

## DEPARTMENT: AI IN INDUSTRY

# Large Language Models and Applications: The Rebirth of Enterprise Knowledge Management and the Rise of Prompt Libraries

Daniel E. O'Leary , University of Southern California, Los Angeles, CA, 90089, USA

*This article investigates how large language systems and the apps developed for them provide a platform for enterprise knowledge management. For those resulting systems to provide consistent and accurate responses for knowledge management, enterprises are using different approaches in their prompts, such as few-shot learning, specification of purpose, and chain-of-thought reasoning. As better and more successful prompts are being built, they are being captured and prompt libraries are being created.*

Increasingly, large language models (LLMs) are becoming platforms that are the basis of the development of applications and services (LLM apps) that can perform activities, such as code development, integrating mathematical reasoning, performing searches, and domain-specific activities. LLM apps allow the integration of tasks so that users can leverage the natural language capabilities of LLMs. In addition, LLM apps can help mitigate some LLM limitations to help guide users to information and knowledge to support their decision making. As a result, LLMs and their apps are becoming the basis for the rebirth of enterprise knowledge management (EKM).

To ensure consistent and correct responses, different approaches are being used to include context in users' prompts. For example, including few-shot learning, purpose, and chain-of-thought information in prompts has been shown to generate better results. Accordingly, enterprises are capturing better prompts and making those prompts available in prompt libraries.

This article examines 1) how LLMs and LLM apps can facilitate the rebirth of knowledge management, mitigating previous knowledge management limitations; 2) how different approaches to prompts are providing better and more consistent results; and 3) how

that is leading to enterprises creating prompt libraries. This article has three purposes. First, this article investigates the growing role of LLMs as a platform for applications that use their capabilities and provide additional functional capabilities to ultimately manage knowledge. Second, this article suggests that the joint LLM and LLM application capability provides a platform solution to the concerns that have been voiced, indicating the death of knowledge management. Third, we drill down into managing prompt libraries and the ability of prompts to capture context to generate more accurate and consistent responses.

## LLMs AND APPLICATIONS

LLMs, such as ChatGPT, BARD, Claude, and others, are based on neural network models of Internet quantities of text with billions of parameters. Although researchers have found that these systems have definite limitations, they still provide what Lu et al.<sup>1</sup> call "an early version of an artificial general intelligence." Because of the computational requirements for their development, these systems are only updated periodically. For example, recently, ChatGPT 3.5 indicated that it was last updated in January 2022. Of course, this means that these systems do not have contemporary information as part of their knowledge. In addition, these LLMs are based on publicly available information and thus do not include information from individual enterprise meetings, policies, procedures, and so on unless that

information is publicly available and captured. Further, as LLMs are text-based models, they generally do not process pictures effectively and have limited knowledge based on visual information. As a result, generally, they are not accurate at capturing information about or providing commentary on visual content. In addition, they have shown limited capabilities to analyze and understand data. However, LLMs have the ability to interact with people and provide human-sounding discussions and summaries of information and, increasingly, applications have been developed to mitigate these limitations. All of this suggests that LLMs are evolving into platforms.

## ENTERPRISE KNOWLEDGE MANAGEMENT

EKM generally refers to two basic management issues: content management and collaboration management (e.g., O’Leary<sup>2</sup> and others). With content management, organizations must decide what knowledge is of interest and then provide an approach to make that knowledge available to their users. For collaboration management, organizations must provide the ability of users to find, contact, and work with networks of others. Many approaches have been proposed and used in knowledge management, ranging from artificial intelligence (AI) to approaches aimed at having people manually contribute knowledge. Ontologies have been used to categorize and search knowledge. Intelligent agents have been used to gather information and connect people to others and to knowledge. However, these many different approaches have led to the use of several technologies and no unifying platform or approach. As a result, it has become costly to generate, index, and make available knowledge available for use. Thus, some have questioned whether knowledge management is still being used. Accordingly, researchers have raised questions such as, “is knowledge management dead or dying?” (O’Leary<sup>3</sup>), “whatever happened to knowledge management?” and “any chance this idea will come back?” (Davenport<sup>4</sup>). This article, argues that LLMs and their apps provide a platform for EKM, and that prompt libraries provide an important role in that knowledge management.

## IS KNOWLEDGE MANAGEMENT DEAD OR DYING? CAN LLMs AND LLM APPS HELP?

As part of the concerns as to whether knowledge management is dead, several rationales have been proposed, indicating why (Davenport<sup>4</sup>). In this section,

I discuss these rationales and how they are largely mitigated by LLMs and LLM app capabilities.

According to Davenport:<sup>4</sup>

“It was too hard to change behavior. Some employees weren’t that interested in acquiring knowledge, others weren’t interested in sharing what they knew. Knowledge is tied up in politics and ego and culture.”

