# Semantic Search

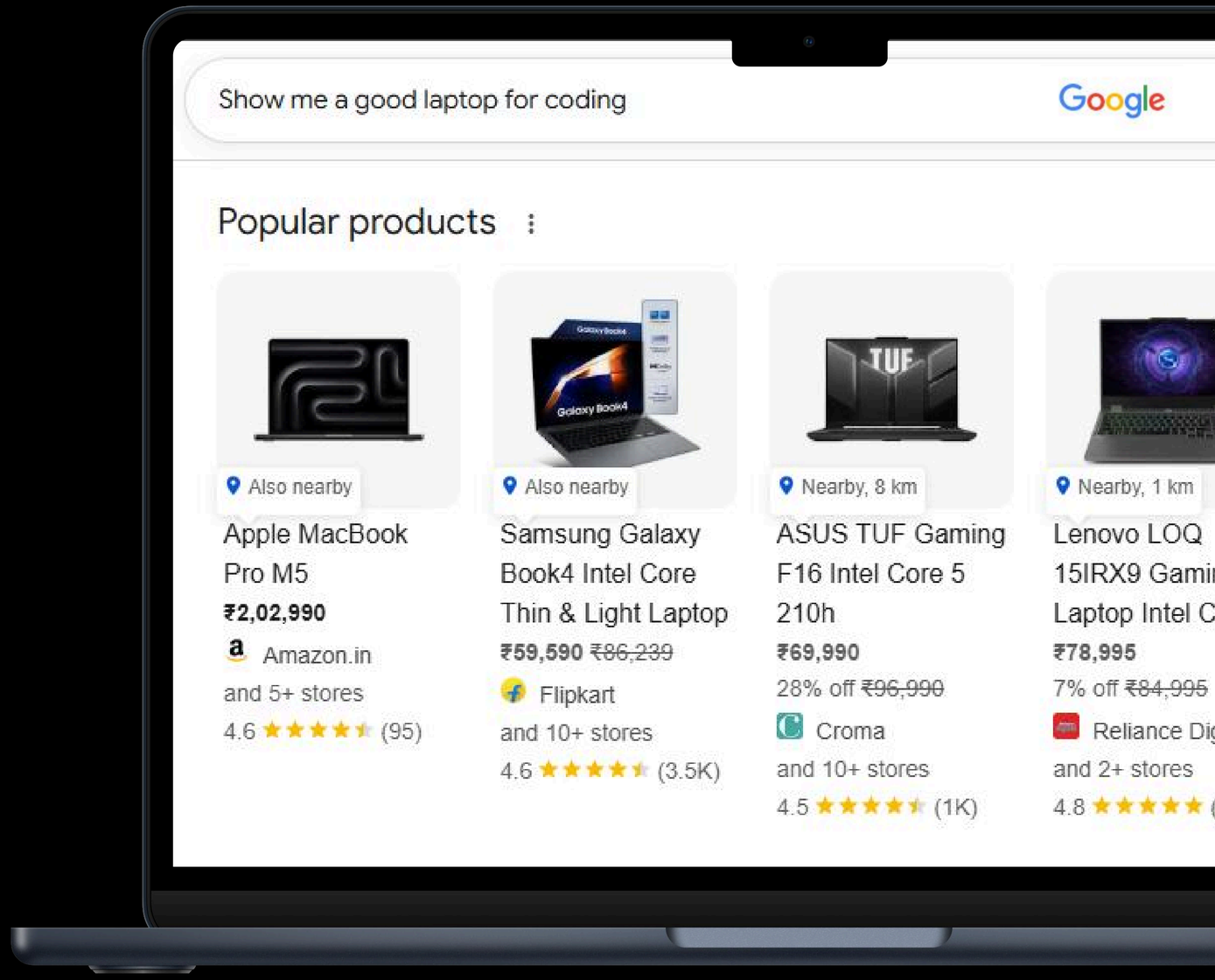
Add a Brain To Your Search Bar

# Why Doesn’t Google Need Exact “Keywords”?

Search anything!

# Go Ahead

Always gives MacBook/Ultrabook or gaming laptop



# Google Search’s 3 (main) Components

### Web Crawler (Googlebot)

Browses the internet, discovering accessible web pages and content

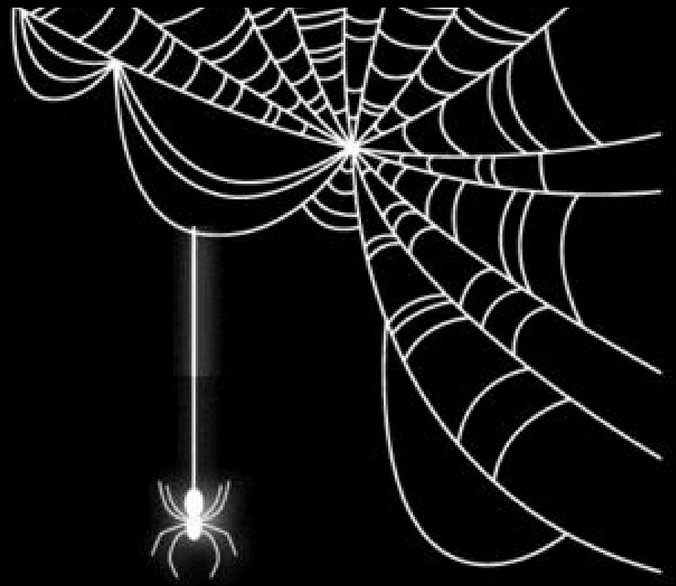
### Index

A massive database containing billions of web pages that crawlers, have discovered and stored

