

Original training dataset

Remaining data



Removed data



Original training



Original Model



Naive
retraining



Unlearning

Similar performance

10 USE-CASES OF RAG

ACROSS DIFFERENT INDUSTRIES

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



Prediction

Aggregation

Training Dataset

\mathcal{D}



Sub-dataset \mathcal{D}_2



Sub-model M_2



<

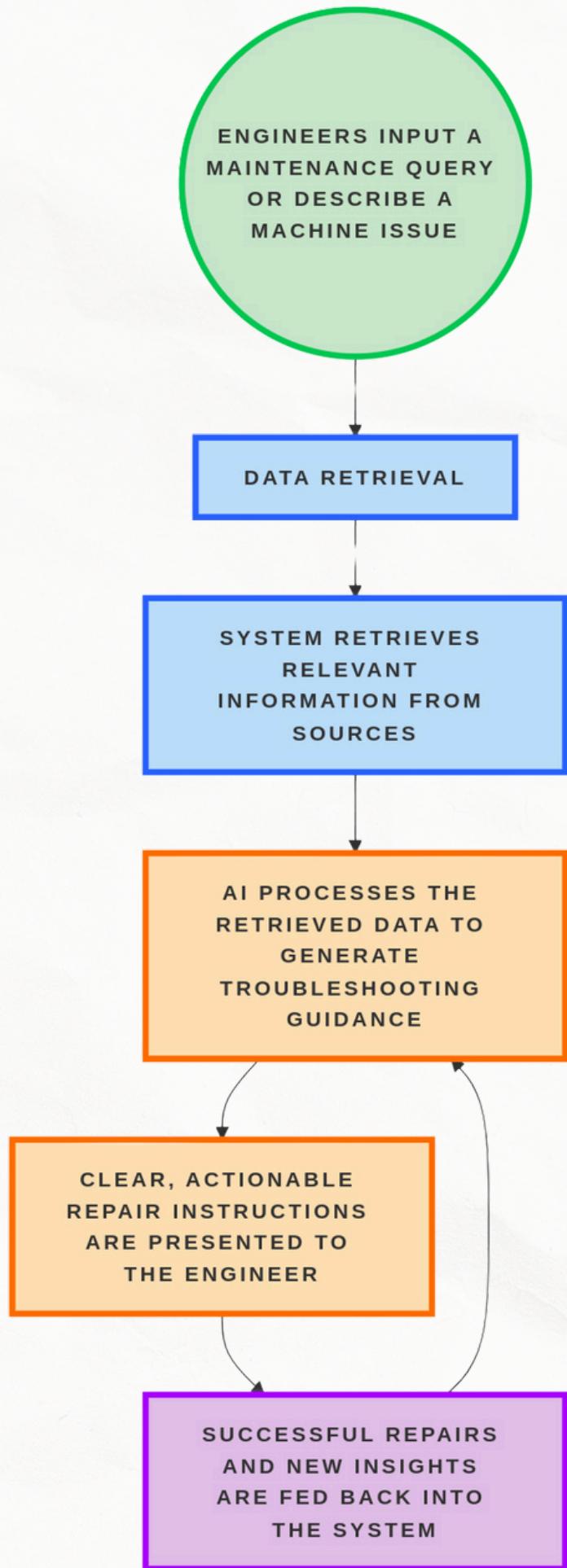
1. EQUIPMENT MAINTENANCE

Retrieval-Augmented Generation (RAG) transforms factory maintenance by acting as an intelligent assistant that pulls insights from machine manuals, sensor data, maintenance logs, and past repair histories.

Engineers can input queries, and the system retrieves relevant information to provide step-by-step troubleshooting guidance with diagrams and proven solutions.

Impact:

- Faster Repairs: Reduced Mean Time To Repair (MTTR)
- Cost Savings: Prevented unplanned downtime
- Efficiency: Lower reliance on senior technicians
- Knowledge Retention: Preserve expertise from the retiring workforce



