THE CYPRUS INSTITUTE

RESEARCH•TECHNOLOGY•INNOVATION

Introduction to Retrieval-Augmented Generation (RAG)

Enhancing AI with External Knowledge

Marios Constantinou

www.cyi.ac.cy

THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

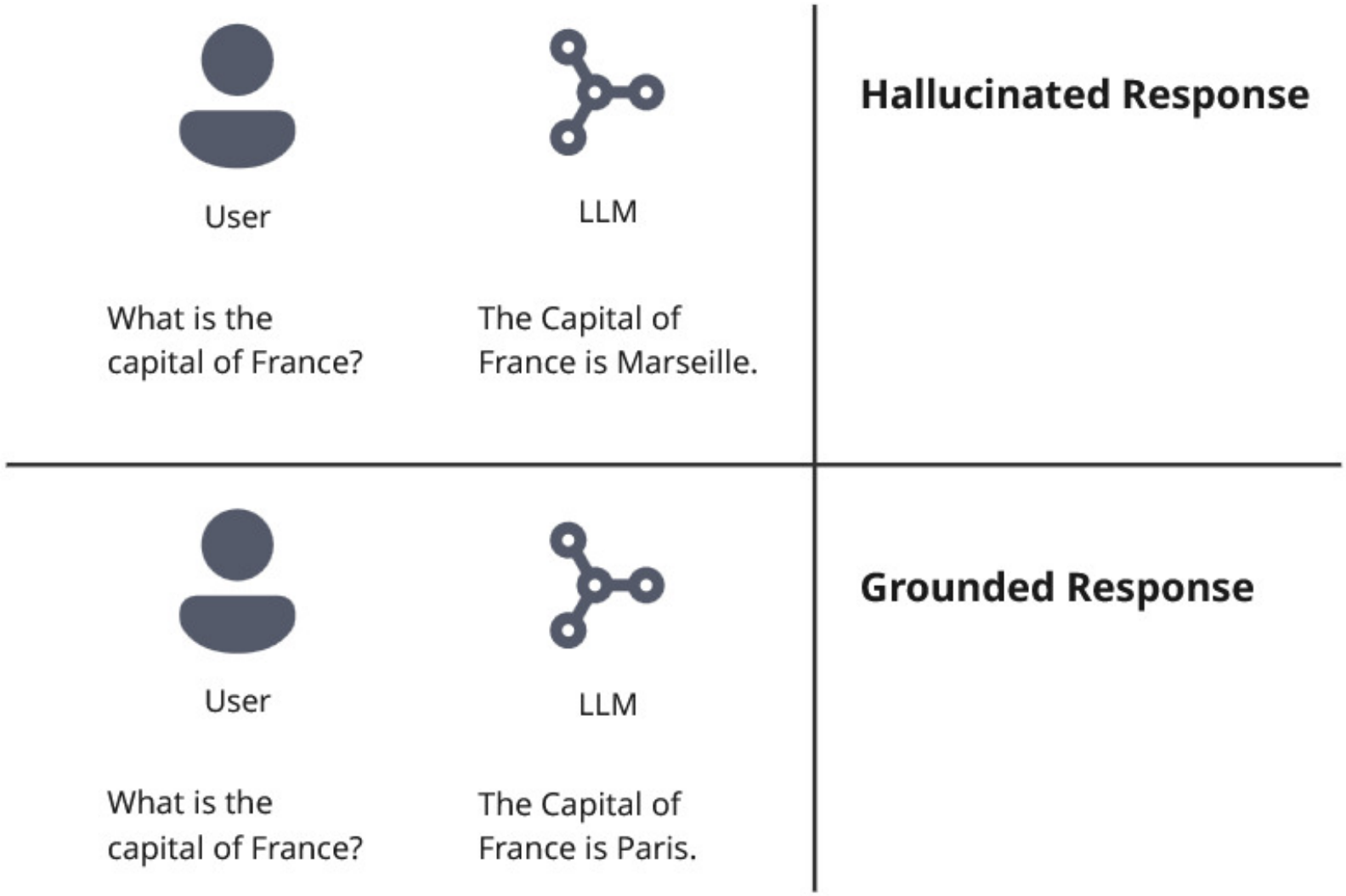
What is Retrieval- Augmented Generation?

