The Map Is Not The Territory

(And That's Good)

Dr. Bryan Bischof – Ayush Chaurasia

11/07/2025

Part I

Our Senses, Language, And Biases Create Mental Maps That We Mistake For Territory Of Real-World Experiences

-- Alfred Korzybski's General Semantics

Part I

“_____ RAG”

\+ *Naive RAG*

Part I

“_____ RAG”

+ *Naive RAG*
+ *Agentic RAG*

Part I

“_____ RAG”

+ *Naive RAG*
+ *Agentic RAG*
+ *Hybrid RAG*

Part I

“_____ RAG”

+ *Naive RAG*
+ *Agentic RAG*
+ *Hybrid RAG*
+ *Graph RAG*

Part I

“_____ RAG”

+ *Naive RAG*
+ *Agentic RAG*
+ *Hybrid RAG*
+ *Graph RAG*
+ *Multi-Modal RAG*

Part I

“_____ RAG”

Leave Me Alone

- *Naive RAG*
- *Agentic RAG*
- *Hybrid RAG*
- *Graph RAG*
- *Multi-Modal RAG*

Part II

No, Not That Multi-Modal

Part II

ONE

Hypothetical Document Embeddings with semantic search is a document enrichment pipeline, not a retrieval technique.

Part II

ONE

Hypothetical Document Embeddings with semantic search is a document enrichment pipeline, not a retrieval technique.

TWO

Agentic RAG is a query enrichment pipeline, not a retrieval technique.

Part II

ONE

Hypothetical Document Embeddings with semantic search is a document enrichment pipeline, not a retrieval technique.

TWO

Agentic RAG is a query enrichment pipeline, not a retrieval technique.

THREE

Rank fusion (kmn) is multi-stage processing, not a retrieval technique.

Part II

— You Could Have Invented ___

The core responsibility of the IR engineer is to understand:

Part II

—— You Could Have Invented ___

The core responsibility of the IR engineer is to understand:

- What is the most likely representation of the user's desire?

Part II

—— You Could Have Invented ___

The core responsibility of the IR engineer is to understand:

- What is the most likely representation of the user's desire?
- What representations can you generate ahead of time from the entities?

Part II

— You Could Have Invented ___

The core responsibility of the IR engineer is to understand:

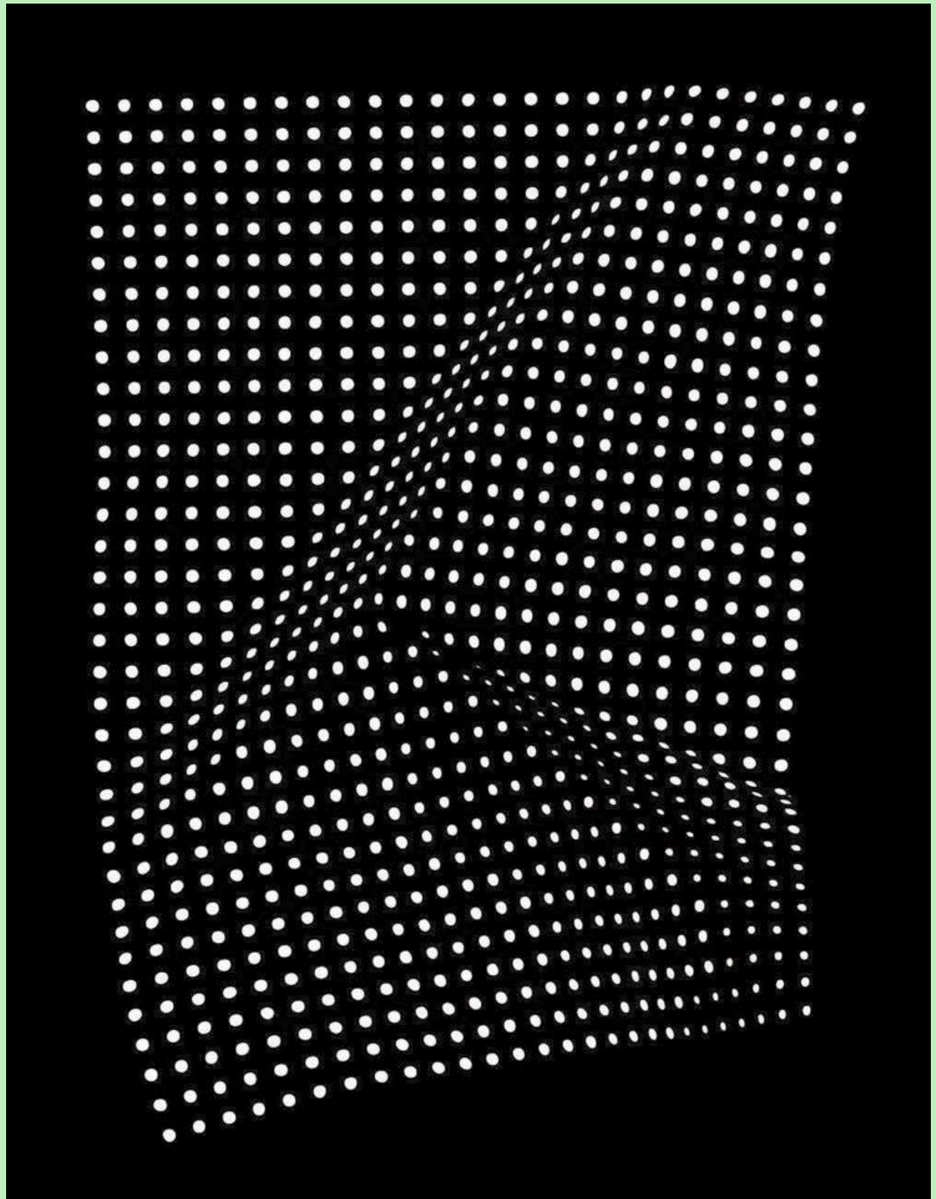
- What is the most likely representation of the user's desire?
- What representations can you generate ahead of time from the entities?
- How can you correctly match these up?