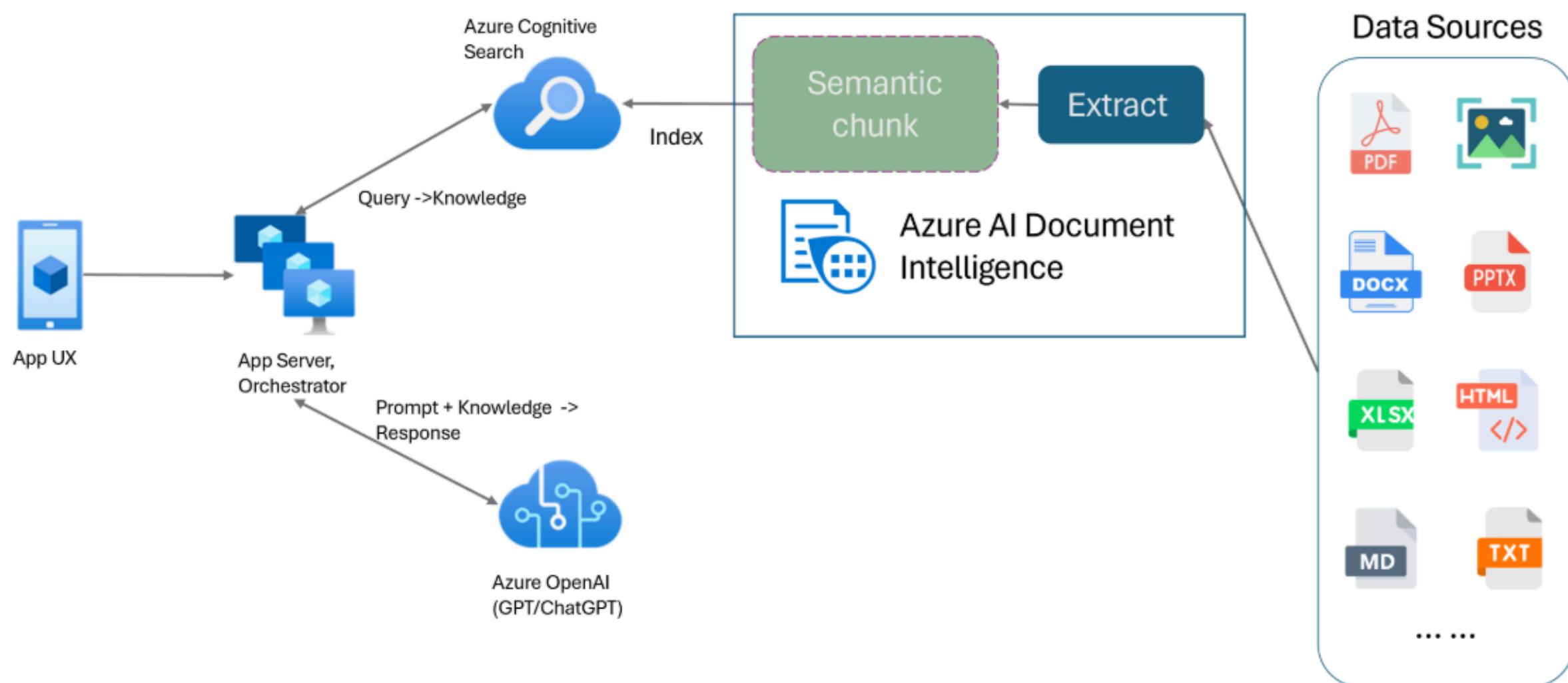
*This diagram is a speculative representation.

2. INTERNAL KNOWLEDGE MANAGEMENT

Retrieval-Augmented Generation (RAG) enhances internal knowledge management by centralizing company documents, policies, and reports.

This allows employees to quickly find the information they need, improving collaboration, decision-making, and overall productivity.

Microsoft's Azure AI Document Intelligence utilizes RAG to give employees seamless access to internal documentation, code repositories, and project details, ensuring efficient knowledge sharing and smoother workflows.