Marios Constantinou

www.cyi.ac.cy

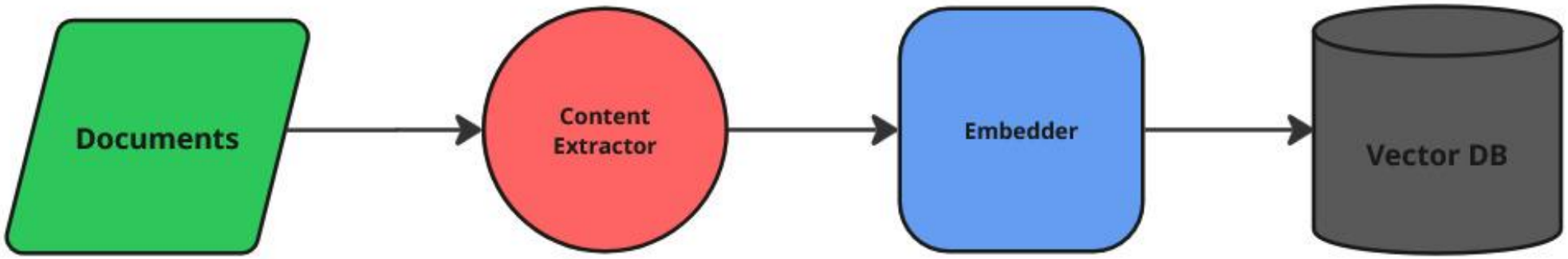
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Why use RAG?

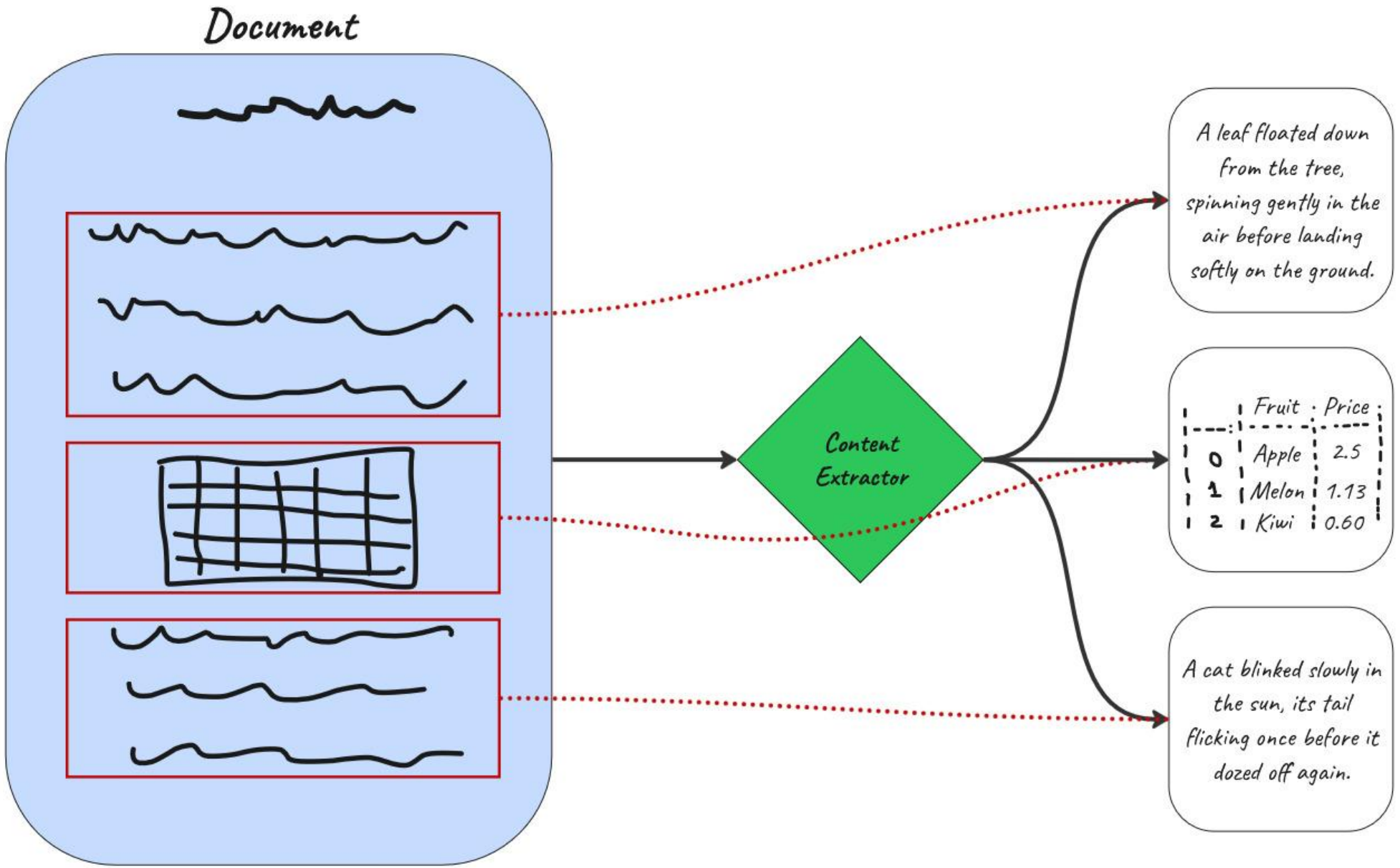


How does it work?

