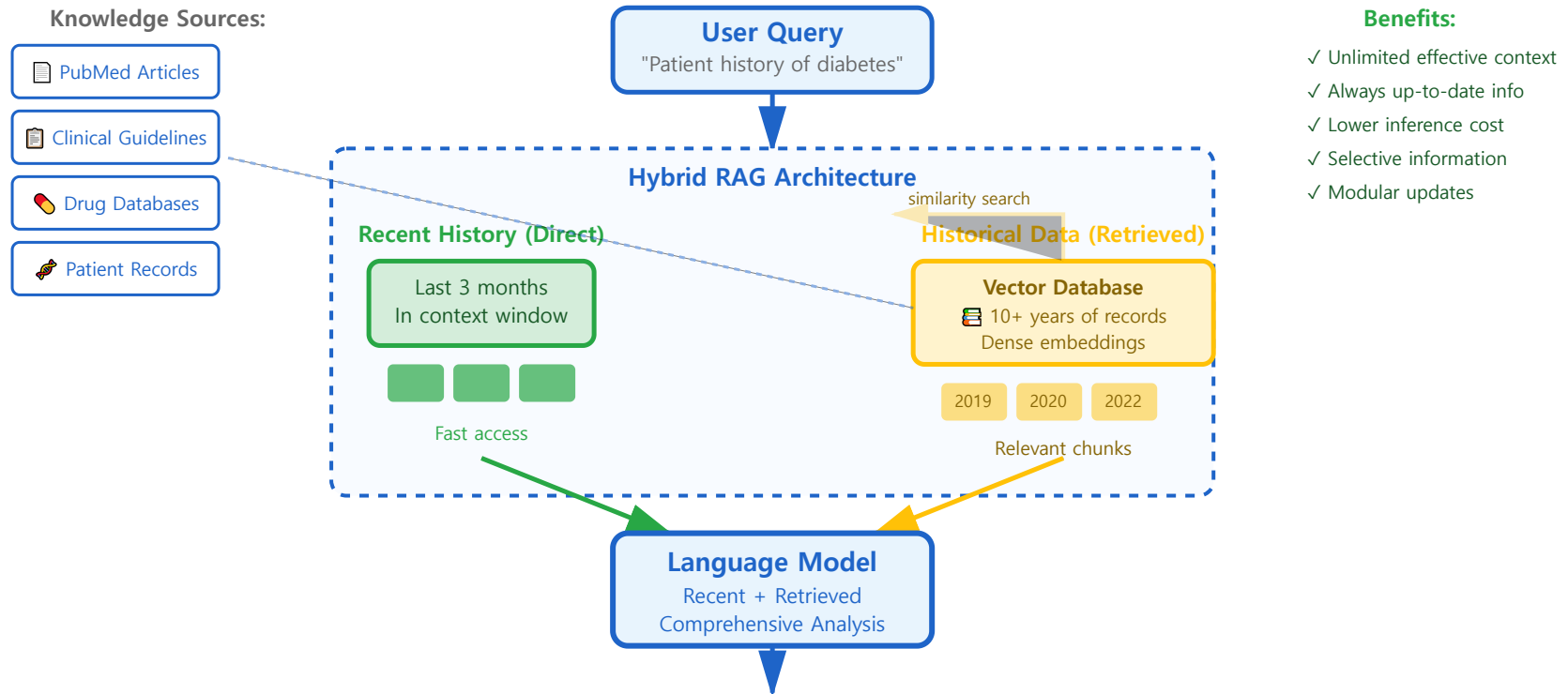


Retrieval-Augmented Context Extension



Best of Both Worlds

RAG combines long-context models with retrieval systems, enabling access to vast medical knowledge while maintaining efficient inference and real-time updates. **Recent data stays in context, historical data retrieved on-demand.**