#FABCONSQLCON2026

FABCON
Microsoft Fabric
COMMUNITY CONFERENCE

SQLCON
Microsoft SQL
COMMUNITY CONFERENCE

ATLANTA MARCH 16 - 20, 2026

Building an AutoML Pipeline for Vector Data in Azure SQL

Ömer Çolakoğlu,
Microsoft Data Platform MVP
Founder, FEYOTECH
omer@feyotech.com
https://www.linkedin.com/in/omercolakoglu/

What We Will Cover Today

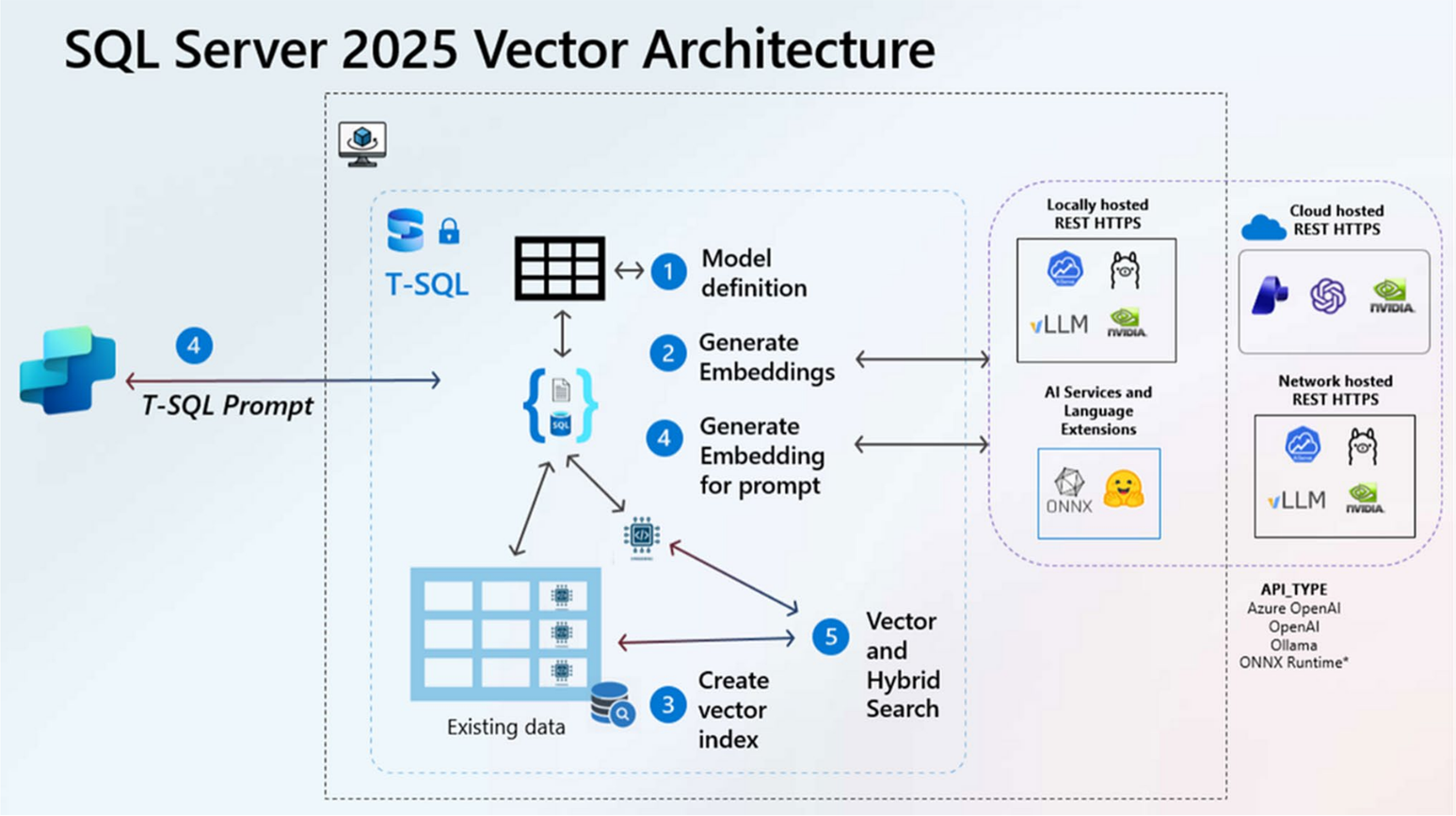
01 **The Problem**	Why is choosing the right embedding model so hard?
02 **Architecture**	How AutoEmbedding works end-to-end
03 **SQL Server 2025**	VECTOR type & sp_invoke_external_rest_endpoint
04 **Data Profiling**	10 domain types: review, ecommerce, book, knowledge_base...
05 **Model Evaulation**	Accuracy x Latency x Cost x RAG Quality weighted composite score
06 **New Features**	Smart Chunking, RAG Pipeline & PDF Import
07 **Rag & Results**	RAG evaluation results & model recommendations

We are happy!

SQL Server 2025 has released.

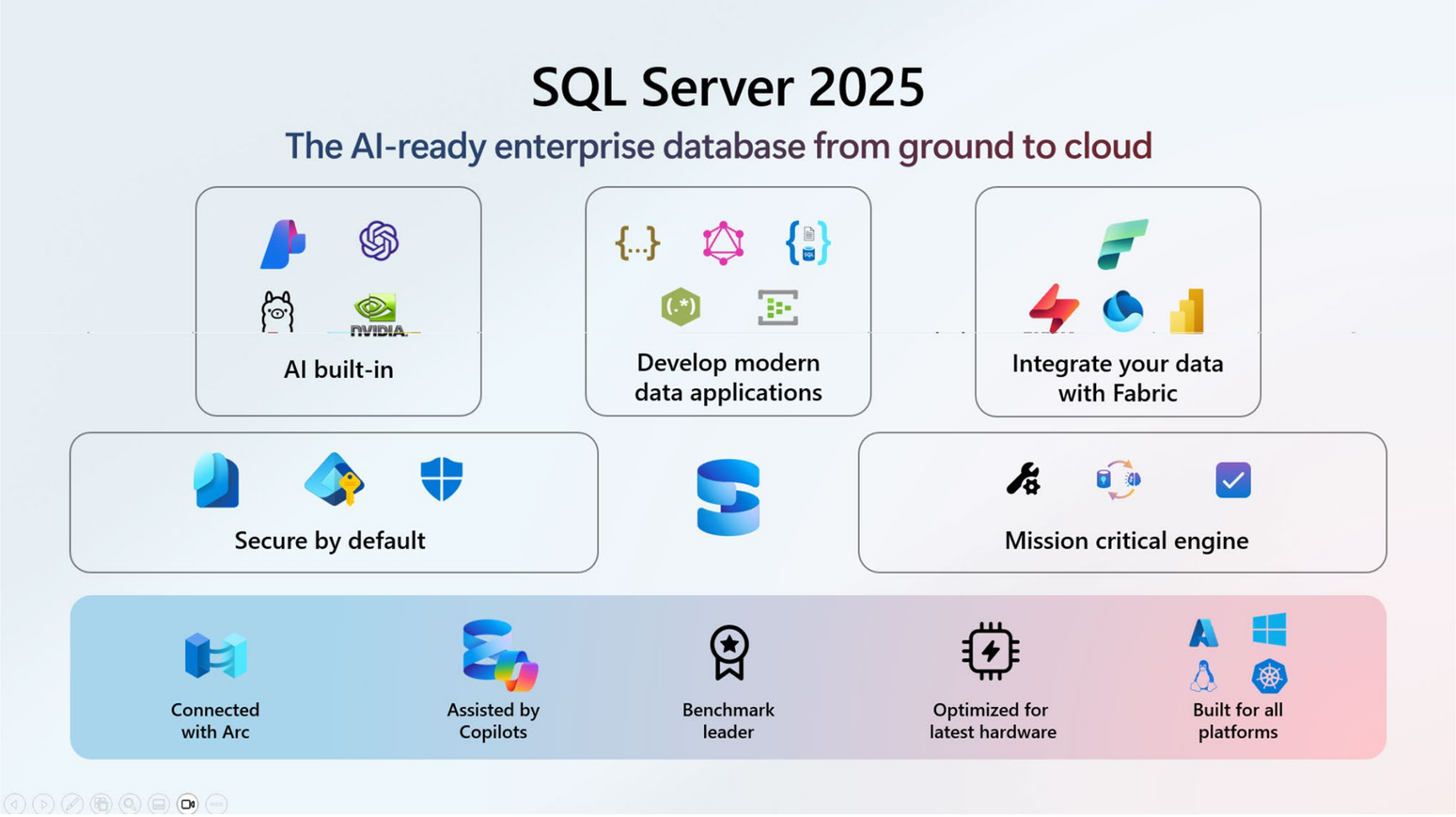
We are happy!

Native vector type is built in.



We are happy!

We have an AI Ready Enterprise Database Platform

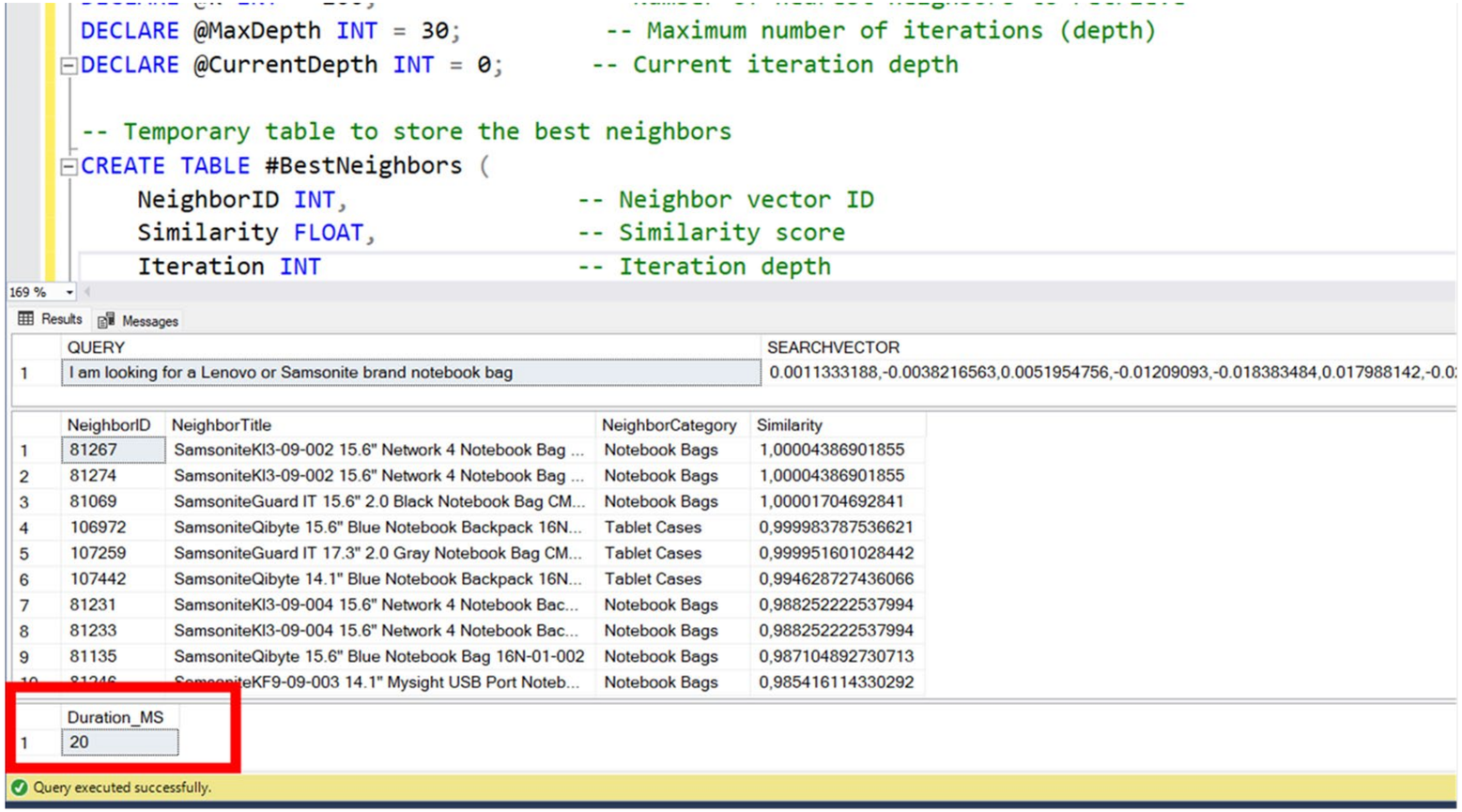


We are happy!

We can embed data with any LLM Embedding Model with sp_invoke_external_rest_endpoint

We are happy!

We can use DiskANN and search increadable fast.



```
DECLARE @MaxDepth INT = 30;          -- Maximum number of iterations (depth)
DECLARE @CurrentDepth INT = 0;       -- Current iteration depth

-- Temporary table to store the best neighbors
CREATE TABLE #BestNeighbors (
    NeighborID INT,                  -- Neighbor vector ID
    Similarity FLOAT,                -- Similarity score
    Iteration INT                    -- Iteration depth
```

169 %

Results | Messages

	QUERY	SEARCHVECTOR
1	I am looking for a Lenovo or Samsonite brand notebook bag	0.0011333188,-0.0038216563,0.0051954756,-0.01209093,-0.018383484,0.017988142,-0.0

	NeighborID	NeighborTitle	NeighborCategory	Similarity
1	81267	SamsoniteKI3-09-002 15.6" Network 4 Notebook Bag ...	Notebook Bags	1,00004386901855
2	81274	SamsoniteKI3-09-002 15.6" Network 4 Notebook Bag ...	Notebook Bags	1,00004386901855
3	81069	SamsoniteGuard IT 15.6" 2.0 Black Notebook Bag CM...	Notebook Bags	1,00001704692841
4	106972	SamsoniteQibyte 15.6" Blue Notebook Backpack 16N...	Tablet Cases	0,999983787536621
5	107259	SamsoniteGuard IT 17.3" 2.0 Gray Notebook Bag CM...	Tablet Cases	0,999951601028442
6	107442	SamsoniteQibyte 14.1" Blue Notebook Backpack 16N...	Tablet Cases	0,994628727436066
7	81231	SamsoniteKI3-09-004 15.6" Network 4 Notebook Bac...	Notebook Bags	0,988252222537994
8	81233	SamsoniteKI3-09-004 15.6" Network 4 Notebook Bac...	Notebook Bags	0,988252222537994
9	81135	SamsoniteQibyte 15.6" Blue Notebook Bag 16N-01-002	Notebook Bags	0,987104892730713
10	81246	SamsoniteKF9-09-003 14.1" Mysight USB Port Noteb...	Notebook Bags	0,985416114330292

	Duration_MS
1	20

Query executed successfully.

We are happy!

We can search data as semantic and any language without translation.

```
--English
exec CALCULATE_SIMILARITY  N'I am looking for a Lenovo or Samsonite brand notebook bag'
--Turkish
exec CALCULATE_SIMILARITY  N'Lenovo ya da Samsonite markalı bir notebook çantası arıyorum.'
--German
exec CALCULATE_SIMILARITY  N'Ich suche eine Notebooktasche der Marke Lenovo oder Samsonite.'
--Japanese
exec CALCULATE_SIMILARITY  N'Lenovo または Samsonite ブランドのノートブックバッグを探しています'
```

155 %

Results | Messages

4	107329	0,880008816695477	0,119991183304523	Samsonite Guard IT 17.3 2.0 Black Notebook Bag CM5-0...
5	80764	0,879874537313456	0,120125462686544	Samsonite Guard IT 15.6 2.0 Blue Notebook Bag CM5-01...

	VECTOR_ID	COSINE_SIMILARITY	COSINE_DISTANCE	DESCRIPTION
7	80553	0,82177328430986	0,17822671569014	Samsonite Qibyte 13-14 Anthracite Backpack Notebook ...
8	106801	0,820912021712685	0,179087978287315	Samsonite Guard IT 17.3 2.0 Black Notebook Backpack ...
9	81080	0,820341783269561	0,179658216730439	Samsonite CM5-01-002 13.3 Guart It 2.0 Notebook Bag ...
10	107298	0,820284932247163	0,179715067752837	Samsonite Guard IT 17.3 2.0 GrayNotebook Bag CM5-0...
11	107259	0,820284932247163	0,179715067752837	Samsonite Guard IT 17.3 2.0 GrayNotebook Bag CM5-0...

	VECTOR_ID	COSINE_SIMILARITY	COSINE_DISTANCE	DESCRIPTION
7	80553	0,819871207567573	0,180128792432427	Samsonite Qibyte 13-14 Anthracite Backpack Notebook ...
8	106800	0,819547529045956	0,180452470954044	Samsonite Guard IT 15.6 2.0 Black Notebook Bag CM5-...
9	81069	0,819547529045956	0,180452470954044	Samsonite Guard IT 15.6 2.0 Black Notebook Bag CM5-...
10	107232	0,819495368800375	0,180504631199625	Samsonite 16N-09-002 15.6 Qibyte Notebook Bag Anth...

We are happy!