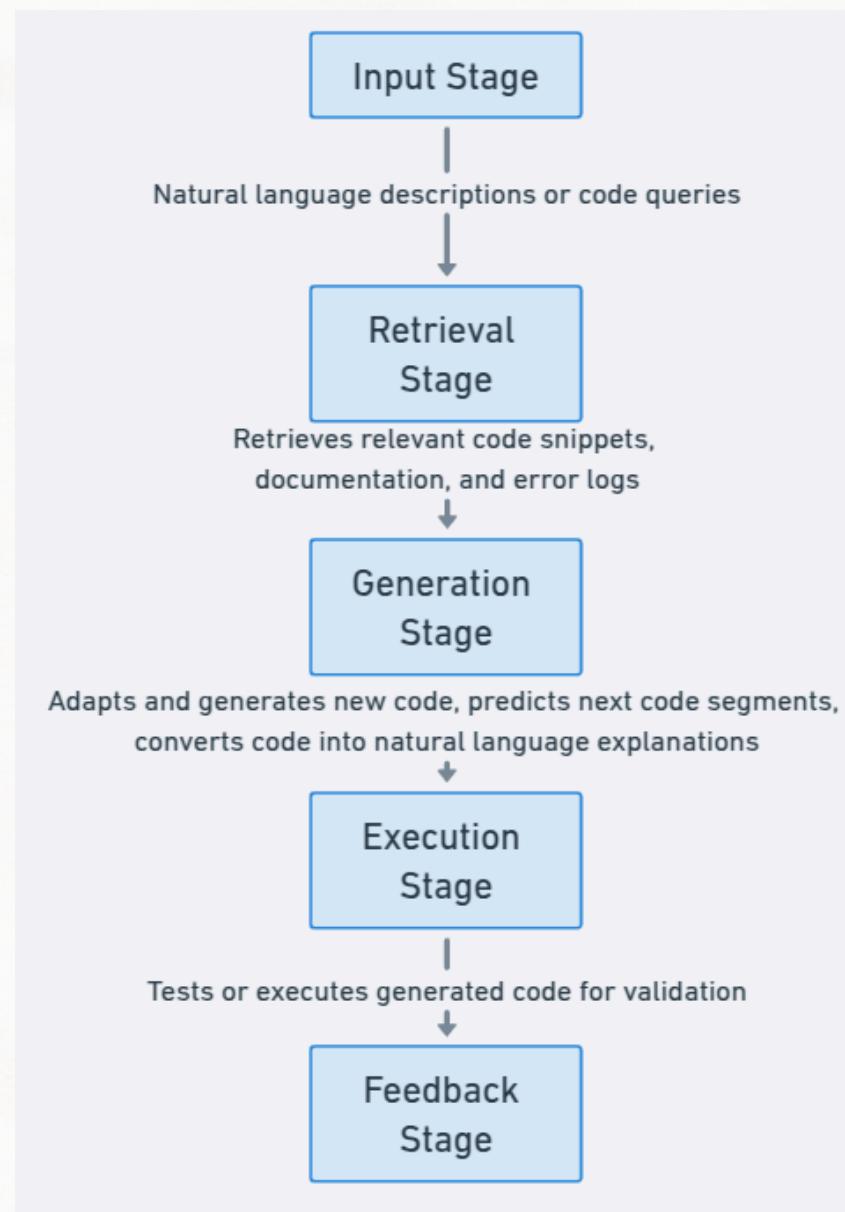


*This diagram is a speculative representation.

3. Code Generation

(RAG) enhances code generation by retrieving relevant code snippets and intelligently adapting them to meet project requirements.

It fetches information from code repositories, documentation, and error logs to generate accurate code, detailed documentation, and even identify and fix errors.



Key Capabilities of RAG in Code Generation:

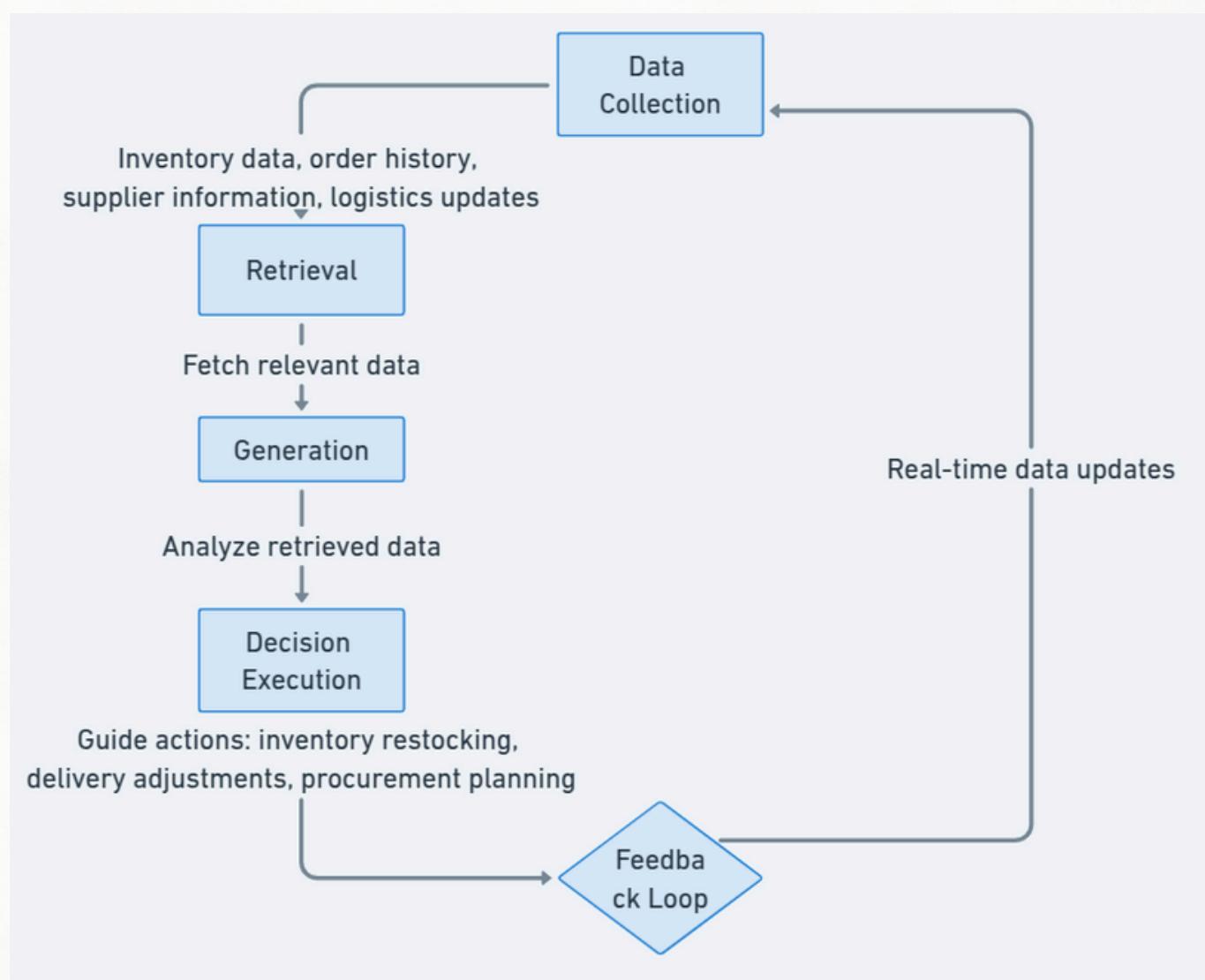
- Converts natural language descriptions into functional code.
- Predicts the next logical code segment.
- Translates code into human-readable documentation.
- Generates and executes new code for analysis and refinement.