Indexing Pipeline

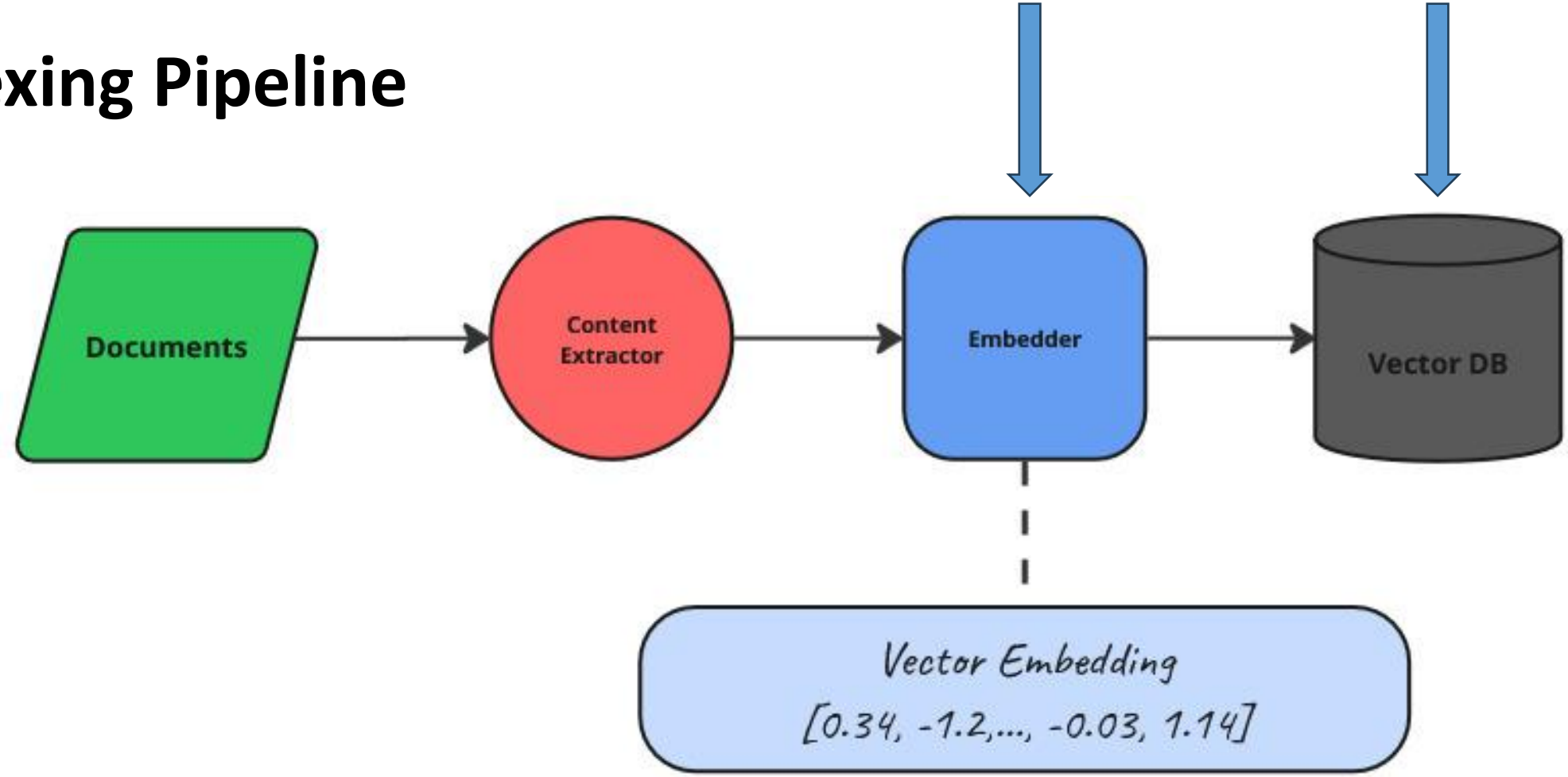


Content Extraction



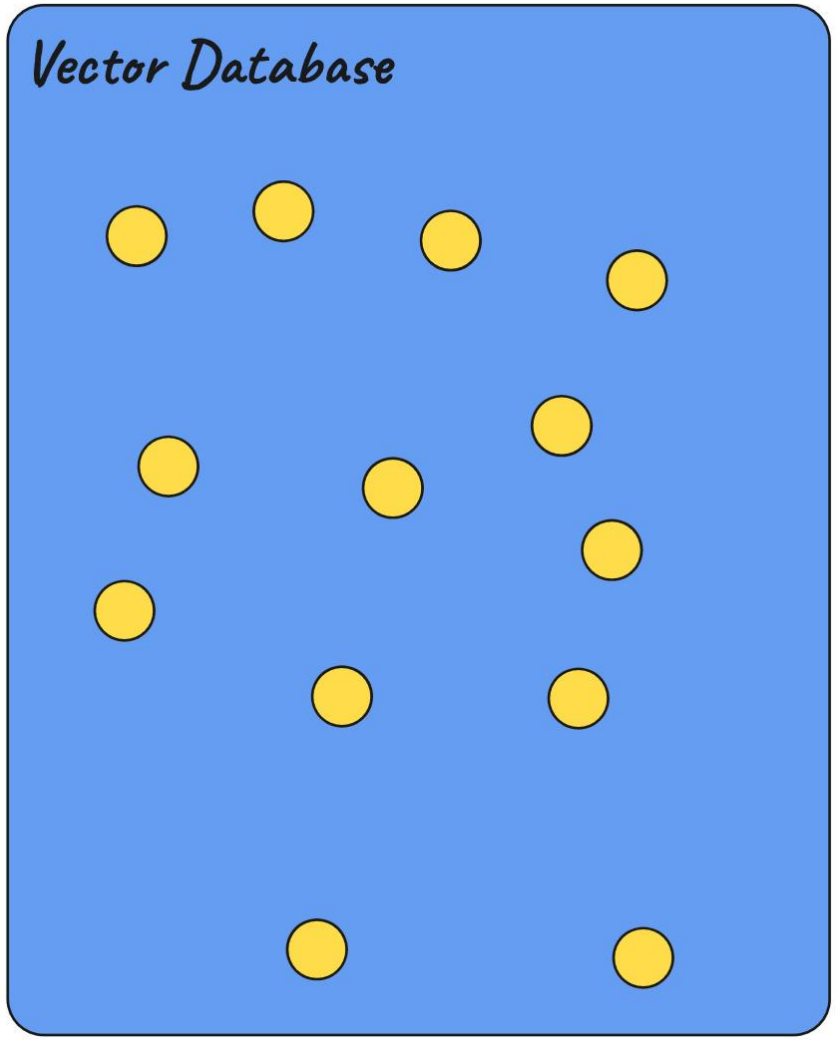
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