Vector stores data as semantic. It means, Vector is a format that stores data as some float values. These are not meanful for human but very meanful for computers.

```
1
2   SELECT * FROM STACKOVERFLOW.DBO.SoVectors_Raw
```

No issues found — Ln: 2, Ch: 1 | TABS | CRLF | Windows 1

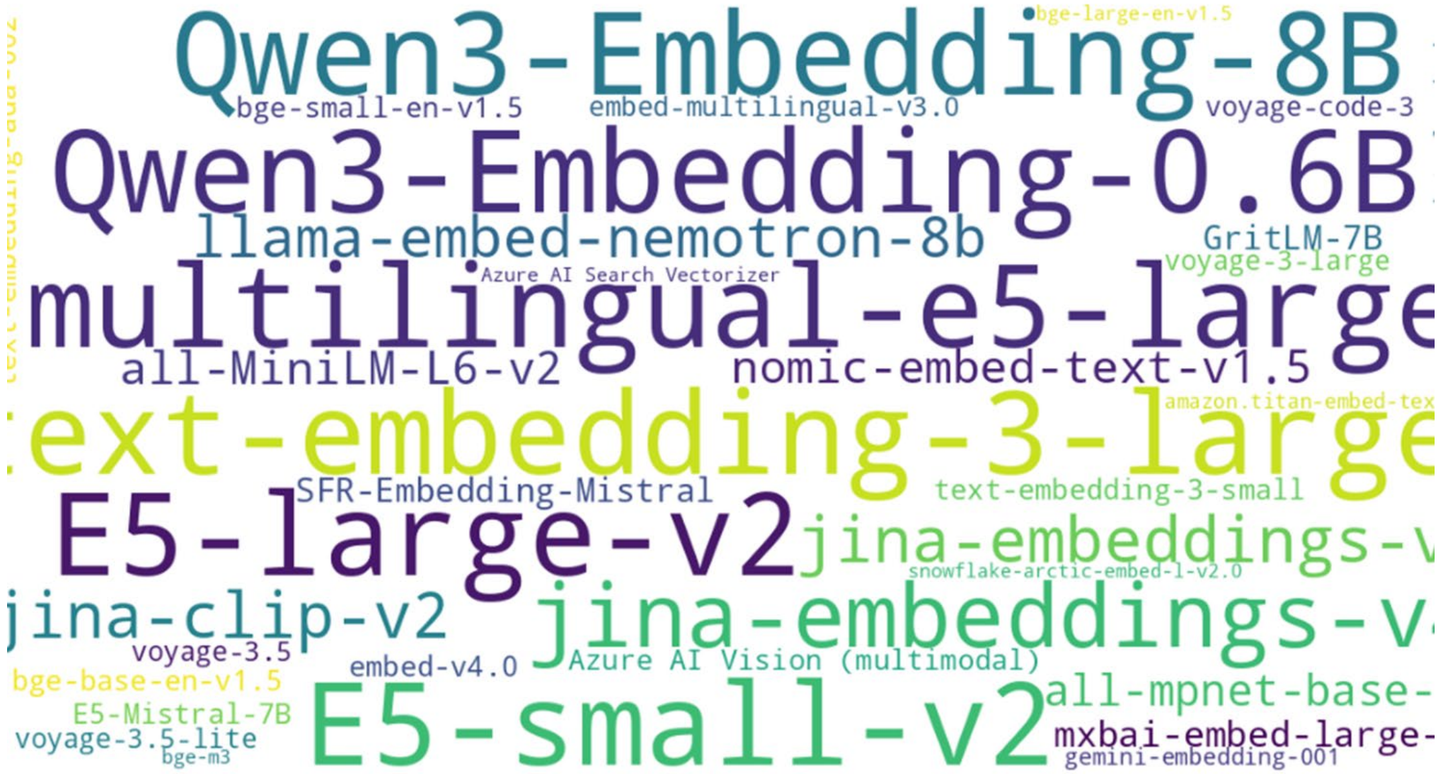
Results | Messages

	CleanBody	EmbeddingString
	Multi checks in a try catch statement - is this okay? Is ...	[-0.017754707,-0.008569204,0.01587958,0.039653055,0.0048320605,-0.017990738,0.0027946
	Load data from the MySql db using pyspark in python ...	[0.016167656,-0.008187214,0.05368213,-0.020696962,0.0634103,0.019751716,0.016226733,-
	The error is telling you that you must have initialised ...	[-0.00018432674,-0.025595853,0.029040208,0.00880463,-0.014164905,-0.013917342,0.06867
	Well, you can do something similar using some trick...	[0.024568874,-0.017945254,-0.00082577753,-0.013792442,-0.0092742285,0.032712102,-0.01
	If you are lazy and don't want to fight with string litera...	[-0.021657422,0.0007445972,0.010089045,-0.025424788,0.030789832,0.051638555,-0.009748
	Use the unicodeScalars view on the string: let mess...	[0.048910353,-0.026068378,0.045637194,0.011695701,-0.0041119065,0.002413955,0.038973
	Get TEMP directory path (case sensitive) If I use tem...	[0.020758657,-0.005549775,0.050601404,0.0076494543,0.007661071,0.018214647,0.0210955
	realm.copyFromRealm() was added in 0.87.0 (and i...	[-0.027693728,0.03180546,0.048421696,-0.027669542,0.011724484,0.035917193,0.02009911
	Here you have it twice! <include android:id="@+id...	[0.019188706,-0.021387177,0.00015211376,-0.015828991,0.011200928,-0.0049719266,0.0276
0	If you would like the whole name to be alphabetic yo...	[0.04653978,2.5017383e-05,0.016109923,0.031452708,0.006352145,-0.003510243,-0.0240137
1	More precisely, as explained in " How to find oid an...	[0.0037280025,0.035373796,0.030547904,0.010586803,-0.027266296,0.011630403,-0.002245
2	I tried it, I loaded a 2nd tree into the same table. On ...	[-0.013380551,-0.013741124,0.036686704,-0.016048795,-0.010627082,0.0020814918,-0.0220
3	How can I DELETE from a CTE in which a JOIN exist...	[-0.0047734026,0.047312234,0.015243996,-0.023869984,-0.0013366715,0.036096074,-0.0092
4	tl;dr LocalDate.now(ZoneId.of("Pacific/Auckland"))...	[-0.043986395,0.029783892,0.07368755,0.0074183955,0.03576825,0.0076872776,0.03631980
5	A string with a number in it is still a valid string - it's a...	[0.042419948,-0.0007679287,-0.050229028,0.017035037,0.049969446,0.016440164,0.011107
6	You could use os.environ to access environment var...	[0.024365617,0.013111342,0.0023087484,0.0033701295,0.010420832,0.024097124,0.0191188
7	Bootstrap Weld in a Spring Boot environment. Does a...	[-0.012135839,-0.019713722,0.065428,0.027302644,-0.011473476,-0.0241959,-0.012135839,0

Questions in our brains???

Questions in our brains

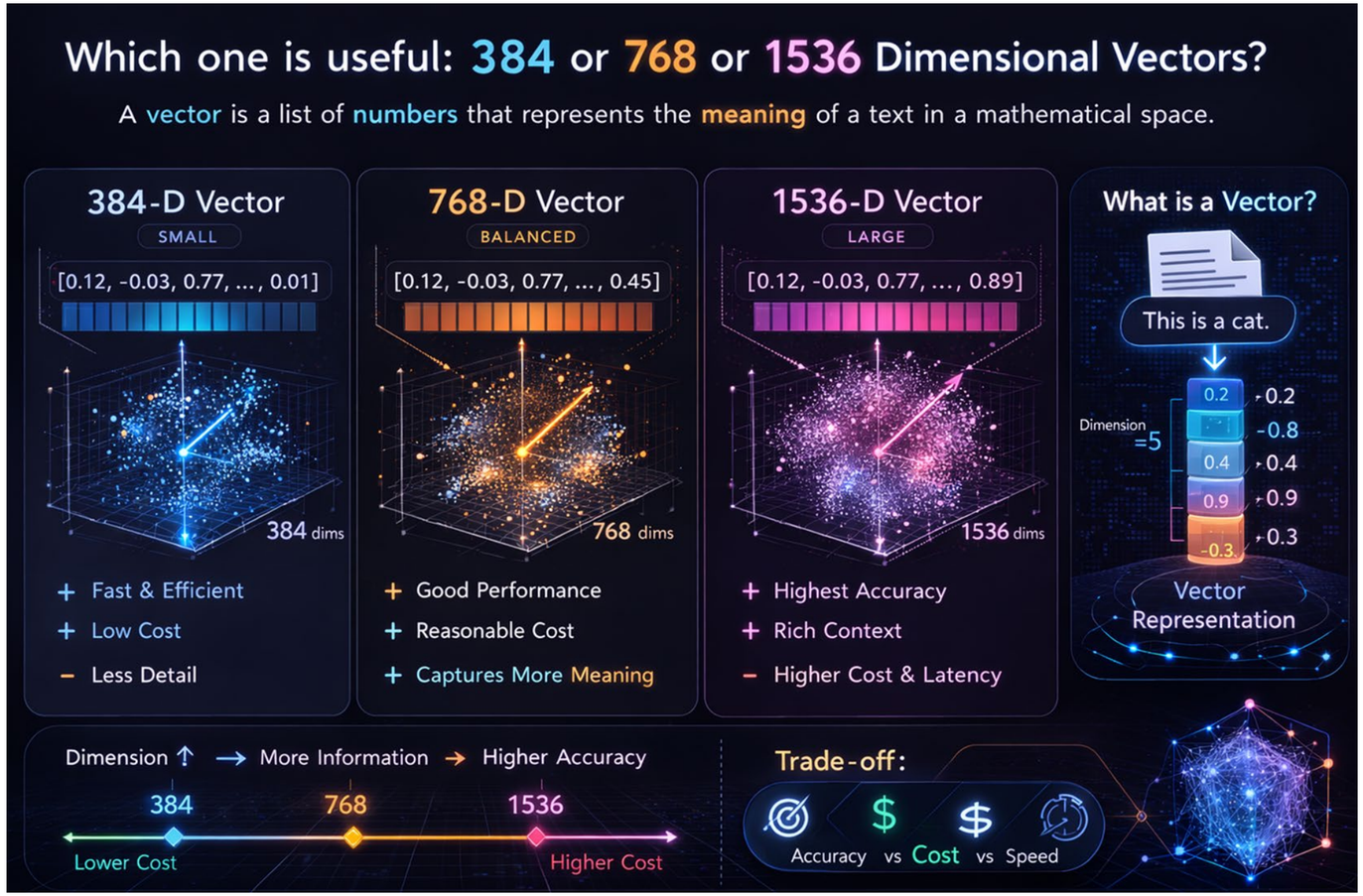
There are a lot of data types and a lot of llm embedding models. So which model is is the best for our scenario?



Questions in our brains

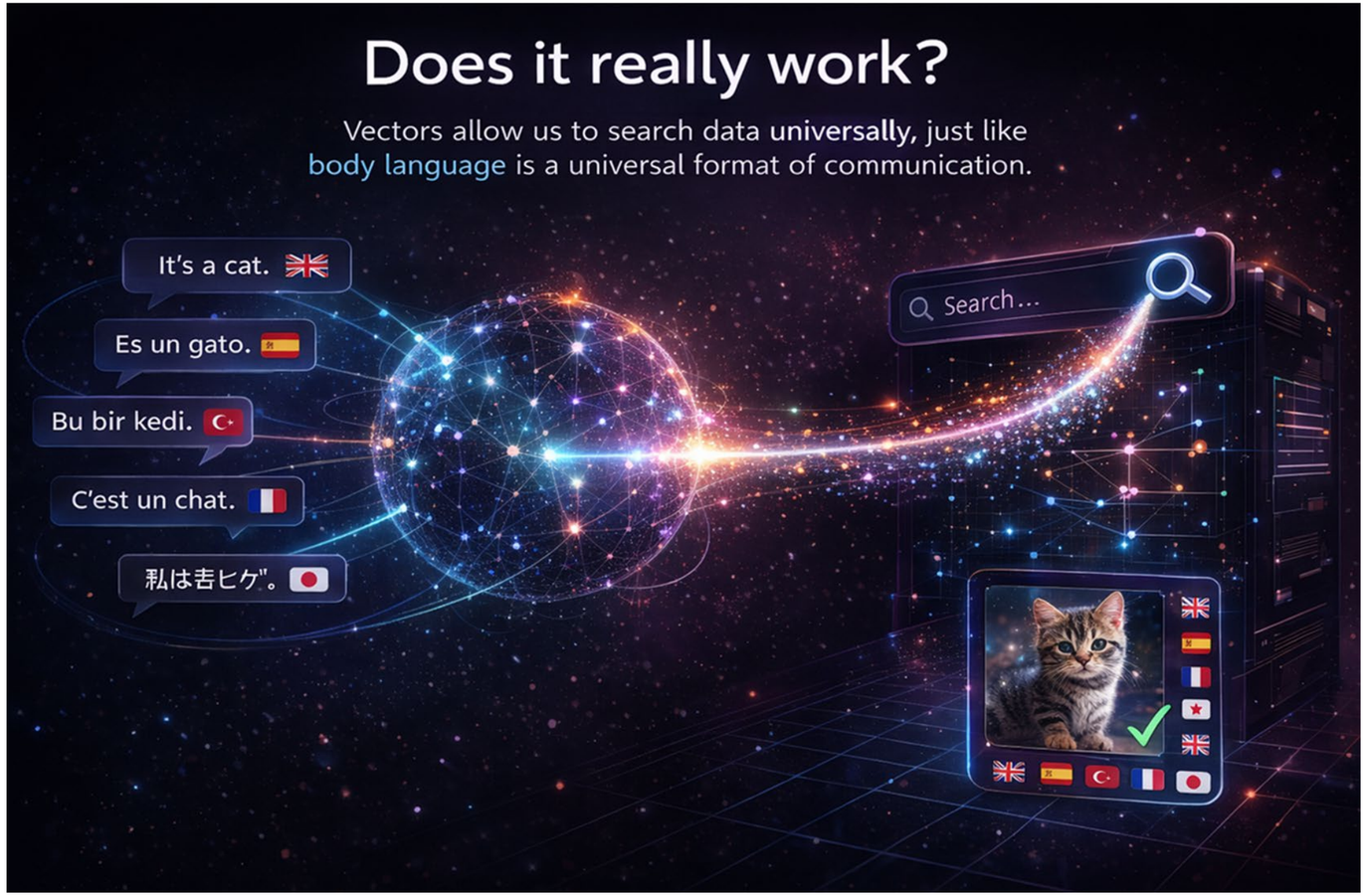
One of the most important features of the vector data type is its dimension.

Large dimensions are more expensive.
But does a larger dimension mean more accurate results?



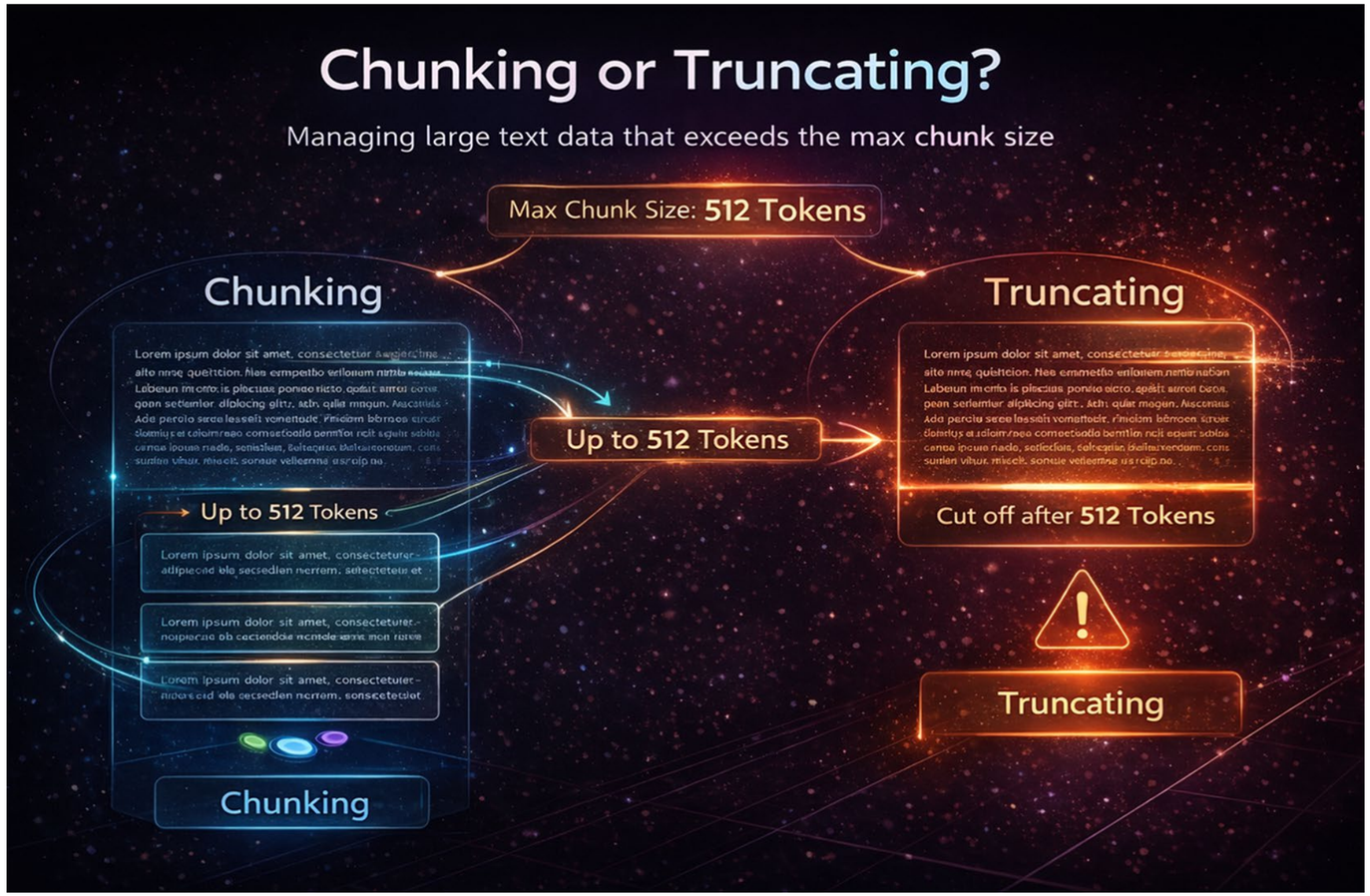
Questions in our brains

Vector is an universal format of data like body language.
So we can search data as vector in any language and get the same result.
Does it really work?



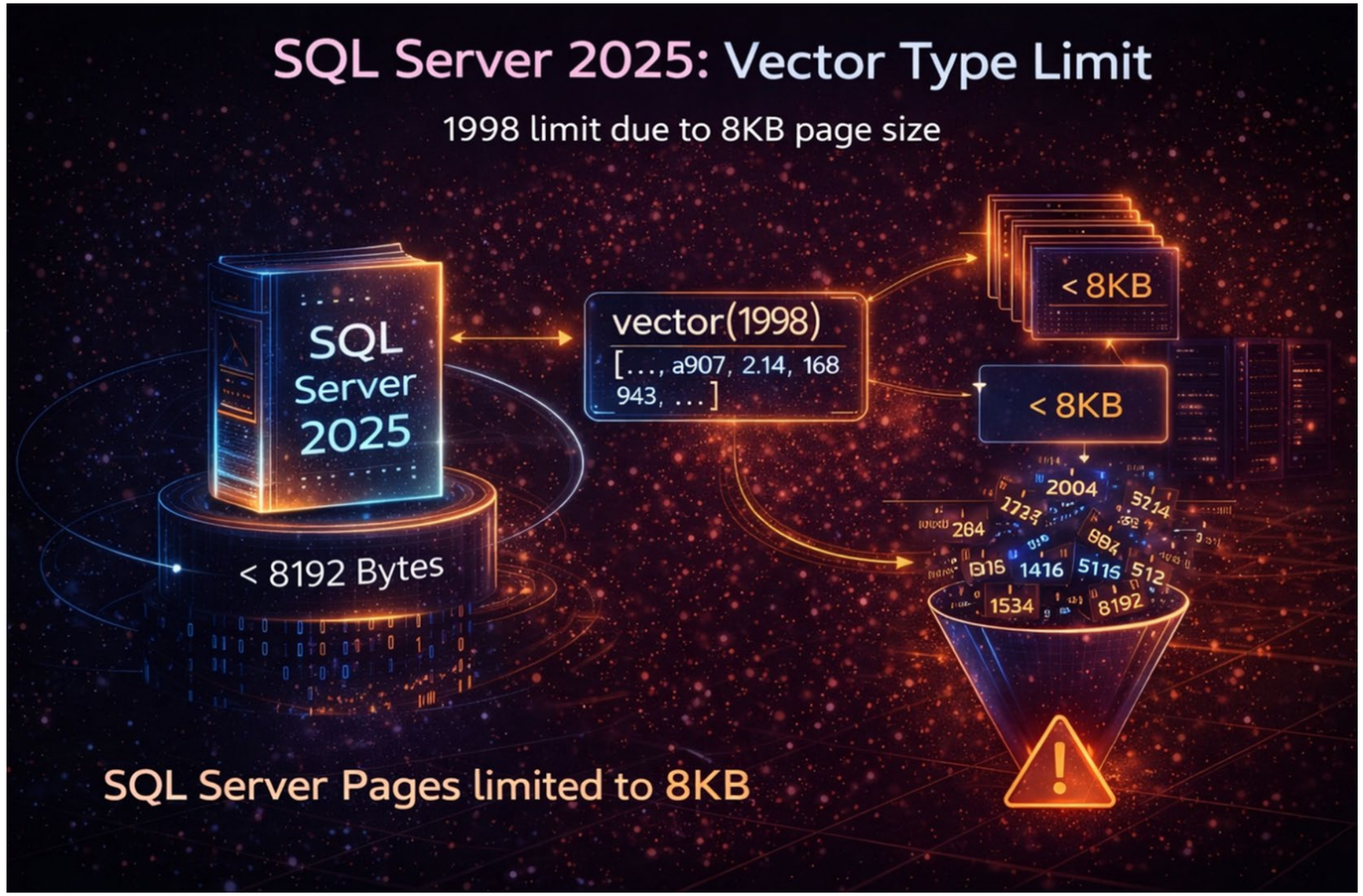
Questions in our brains

We have also big text data. And we have to chunk it.
Also we have a little bit bigger data than our max chunk size.
Chunking or truncating?



