The state of artificial intelligence (AI) is ever advancing in its potential, becoming increasingly relied upon in applications from search engines to cloud services. At the same time, contemporary AI systems continue to face limitations in processing and contextual understanding. AI models can be subject to hallucinations, content that seems plausible but is in reality factually untrue or fantastical. Pre-trained neural large language models (LLMs) are trained on a dataset that becomes a working parameterized "memory." While models are able to learn a significant amount of knowledge through this method, and can subsequently recall information in responses without access to an external information source, the lack of flexibility and transparency of the parameterized memory can contribute to hallucinations. Retrieval augmented generation (RAG) is a hybridized AI technique that combines the standard parametric with a retrieval-based fact source that serves as a non-parametric memory. Here we discuss Poseidon-Bot, a model developed with the goal of increasing user oversight and control over the non-parametric memory of standard RAG. A typical RAG model utilizes publicly accessible large vector-formatted data sources for its non-parametric memory, which limits the scope of the memory [8]. Poseidon-Bot uses specific fact sources uploaded by the user into the working non-parametric memory, providing the model with retrievable data in a more secure format. The increased user oversight is intended to enhance its potential for usage in processing sensitive or non publicly accessible information. We examine the performance of Poseidon-Bot in comparison to a standard LLM to assess the response quality, accuracy, and efficiency of its modified RAG approach.