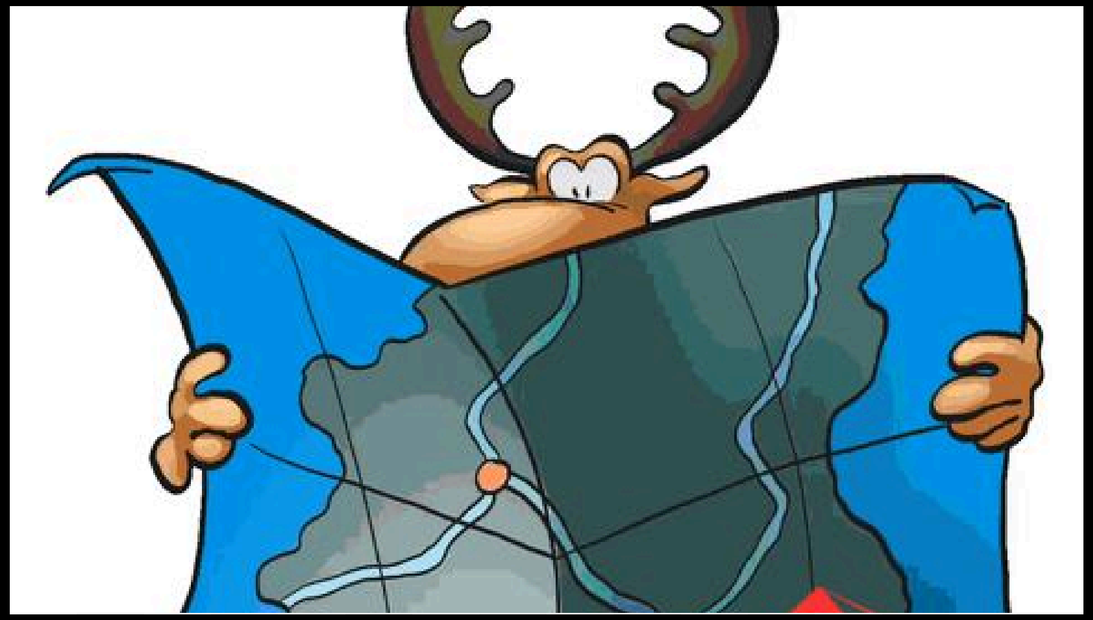
### Ranking Algorithm

determine the relevance, quality, and authority of webpages in response to a user's search query

# Web Crawler



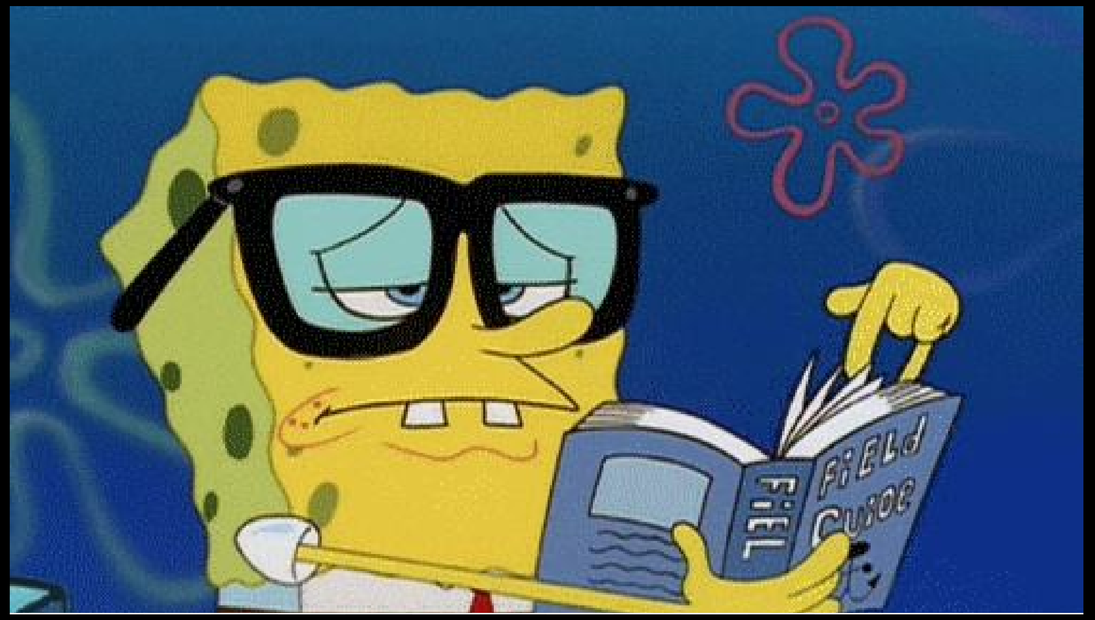
Finds Neighbor



Goes To Every Node

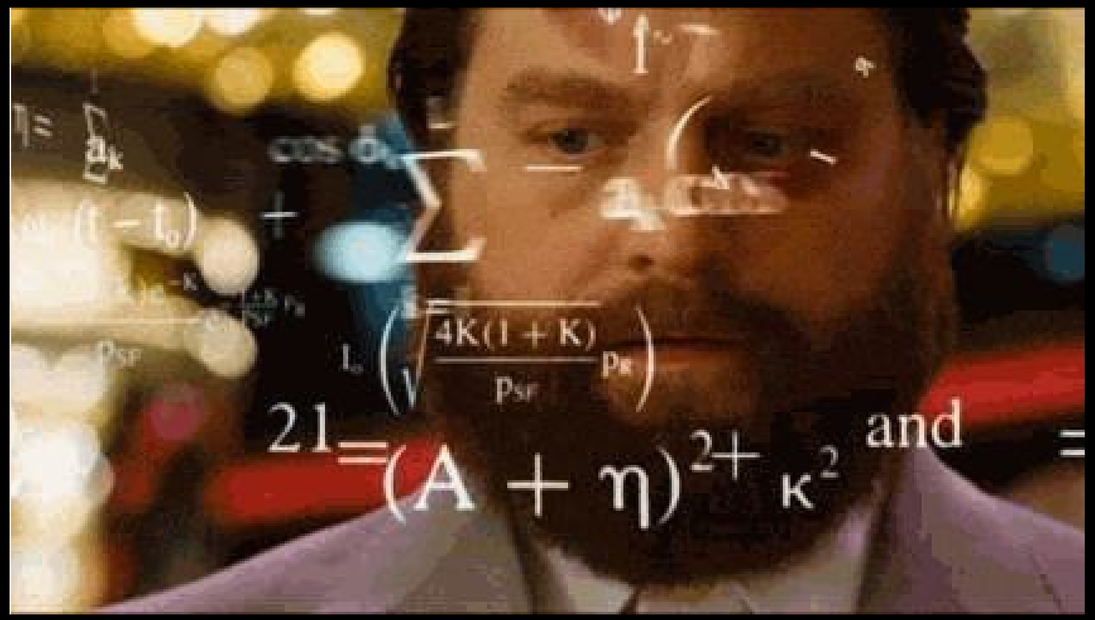
# Index

Stores What it gets




Reads and Analyse

# Ranker




Applies MATHS ! ! !