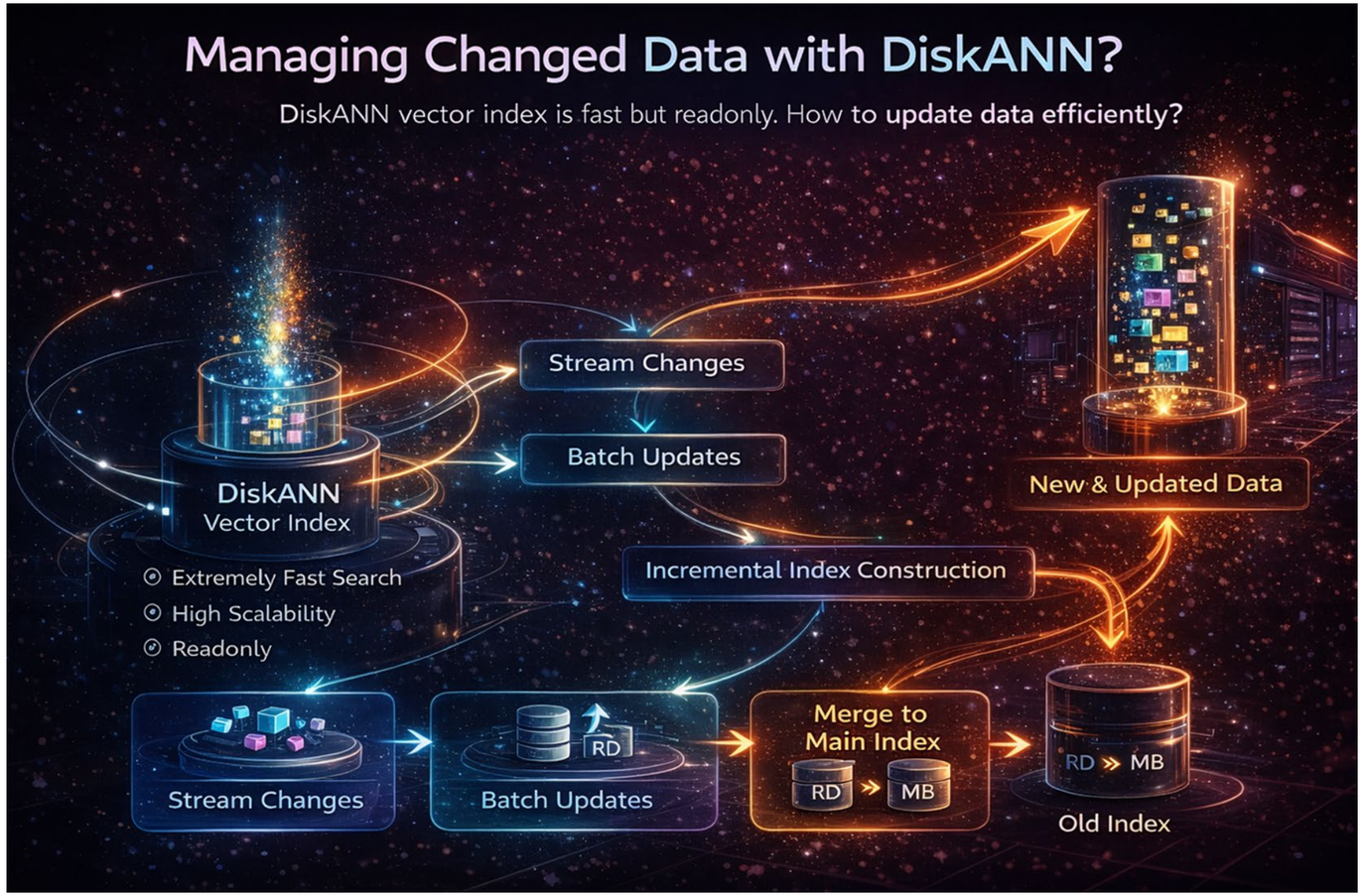
Questions in our brains

SQL Server 2025 has 1998 limit for the vector type, because of the 8KB page size.



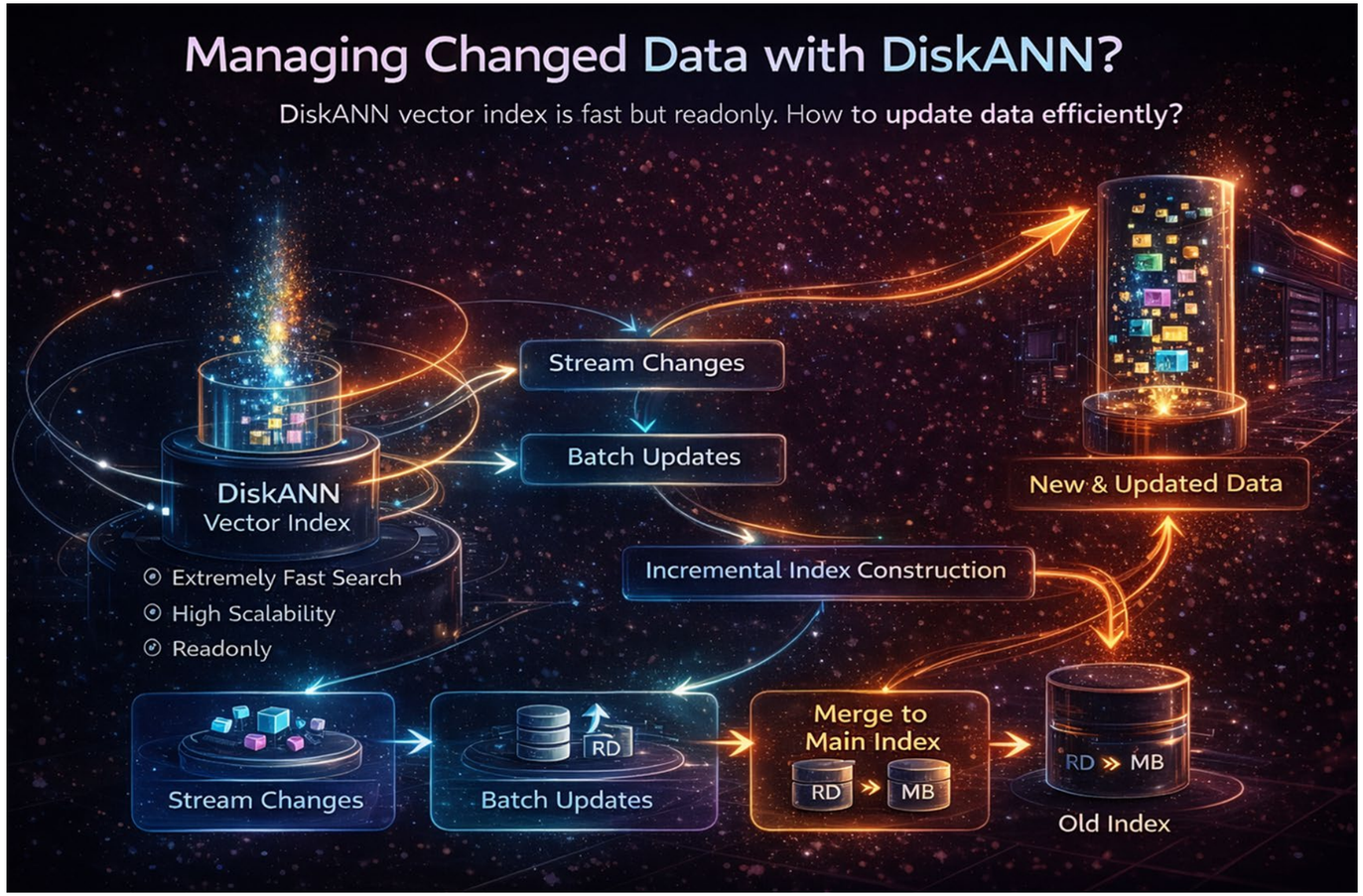
Questions in our brains

DiskANN vector index is really fast but it is readonly.
So, what should we do for the changed data.



Questions in our brains

DiskANN vector index is really fast but it is readonly.
So, what should we do for the changed data.



Why Is Embedding Model Selection So Hard?

Why Is Embedding Model Selection So Hard?

Real-World Challenges

01	30+ available models with different dimensions (768-3072)
02	Performance varies dramatically per domain
03	Cost: $0.00002 to $0.001 per 1K tokens
04	Latency: 50ms to 2000ms per request
05	Azure vs OpenAI vs Voyage vs Jina vs Ollama...
06	SQL Server integration requires custom code each time

How It's Done Today

- Manual benchmarking takes days
- Repeat for every new model release
- Excel spreadsheets for comparison
- Which metric matters most? How to weight?
- SQL integration is always custom per project
- Customer reviews != Product catalog

So, we need an intelligent pipeline.

End-to-End Pipeline

1. Datasource Management

- SQL Server connection / PDF import
- Datasource CRUD (add, edit, test)

2. Data Profile

- Language detection and distribution
- Type classification (10 domains, score-based)
- Text statistics (length, duplicate rate)
- Category extraction
- Column profiling (Text, Filter, Date, ID)
- Noise detection and cleaning (HTML, noise)
- Recommended models and strategy
- Text structure optimization (field detection, LLM suggestion)
- VectorStr formula definition

3. Question Management

- Multilingual question pool (TR, EN, DE, FR)
- Automatic question generation via LLM
- Expected keywords / ground truth

4.Experiment

- Model selection (multi-select)
- Strategy selection (recommended by data type)
- Weight settings (accuracy, latency, cost, RAG quality)
- Sampling count
- Chunk settings (size, overlap, semantic/fixed)
- RAG settings (enable/disable, top-K, max queries)
- Live progress tracking + stop control

5.Analysis

- Model metrics (overall score, accuracy, latency, cost, memory)
- Consistency test results (similar/different pair analysis)
- Multilingual test (language pair similarity)
- Retrieval test (precision, top-K ranking)
- RAG quality details (context relevance, faithfulness, completeness)
- Latency performance indicator
- Full model comparison table
- Strategy simulation (change weights → see score impact)
- Live Query Test (ad-hoc search, LLM evaluation)

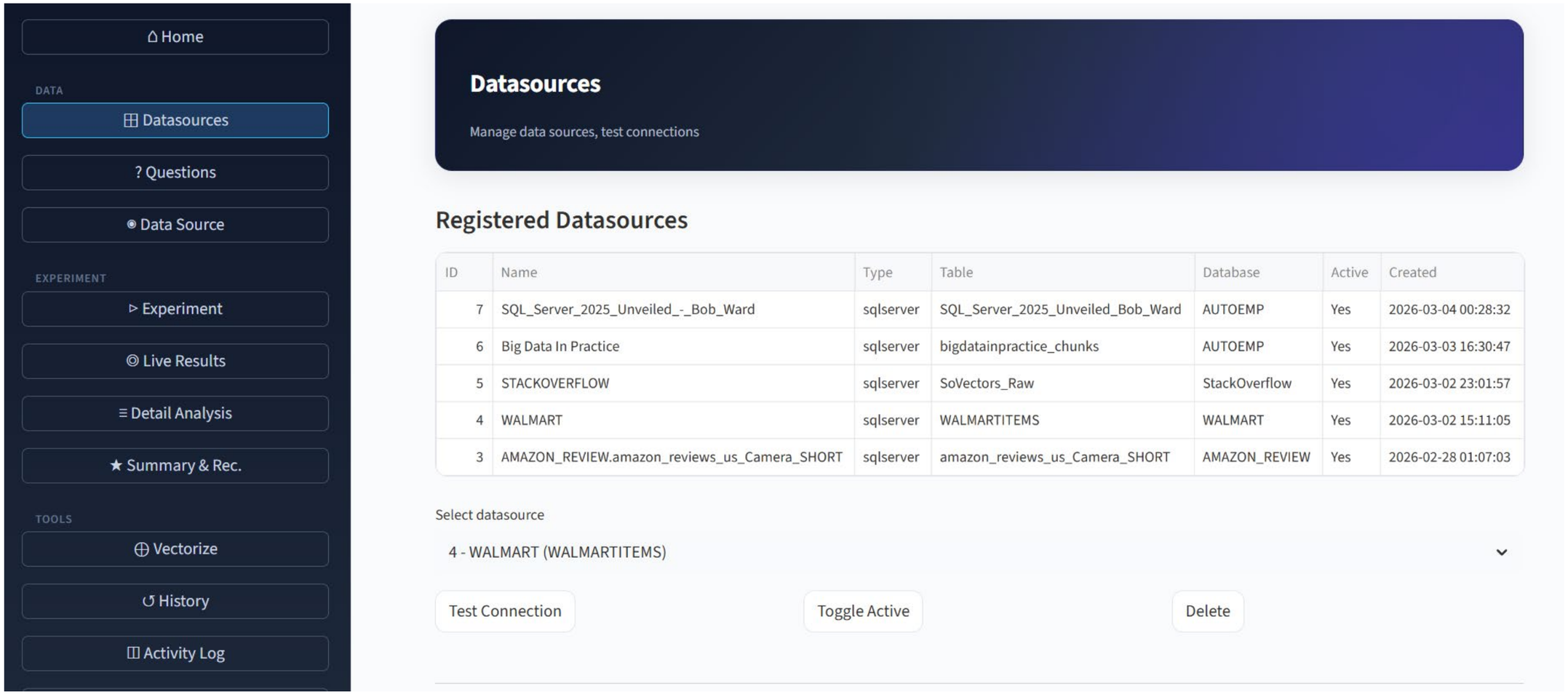
6.Results and Evaulation

- Winning model and score summary
- AI recommendation (LLM explanation)
- Key decision factors
- Model ranking table
- Vectorizing all data with the selected model

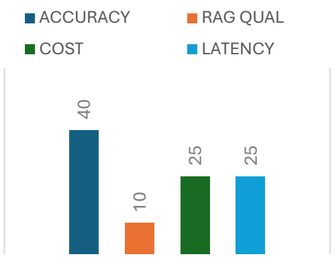
7.Tools

- A/B Experiment comparison
- Experiment history
- Activity log (audit trail)
- SQL Server status panel
- Multilingual query test (cross-language analysis)

Data Source Management

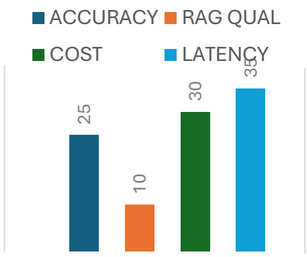


Data Profiling-Automatic Domain Detection



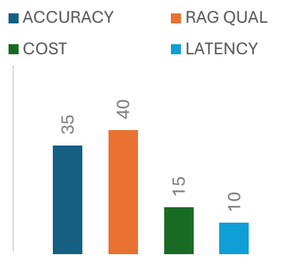
Reviews

- ***Amazon Camera Reviews (Used this Project)***
- Amazon Book Reviews
- Yelp Restaurant Reviews



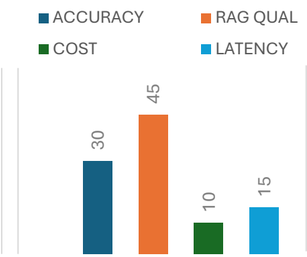
E Commerce

- ***Walmart Products (Used in this Project)***
- Shopify Store Inventory
- eBay Listings



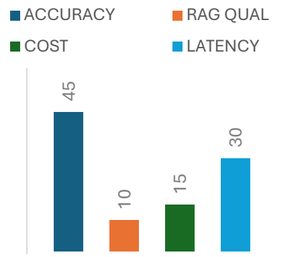
Book

- ***Big Data in Practice (Used in this Project)***
- ***SQL Server 2025 Unveiled (Used in this Project)***



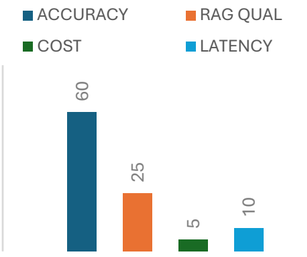
Knowledge Base

- StackOverflow Q&A (Used in this Project)
- Wikipedia Articles



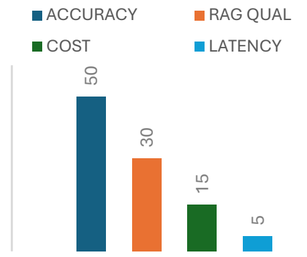
Code

- GitHub Repository Docs
- API Documentation
- Jupyter Notebook Cells



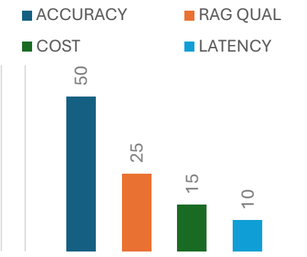
Medical

- PubMed Abstracts
- Clinical Trial Reports
- Drug Interaction Database



Scientific

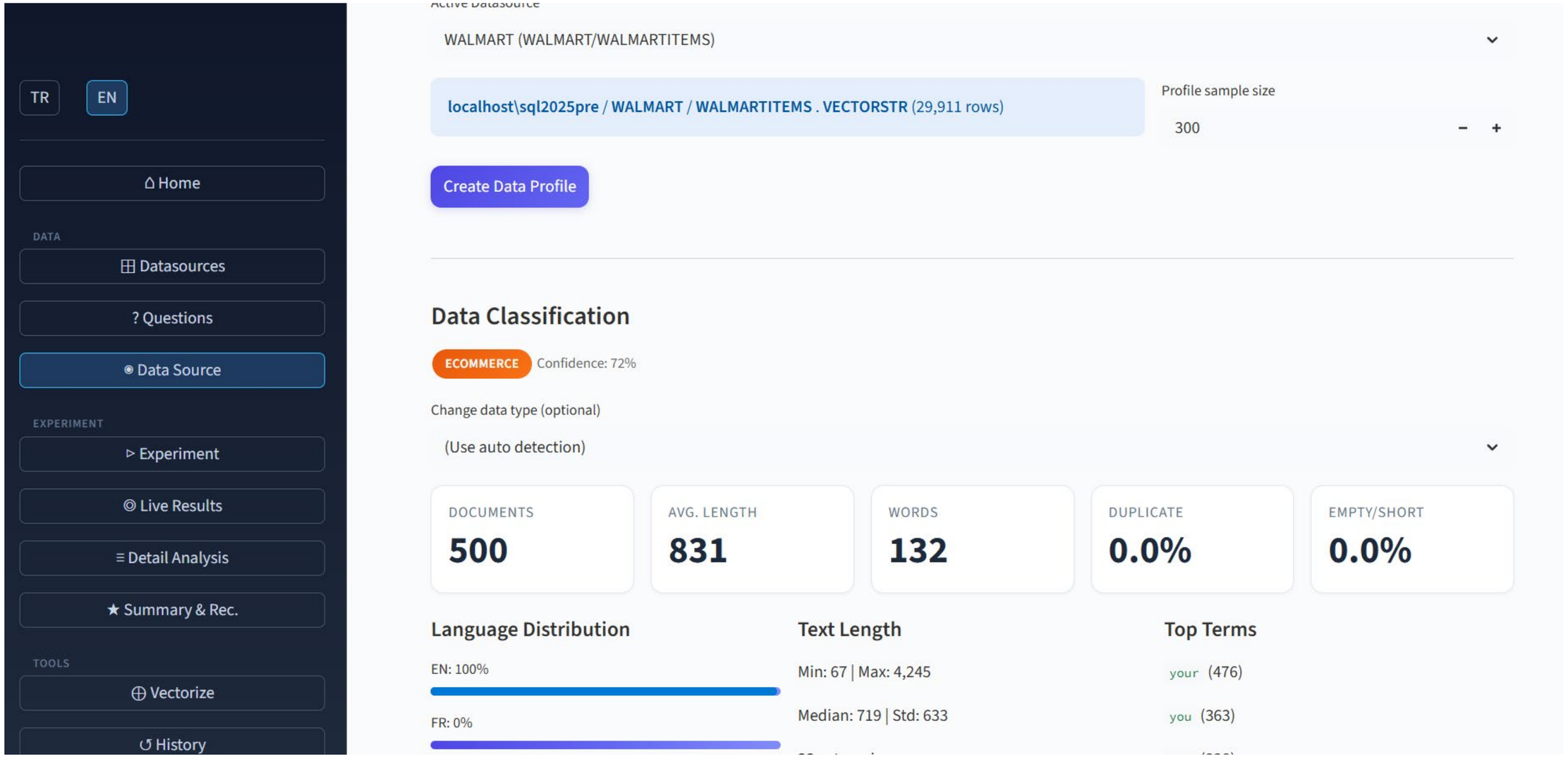
- Nature Publications
- Research Datasets (UCI)
- Physics Textbook



