**Known limitations: (excluded from brief to avoid bias)**

## Current Features (What Works Well)

The chatbot includes several implemented features:

✓ **Real-time status indicators** during query processing (Initializing → Classifying → Searching → Generating)

✓ **Streaming responses** (text appears as it's generated, not all at once)

✓ **Session persistence** across widget and full page views

✓ **Conversation transfer** between widget popup and full page interface

✓ **Citation system** with clickable links to source documents

✓ **PostHog analytics** tracking query performance and user interactions

## Known Limitations

* Basic interface lacking user guidance and expectation management
* Initial embedding service startup can take 15-30 seconds on first query of the day
* No onboarding or example queries shown to new users to explain chatbot capabilities
* Minimal formatting and visual hierarchy in response presentation
* Unclear link styling and interaction patterns
* Widget may be overlooked by users unfamiliar with floating chat buttons
* Session management requires manual copy/paste of session ID for sharing
* Citation links open in same tab (may disrupt conversation flow)
* No conversation export functionality yet
* Limited mobile optimization (desktop-first design)

**Performance Note:**

Typical response times range from 2-10 seconds depending on query complexity. The system provides real-time status updates during processing:

* 🔄 Initializing the request
* 🧠 Classifying query intent
* 🔍 Searching the repository
* ✨ Generating response

First-time queries may experience 15-30 second delays while the embedding service initializes. Timeout warnings appear after 15 seconds for unusually long operations.

**Technical Considerations**

System Architecture: The chatbot employs advanced retrieval techniques including:

* Field-aware hybrid search combining vector similarity and keyword matching
* Multi-model fallback for enhanced reliability
* Semantic intent classification to optimize response relevance
* Real-time citation consistency validation