

Argonne National Laboratory, a federal research center, relies on secure AI systems to support operations and scientific research. Its generative AI tool, Argo, provides contextual responses from internal documents that ensure data privacy. While Argo can retrieve Argonne-specific knowledge relevant to user queries, it sometimes generates incomplete or irrelevant responses. To address this, our team integrated AI agents to enhance response quality and reliability.

AI agents identify the information needed to answer a request through query interpretation and retrieve the necessary information from external expert sources to generate better responses. The team explored techniques for integrating agents into Argo's existing architecture. We designed a rudimentary agent capable of interpreting user queries, accessing internal databases, performing external Google searches, assessing quality with feedback loops, and synthesizing this information into digestible outputs.

In testing the agent's functionality, sample queries triggered distinct logic paths based on its reasoning and tools, enabling the system to answer previously unresolvable questions like "What day of the week is the final report due?" (Figure 2). This required both the document's specified date and internet information to determine the day of the week, demonstrating significant improvements in Argo's accuracy and adaptability.



Figure 1: Schema used to design MRKL agent rationale



Figure 2: Example MRKL Agent Chain responding to the query: "What day of the week was the SSR report due in the year 2024?