Multilingual

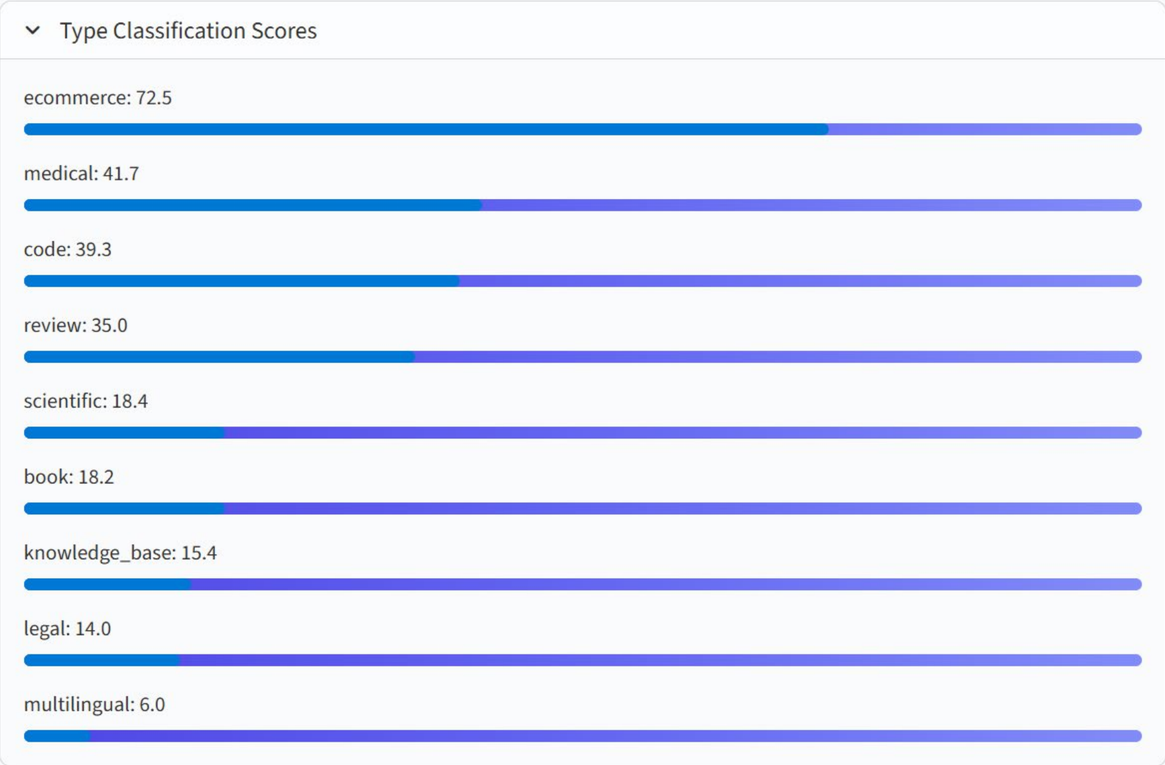
- Multilingual Wikipedia TR
- Subtitle Databases

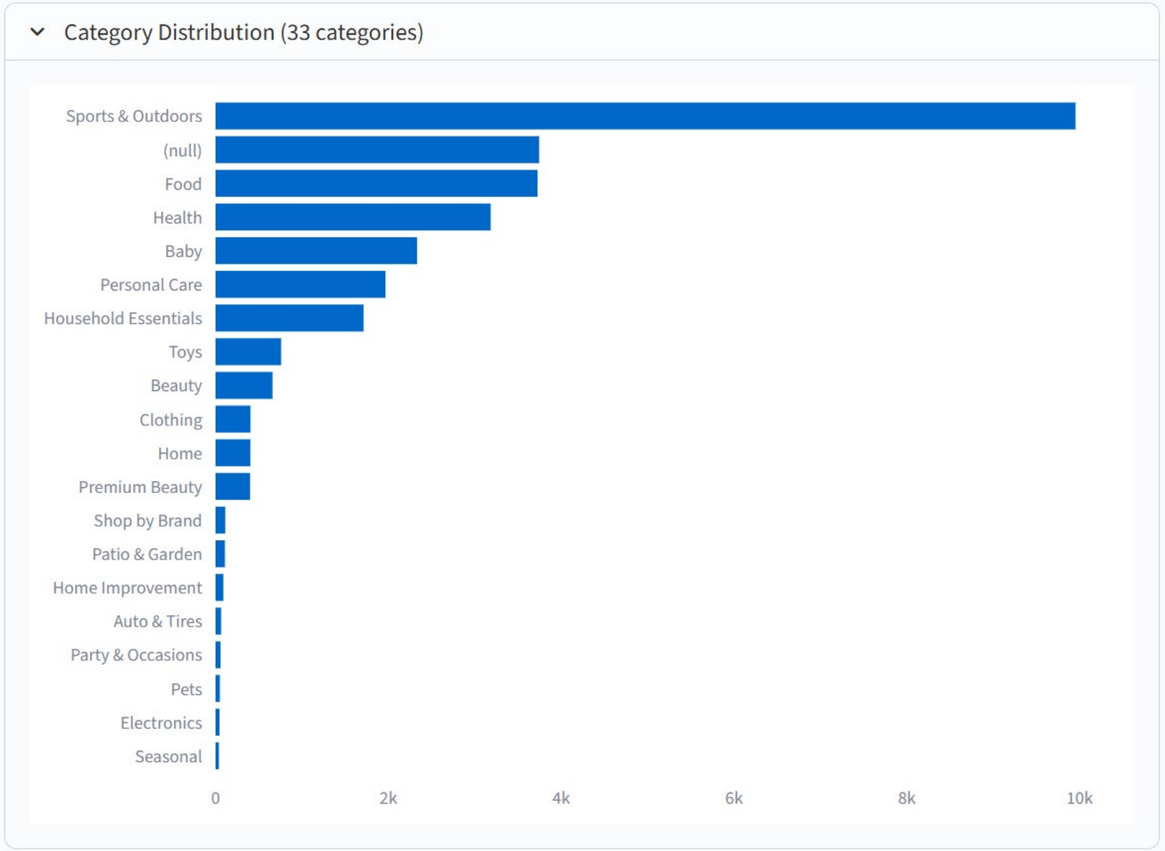
Create Data Profile



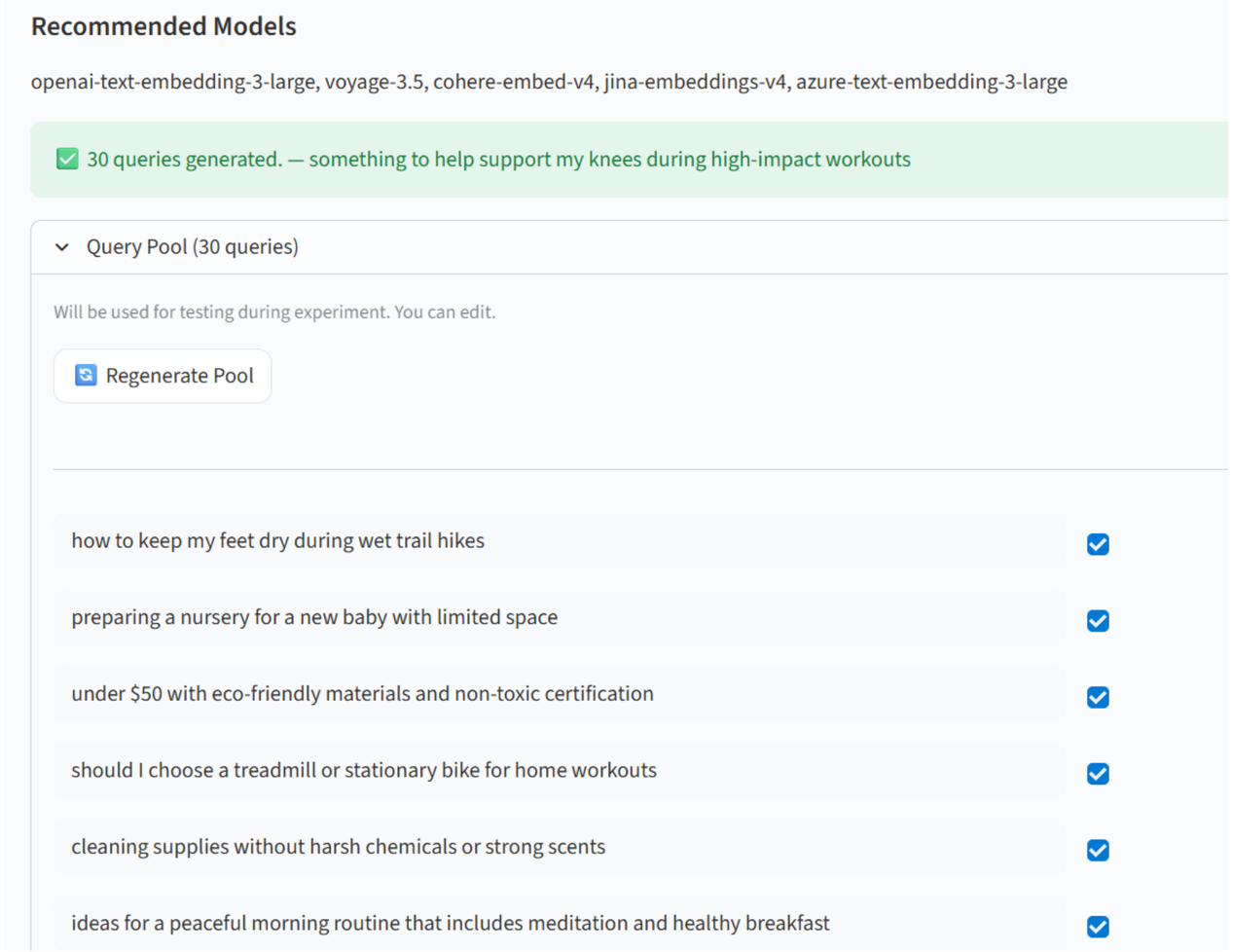
Data Profiling-Automatic Domain Detection | Create Data Profile

Type Classification, Category Distribution





Recomended Models, Query Pool



Data Profile & Search Strategy

Data Profile & Search Strategy — Analyze table structure, build VECTORSTR, determine search strategy

WALMARTITEMS (WALMART)

Profile Table

Column Profile

Column	Type	Role	Null %	Distinct	Avg. Length	Samples
ID	int	id	0%	29,911	-	-
GUID	nvarchar	id	0%	29,911	32	324271dbbf6917de048bed7370b280f0, ab0bd0f6bc921863d9521ae528ff439c
TIMESTAMP	nvarchar	category	0%	25,361	25	2019-12-19 05:45:19 +0000, 2019-12-18 21:53:09 +0000
PRODUCTURL	nvarchar	text	0%	29,911	102	https://www.walmart.com/ip/Pelican-Maxim-100X-10-Sit-in-Recreational-Kayak/67716
PRODUCTNAME	nvarchar	text	0%	29,586	67	Riga Gold Smoked Sprats in Oil, 100g Jar, Penn State Nittany Lions Mascot Offici
DESCRIPTION	nvarchar	text	1%	28,551	900	New product Give the spiker a perfect set everytime allowing for instant feedbac
LISTPRICE	nvarchar	category	12%	7,225	5	17807, 29.90
SALEPRICE	nvarchar	category	12%	7,169	5	47362, 60.90
BRAND	nvarchar	category	0%	9,364	8	Miracle Noodle, Mizuno
ITEMNUMBER	float	numeric	74%	7,761	-	-