Serves You The DISH

# Veritasium


Markov Chains
32:32
The Strange Math That Predicts (Almost) Anything
Veritasium
19.6m subscribers
244k
Share
Ask
Download

## The Funnel

To see what’s relevant

# How About Our Brains

Ask anything!

# It uses 3 things

## Meaning

The sense or lexical content of a word or phrase.

## Context

Words coming before or after a word (for understanding its meaning)

## Relations

The connection between TWO or MORE things

# Show me a good laptop for coding

## Good: Must Mean

A quality product, sharp screen, etc

## For Coding

Need a great processor, must be fast, comfortable for eyes...

## Bonus: Its MacBook

Easily recommended, for each use case

# But What If It Did Use Keyword In Our Brains

You would never find your socks in time!!

# But What If It Did Use Keyword In Our Brains

# Semantic Search

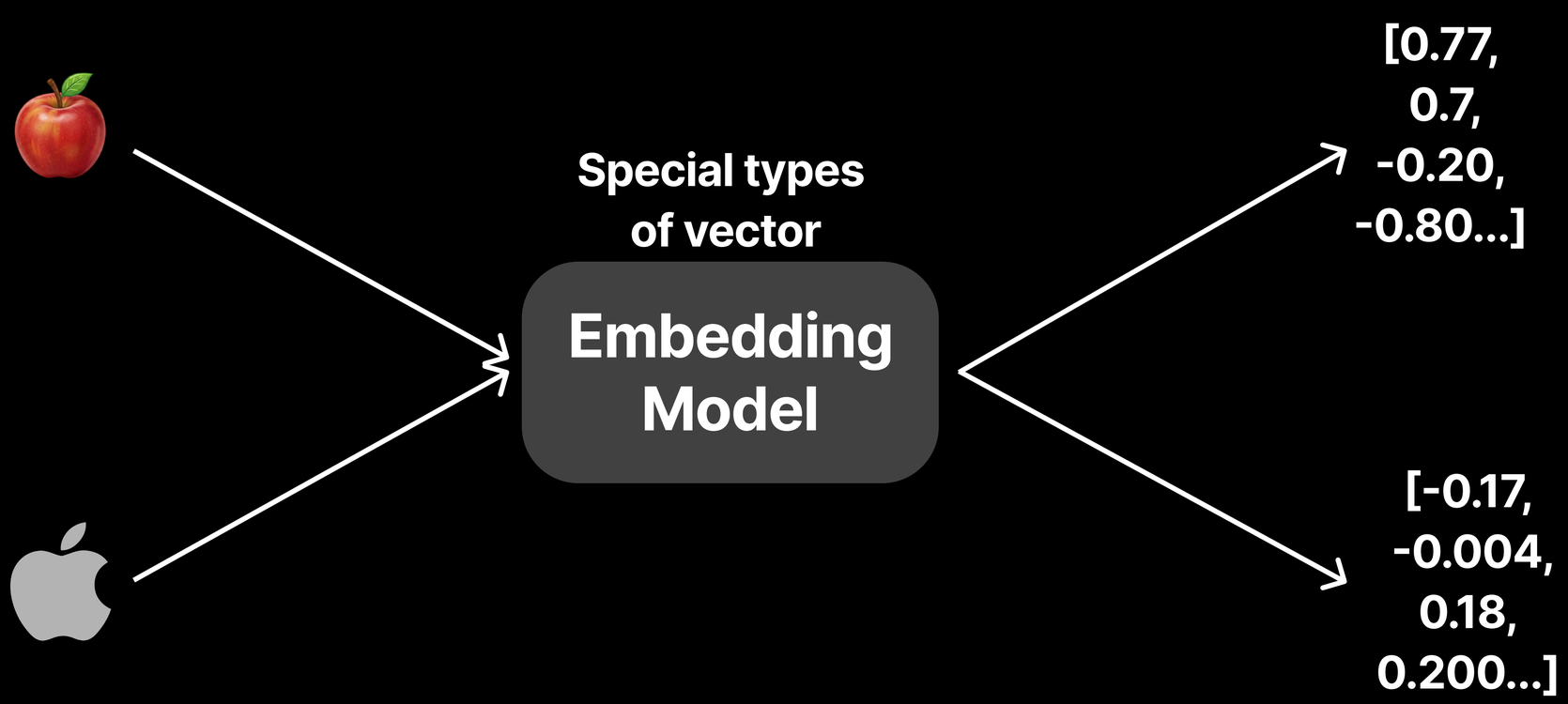
Add a Brain To Your Search Bar

# Akshat Sharma

AI/ML Developer, Futuresoft India

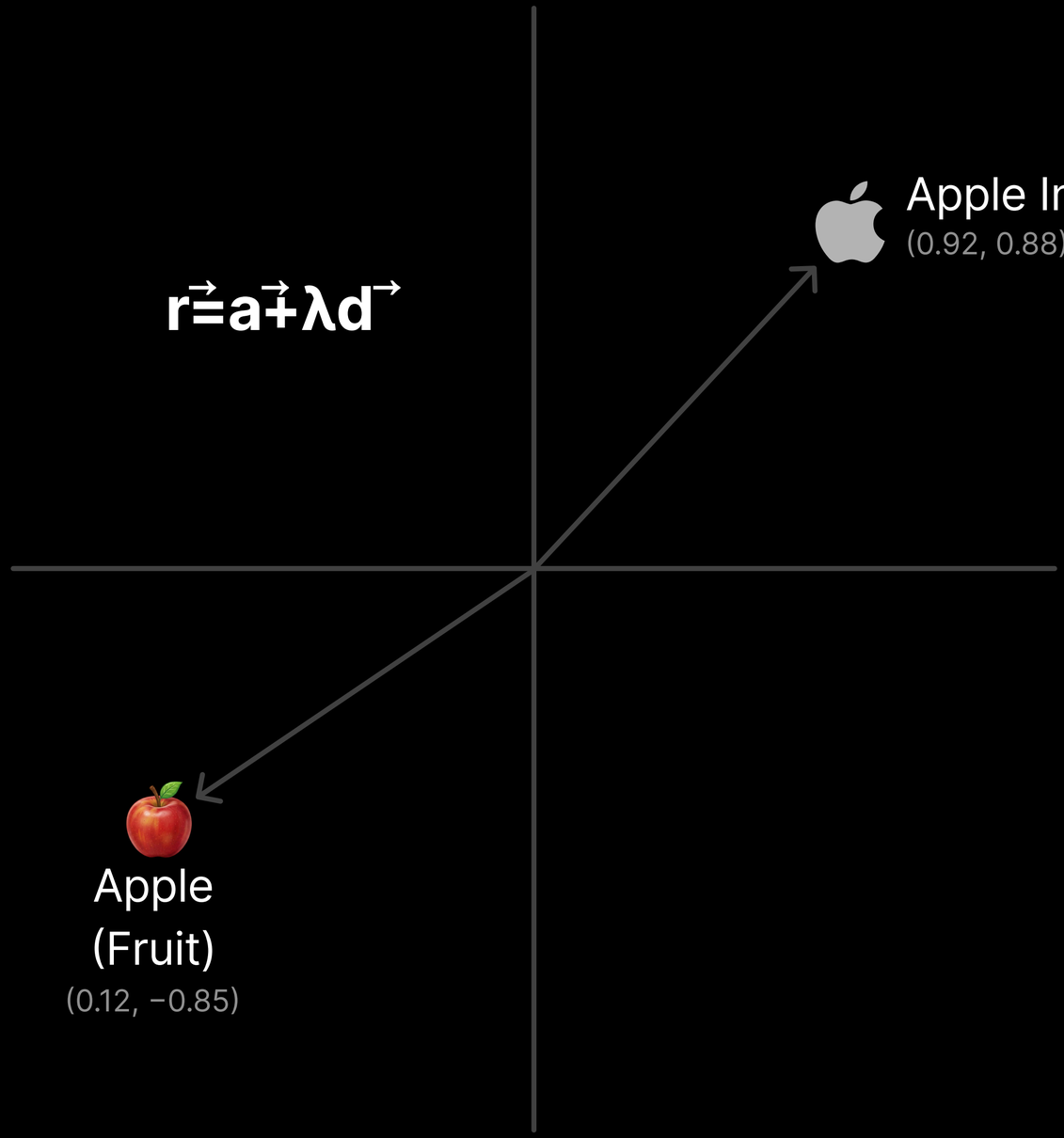
# So How Do Machines Replicate This?

## How Do We Add Brains to It


Special types
of vector
Embedding
Model
[0.77,
0.7,
-0.20,
-0.80...]
[-0.17,
-0.004,
0.18,
0.200...]