*This diagram is a speculative representation.

4. Supply Chain Management

RAG enhances finance, accounting, and supply chain management by providing instant access to financial data, inventory levels, and vendor info. This empowers analysts, auditors, and CFOs to make data-driven decisions, optimize resources, and boost efficiency.

Users can quickly gain insights by asking meaningful questions, enabling a fast understanding of industry trends and smarter decisions.



Amazon uses RAG to manage its supply chain by analyzing real-time data from inventory, vendor contracts, and logistics. It predicts demand, optimizes stock levels, and reroutes shipments to avoid delays. This ensures accurate inventory tracking, efficient vendor coordination, and smooth logistics, reducing costs and boosting efficiency.

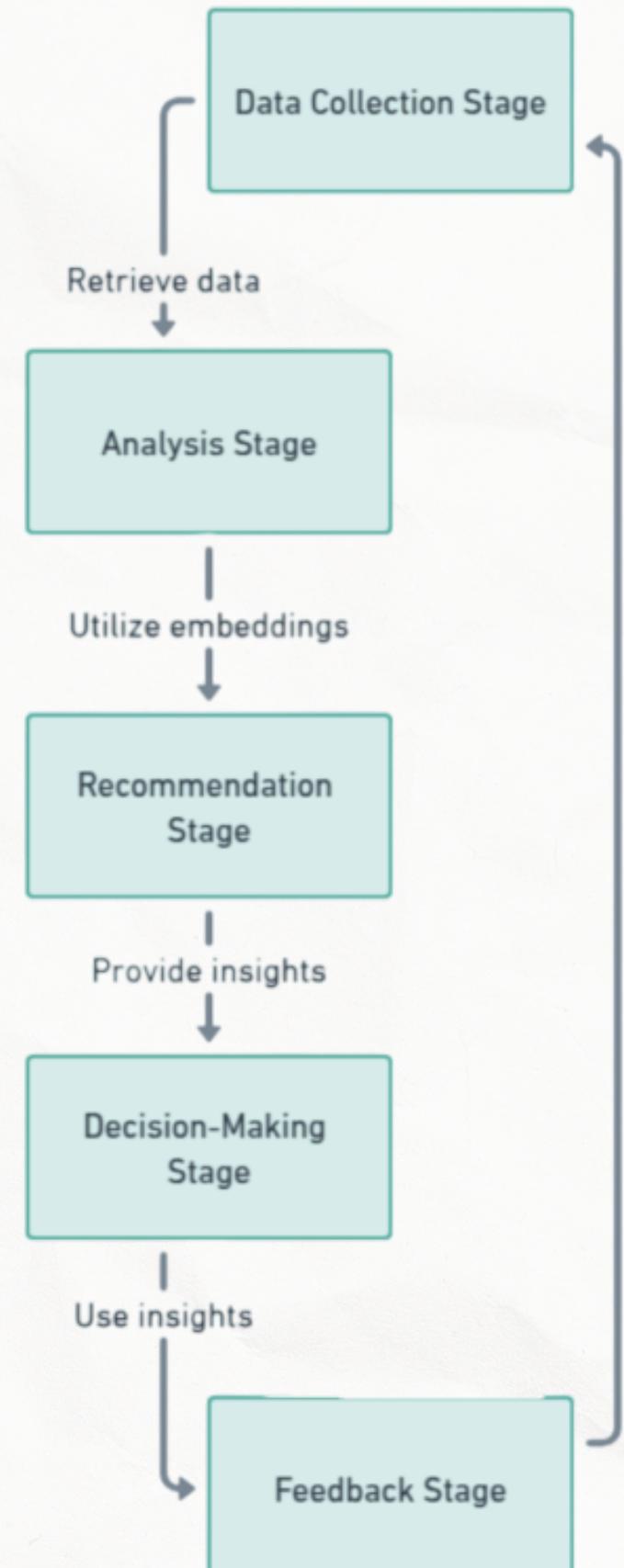
*This diagram is a speculative representation.