Part III

How To Curve Space

Part III

Let’s retrieve Financial Documents

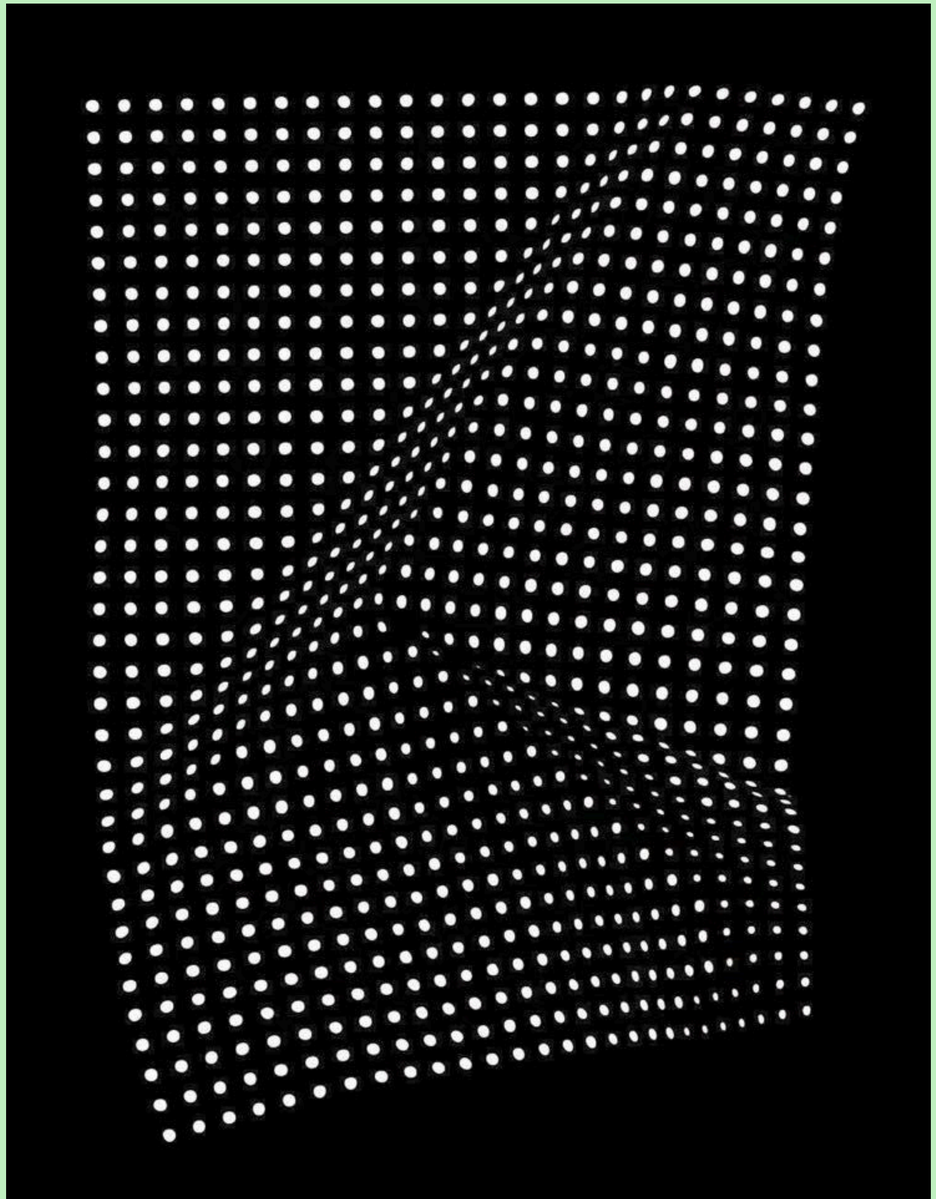


ONE

What are the different concepts one may construct from these docs? Summaries? Tables of Data? Lists of Named Entities? Form types/categories? What else?