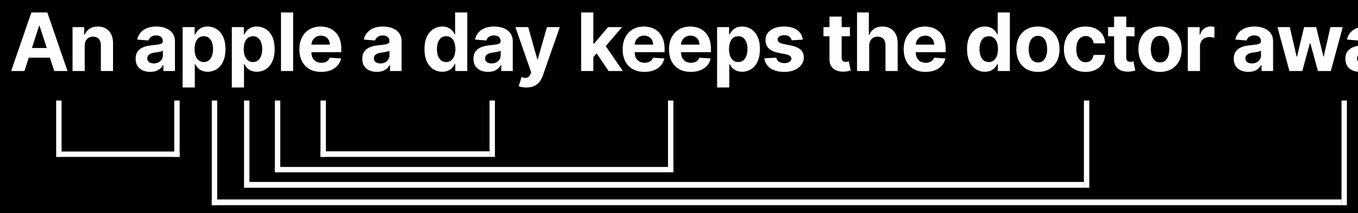
# Vector Space

**Everything is expressed in numbers**



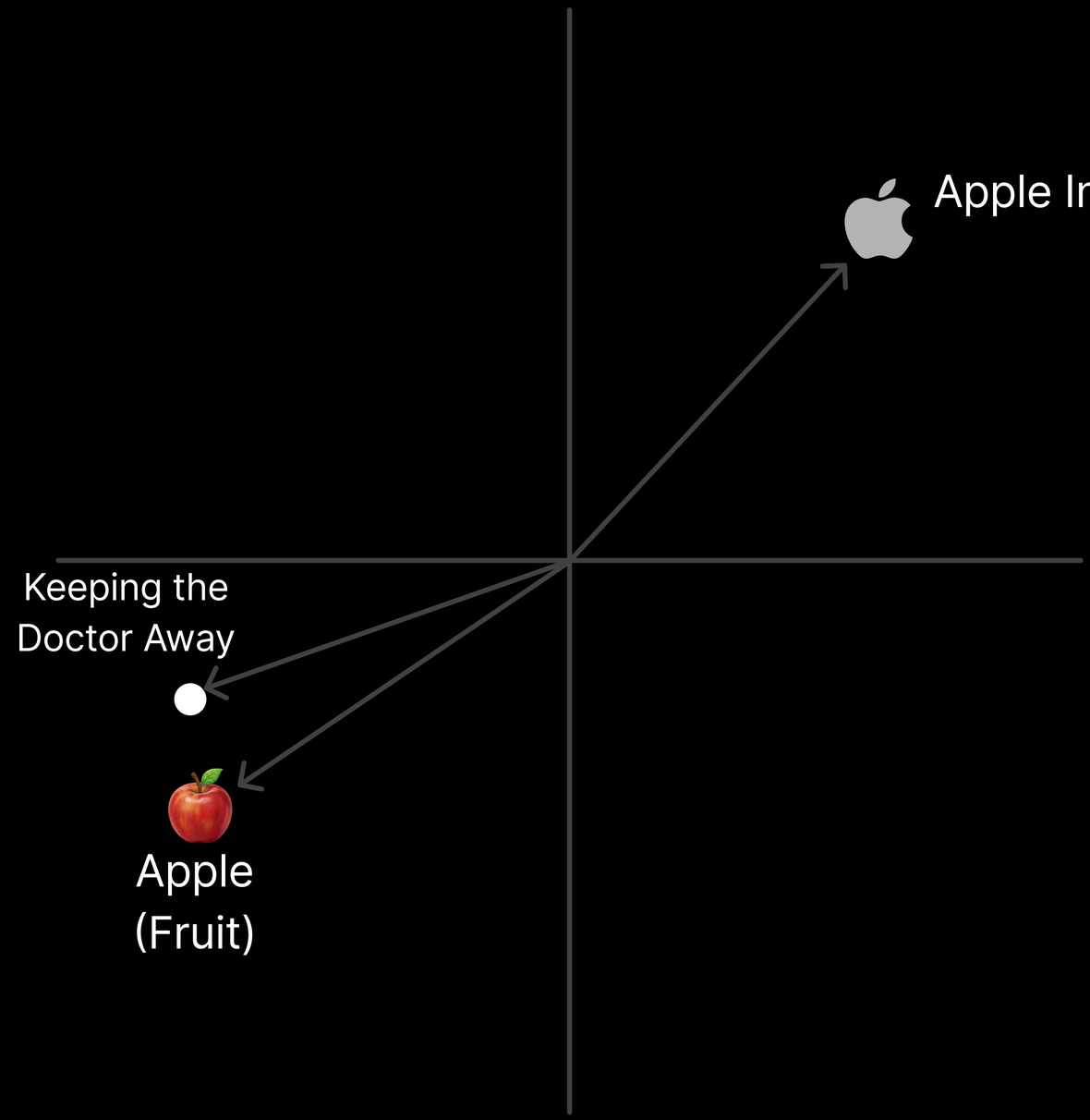

https://huggingface.co/spaces/jphwang/colorful_vectors

# Context in Vectors



# Vector Space

**See how close the new vector came to the fruit Apple**




https://huggingface.co/spaces/jphwang/colorful_vectors

# How Do We Find Similarity Then

## Euclidean Distance

How far apart two vectors are in the embedding space.

## Dot Product

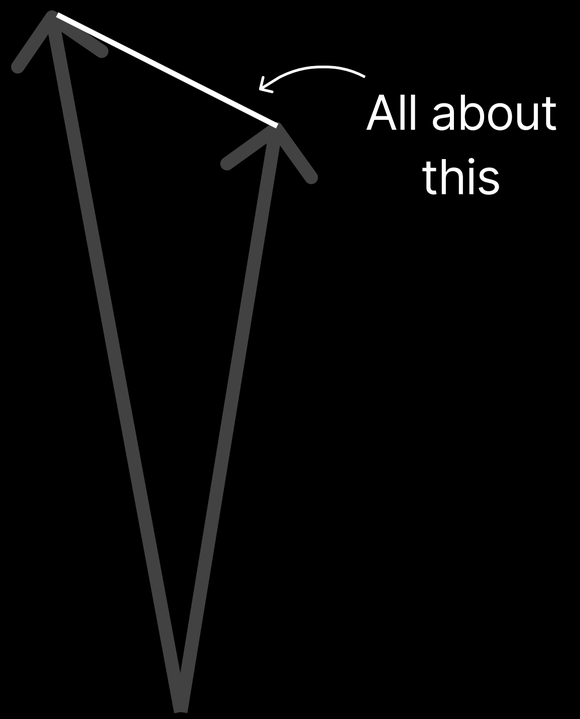
How much one vector "points" in the same direction as the other.

## Cosine Similarity

How aligned two vectors are — the smaller the angle, the more similar they are.