5. Risk Pattern Recognition

RAG retrieves data using hybrid semantic and structured retrievers and applies temporal-aware embeddings to analyze historical claims trends. It then provides risk assessment recommendations, highlights similar case examples, and suggests optimal policy terms and conditions.

Impact:

- Automated underwriting for routine tasks.
- Consistent risk assessment across evaluations.
- Reduced loss ratios through accurate risk identification.
- Enhanced policy customization for clients.



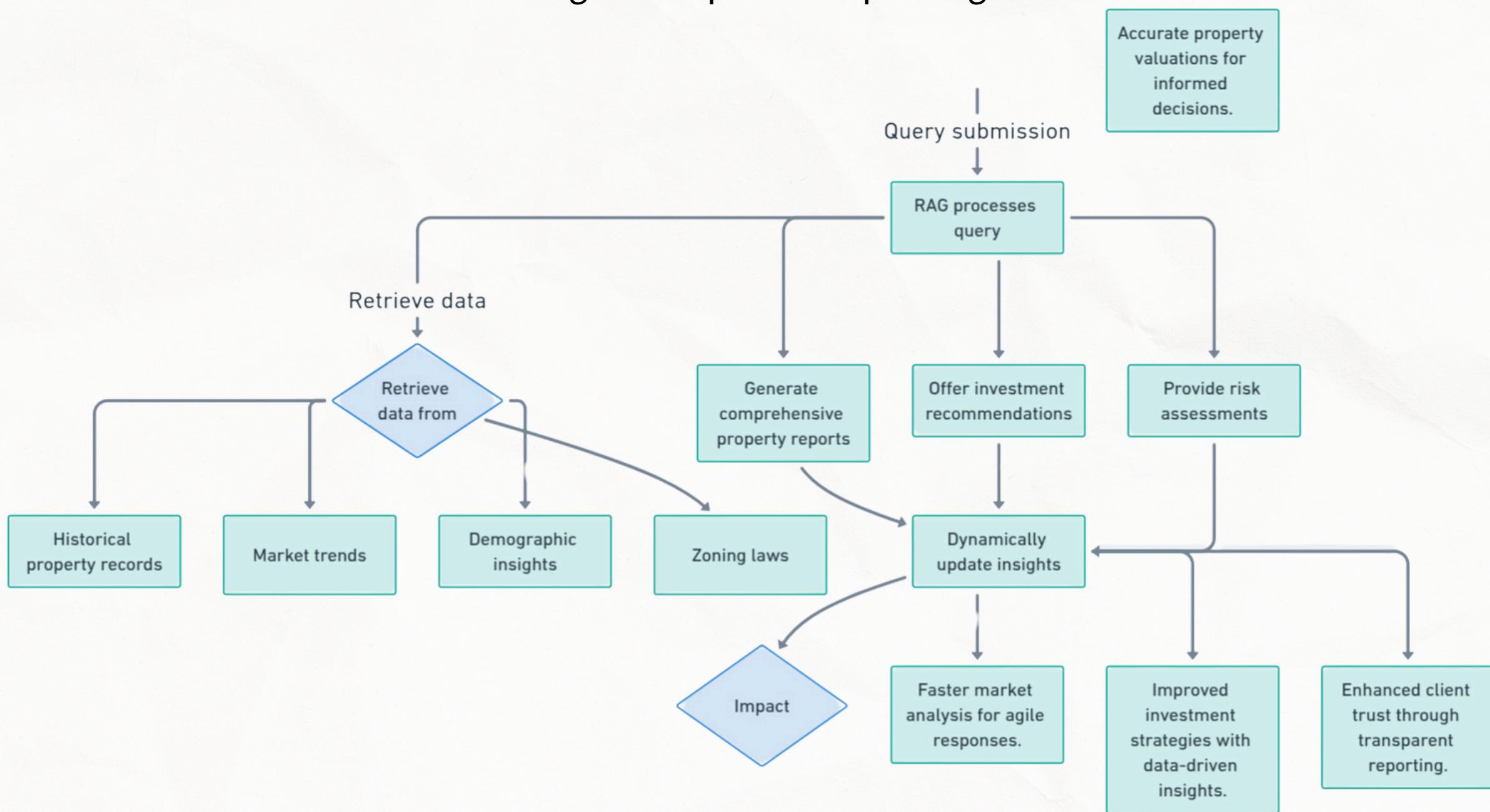
*This diagram is a speculative representation.

6. Technical Support

Agents or clients submit property queries, which RAG processes by retrieving data from historical property records, market trends, demographic insights, and zoning laws. It then generates comprehensive property reports, offers investment recommendations, and provides risk assessments, dynamically updating insights as market conditions change.

