

## 2026-27 Roadmap:

Tasks	Month 1	Month 2	Month 3
<b>RAG Database</b>	Implement the first version of the unified text parser to accept output from Data Cleaning/Scraping. Begin implementing metadata filtering for chunks.	Start Vector DB scaling research. Integrate basic sentence-window chunking.	Prototype a simplified solution for paid models using API keys. Complete research on Graph DB architecture integration.
<b>Data Scraping</b>	Implement the "smart router" that recognizes basic input types (link, text, CSV) and routes them. Finalize the first version of the consistent JSON output format.	Build and test an API-first connector for a new, specific source (e.g., Twitter/X). Fully integrate the smart router with API preference logic.	Integrate More Scraping (new, compliant sources). Collaborate with Data Cleaning to refine and stabilize the unified output format.
<b>Data Cleaning</b>	Establish Full integration with Scraping Output by accepting and processing the new consistent JSON format. Refine current scraping awareness of HTML classes/elements.	Develop higher-level solutions for content quality and consistency (e.g., new parsing logic). Investigate fine-tuning scraping with a small AI model.	Work with RAG to ensure the cleaned output integrates correctly for chunking. Initial design and implementation of Advanced Metadata Extraction (Attribution) logic.

<b>Agentic Systems</b>	Implement more MCP servers (moving beyond the 2-3 proof-of-concept ones). Finalize Basic Agent UI design.	Implement the core Basic Agent UI features (LLM and Tool Selection). Begin prototyping the custom tool creation flow.	Develop the language model-assisted flow for coding custom tools (e.g., "Make me a calculator..."). Start implementation of the Custom Tool Upload feature.
<b>Fine-tuning</b>	Finalize the UI design and implementation, including connecting pages to the software. Implement graphs below the log terminal in the UI for better understanding.	Implement Model Quantization (LORA/QLORA). Work on connecting the existing BERT/SQuAD encoder form to the live software interface.	Expand encoder options beyond BERT. Implement an easier way to enter the LMForge for non-technical users (e.g., business majors).