# Euclidean Distance

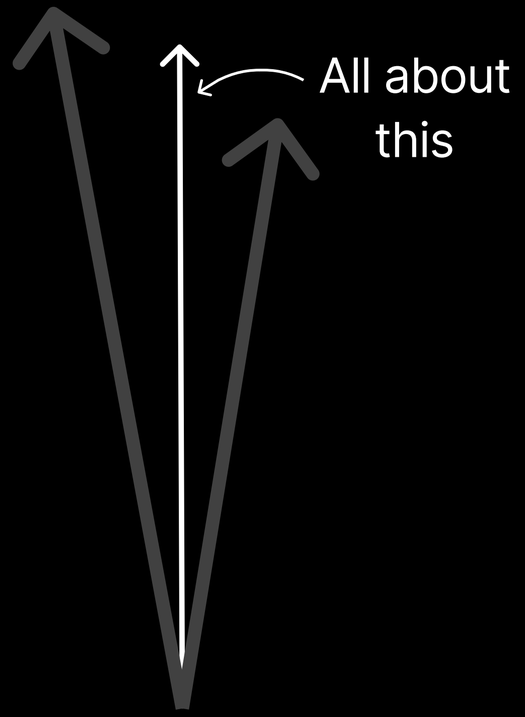
Distance between vectors




$$d(A,B)=\sqrt{[\sum(A_i-B_i)^2]}$$

# Dot Product

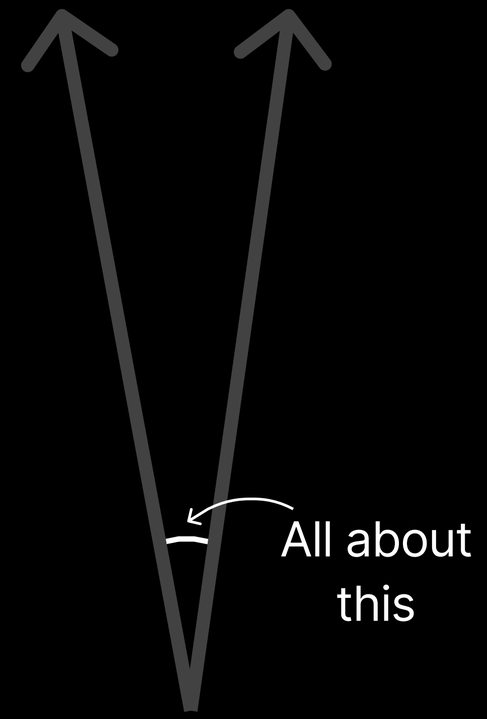
How much one vector "points" in the same direction as the other.




$$A.B=\sum(A_i*B_i)]$$

# Cosine Similarity

How aligned two vectors are.
Smaller the better




$$\text{sim}(A, B) = \frac{A \,.\, B}{\|A\| \,.\, \|B\|}$$

# A Recap

## How Google Search

And its 3 components

## How Our Brain Thinks

And the 3 things it uses

## What is Vector Search

And Minimum 3 ways to search in vector space

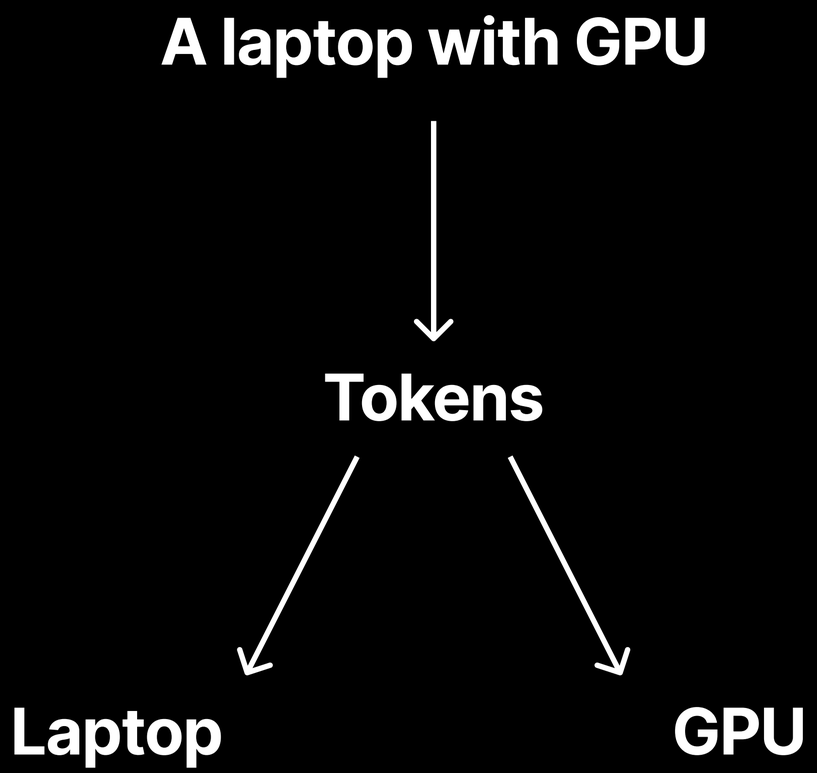
# Keyvord Besad Sarch

No results found! Try Again

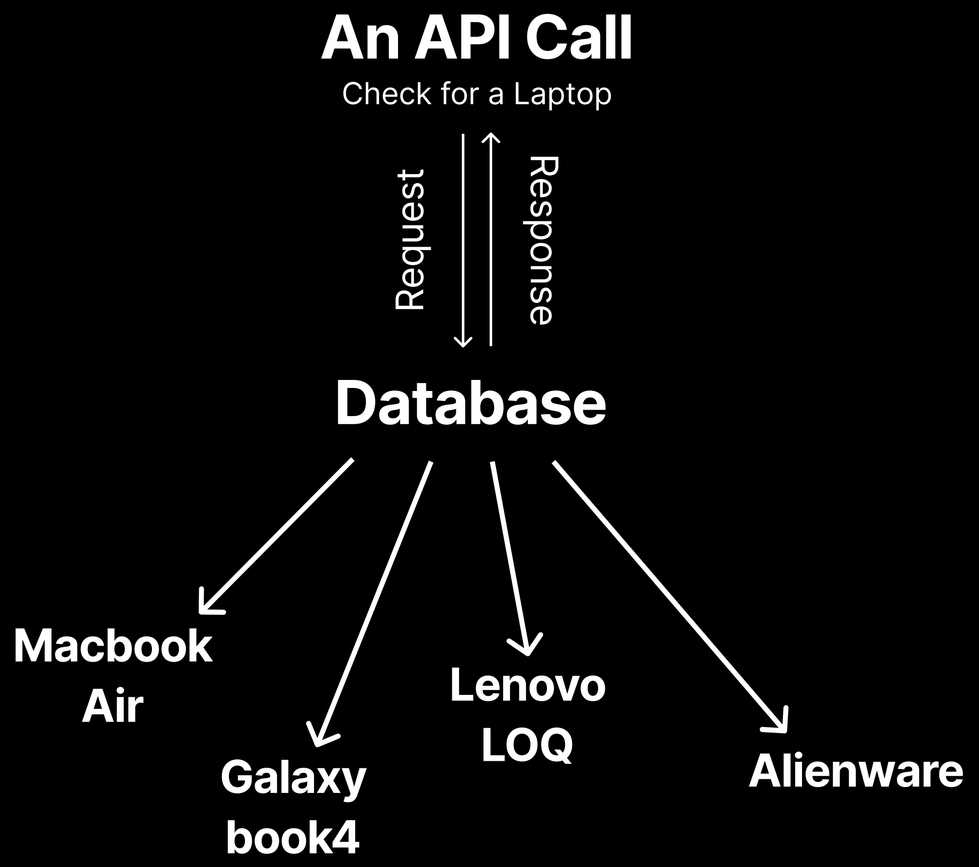
# E-Commerce Website

(Over Simplified)