Impact:

- Accurate property valuations for informed decisions.
- Faster market analysis for agile responses.
- Improved investment strategies with data-driven insights.
- Enhanced client trust through transparent reporting



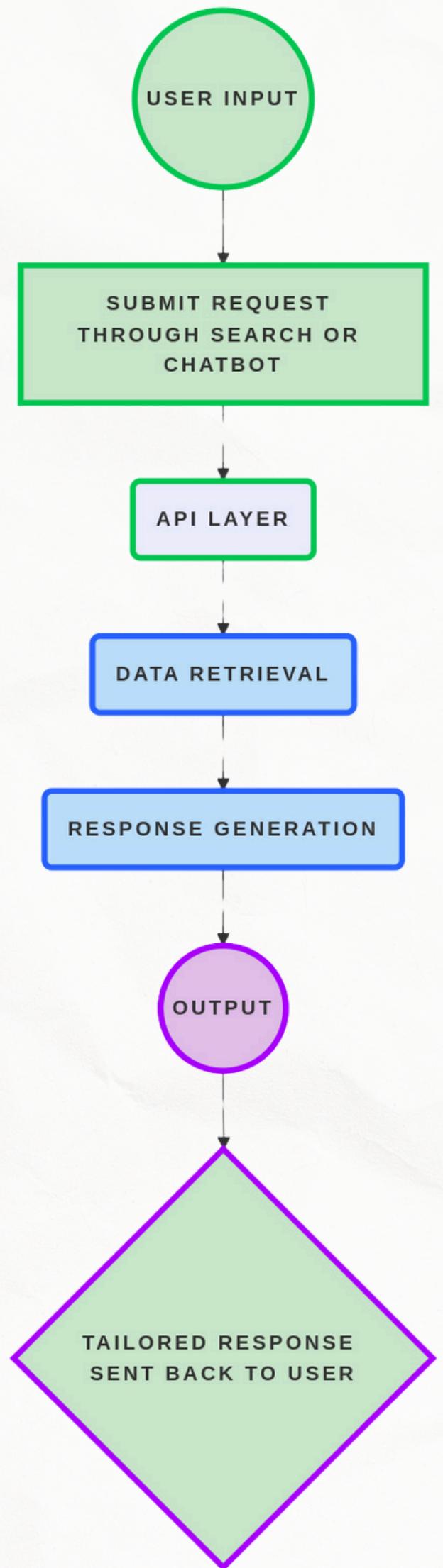
*This diagram is a speculative representation.

7. Virtual Assistant

RAG enhances virtual assistants by accessing up-to-date information on events, weather, and news, and generating natural language responses tailored to user queries. This approach ensures accurate, context-aware answers, significantly improving user experience.

Workflow:

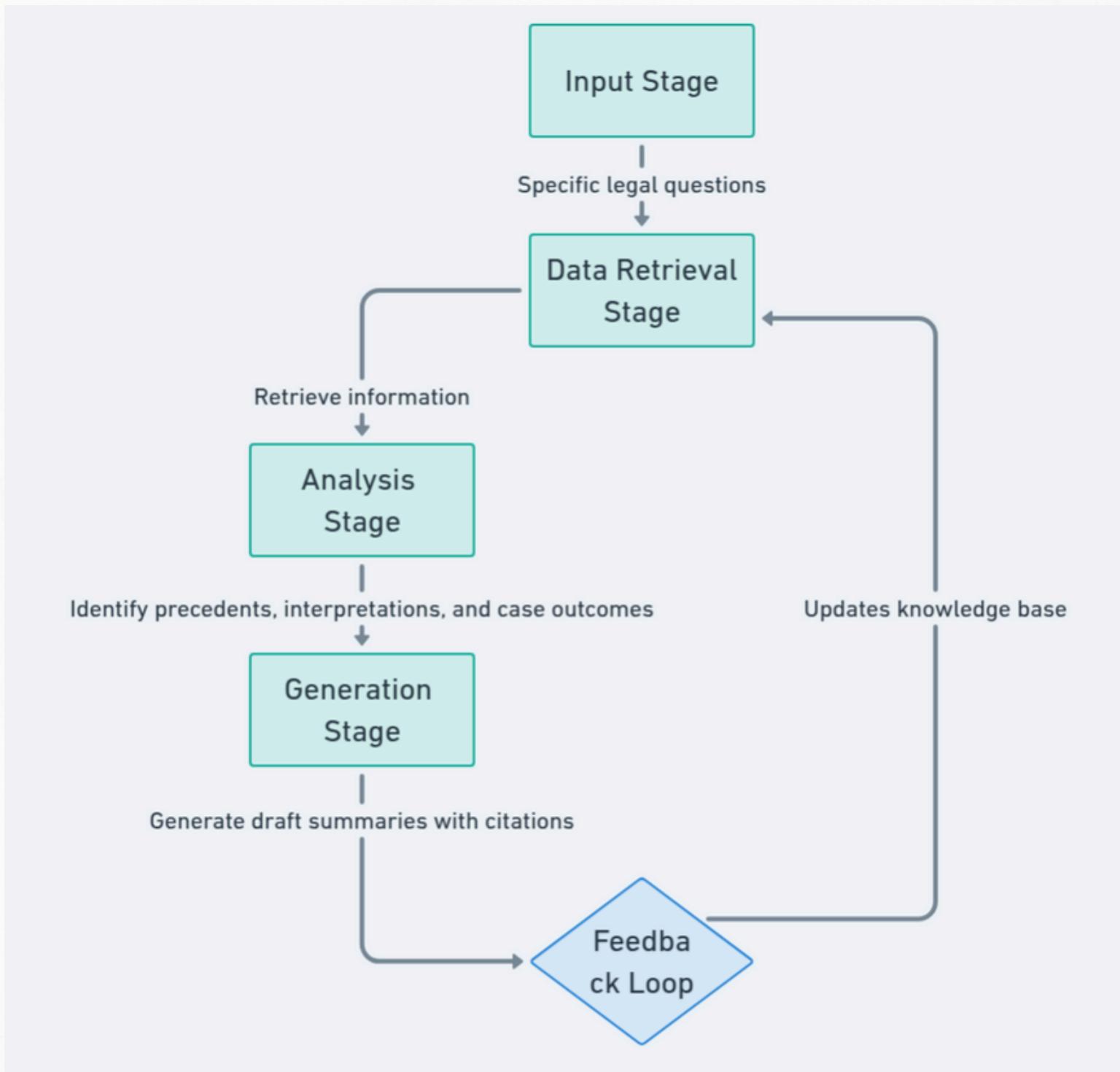
1. The user submits a query via search or chatbot.
2. The API layer interacts with the RAG model, which retrieves context-specific information using a custom retriever.
3. The generative model processes the query, retrieved data, and user context to generate a precise response.
4. The final response is delivered in a user-friendly format, tailored to the interaction channel.