LLMs can unobtrusively capture large quantities of enterprise knowledge. In addition to published text information, LLMs can be used to gather and summarize information from meetings, presentations, and other sources that normally would be difficult to capture. LLMs can also capture information and queries entered into the system as part of the use of those systems. For example, a user could enter data into a system and ask the system to analyze it, and that query and data could be captured for others to use. As a result, using an LLM and apps can work to minimize effort required of employees to share knowledge. In addition, any new knowledge can be readily integrated with other information.

“Everything devolved to technology. KM [knowledge management] is a complex idea, but most organizations just wanted to put in a system to manage knowledge, and that wasn’t enough to make knowledge flow and be applied.”

LLMs provide a browser-based platform that can be the basis of interacting with employees. For those organizations interested in a technology solution, LLMs and a growing number of LLM apps can provide platform-based approach that allows interrogation of the LLM. Further, besides being easy to use, LLMs and apps can be used to integrate with other types of systems, such as help desk or transaction systems.

“It was too time-consuming to search for and digest stored knowledge. Even in organizations where a lot of knowledge was contributed to KM systems ... there was often too much knowledge to sort through. Many people didn’t have the patience or time to find everything they needed. Ironically, the greater the amount of knowledge, the more difficult it was to find and use.”

With robust search and presentation capabilities, LLMs provide a platform that provides users with information in a “consumed” manner, rather than as raw information gathered from different websites. LLMs summarize large quantities of text. Further, the responses from LLMs are text explanations in response

to the query of interest. Some LLMs, such as BARD, also provided Internet links with the queries, or the LLM can also be asked to provide references so that users can gain access to additional supporting knowledge.

“Google also helped kill KM. When people saw how easy it was to search external knowledge, they were no longer interested in the more difficult process for searching out internal knowledge. ...”

Using LLMs and LLM apps may actually be easier than doing a Google search. Interviews that I have done with those in practice, suggest that some keep their chatbot open and use it rather than Google to look for knowledge and find answers to their questions. In addition, in some chatbots, such as Google’s BARD, Internet links to articles and websites are sometimes provided as a default, or they can be requested. Further, as noted by O’Leary,<sup>5</sup> there are emerging approaches for enterprises to use LLMs to capture and use internal information with LLMs.

“KM never incorporated knowledge derived from data and analytics. ...”

Although LLMs can provide guidance and information about how to do data and analytic analysis, LLMs are not known for their ability to use data and analytics. However, LLM apps have been and are being developed to facilitate mathematical and statistical analyses. As an example, Schick et al.<sup>6</sup> developed “Toolformer,” which allows LLMs to teach themselves to use different tools from examples and sets of instructions. The tools include a calculator, translation capabilities, a calendar, and other capabilities. As another example, Lu et al.<sup>1</sup> examine the use of an app to improve the ability of LLMs to improve performance on mathematical reasoning. As still another example, LLMs can be used to summarize information about knowledge from data and analytics projects. Accordingly, LLMs and LLM apps can provide responses about knowledge derived from data.

## KNOWLEDGE IN CONTEXT FOR LLMs

Researchers have found that the context information in the queries to an LLM can make a difference in its response accuracy and reproducibility. Thus, if enterprises are concerned with using LLMs and apps for knowledge management, there is a need for users to be able to generate similar and accurate results. Research has found several approaches to capturing

that context by providing different types of information in the queries.

## EXAMPLES—FEW-SHOT LEARNING

Examples of correct answers, so-called “few-shot” reasoning, can help an LLM determine which kinds of answers that a user is concerned with and potentially improve performance (Brown et al.<sup>7</sup>). For example, I told ChatGPT to “Translate French to English. As an example, Je means I. Translate Bonjour.” ChatGPT indicated, “Certainly! ‘Bonjour’ in French translates to ‘Hello’ in English.”

## PURPOSE

In an analysis of *Jeopardy!* questions by ChatGPT, O’Leary<sup>8</sup> found that by letting ChatGPT know that the purpose was to play a game of *Jeopardy!*, the results generated by the system were statistically significantly better than when no game context was provided. In fact, at one point in the analysis, as questions were being fed to the system, ChatGPT simultaneously started playing *Jeopardy!* with the users and had to be told when to revert to a “long form” of answering queries.

## CHAIN-OF-THOUGHT REASONING

Wei et al.<sup>9</sup> found that an approach that provides chain-of-thought reasoning as part of the prompt information to the chatbot can improve system accuracy when solving mathematical word problems. Chain-of-thought reasoning could include information about how to solve a related problem as part of a query.

