Tooling and Infrastructure: Software Ecosystem, Knowledge Graph Deployment, and Quality Assurance

The practical utility of a computational meta-analysis depends on robust tooling at each pipeline stage: assertion extraction, modeling and simulation, knowledge graph infrastructure, and quality assurance.

LLM-Based Assertion Extraction

Extracting structured assertions from unstructured text is the most labor-intensive component of knowledge graph construction. Manual annotation produces high-quality results but does not scale to corpora of thousands of papers—a constraint demonstrated by Knight et al. knight2022fep, whose systematic literature analysis of FEP and Active Inference publications required manual coding of structural, visual, and mathematical features for hundreds of annotated papers. We implement a hybrid approach: LLMs perform initial extraction, with human review for validation and correction.

Our extraction pipeline deploys a locally hosted LLM through Ollama ollama2024. Each paper's abstract is assessed against the