*This diagram is a speculative representation.

8. Legal Processes

In this use case, the RAG model enhances legal research by efficiently retrieving relevant contexts across multiple sources. When lawyers ask questions, the AI swiftly identifies key past cases, connects insights from legal documents, and explains their relevance. With RAG, legal professionals can streamline research, improve accuracy, and optimize case preparation.



Impact:

- Faster legal research.
- Enhanced case preparation quality.
- Consistent legal insights.
- Optimized productivity for junior lawyers.

*This diagram is a speculative representation.

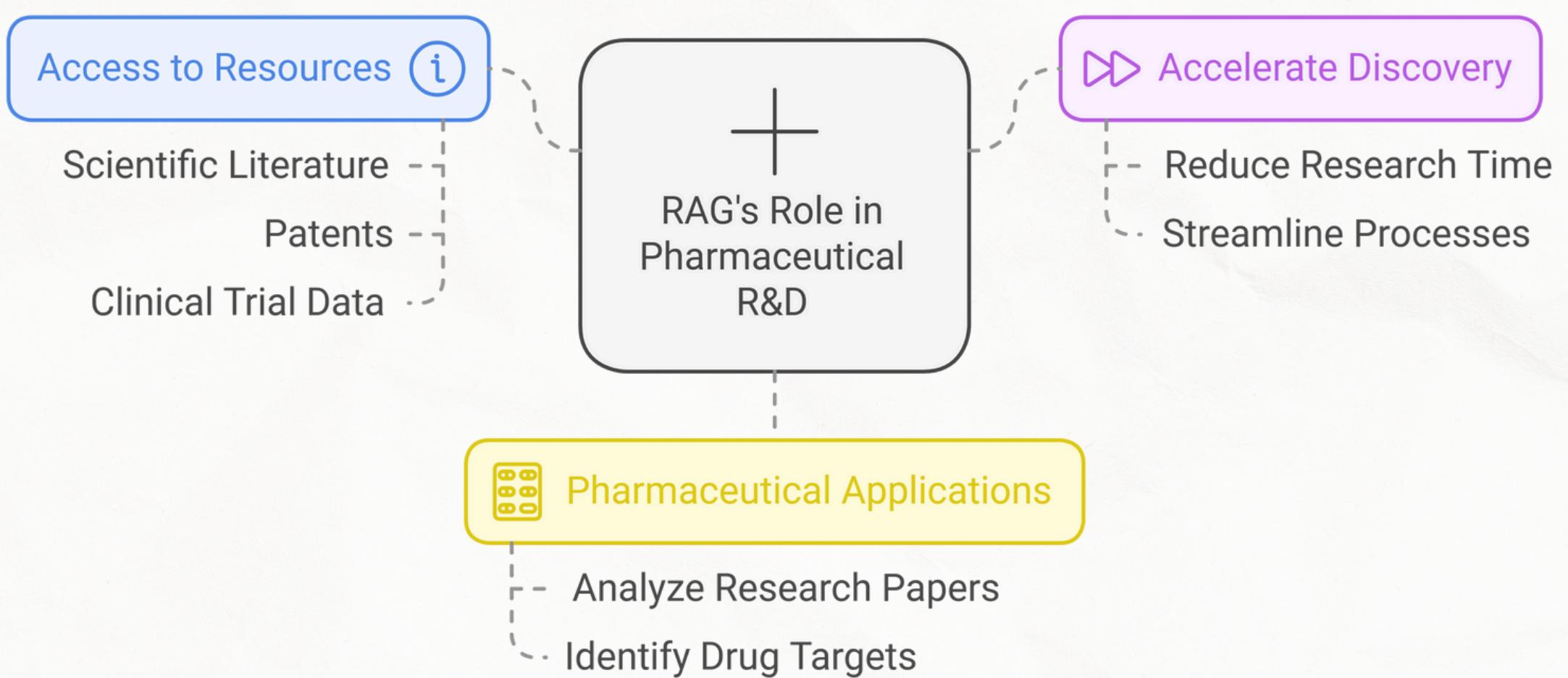
9. Research & Development

RAG empowers R&D by offering rapid access to scientific literature, patents, clinical trial data, and other relevant resources, streamlining the research process.