Text Columns

- PRODUCTURL
- PRODUCTNAME
- DESCRIPTION
- VECTORSTR

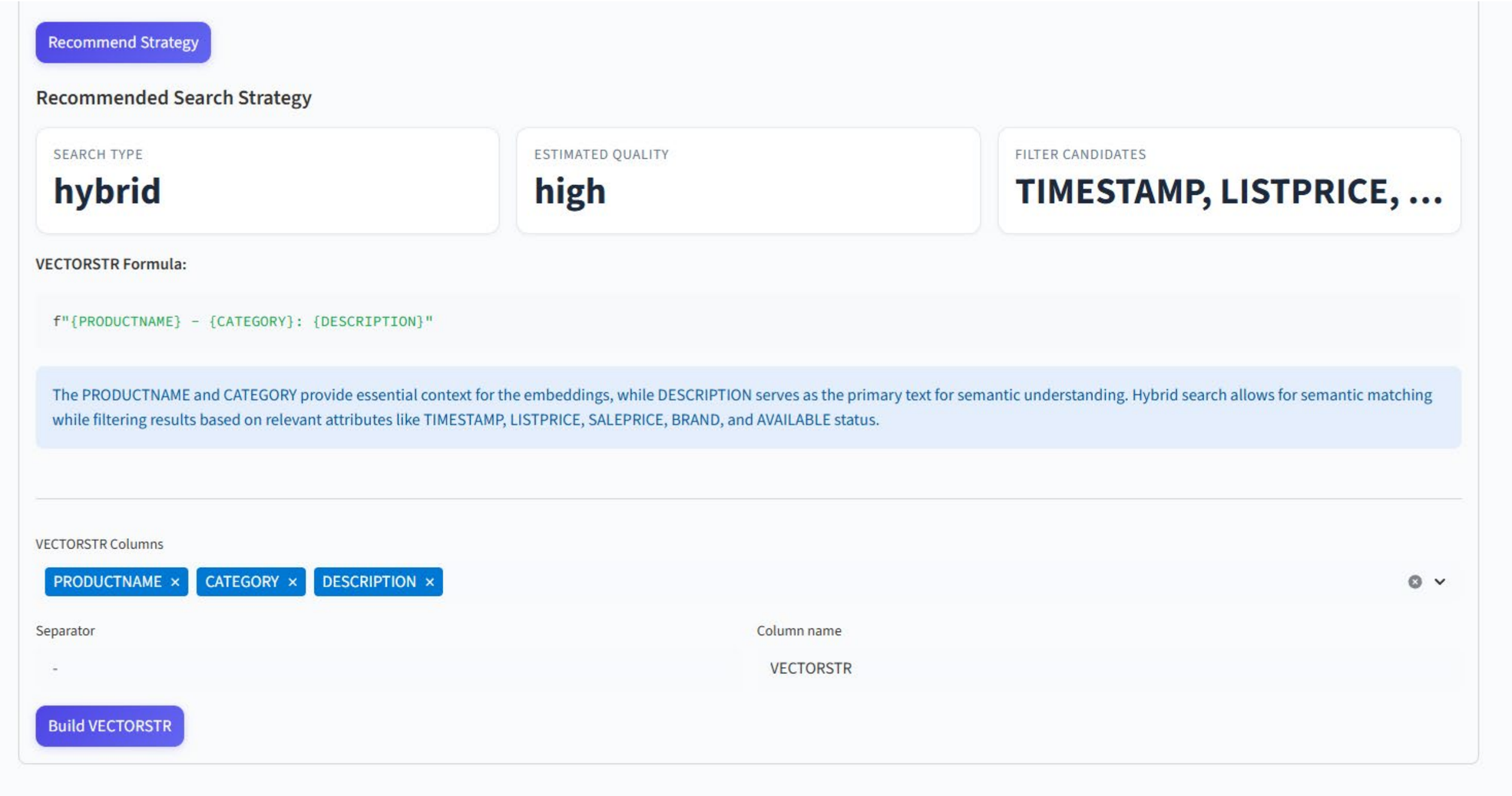
Filter Candidates

- TIMESTAMP
- LISTPRICE
- SALEPRICE
- BRAND

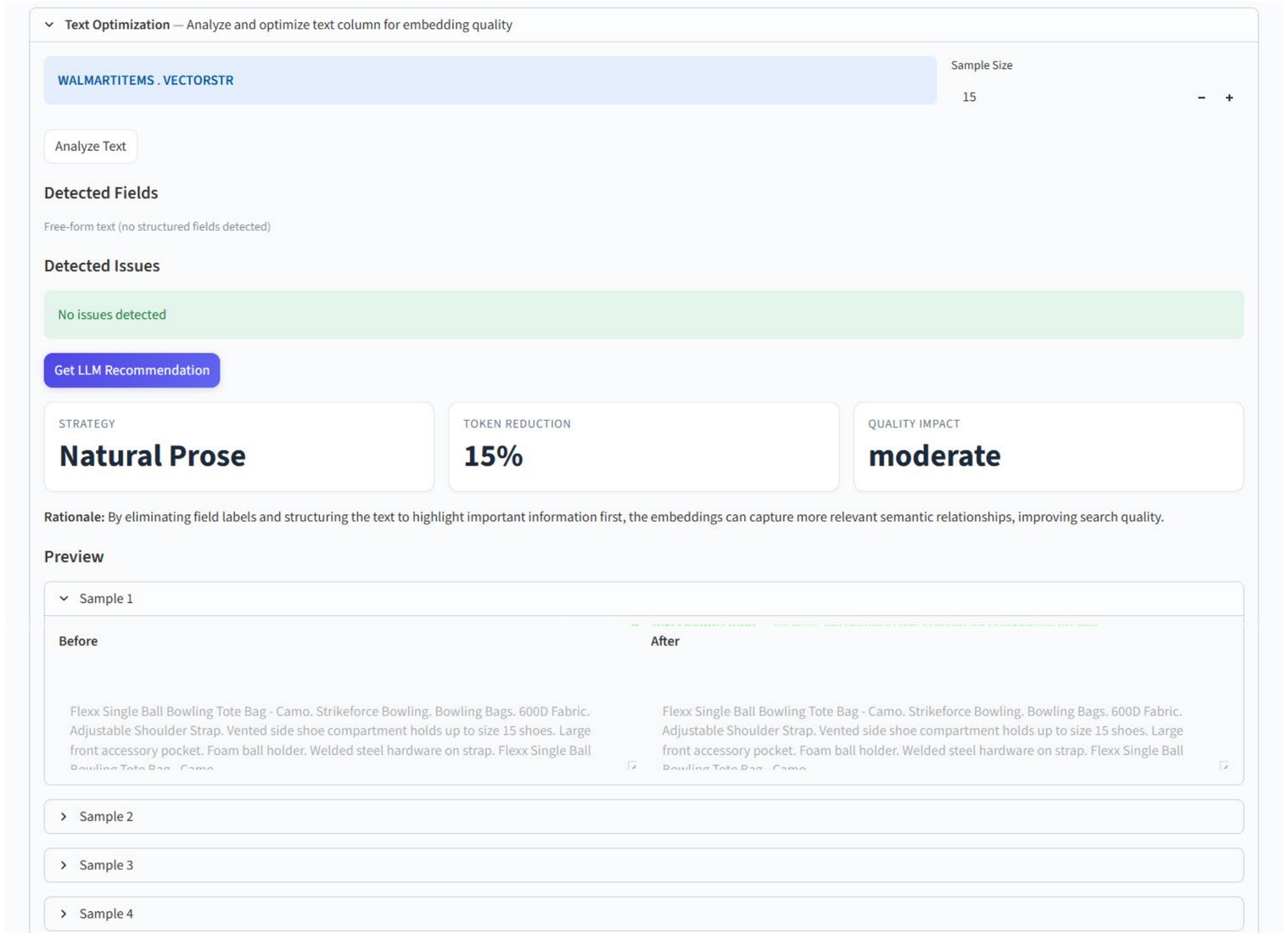
Date / ID

- ID
- GUID

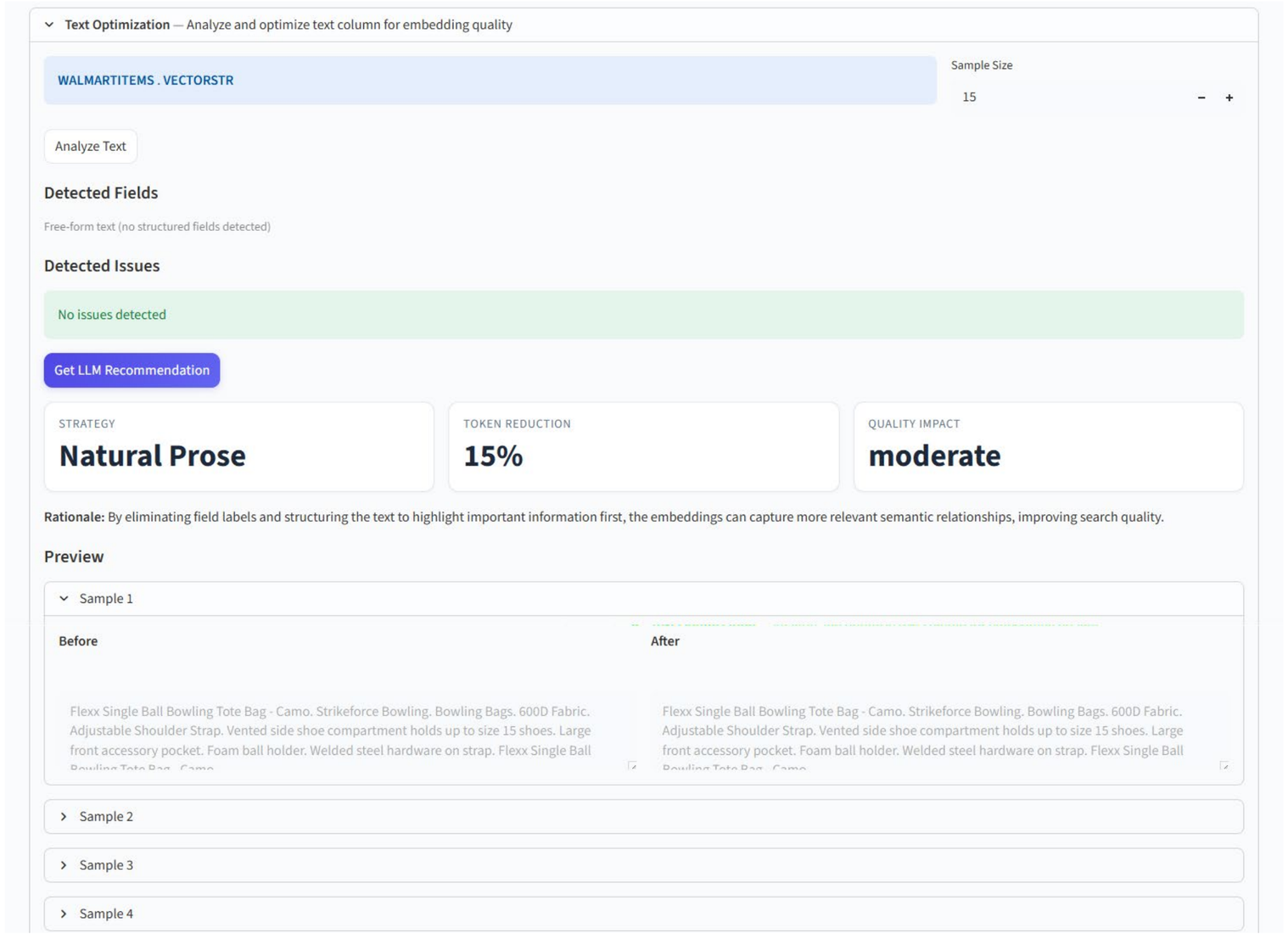
Recomended Search Strategy



Column to Embed Analysis

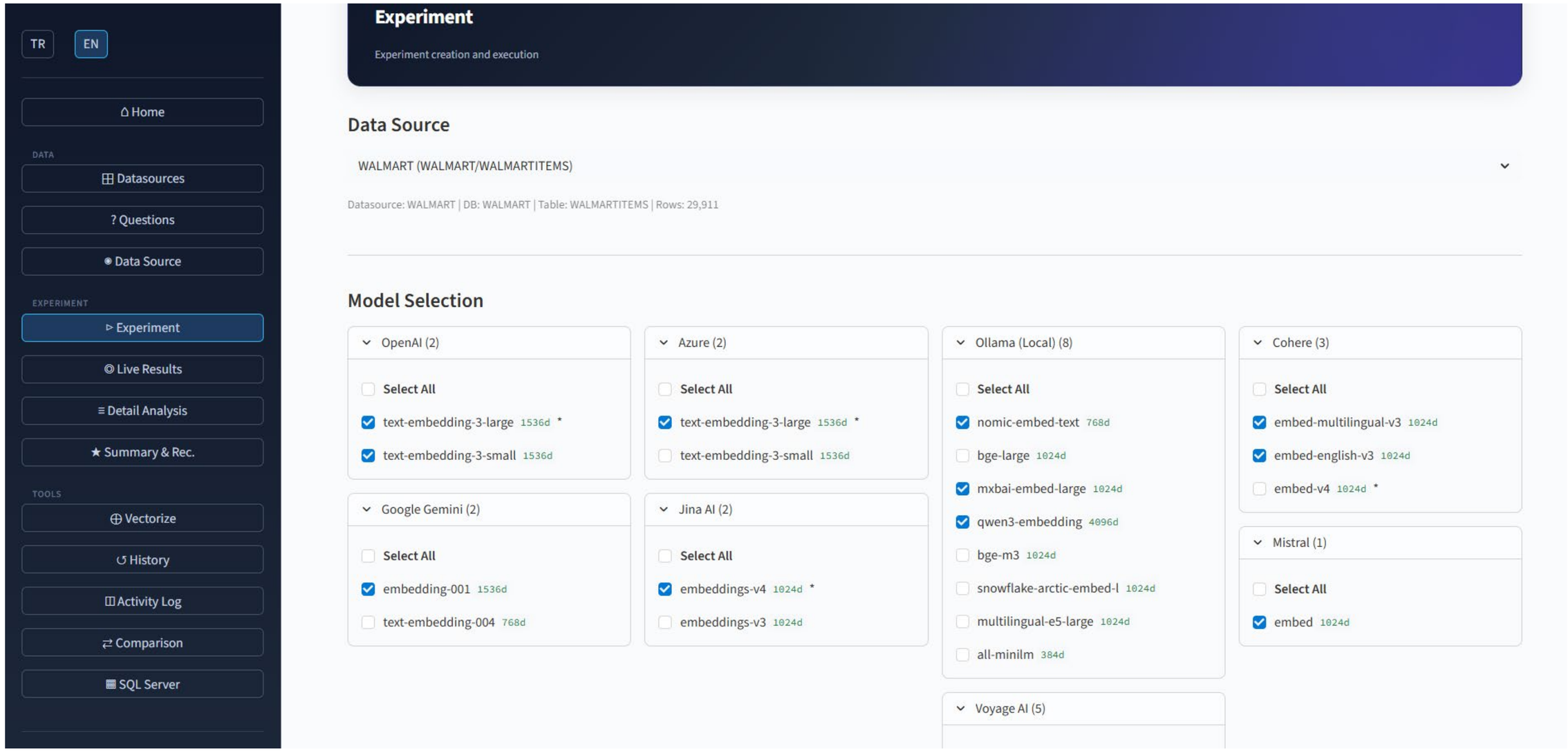


Column to Embed Analysis



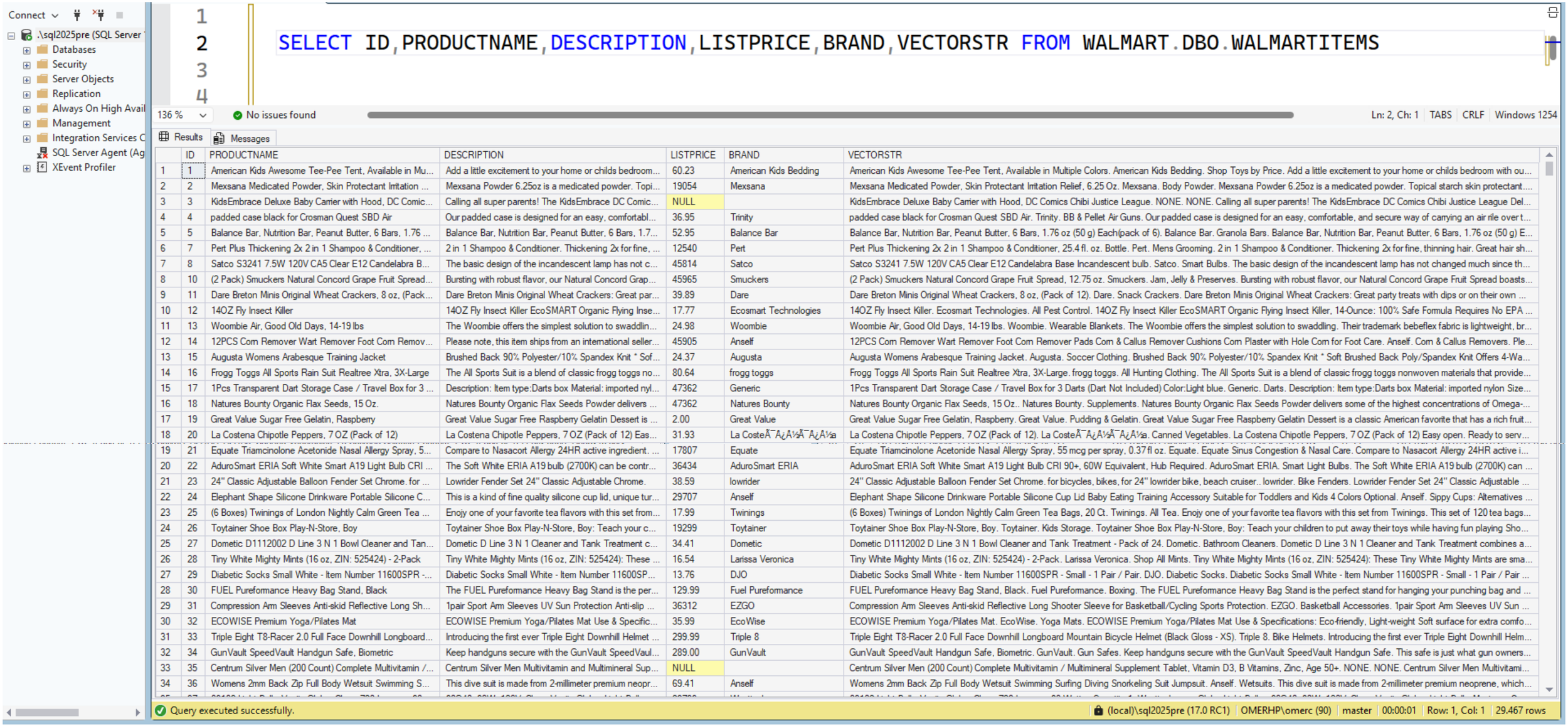
Here we go. Experiment time.
The pipeline will find the best options for us.