Part III

Let’s retrieve Financial Documents



ONE

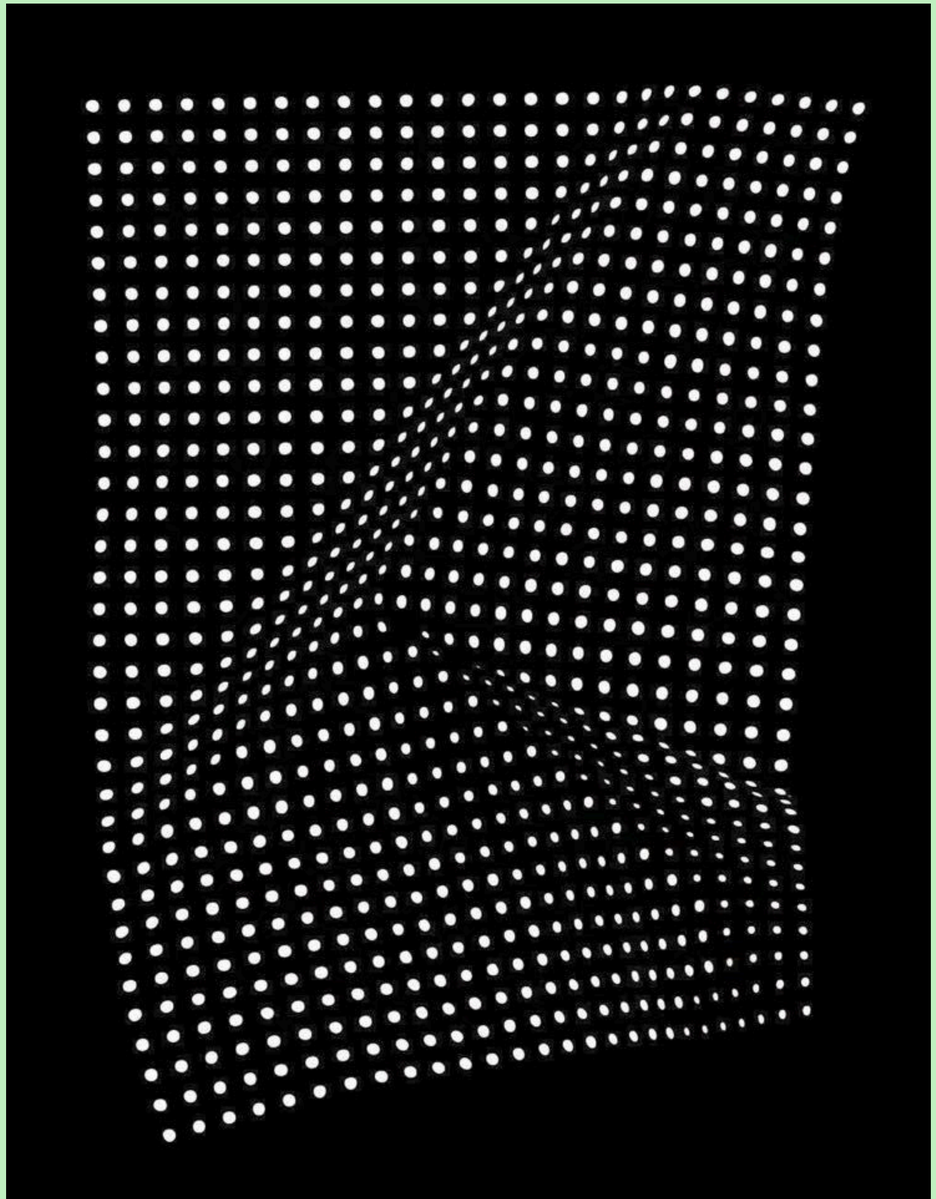
What are the different concepts one may construct from these docs? Summaries? Tables of Data? Lists of Named Entities? Form types/categories? What else?