Indexing Pipeline



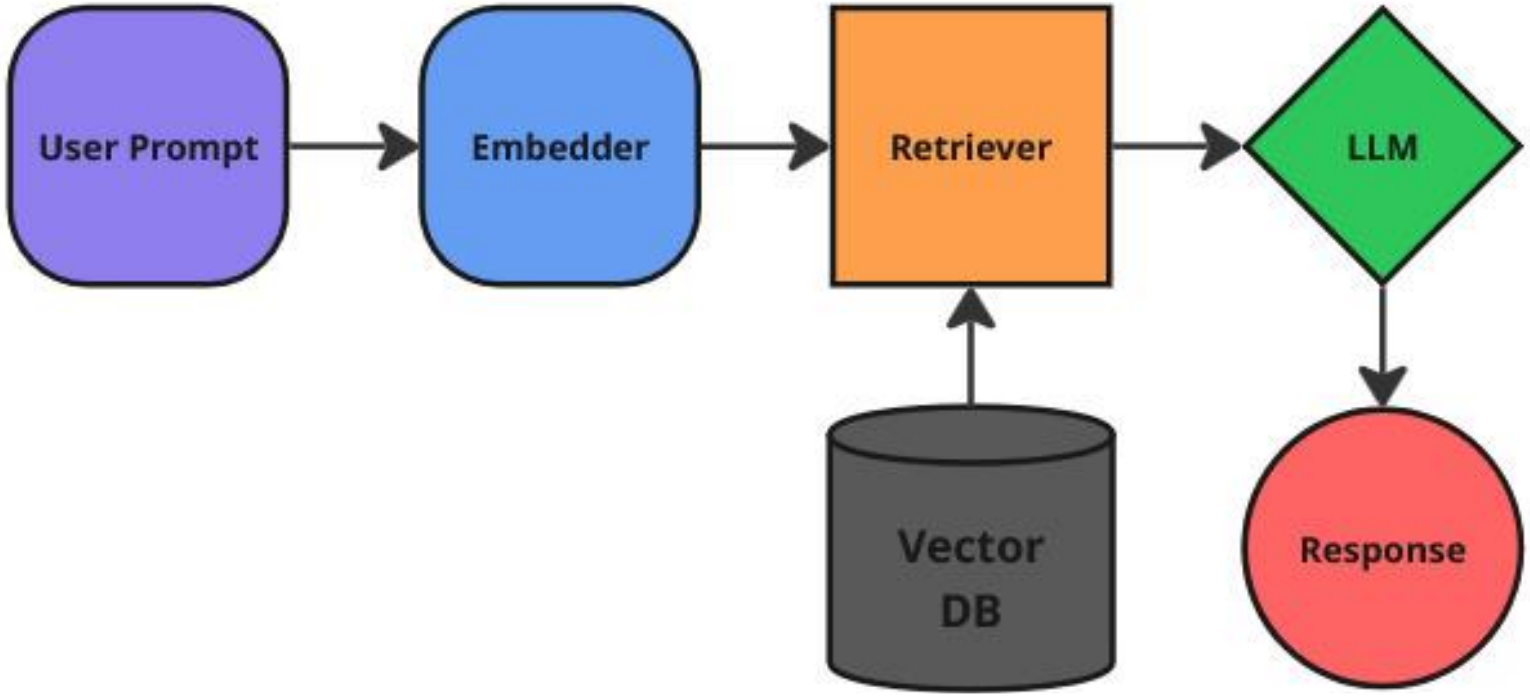
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Vector Database