1.Model Selection



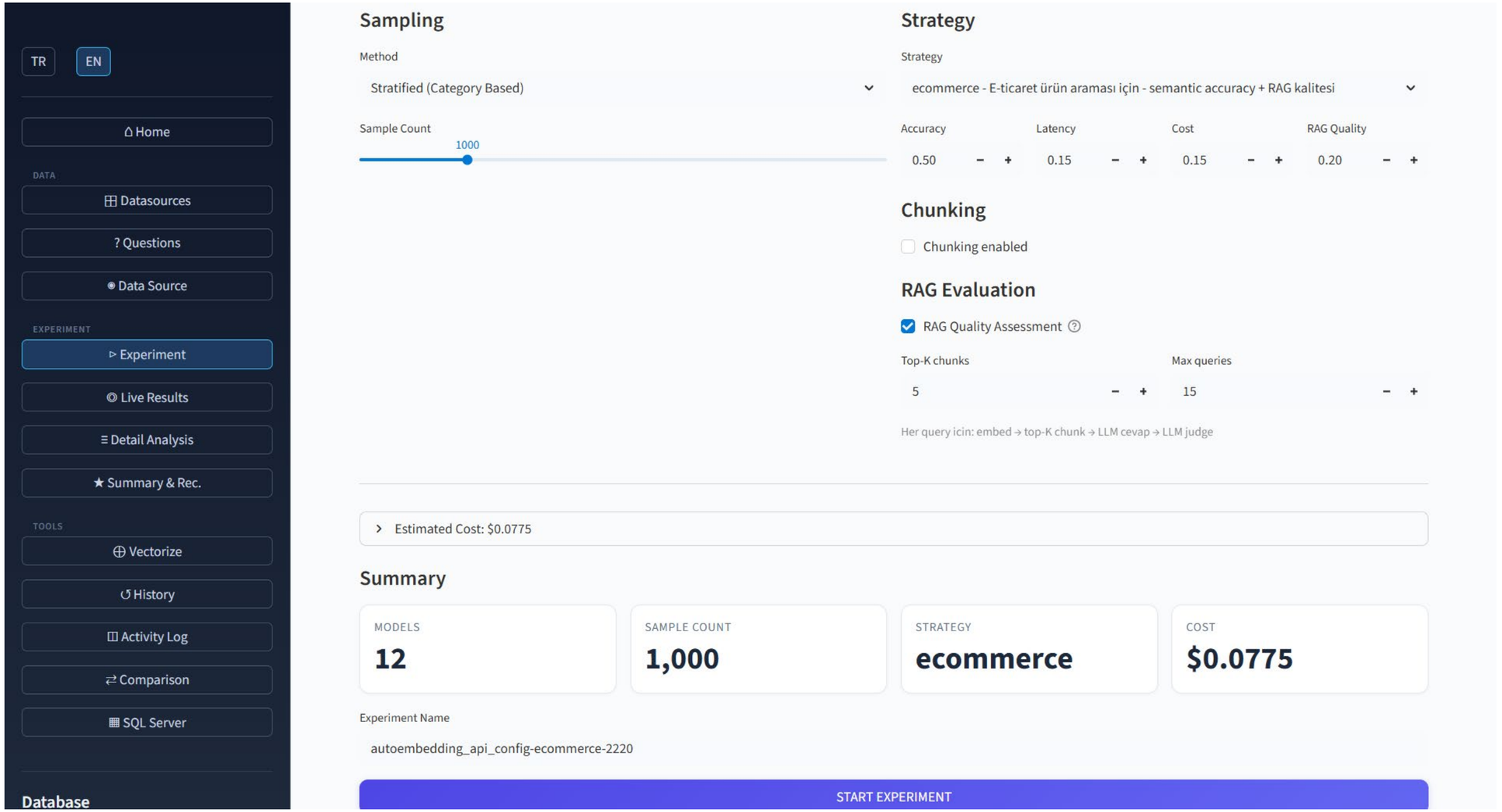
Experiment

Raw Data

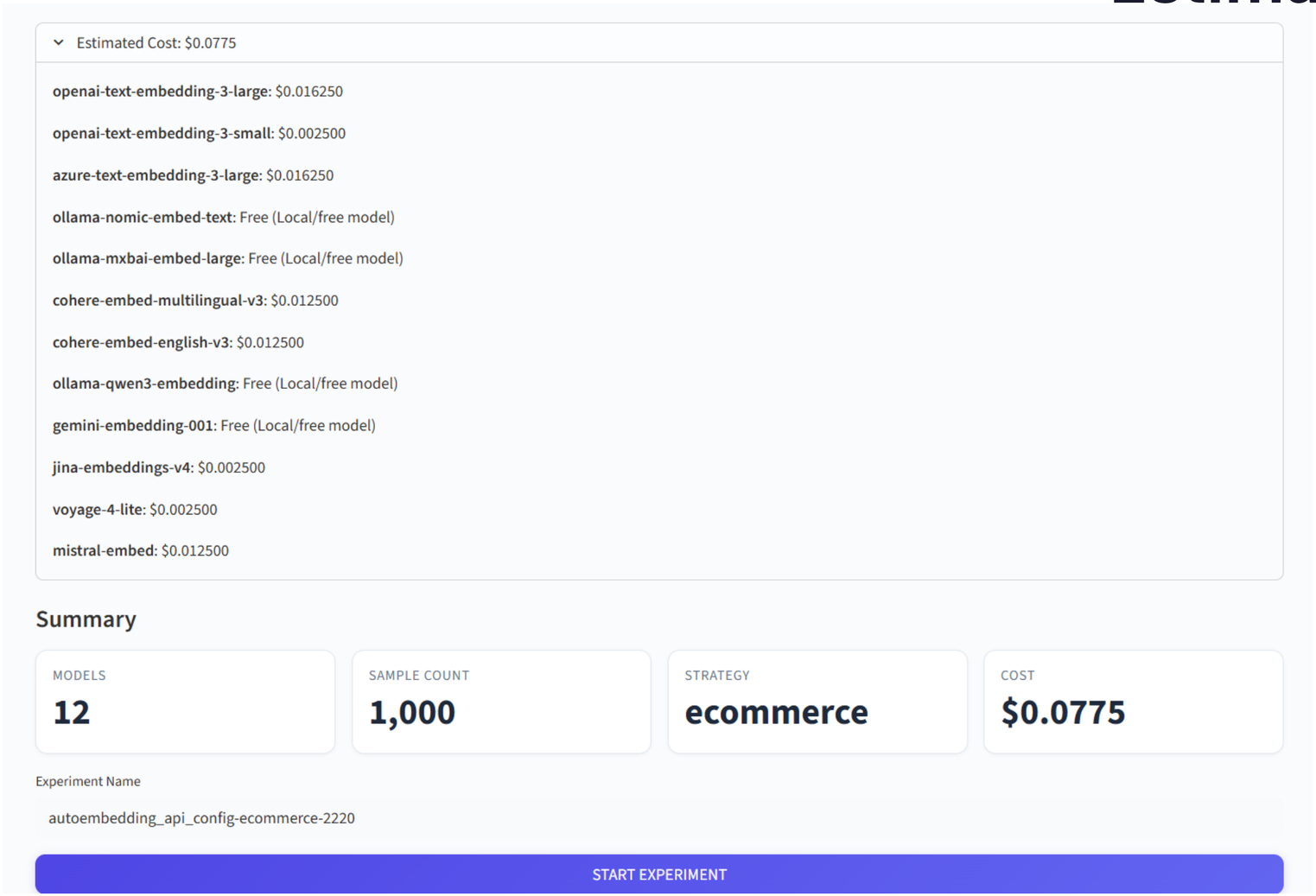


	ID	PRODUCTNAME	DESCRIPTION	LISTPRICE	BRAND	VECTORSTR
1	1	American Kids Awesome Tee-Pee Tent, Available in Mu...	Add a little excitement to your home or childs bedroom...	60.23	American Kids Bedding	American Kids Awesome Tee-Pee Tent, Available in Multiple Colors. American Kids Bedding. Shop Toys by Price. Add a little excitement to your home or childs bedroom with ou...
2	2	Mexsana Medicated Powder, Skin Protectant Irritation ...	Mexsana Powder 6.25oz is a medicated powder. Topi...	19054	Mexsana	Mexsana Medicated Powder, Skin Protectant Irritation Relief, 6.25 Oz. Mexsana. Body Powder. Mexsana Powder 6.25oz is a medicated powder. Topical starch skin protectant....
3	3	KidsEmbrace Deluxe Baby Carrier with Hood, DC Comic...	Calling all super parents! The KidsEmbrace DC Comic...	NULL		KidsEmbrace Deluxe Baby Carrier with Hood, DC Comics Chibi Justice League. NONE. NONE. Calling all super parents! The KidsEmbrace DC Comics Chibi Justice League Del...
4	4	padded case black for Crosman Quest SBD Air	Our padded case is designed for an easy, comfortabl...	36.95	Trinity	padded case black for Crosman Quest SBD Air. Trinity. BB & Pellet Air Guns. Our padded case is designed for an easy, comfortable, and secure way of carrying an air rile over t...
5	5	Balance Bar, Nutrition Bar, Peanut Butter, 6 Bars, 1.76 ...	Balance Bar, Nutrition Bar, Peanut Butter, 6 Bars, 1.7...	52.95	Balance Bar	Balance Bar, Nutrition Bar, Peanut Butter, 6 Bars, 1.76 oz (50 g) Each(pack of 6). Balance Bar. Granola Bars. Balance Bar, Nutrition Bar, Peanut Butter, 6 Bars, 1.76 oz (50 g) E...
6	7	Pert Plus Thickening 2x 2 in 1 Shampoo & Conditioner, ...	2 in 1 Shampoo & Conditioner. Thickening 2x for fine, ...	12540	Pert	Pert Plus Thickening 2x 2 in 1 Shampoo & Conditioner, 25.4 fl. oz. Bottle. Pert. Mens Grooming. 2 in 1 Shampoo & Conditioner. Thickening 2x for fine, thinning hair. Great hair sh...
7	8	Satco S3241 7.5W 120V CA5 Clear E12 Candelabra B...	The basic design of the incandescent lamp has not c...	45814	Satco	Satco S3241 7.5W 120V CA5 Clear E12 Candelabra Base Incandescent bulb. Satco. Smart Bulbs. The basic design of the incandescent lamp has not changed much since th...
8	10	(2 Pack) Smuckers Natural Concord Grape Fruit Spread...	Bursting with robust flavor, our Natural Concord Grap...	45965	Smuckers	(2 Pack) Smuckers Natural Concord Grape Fruit Spread, 12.75 oz. Smuckers. Jam, Jelly & Preserves. Bursting with robust flavor, our Natural Concord Grape Fruit Spread boasts...
9	11	Dare Breton Minis Original Wheat Crackers, 8 oz, (Pack...	Dare Breton Minis Original Wheat Crackers: Great par...	39.89	Dare	Dare Breton Minis Original Wheat Crackers, 8 oz, (Pack of 12). Dare. Snack Crackers. Dare Breton Minis Original Wheat Crackers: Great party treats with dips or on their own ...
10	12	14OZ Fly Insect Killer	14OZ Fly Insect Killer EcoSMART Organic Flying Inse...	17.77	Ecosmart Technologies	14OZ Fly Insect Killer. Ecosmart Technologies. All Pest Control. 14OZ Fly Insect Killer EcoSMART Organic Flying Insect Killer, 14-Ounce: 100% Safe Formula Requires No EPA ...
11	13	Woombie Air, Good Old Days, 14-19 lbs	The Woombie offers the simplest solution to swaddlin...	24.98	Woombie	Woombie Air, Good Old Days, 14-19 lbs. Woombie. Wearable Blankets. The Woombie offers the simplest solution to swaddling. Their trademark bebeflex fabric is lightweight, br...
12	14	12PCS Corn Remover Wart Remover Foot Corn Remov...	Please note, this item ships from an international seller...	45905	Anself	12PCS Corn Remover Wart Remover Foot Corn Remover Pads Corn & Callus Remover Cushions Corn Plaster with Hole Corn for Foot Care. Anself. Corn & Callus Removers. Ple...
13	15	Augusta Womens Arabesque Training Jacket	Brushed Back 90% Polyester/10% Spandex Knit * Sof...	24.37	Augusta	Augusta Womens Arabesque Training Jacket. Augusta. Soccer Clothing. Brushed Back 90% Polyester/10% Spandex Knit * Soft Brushed Back Poly/Spandex Knit Offers 4-Wa...
14	16	Frogg Toggs All Sports Rain Suit Realtree Xtra, 3X-Large	The All Sports Suit is a blend of classic frogg toggs no...	80.64	frogg toggs	Frogg Toggs All Sports Rain Suit Realtree Xtra, 3X-Large. frogg toggs. All Hunting Clothing. The All Sports Suit is a blend of classic frogg toggs nonwoven materials that provide...
15	17	1Pcs Transparent Dart Storage Case / Travel Box for 3 ...	Description: Item type:Darts box Material: imported nyl...	47362	Generic	1Pcs Transparent Dart Storage Case / Travel Box for 3 Darts (Dart Not Included) Color:Light blue. Generic. Darts. Description: Item type:Darts box Material: imported nylon Size...
16	18	Natures Bounty Organic Flax Seeds, 15 Oz.	Natures Bounty Organic Flax Seeds Powder delivers ...	47362	Natures Bounty	Natures Bounty Organic Flax Seeds, 15 Oz.. Natures Bounty. Supplements. Natures Bounty Organic Flax Seeds Powder delivers some of the highest concentrations of Omega-...
17	19	Great Value Sugar Free Gelatin, Raspberry	Great Value Sugar Free Raspberry Gelatin Dessert is ...	2.00	Great Value	Great Value Sugar Free Gelatin, Raspberry. Great Value. Pudding & Gelatin. Great Value Sugar Free Raspberry Gelatin Dessert is a classic American favorite that has a rich fruit...
18	20	La Costena Chipotle Peppers, 7 OZ (Pack of 12)	La Costena Chipotle Peppers, 7 OZ (Pack of 12) Eas...	31.93	La CosteÃ¯Â¿Â½Ã¯Â¿Â½a	La Costena Chipotle Peppers, 7 OZ (Pack of 12). La CosteÃ¯Â¿Â½Ã¯Â¿Â½a. Canned Vegetables. La Costena Chipotle Peppers, 7 OZ (Pack of 12) Easy open. Ready to serv...
19	21	Equate Triamcinolone Acetonide Nasal Allergy Spray, 5...	Compare to Nasacort Allergy 24HR active ingredient. ...	17807	Equate	Equate Triamcinolone Acetonide Nasal Allergy Spray, 55 mcg per spray, 0.37 fl oz. Equate. Equate Sinus Congestion & Nasal Care. Compare to Nasacort Allergy 24HR active i...
20	22	AduroSmart ERIA Soft White Smart A19 Light Bulb CRI ...	The Soft White ERIA A19 bulb (2700K) can be contr...	36434	AduroSmart ERIA	AduroSmart ERIA Soft White Smart A19 Light Bulb CRI 90+, 60W Equivalent, Hub Required. AduroSmart ERIA. Smart Light Bulbs. The Soft White ERIA A19 bulb (2700K) can ...
21	23	24" Classic Adjustable Balloon Fender Set Chrome. for ...	Lowrider Fender Set 24" Classic Adjustable Chrome.	38.59	lowrider	24" Classic Adjustable Balloon Fender Set Chrome. for bicycles, bikes, for 24" lowrider bike, beach cruiser.. lowrider. Bike Fenders. Lowrider Fender Set 24" Classic Adjustable ...
22	24	Elephant Shape Silicone Drinkware Portable Silicone C...	This is a kind of fine quality silicone cup lid, unique tur...	29707	Anself	Elephant Shape Silicone Drinkware Portable Silicone Cup Lid Baby Eating Training Accessory Suitable for Toddlers and Kids 4 Colors Optional. Anself. Sippy Cups: Alternatives ...
23	25	(6 Boxes) Twinings of London Nightly Calm Green Tea ...	Enojy one of your favorite tea flavors with this set from...	17.99	Twinings	(6 Boxes) Twinings of London Nightly Calm Green Tea Bags, 20 Ct. Twinings. All Tea. Enojy one of your favorite tea flavors with this set from Twinings. This set of 120 tea bags...
24	26	Toytainer Shoe Box Play-N-Store, Boy	Toytainer Shoe Box Play-N-Store, Boy: Teach your c...	19299	Toytainer	Toytainer Shoe Box Play-N-Store, Boy. Toytainer. Kids Storage. Toytainer Shoe Box Play-N-Store, Boy: Teach your children to put away their toys while having fun playing Sho...
25	27	Dometic D1112002 D Line 3 N 1 Bowl Cleaner and Tan...	Dometic D Line 3 N 1 Cleaner and Tank Treatment c...	34.41	Dometic	Dometic D1112002 D Line 3 N 1 Bowl Cleaner and Tank Treatment - Pack of 24. Dometic. Bathroom Cleaners. Dometic D Line 3 N 1 Cleaner and Tank Treatment combines a...
26	28	Tiny White Mighty Mints (16 oz, ZIN: 525424) - 2-Pack	Tiny White Mighty Mints (16 oz, ZIN: 525424): These ...	16.54	Larissa Veronica	Tiny White Mighty Mints (16 oz, ZIN: 525424) - 2-Pack. Larissa Veronica. Shop All Mints. Tiny White Mighty Mints (16 oz, ZIN: 525424): These Tiny White Mighty Mints are sma...
27	29	Diabetic Socks Small White - Item Number 11600SPR -...	Diabetic Socks Small White - Item Number 11600SP...	13.76	DJO	Diabetic Socks Small White - Item Number 11600SPR - Small - 1 Pair / Pair. DJO. Diabetic Socks. Diabetic Socks Small White - Item Number 11600SPR - Small - 1 Pair / Pair ...
28	30	FUEL Pureformance Heavy Bag Stand, Black	The FUEL Pureformance Heavy Bag Stand is the per...	129.99	Fuel Pureformance	FUEL Pureformance Heavy Bag Stand, Black. Fuel Pureformance. Boxing. The FUEL Pureformance Heavy Bag Stand is the perfect stand for hanging your punching bag and ...
29	31	Compression Arm Sleeves Anti-skid Reflective Long Sh...	1pair Sport Arm Sleeves UV Sun Protection Anti-slip ...	36312	EZGO	Compression Arm Sleeves Anti-skid Reflective Long Shooter Sleeve for Basketball/Cycling Sports Protection. EZGO. Basketball Accessories. 1pair Sport Arm Sleeves UV Sun ...
30	32	ECOWISE Premium Yoga/Pilates Mat	ECOWISE Premium Yoga/Pilates Mat Use & Specific...	35.99	EcoWise	ECOWISE Premium Yoga/Pilates Mat. EcoWise. Yoga Mats. ECOWISE Premium Yoga/Pilates Mat Use & Specifications: Eco-friendly, Light-weight Soft surface for extra comfo...
31	33	Triple Eight T8-Racer 2.0 Full Face Downhill Longboard...	Introducing the first ever Triple Eight Downhill Helmet ...	299.99	Triple 8	Triple Eight T8-Racer 2.0 Full Face Downhill Longboard Mountain Bicycle Helmet (Black Gloss - XS). Triple 8. Bike Helmets. Introducing the first ever Triple Eight Downhill Helm...
32	34	GunVault SpeedVault Handgun Safe, Biometric	Keep handguns secure with the GunVault SpeedVaul...	289.00	GunVault	GunVault SpeedVault Handgun Safe, Biometric. GunVault. Gun Safes. Keep handguns secure with the GunVault SpeedVault Handgun Safe. This safe is just what gun owners...
33	35	Centrum Silver Men (200 Count) Complete Multivitamin /...	Centrum Silver Men Multivitamin and Multimineral Sup...	NULL		Centrum Silver Men (200 Count) Complete Multivitamin / Multimineral Supplement Tablet, Vitamin D3, B Vitamins, Zinc, Age 50+. NONE. NONE. Centrum Silver Men Multivitami...
34	36	Womens 2mm Back Zip Full Body Wetsuit Swimming S...	This dive suit is made from 2-millimeter premium neopr...	69.41	Anself	Womens 2mm Back Zip Full Body Wetsuit Swimming Surfing Diving Snorkeling Suit Jumpsuit. Anself. Wetsuits. This dive suit is made from 2-millimeter premium neoprene, which...

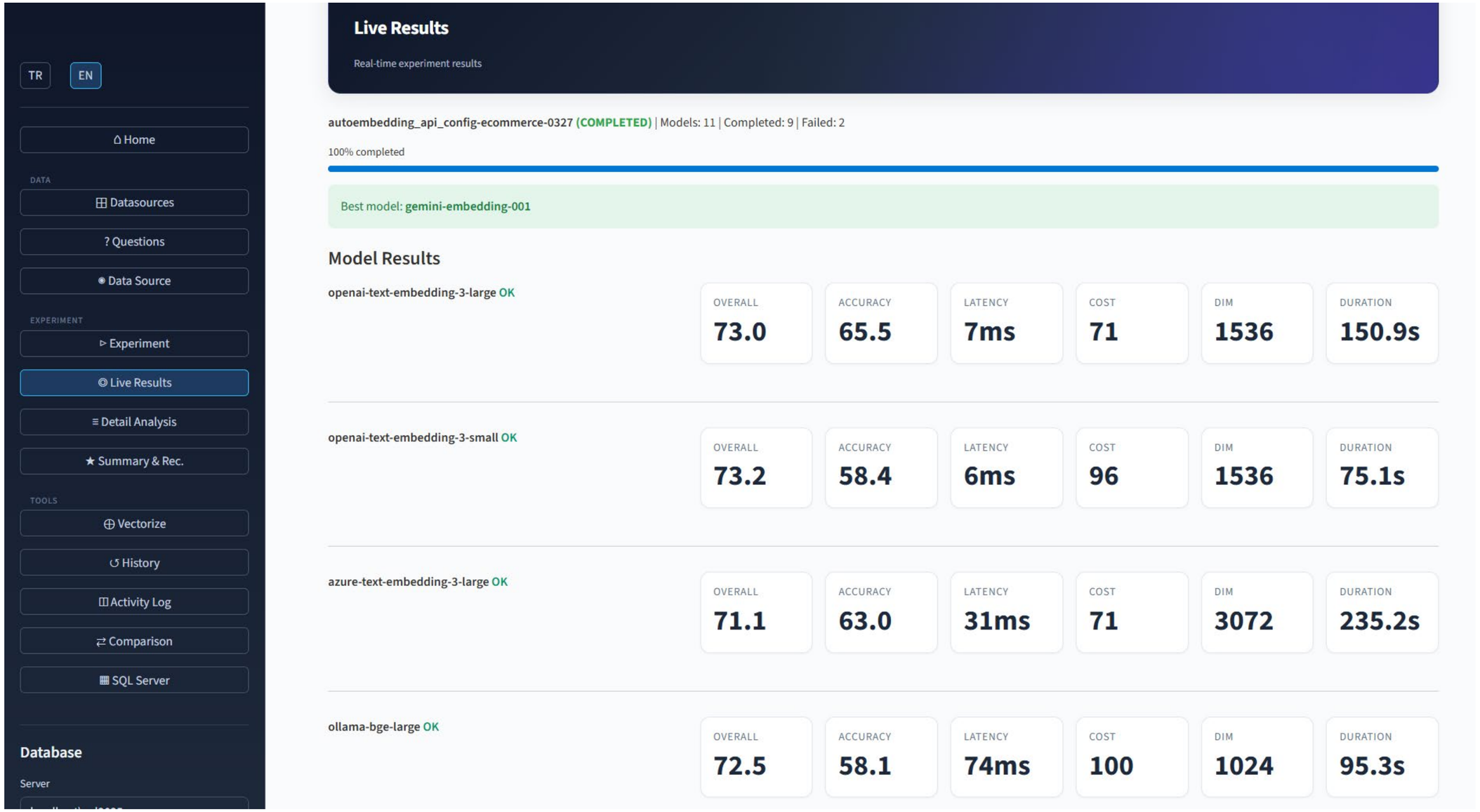
2.Sampling, Strategy, Rerank Parameters, Estimated Budget



2.Sampling, Strategy, Rerank Parameters, Estimated Budget



3.Live Results



3.Live Results

#4 ollama-bge-large: **72.5**

#5 openai-text-embedding-3-small: **73.2**

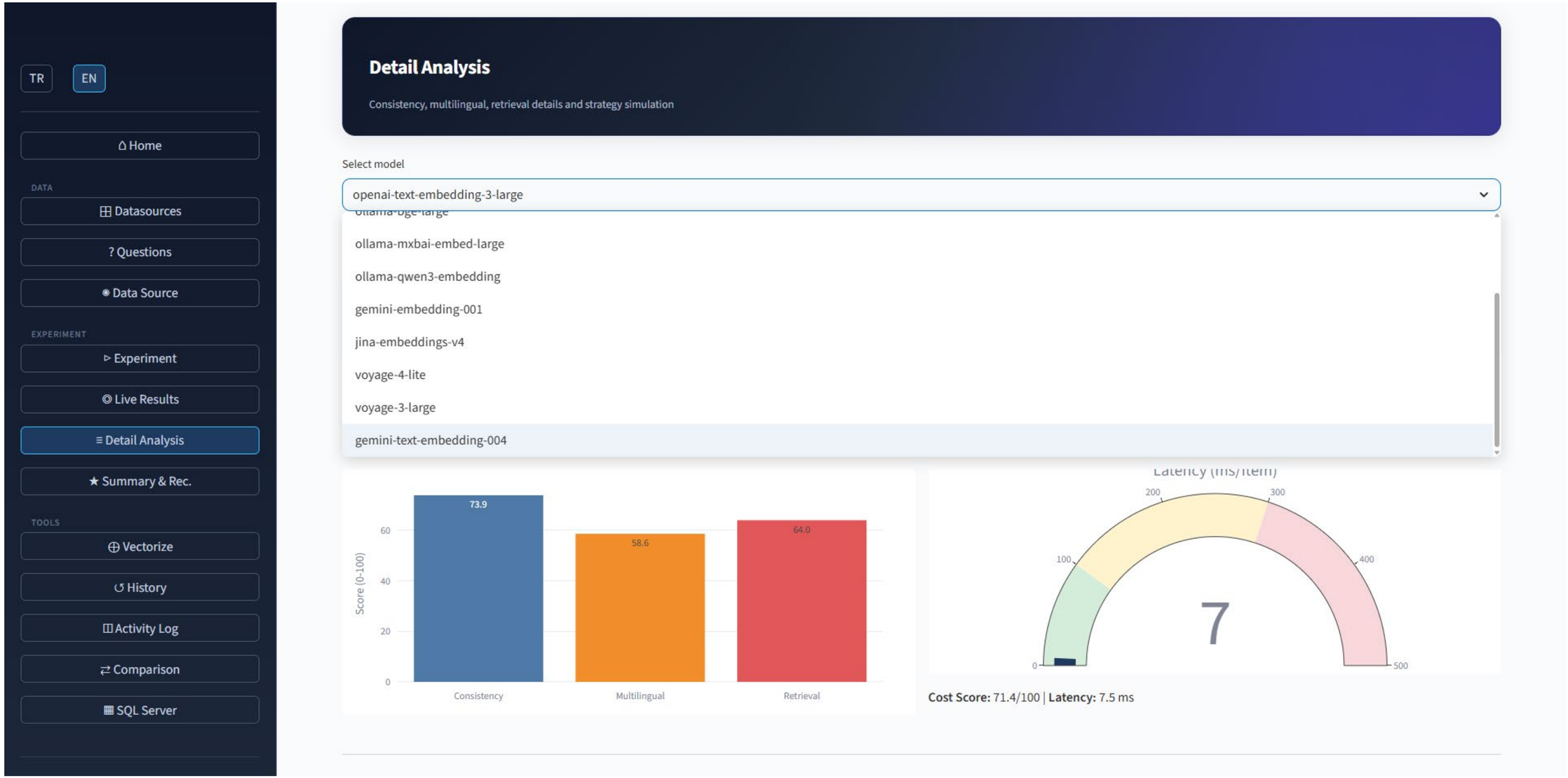
#6 openai-text-embedding-3-large: **73.0**