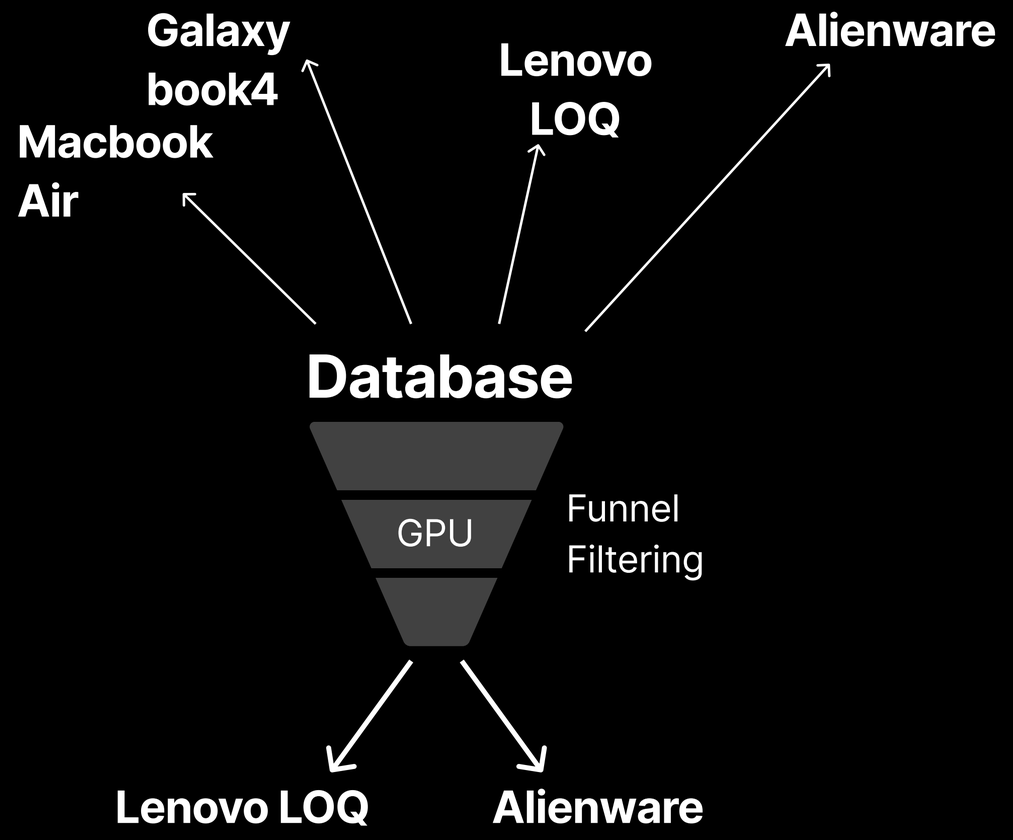
# Query in *Broken* Words



# Check in Database



# Filter Out The Laptops



# What Yahoo Did?

Just repeat it a few times

# Show me a good laptop for coding

## Breaking Words

[Good, Laptop, Developers]

## Frequencies

list of product containing x word
**"laptop"** → [P 1, P 7, P 120...]
**"coding"** → [P 4, P 7, P 99...]

