selection while delivering substantial efficiency gains.
- > The agent enables analysts to dedicate more capacity to analytical and interpretive aspects of their work while maintaining methodological rigor.
- > While NICE TSDs served as the pilot use case, this approach demonstrates broader potential for multi-agent systems applicable to other guidance repositories, including ISPOR methodological frameworks, health technology assessment body guidelines, and organization-specific documentation.
- > Future research and development should focus on expanding specialized tools for enhanced AI-agent capabilities and implementing human-in-the-loop approaches to ensure expert judgment remains central while maximizing efficiency

References

1. Reason T, Rawlinson W, Langham J, Gimblett A, Malcolm B, Klijn S. Artificial Intelligence to Automate Health Economic Modelling: A Case Study to Evaluate the Potential Application of Large Language Models. PharmacoEconomics – Open. 2024/03/01 2024;8(2):191-203; **2.** NICE Decision Support Unit. Full list of technical support documents (TSDs). Available at: https://www.sheffield.ac.uk/nice-dsu/tsds/full-list [Accessed October 2025]; **3.** Gutowska, Anna. "What Are AI Agents? | IBM." Www.ibm.com, 3 July 2024, www.ibm.com/think/topics/ai-agents. [Accessed October 2025]; **4.** OpenAI. O3-mini | openai. Available at: https://openai.com/index/openai-o3-mini/ [Accessed: 20 October 2025]; **5.** J. Wang and Z. Duan, "Agent ai with langgraph: A modular framework for enhancing machine translation using large language models," arXiv preprint arXiv:2412.03801, 2024; **6.** Sapkota R, Shrestha R, Rijal M, Karkee M. LangChain vs. LangGraph vs. LangSmith: Taxonomies of Agentic AI Toolchains for End-to-End Orchestration. Authorea Preprints;