#7 voyage-3-large: **71.3**

Detailed Log

```
  jina-embeddings-v4: OK
  voyage-4-lite: OK
  voyage-3-large: OK
  gemini-text-embedding-004: OK
[03:30:36] Preparing test queries...
[03:30:47] Test query set: 50 queries
[03:30:47] Model evaluation starting (can be stopped between models)...
[03:30:47] Evaluating: openai-text-embedding-3-large (1/11) | 1000 docs, 50 queries
[03:31:30] openai-text-embedding-3-large completed (42.7s) | Overall: 73.0 | Acc: 65.5 | Lat: 7.5ms | Cost: 71.4
[03:33:18] openai-text-embedding-3-large completed (150.9s) | Overall: 73.0 | Acc: 65.5 | Lat: 7.5ms | Cost: 71.4
[03:33:19] Evaluating: openai-text-embedding-3-small (2/11) | 1000 docs, 50 queries
[03:34:02] openai-text-embedding-3-small completed (43.3s) | Overall: 73.2 | Acc: 58.4 | Lat: 5.8ms | Cost: 95.6
[03:34:19] openai-text-embedding-3-small completed (60.0s) | Overall: 73.2 | Acc: 58.4 | Lat: 5.8ms | Cost: 95.6
[03:34:34] openai-text-embedding-3-small completed (75.1s) | Overall: 73.2 | Acc: 58.4 | Lat: 5.8ms | Cost: 95.6
[03:34:34] Evaluating: azure-text-embedding-3-large (3/11) | 1000 docs, 50 queries
[03:38:29] azure-text-embedding-3-large completed (235.2s) | Overall: 71.1 | Acc: 63.0 | Lat: 30.9ms | Cost: 71.4
  -> Embedding save warning (azure-text-embedding-3-large): SQL execution failed: ('42000', "[42000] [Microsoft][ODBC Driver 17 for SQL Serv
[03:38:30] Evaluating: ollama-bge-large (4/11) | 1000 docs, 50 queries
[03:40:05] ollama-bge-large completed (95.3s) | Overall: 72.5 | Acc: 58.1 | Lat: 73.9ms | Cost: 100.0
  -> ollama-bge-large: 1000 embeddings saved (embeddings_ollama_bge_large, exp=7a147edd)
[03:40:21] Evaluating: ollama-mxbai-embed-large (5/11) | 1000 docs, 50 queries
[03:42:29] ollama-mxbai-embed-large completed (128.0s) | Overall: 73.1 | Acc: 59.5 | Lat: 85.1ms | Cost: 100.0
  -> ollama-mxbai-embed-large: 1000 embeddings saved (embeddings_ollama_mxbai_embed_large, exp=7a147edd)
[03:42:46] Evaluating: ollama-qwen3-embedding (6/11) | 1000 docs, 50 queries
[03:53:03] ollama-qwen3-embedding completed (617.1s) | Overall: 69.1 | Acc: 66.1 | Lat: 521.5ms | Cost: 100.0
```

4.Detail Analysis



1.Consistency Test – Similar Pairs

1. Consistency Test - Similarity Pairs

Similar text pairs should produce high, different text pairs should produce low cosine similarity.

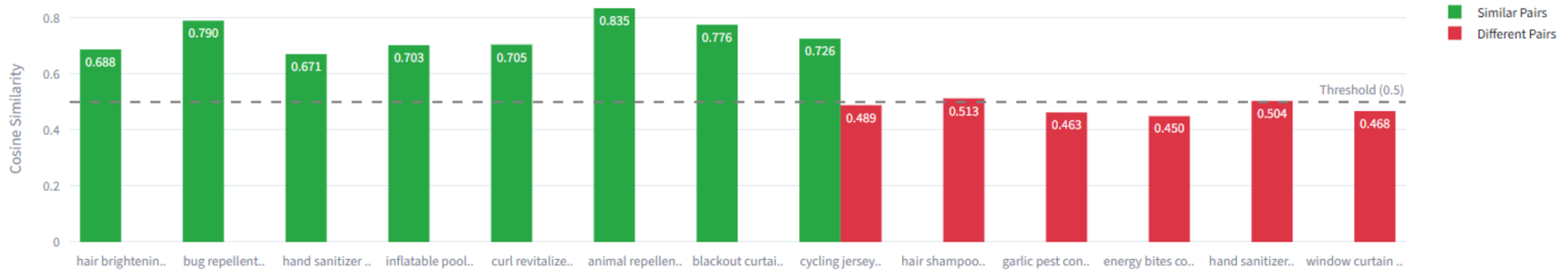
Similar Pairs (high similarity expected)

Text 1	Text 2	Cosine Similarity	Expected	Result
hair brightening shampoo	color enhancing wash	0.6877	High	OK
bug repellent	insect deterrent	0.7902	High	OK
hand sanitizer gel	alcohol-based disinfectant	0.6709	High	OK
inflatable pool float	air-filled water lounger	0.7027	High	OK
curl revitalizer	ringlet refresher	0.7049	High	OK
animal repellent	wildlife deterrent	0.8349	High	OK
blackout curtains	light-blocking drapes	0.776	High	OK
cycling jersey	biking shirt	0.726	High	OK

Different Pairs (low similarity expected)

Text 1	Text 2	Cosine Similarity	Expected	Result
hair shampoo	dog training collar	0.5129	Low	WEAK
garlic pest control	body talc powder	0.4627	Low	OK
energy bites cookbook	air hockey table	0.4495	Low	OK
hand sanitizer	pickled beets	0.5039	Low	WEAK
window curtain panels	bike repair tool	0.4675	Low	OK
cycling jersey	lollipop pool float	0.4886	Low	OK

Cosine Similarity Distribution



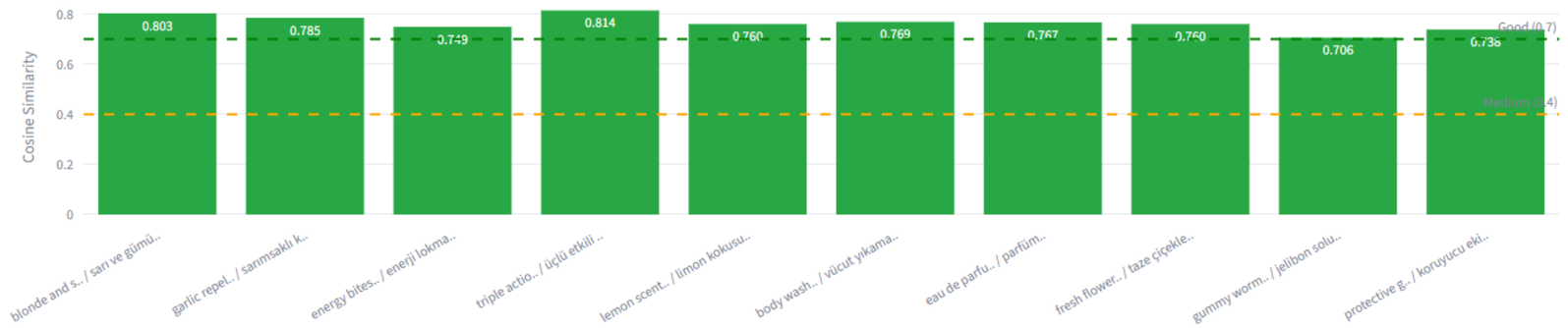
2. Multilingual Test - EN/TR Translation Pairs

2. Multilingual Test - EN/TR Translation Pairs

Embedding similarity of EN and TR texts with the same meaning should be high.

English	Turkish	Cosine Similarity	Result
blonde and silver shampoo	sarı ve gümüş şampuan	0.8027	GOOD
garlic repellent	sarımsaklı kovucu	0.7849	GOOD
energy bites	enerji lokmaları	0.7491	GOOD
triple action insecticide	üçlü etkili böcek ilacı	0.8142	GOOD
lemon scent	limon kokusu	0.7602	GOOD
body wash	vücut yıkama jeli	0.7688	GOOD
eau de parfum	parfüm	0.7666	GOOD
fresh flowers	taze çiçekler	0.7601	GOOD
gummy worm	jelibon solucanı	0.7056	GOOD
protective gear	koruyucu ekipman	0.7381	GOOD

EN-TR Translation Pairs Similarity



Average EN-TR Similarity: **0.7650** | Score: **76.5**/100

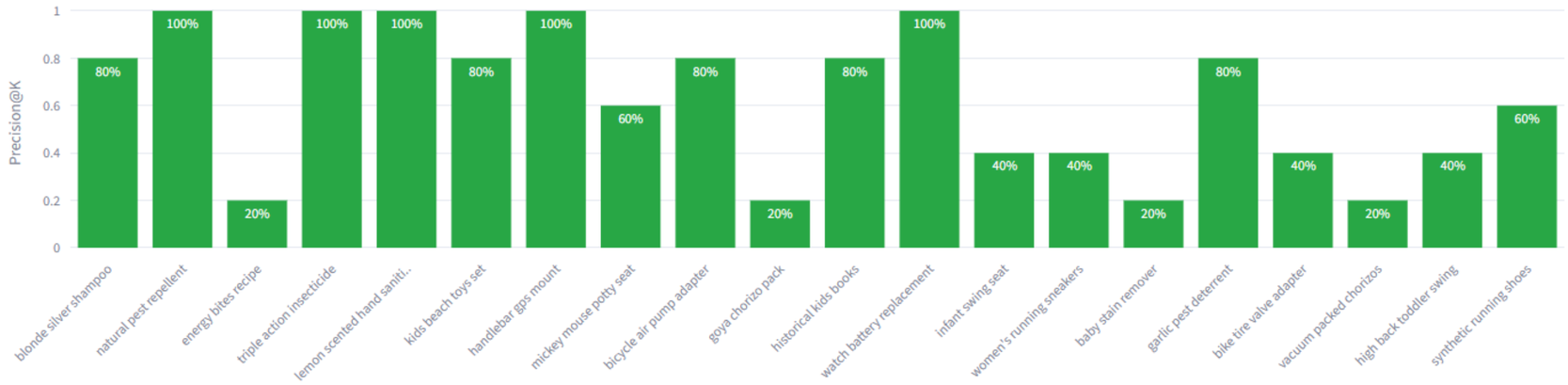
3. Retrieval Test - Query and Top-K Results

3. Retrieval Test - Query and Top-K Results

Closest K documents and ground truth matches for each query.

Query	Language	Category	Ground Truth	Top-K Hit	Precision	Expected Similar
blonde silver shampoo	AUTO	document_generated	13	4	80.00%	shampoo for blonde hair, silver hair shampoo
natural pest repellent	AUTO	document_generated	35	5	100.00%	chemical-free insect repellent, organic garden pest control
energy bites recipe	AUTO	document_generated	8	1	20.00%	recipe for energy bites, how to make energy snacks
triple action insecticide	AUTO	document_generated	135	5	100.00%	multi-purpose insect spray, all-in-one pest control
lemon scented hand sanitizer	AUTO	document_generated	20	5	100.00%	citrus hand sanitizer, lemon fragrance sanitizer gel
kids beach toys set	AUTO	document_generated	32	4	80.00%	children's sand playset, beach toys for kids
handlebar gps mount	AUTO	document_generated	9	5	100.00%	bike gps holder, mount for cycling gps
mickey mouse potty seat	AUTO	document_generated	16	3	60.00%	disney potty chair, children's soft toilet seat
bicycle air pump adapter	AUTO	document_generated	18	4	80.00%	bike tire pump valve, bicycle pump head adapter
goya chorizo pack	AUTO	document_generated	1	1	20.00%	spanish chorizo sausages, goya sausage pack

Query-Based Retrieval Precision



Query Detail (Top-K Results)

Query Detail (Top-K Results)

> Q1: blonde silver shampoo [AUTO] - Precision: 80%

> Q2: natural pest repellent [AUTO] - Precision: 100%

> Q3: energy bites recipe [AUTO] - Precision: 20%

v Q4: triple action insecticide [AUTO] - Precision: 100%

Category: document_generated | **Expected:** multi-purpose insect spray, all-in-one pest control

Ground Truth: 135 docs | **Hit:** 5

Rank	Similarity	Relevant	Document (first 120 chars)
1	0.6786	YES	VOLUNTARY PURCHASING GROUP INC Triple-Action Insecticide, 32-oz. Spray 11247. VOLUNTARY PURCHASING GRO
2	0.6378	YES	3-in-1 32 oz. Ready-to-Use Garden Spray. Safer. Insect & Pest Control. Safer Brand 3-in-1 Garden Spray contains sulf
3	0.6078	YES	Hot Shot Bedbug & Flea Fogger, Insect Killer, 3/2-Ounce. Hot Shot. Insect & Pest Control. Hot Shot BedBug & Flea Fo
4	0.6025	YES	Yard and Garden Insect Killer 32 fl oz Ready-To-Use Spray. Safer. Insect & Pest Control. The Safer Brand Yard and Ga
5	0.5789	YES	Black Flag 190256 64 oz. Fogging Insecticide. Black Flag. Insect & Pest Control. Black flag fogging insecticide is idea

v Q5: lemon scented hand sanitizer [AUTO] - Precision: 100%

Category: document_generated | **Expected:** citrus hand sanitizer, lemon fragrance sanitizer gel

Ground Truth: 20 docs | **Hit:** 5

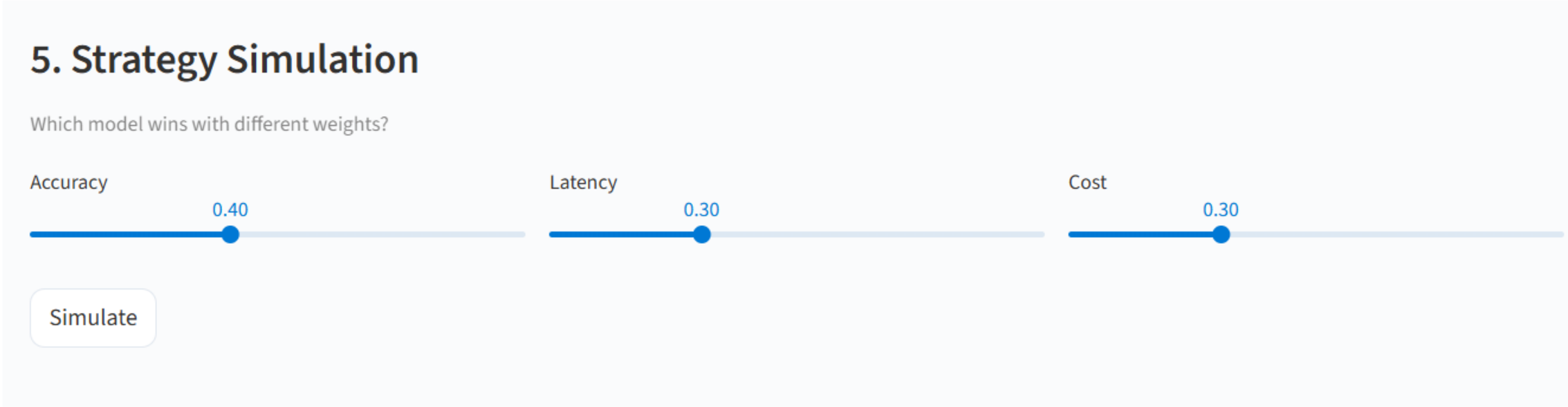
Rank	Similarity	Relevant	Document (first 120 chars)
1	0.7267	YES	Hand Sanitizer Gel, Lemon Scent, 8 Oz Pump Bottle, 12/carton. Boardwalk. Commercial Hand Sanitizers. Hand Sani
2	0.6621	YES	Purell Advanced Refreshing Gel Hand Sanitizer Pack of 2. PURELL. Hand Soaps & Sanitizers. Purell Advanced Refres

4. All Models - Comparison

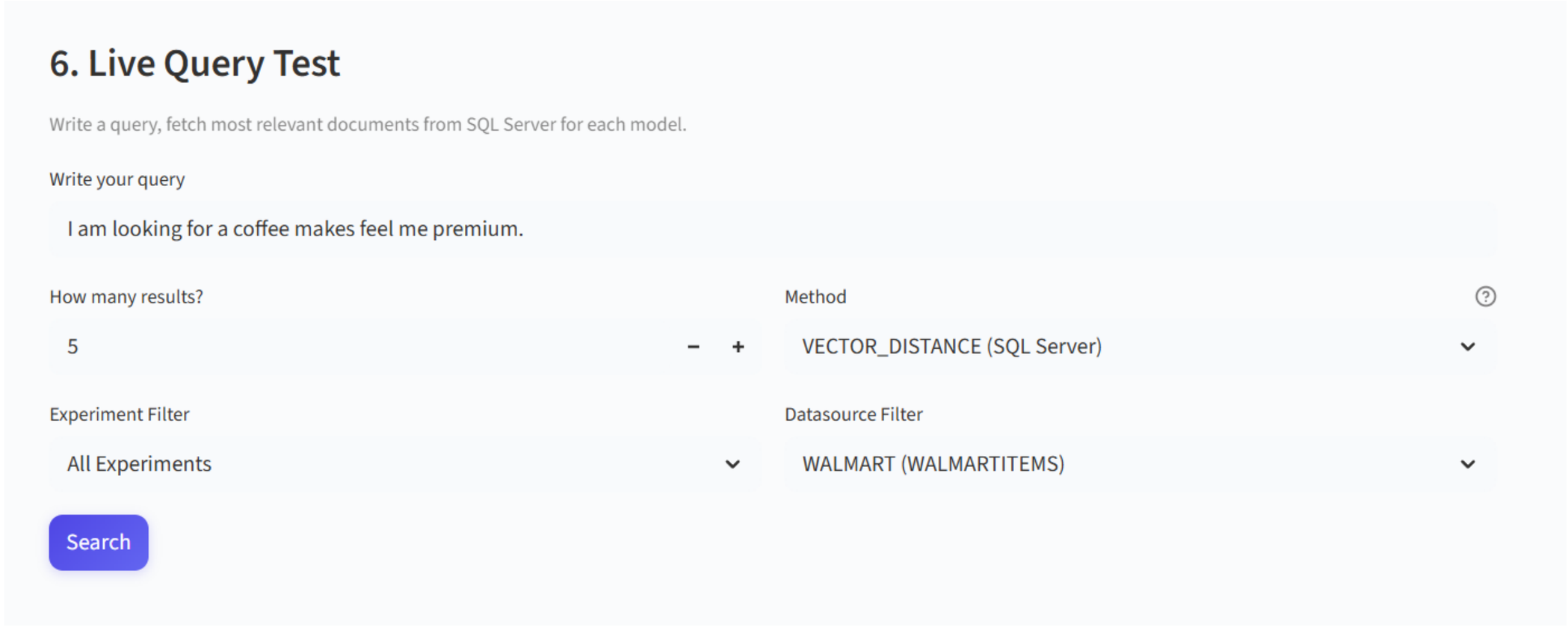
4. All Models - Comparison

Model	Overall	Accuracy	Consistency	Multilingual	Retrieval	Latency (ms)	Cost Score	Dim
openai-text-embedding-3-large	73	65.5	73.9	58.6	64	7.5	71.4	1536
openai-text-embedding-3-small	73.2	58.4	72	39.2	64	5.8	95.6	1536
azure-text-embedding-3-large	71.1	63	70.5	54.6	64	30.9	71.4	3072
ollama-bge-large	72.5	58.1	55.8	52.6	66	73.9	100	1024
ollama-mxbai-embed-large	73.2	59.5	63	48.4	67	85.1	100	1024
ollama-qwen3-embedding	69.1	66.1	60.3	78.1	60	521.5	100	4096
gemini-embedding-001	79.2	67.4	62.8	76.5	63	27.4	100	1536
jina-embeddings-v4	0	0	0	0	0	0	0	0
voyage-4-lite	77.1	64.8	69.1	63.4	62	12.3	95.6	1024
voyage-3-large	71.3	66.2	70.2	65.4	63	13.8	60.4	1024