## Ranking

The connection between TWO or MORE things

# Yahoo Would Give You

**DNA Coding**

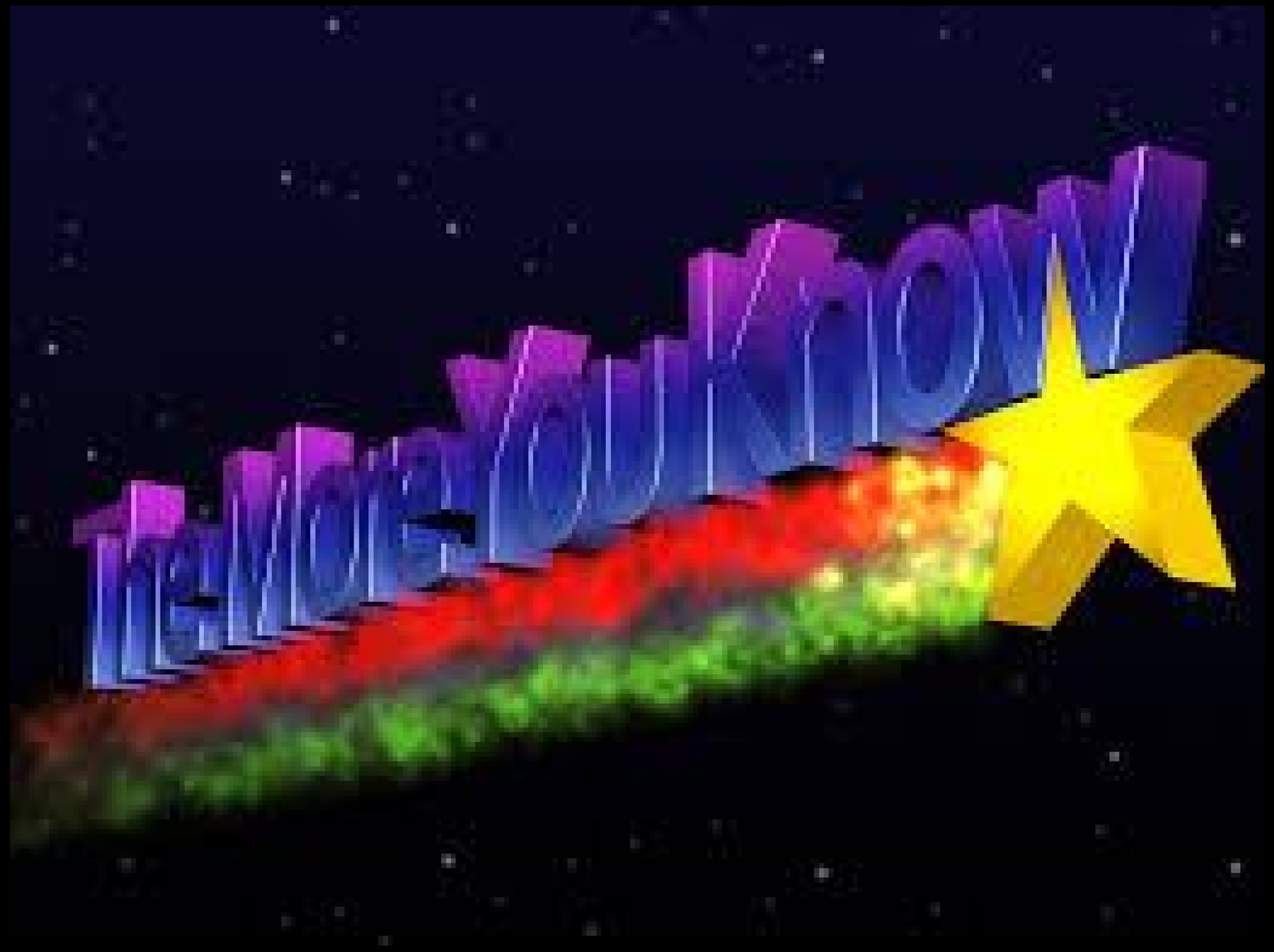
Coding: 200

1

**Macbook**

Coding: 50

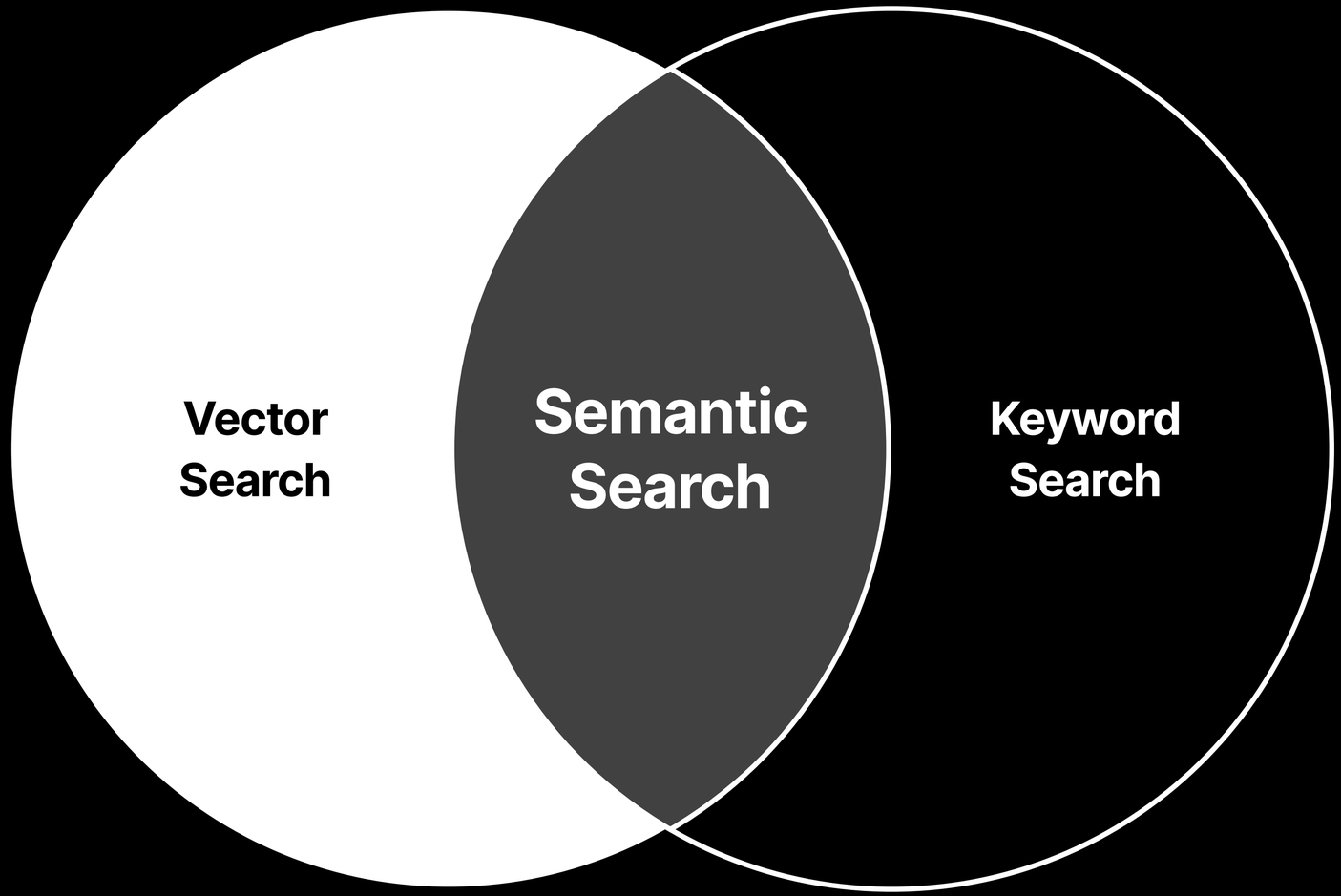
2


The More You Know

# The Conclusion

# Then What’s The Way To Go

We have to use the best of both world



So

# Be a Chef


WHERE'S THE LAMB SAUCE!?

# Thank You All!

# Feel Free to Connect

**Akshat Sharma**

Futuresoft India | Ex Intern @ DRDO | GDG Noida | ML and GenAI |