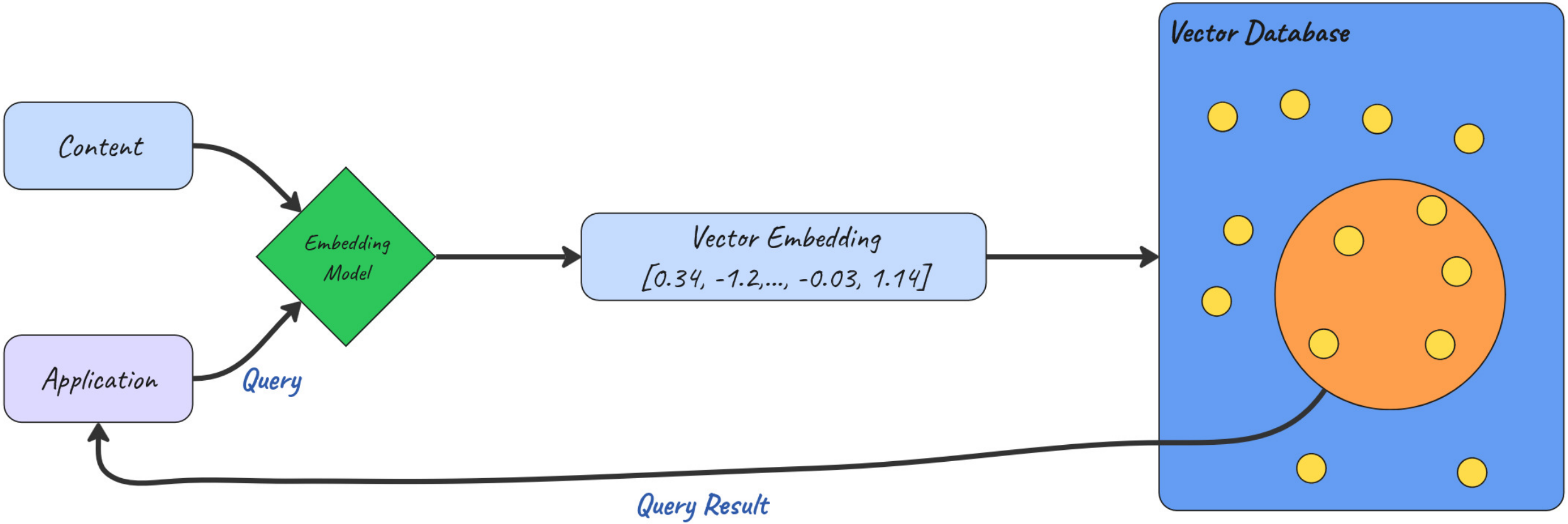
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