Accelerates discovery by reducing research time, facilitating quicker development of new products and technologies.

Pharmaceutical companies can use RAG to:

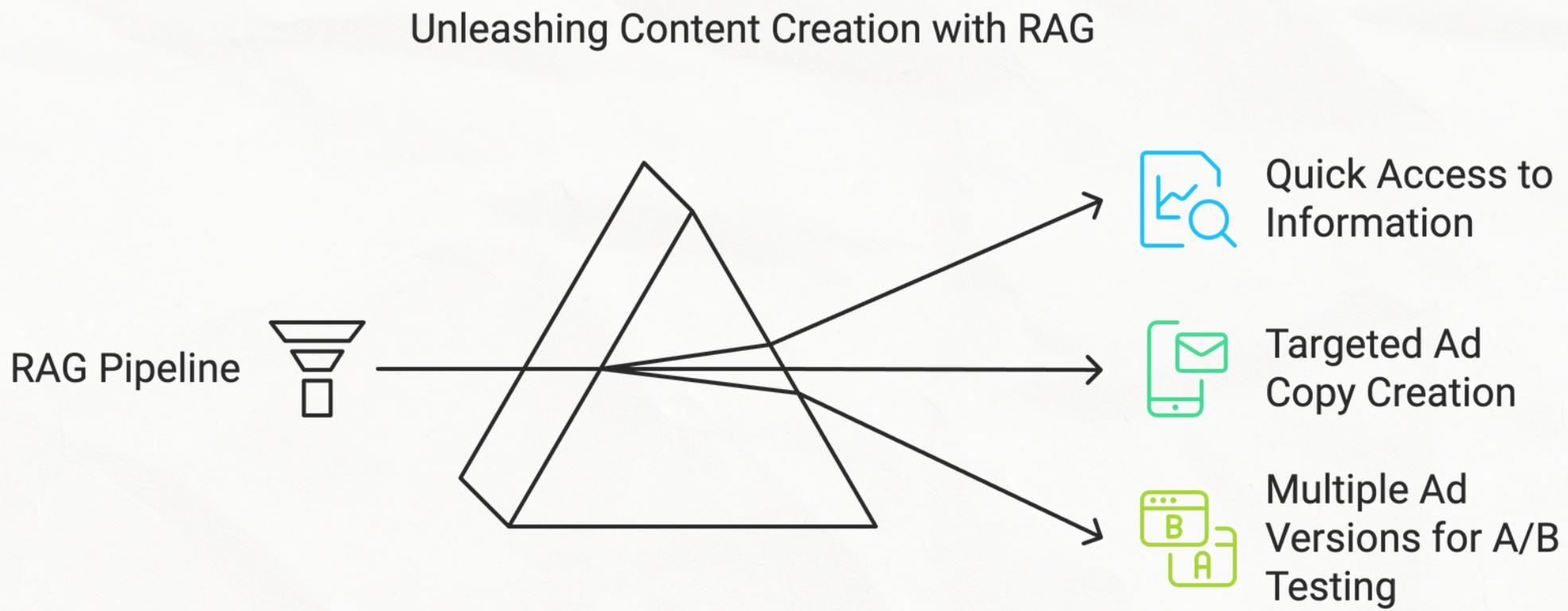
- Analyze thousands of research papers on specific diseases.
- Identify potential drug targets, speeding up the drug discovery process.



*This diagram is a speculative representation.

10. Content Creation

RAG enhances content creation by providing quick access to relevant information, data, and examples, streamlining the generation of blog posts, marketing copy, social media updates, and more.



Marketing agencies or content platforms can use RAG to:

- Create targeted ad copy by accessing product details and customer data.
- Generate multiple ad versions for A/B testing, optimizing campaign performance.

*This diagram is a speculative representation.



**Follow to stay updated on
Generative AI**



LIKE



COMMENT



REPOST