## CHAIN-OF-CODE REASONING

Li et al.<sup>10</sup> examined another related approach, where code writing is used to improve that chain-of-thought reasoning. The primary idea is to provide LLMs with different semantic subtasks, provided as pseudocode to specify behaviors and limit undefined behaviors.

## GOOD OLD-FASHIONED AI

The last approach that we will include here is embedding good old-fashioned AI (GOFAI) into queries. Classic knowledge engineering approaches generate substantial knowledge that may be captured as plans, case-based reasoning, and other approaches that can be embedded in LLM queries to guide an LLM to consistent and more accurate responses. By providing the LLM with process knowledge or additional context information, it is given greater specificity of the user’s

## AI IN INDUSTRY

**T**he new AI in Industry department aims to feature AI innovations in industry, case studies, issues and challenges in AI adoption, and more. If you would like to contribute a short article, contact Daniel O'Leary at [oleary@usc.edu](mailto:oleary@usc.edu).

interest and is more likely to answer the query correctly and consistently.

## KNOWLEDGE MANAGEMENT OF PROMPT CONTEXTS

Although each of these forms of capturing context can improve response accuracy and consistency, not all users would necessarily be able to formulate queries using these approaches or even be aware of them. As a result, an important part of LLM knowledge management is creating prompt libraries to provide users with prompt guidance, provide the right context, and guide the system to the right results with what might be called *best practice prompts*. In addition, prompt libraries can help make the process of interacting with a chatbot more efficient by allowing a reuse of the previously successful prompt.

Prompt libraries can be designed with examples (few-shot prompting) in the query to provide context information an indication as to whether a game is being played, the chain of thought or reasoning, or other approaches. One approach to managing these different approaches in prompt libraries is case-based reasoning, which helps users find the right or best query. Using this approach, various constructs can help users find the right prompt. Historically, chatbots, such as ChatGPT or BARD, index prompts by time and user. However, prompt knowledge management systems would also use factors such as subject, context approach used in the query, reason for the query, quality of the results, role of the person using the query, role of the person developing the query, and other factors.

## SUMMARY

We suggest using LLM as an EKM platform and using LLM apps to mitigate LLM limitations. Taken together, LLMs and LLM apps provide the opportunity to eliminate

many of the rationales for why researchers think that "knowledge management is dead." In addition, to ensure that users generate correct and consistent responses from LLMs, best practices can be used to manage user prompts. Emerging approaches such as providing examples that indicate the purpose and provide chain-of-thought and chain-of-code reasonings can be captured in prompt libraries to facilitate prompt reuse and prompt management.

## ACKNOWLEDGMENTS

The author would like to acknowledge the comments of San Murugesan on an earlier version of this article.

## REFERENCES

1. P. Lu et al., "Dynamic prompt learning via policy gradient for semi-structured mathematical reasoning," 2022, *arXiv:2209.14610*.
2. D. E. O'Leary, "Enterprise knowledge management," *Computer*, vol. 31, no. 3, pp. 54–61, Mar. 1998, doi: [10.1109/2.660190](https://doi.org/10.1109/2.660190).
3. D. E. O'Leary, "Is knowledge management dead (or dying)?" *J. Decis. Syst.*, vol. 25, no. sup1, pp. 512–526, 2016, doi: [10.1080/12460125.2016.1193930](https://doi.org/10.1080/12460125.2016.1193930).
4. T. Davenport, "Whatever happened to knowledge management?" *Wall Street J.*, Jan. 2015. [Online]. Available: <https://www.wsj.com/articles/BL-CIOB-7428>
5. D. E. O'Leary, "The rise and design of enterprise large language models," *IEEE Intell. Syst.*, vol. 39, no. 1, pp. 60–63, Jan./Feb. 2024, doi: [10.1109/MIS.2023.3345591](https://doi.org/10.1109/MIS.2023.3345591).
6. T. Schick et al., "Toolformer: Language models can teach themselves to use tools," 2023, *arXiv:2302.04761*.
7. T. Brown et al., "Language models are few-shot learners," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2020, vol. 33, pp. 1877–1901.
8. D. E. O'Leary, "An analysis of Watson vs. BARD vs. ChatGPT: The Jeopardy! Challenge," *AI Mag.*, vol. 44, no. 3, pp. 282–295, Fall 2023, doi: [10.1002/aaai.12118](https://doi.org/10.1002/aaai.12118).
9. J. Wei et al., "Chain-of-thought prompting elicits reasoning in large language models," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, vol. 35, pp. 24824–24837.
10. C. Li et al., "Chain of code: Reasoning with a language model-augmented code emulator," 2023, *arXiv:2312.04474*.

**DANIEL E. O'LEARY** is a professor at the Marshall School of Business, University of Southern California, Los Angeles, CA, 90089, USA. Contact him at [oleary@usc.edu](mailto:oleary@usc.edu).