RAG Pipeline



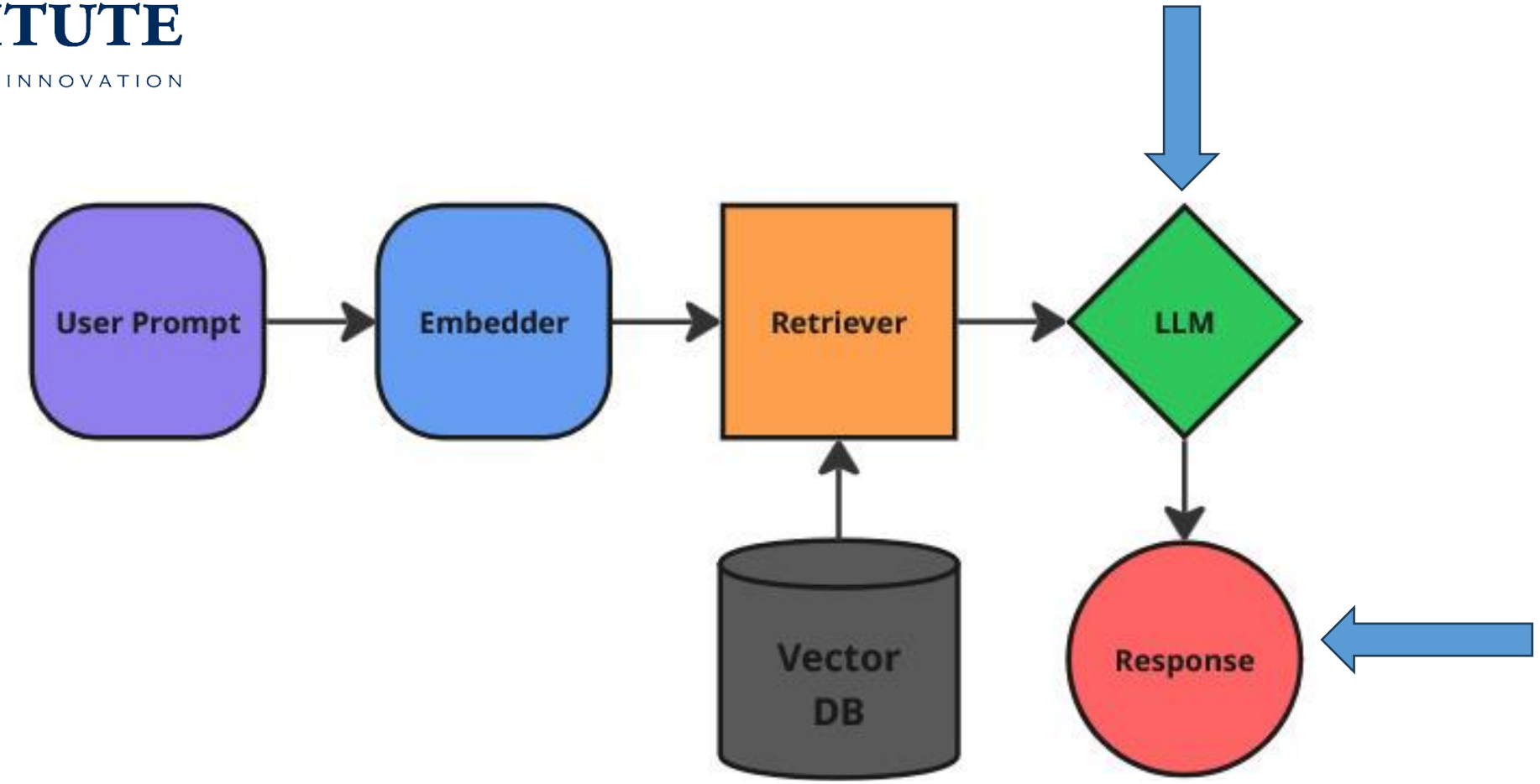
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Embedding Retriever



THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

RAG Pipeline



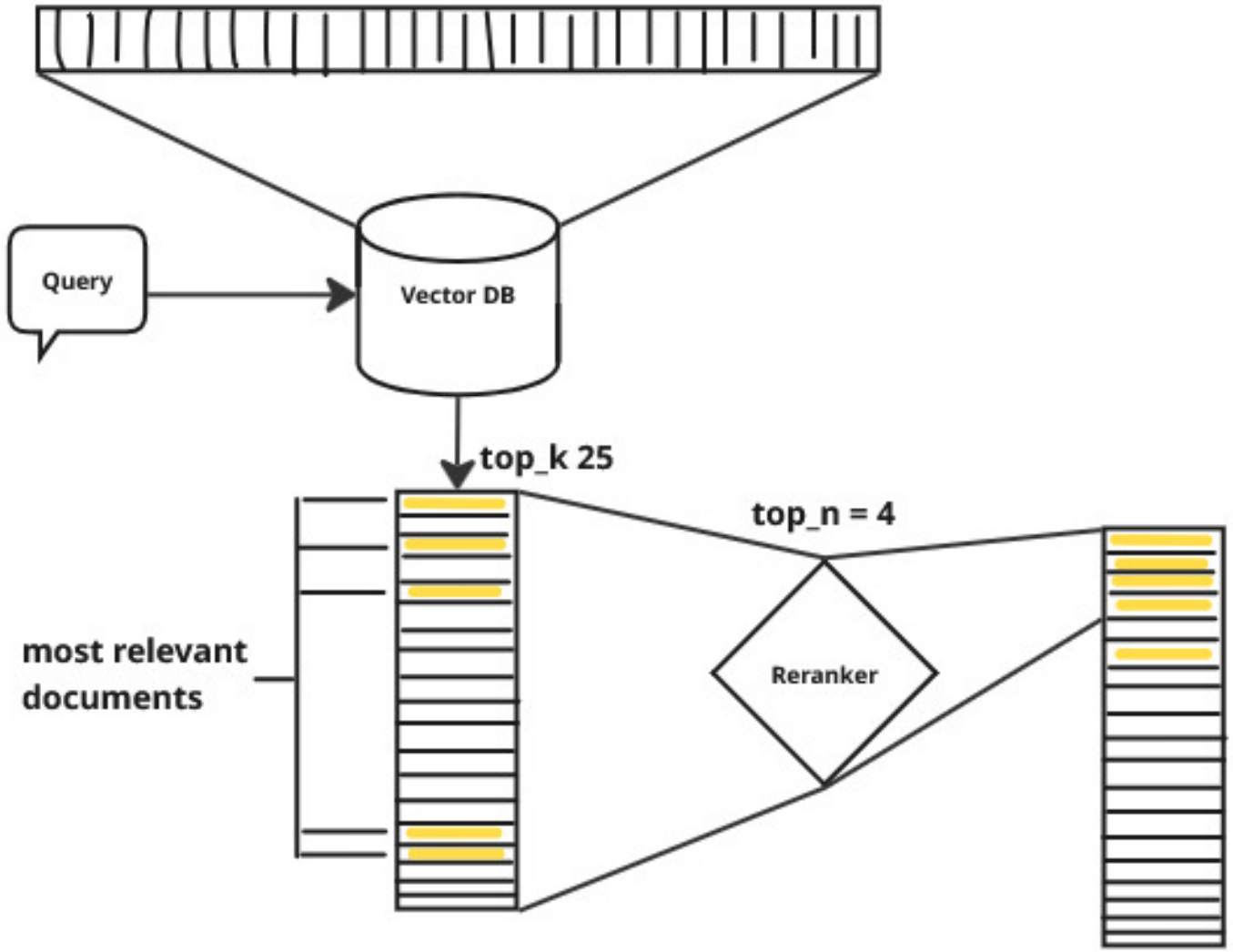
Tricks to boost RAG performance

Marios Constantinou

www.cyi.ac.cy

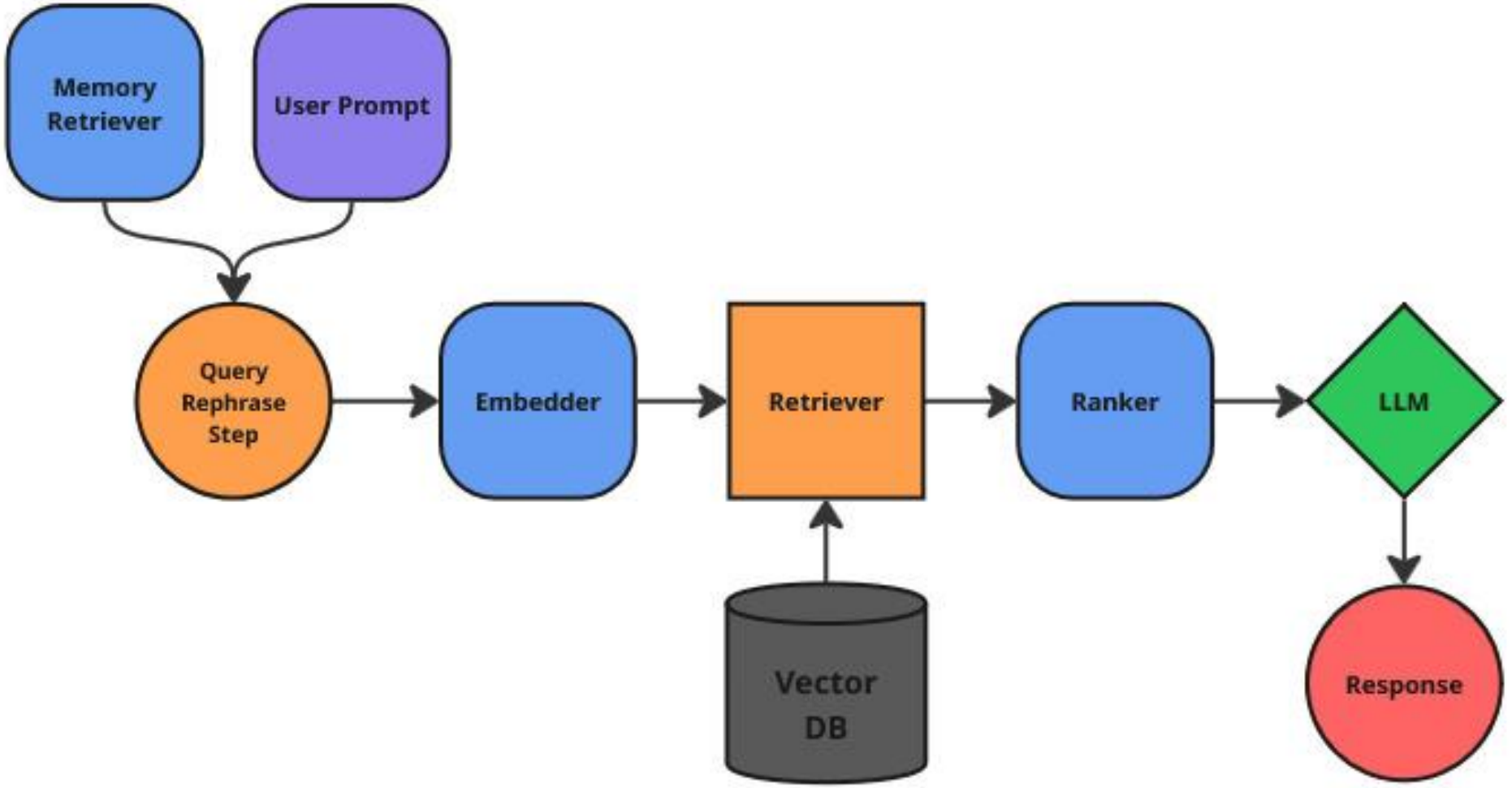
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Re-Ranking Documents