TWO

What is the appropriate index strategy for each? Semantic Search? BM25? Keyword Matching? Pre-filters? When should you use multiple on one index? When not?

Part III

Let's retrieve Financial Documents



ONE

What are the different concepts one may construct from these docs? Summaries? Tables of Data? Lists of Named Entities? Form types/categories? What else?

TWO

What is the appropriate index strategy for each? Semantic Search? BM25? Keyword Matching? Pre-filters? When should you use multiple on one index? When not?

THREE

When is another "stage" required? Is the easiest thing to retrieve what the agent needs? Can one retrieval enable another?

Section IV

Just Say ‘agents’, Dammit!

Part IV

Agents Are Transformers

(no, not like that)

Part IV

Agents Are Transformers

(no, not like that)

In both Agentic RAG and Multi-Agent RAG, the core idea is to use general LM's to transform incoming data and instructions into a structured output.

Thinking of them like 'routers' can be very helpful: route the data to different indices based on instructions, or route upstream stages to more retrieval stages.

Part V

No Map Is Ever Final

Part V

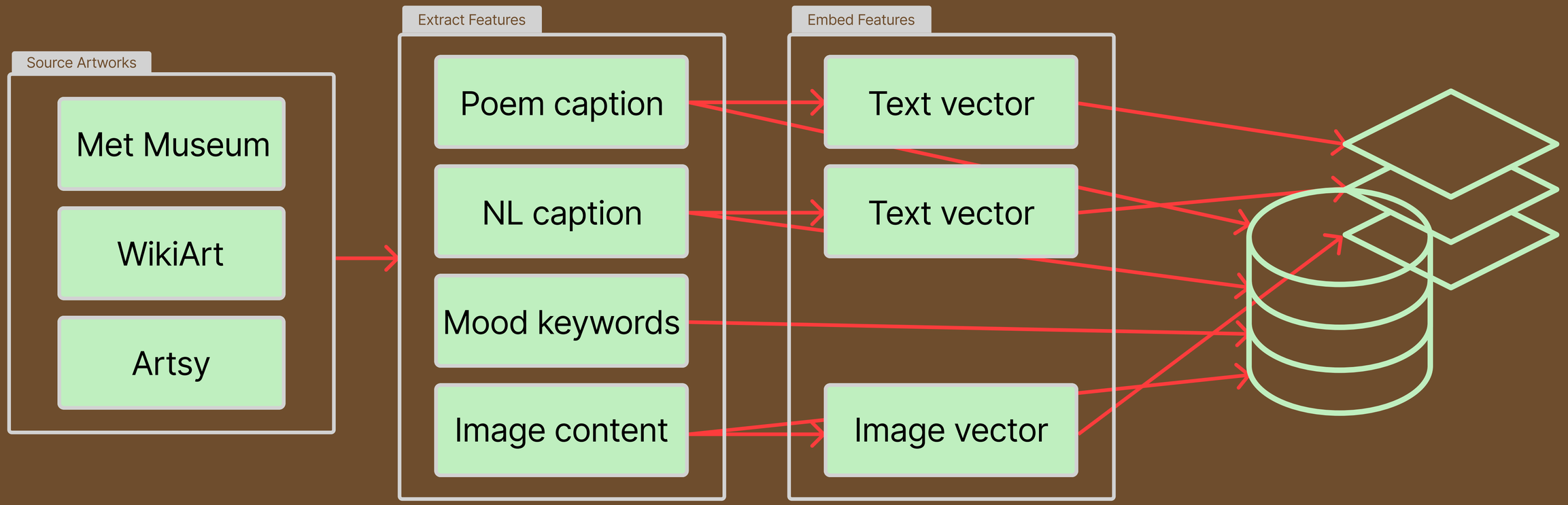
The Danger In All Of This, Is That Many Documents Are Dynamic. This Means That The Representations Get Out Of Date.

Part V

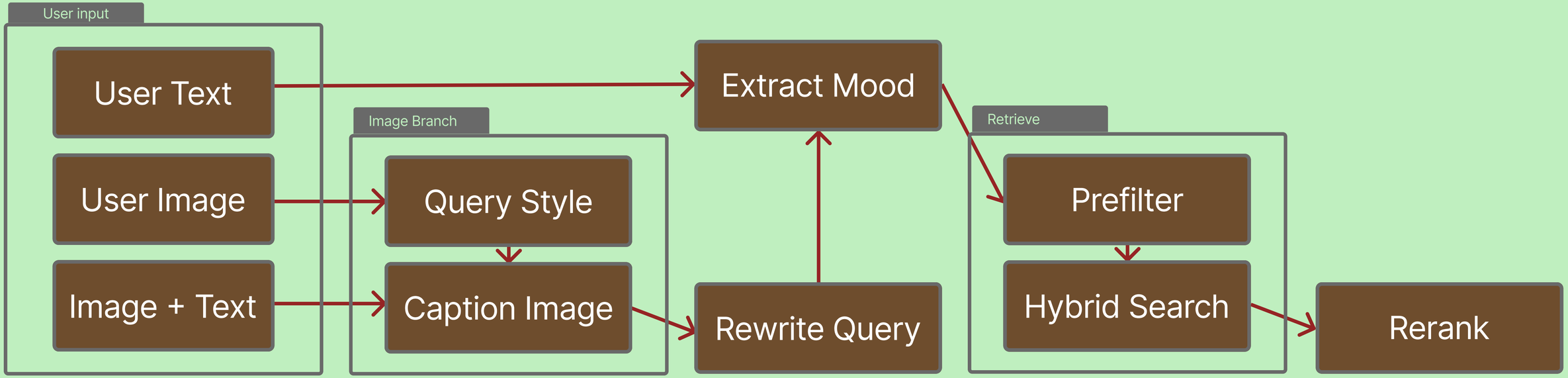
The Danger In All Of This, Is That Many Documents Are Dynamic. This Means That The Representations Get Out Of Date.

The Upshot Here Is That You Need To Design The System To Shard Each Entity In The Easiest Way Possible To Understand What's Changed, And Then Re-Index That Which Has.

Represent!


Source Artworks
Met Museum
WikiArt
Artsy
Extract Features
Poem caption
NL caption
Mood keywords
Image content
Embed Features
Text vector
Text vector
Image vector

Discover!


User input
User Text
User Image
Image + Text
Image Branch
Query Style
Caption Image
Extract Mood
Rewrite Query
Retrieve
Prefilter
Hybrid Search
Rerank

Thank You!