RAG_BOT architecture

Ser

Formatted Response

Core Application Logic

MessageHandler

Use

RAG Agent (LangGraph -
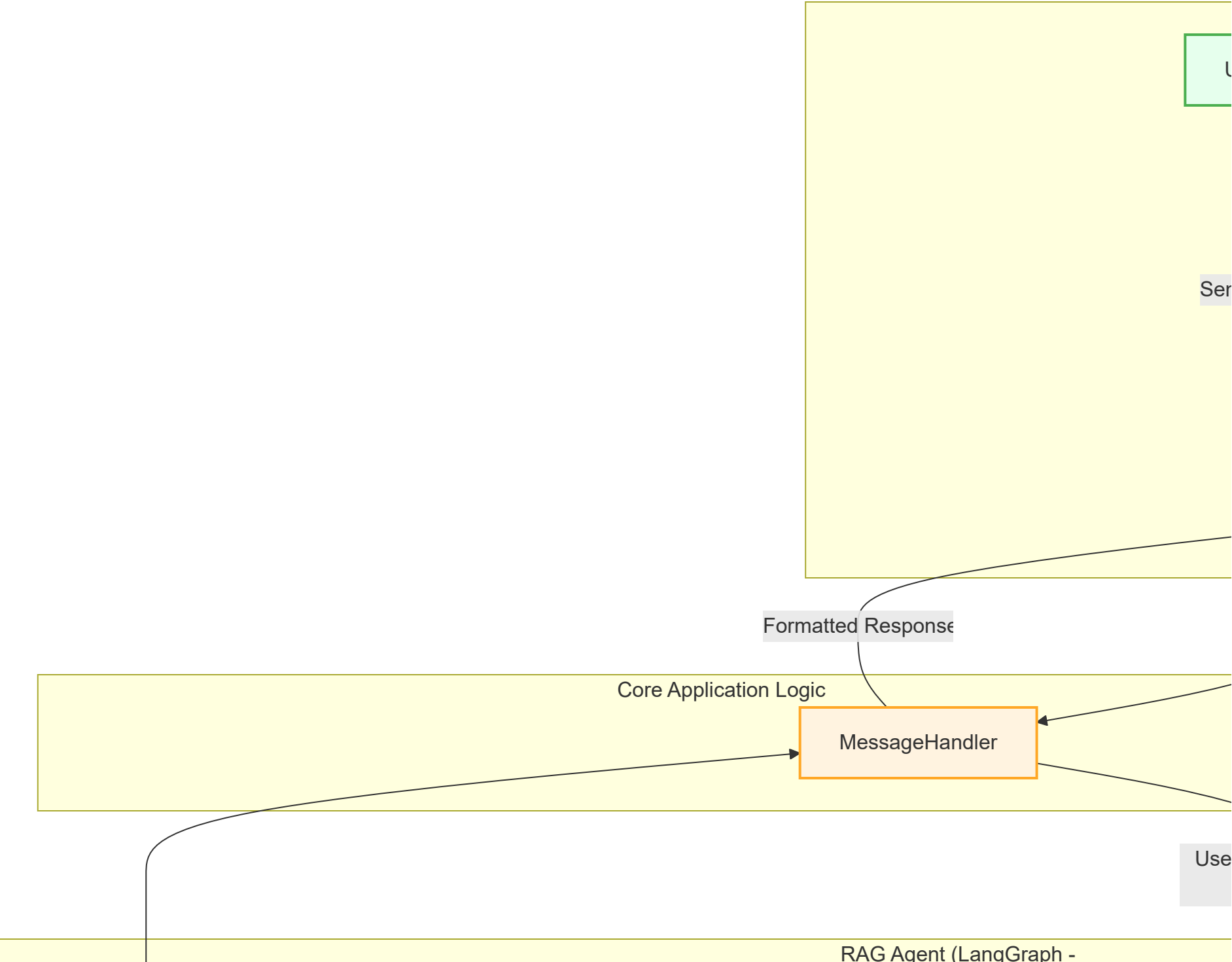
agent/)

Age
(g
De

Invokes

Agent Workflow Nodes

Agent Initial Node
(agent_node.py)\n-
Analyzes query\n- Decides
tool use or direct answer

JSON Respons

Retri
(Too
Cor

Retrieve

Rerank Cor
(retrieval_ne
Reranks
docur

Direct Answer Path

Reframed Query

Reranked Docs

Evaluate Context Node
(evaluation_nodes.py)\n-
Assesses context relevance

Evaluation: Sufficient

Evaluation: Insufficient

Agent Final Answer Node
(agent_node.py)\n-
Generates final JSON
response

Reframe Query Node
(evaluation_nodes.py)\n-
Reformulates query if
needed

External AI M
Servic

Google Gemini LLM
(Query Analysis,
Evaluation, Reframing,
Generation)