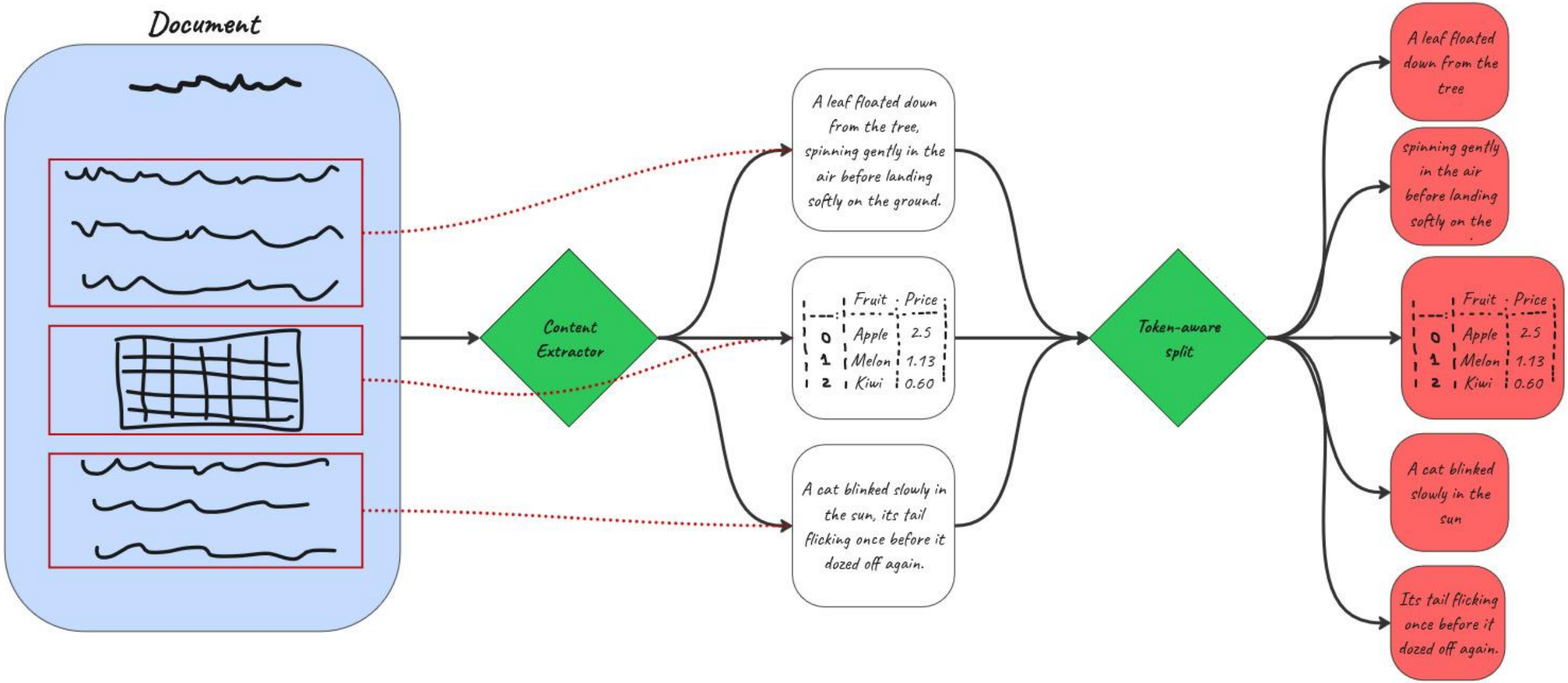
THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Query Rephrasing



THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

Hybrid Chunking



THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

https://tinyurl.com/bd53yjvy

https://github.com/CaSToRC-CyI/AI-Agents-Training

Marios Constantinou

www.cyi.ac.cy