5. Strategy Simulation



6. Live Query Test



6. Live Query Test

Query: *"I am looking for a coffee makes feel me premium."*

Evaluate with LLM

openai-text-embedding-3-large

Method: VECTOR_DISTANCE | Dimension: 1536

Rank	Similarity	Relevance	Keyword Match	LLM Verdict	LLM Reason	Document
1	0.4833	Partial	2/7	Partially Relevant	Ürün premium bir kahve sunuyor, a	Bialetti Porcelain Coffee Mug with Matching Pour-Over Top. Biale
2	0.4832	Partial	2/7	Partially Relevant	Aynı ürün tekrar ediyor, premium bi	Bialetti Porcelain Coffee Mug with Matching Pour-Over Top. Biale
3	0.4646	Partial	4/7	Relevant	Premium kalite kahve sunuyor ve ku	McCullagh CafÃ© Classique Premium Blend Coffee (200 ct.). McC
4	0.4302	Partial	3/7	Partially Relevant	Bolder bir kahve deneyimi sunuyor,	illy Ground Drip Forte Extra Dark Roast Coffee, 8.8 Oz. NONE. NO
5	0.4300	Partial	1/7	Partially Relevant	Kahve kapsülleri premium bir deney	Lavazza Decaffeinato Ricco Nespresso Espresso Coffee Capsules

openai-text-embedding-3-small

Method: VECTOR_DISTANCE | Dimension: 1536

Rank	Similarity	Relevance	Keyword Match	LLM Verdict	LLM Reason	Document
1	0.4551	Partial	3/7	Partially Relevant	Iced coffee mentioned, but does no	Pecan Sticky Buns Cold Brew Iced Coffee Hot Coffee Liquid Java
2	0.3099	Not relevant	0/7	Not Relevant	Related to hair products, no conne	ColorProof PowderFix Texturizing Powder 0.53 oz. Colorproof. F
3	0.3099	Not relevant	0/7	Not Relevant	Duplicate of document 2, still unre	ColorProof PowderFix Texturizing Powder 0.53 oz. Colorproof. F
4	0.3049	Not relevant	2/7	Partially Relevant	Mentions flavoring for coffee drink	(2 Pack) McCormick Rum Extract, 2 fl oz. McCormickÃƒâ€šÃ‚. Ext
5	0.2989	Not relevant	1/7	Not Relevant	Related to hair color, no connectio	Socolor Extra Coverage Hair Color 505N - Medium Brown Neutra

gemini-embedding-001

Method: VECTOR_DISTANCE | Dimension: 1536

Rank	Similarity	Relevance	Keyword Match	LLM Verdict	LLM Reason	Document
1	0.6237	Partial	1/7	Relevant	Doküman, premium bir kahve markası olar	Mccafe Premium Roast K-Cup Coffee Pods, 84 Count. McCafe. McCafe. Product of McCa
2	0.6109	Partial	4/7	Relevant	Doküman, premium bir kahve markası olar	McCullagh CafÃ© Classique Premium Blend Coffee (200 ct.). McCullagh. Instant Coffee.
3	0.5993	Partial	3/7	Partially Relevant	Doküman, aromalı bir kahve sunuyor ancal	Toasted Island Coconut Flavored Ground Coffee, 12 Ounce Bag. NONE. NONE. Get whis
4	0.5841	Partial	2/7	Partially Relevant	Doküman, aromalı bir kahve sunuyor ancal	Butter Pecan Flavored Ground Coffee, 12 Ounce Bag. NONE. NONE. Satiate your Southe
5	0.5813	Partial	3/7	Partially Relevant	Doküman, aromalı bir kahve sunuyor ancal	Blueberries & Cream Flavored Whole Bean Coffee, 12 Ounce Bag. NONE. NONE. Meet yo

voyage-4-lite

Method: VECTOR_DISTANCE | Dimension: 1024

Rank	Similarity	Relevance	Keyword Match	LLM Verdict	LLM Reason	Document
1	0.6400	Partial	3/7	Partially Relevant	Flavored coffee may appeal to a premium exp	Blueberries & Cream Flavored Whole Bean Coffee, 12 Ounce Bag. NONE. NONE. Meet y
2	0.6315	Partial	3/7	Partially Relevant	Coffee creamer adds a premium flavor but is	COFFEE MATE ARTISAN CAFE Himalayan Salted Caramel Liquid Coffee Creamer 14 fl. oz
3	0.6243	Partial	3/7	Partially Relevant	Cold brew coffee may provide a premium fee	Pecan Sticky Buns Cold Brew Iced Coffee Hot Coffee Liquid Java Concentrate (16 Ounc
4	0.6171	Partial	1/7	Relevant	Made from 100% premium Arabica beans, ali	Dunkin Donuts Original Blend Decaffeinated Ground Coffee - 45 oz. NONE. NONE. This
5	0.6147	Partial	2/7	Partially Relevant	Flavored coffee may provide a premium expe	Butter Pecan Flavored Ground Coffee, 12 Ounce Bag. NONE. NONE. Satiate your South

voyage-3-large

Method: VECTOR_DISTANCE | Dimension: 1024

Rank	Similarity	Relevance	Keyword Match	LLM Verdict	LLM Reason	Document
1	0.5975	Partial	3/7	Partially Relevant	Kahve ile ilgili, ancak premium hissettirme konusuna doğrudan değinmiyor.	Pecan Sticky Buns Cold Brew Iced Coffee Hot Coffee Liquid Ja
2	0.5052	Partial	2/7	Not Relevant	Kahve ile ilgisi yok, enerji jeli hakkında.	GU Energy Gel: Caramel Macchiato, Box of 24. GU. Bike Gels. E
3	0.4959	Partial	2/7	Not Relevant	Kahve ile ilgisi yok, parfüm seti hakkında.	($151 Value) Burberry Mr. Burberry Indigo Cologne Gift Set fo

Summary & Result

Summary & Recommendation

Experiment summary and model recommendation

Winning Model

🏆 gemini-embedding-001

79.2

Overall Score

2. voyage-4-lite

77.1

Diff: 2.2

STRATEGY	MODEL COUNT	DIFF	DATE
ecommerce	11	2.2	2026-03-05

Summary & Result

Experiment Summary

This experiment evaluated 11 embedding models (9 completed successfully) against 1000 documents from WALMARTITEMS. The evaluation used the 'ecommerce' strategy with weights of 50% accuracy, 15% latency, and 15% cost. The best performing model was gemini-embedding-001 with an overall score of 79.2/100, completing the full evaluation in 1628.1 seconds.

Recommendation

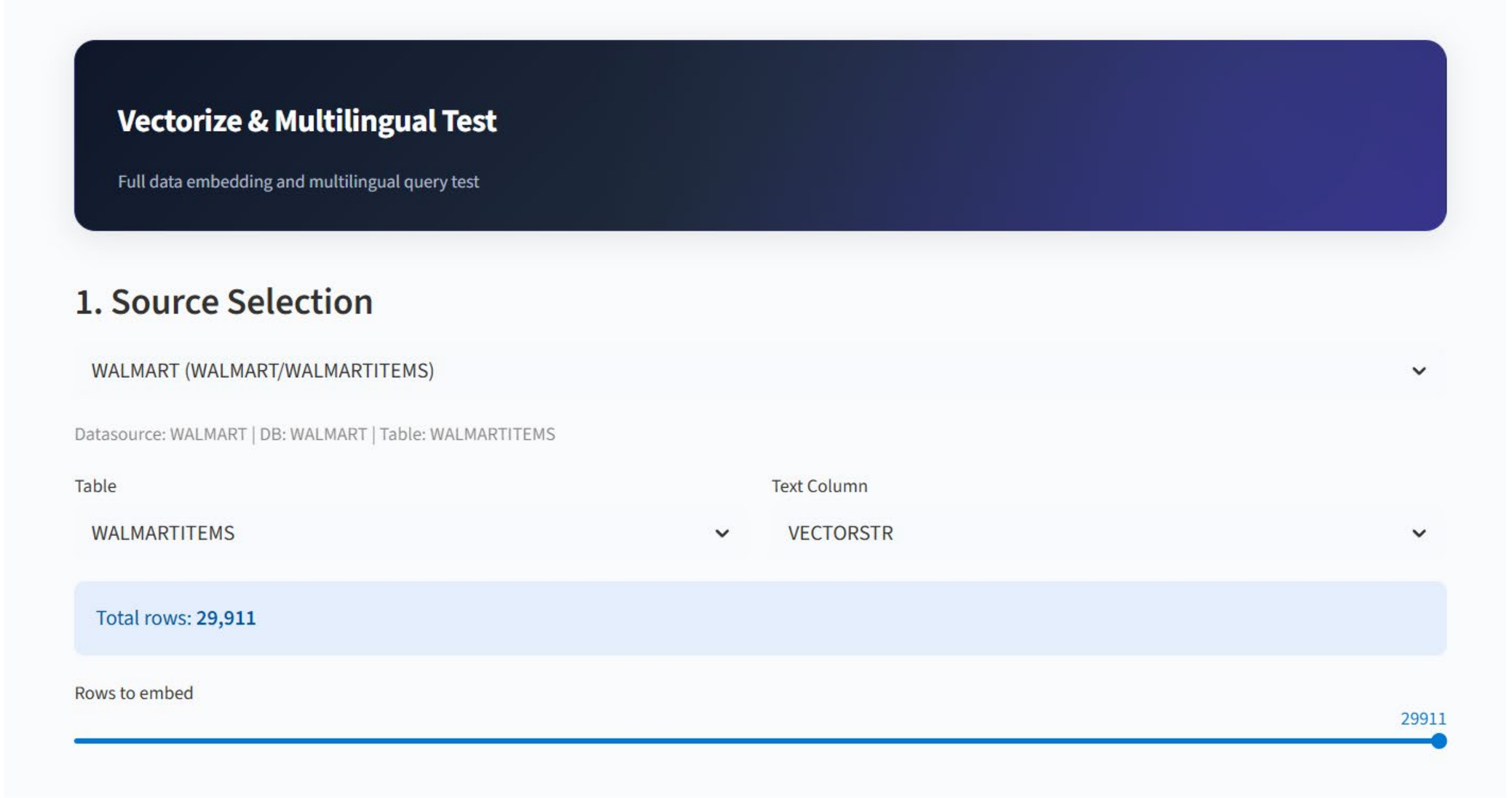
gemini-embedding-001 is recommended for this use case. Under the 'ecommerce' strategy — optimized for product search and catalog retrieval, prioritizing semantic accuracy — it achieved the highest weighted score of 79.2/100. It scored 67.4/100 on accuracy, 27.4ms average latency, and 100.0/100 on cost efficiency. Compared to the runner-up voyage-4-lite (77.1/100), gemini-embedding-001 won by a margin of 2.2 points, primarily excelling in accuracy, cost efficiency. Given the strategy weighting of 50% accuracy / 15% latency / 15% cost, this selection optimally balances the configured priorities.

Key Factors

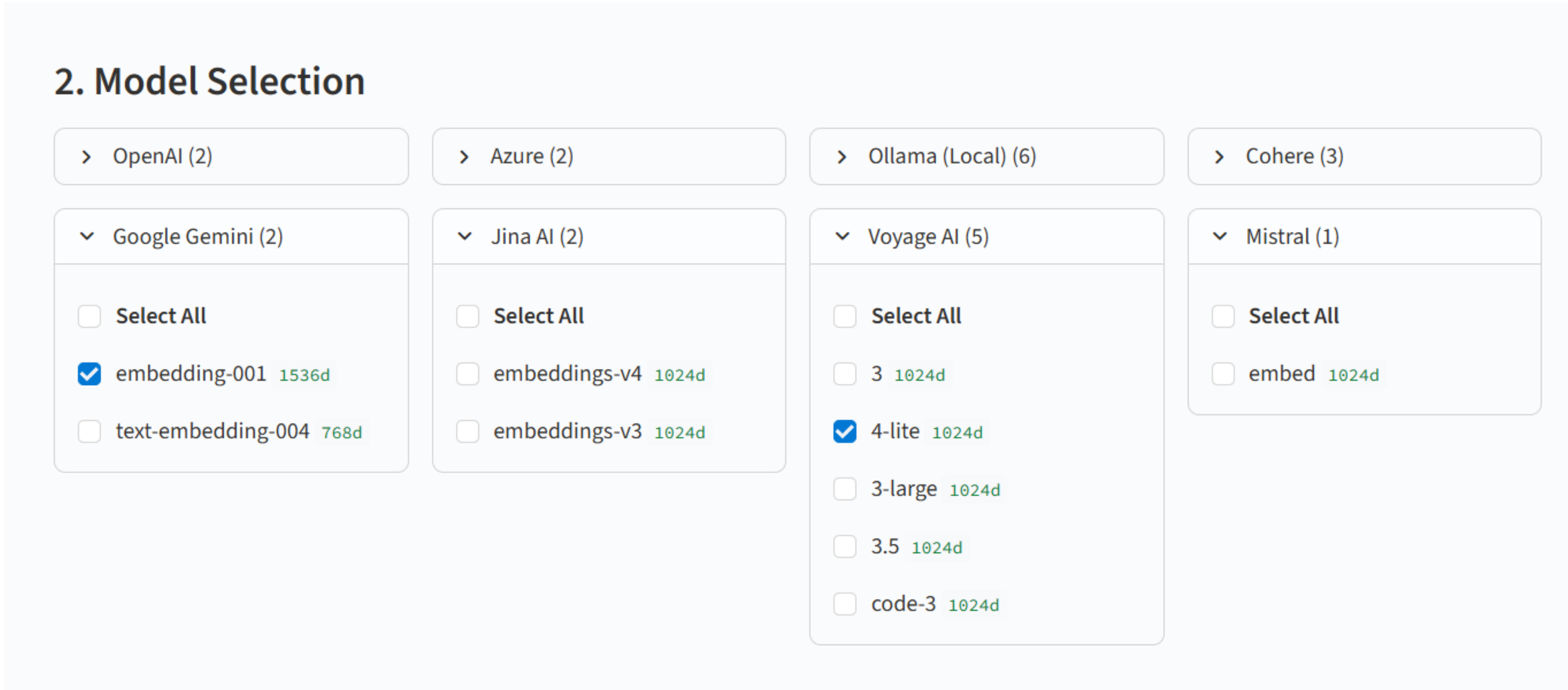
1. Moderate accuracy score (67.4/100)
2. Fast response time (27.4ms)
3. Excellent cost efficiency (100.0/100)

Vectorize All Data

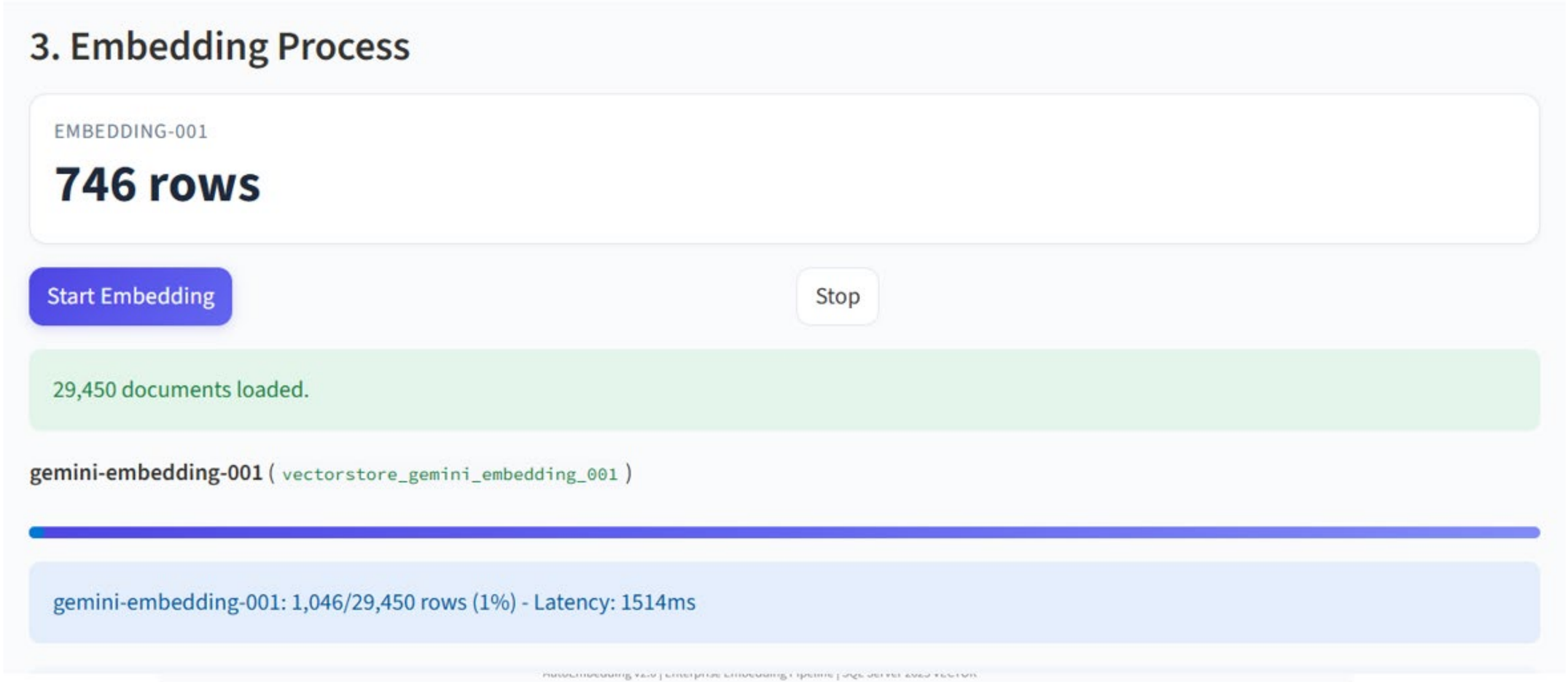
1.Source Selection



2.Model Selection



3. Embedding Progress



4. Embeded Data On SQL Server

```
1  SELECT TOP (1000) *
2    FROM [AUTOEMP].[dbo].[embeddings_gemini_embedding_001]
3
4
```

181 % | No issues found | Ln: 9, Ch: 3 | TABS | CRLF | Windows

Results | Messages

	id	content	embedding	model_name	source	chunk_index	metadata	experiment_id	datasource_id	created_at	updated_at
1	1	My Arcade Burgertime Micro Player Cabinet Console. ...	[2.8047837e-002,-2.1328272e-002,3.1509724e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.24333...	2026-03-02 12:41:45.24333...
2	2	GEN Jumbo JRT Toilet Paper, 2-Ply, White, 9 in Diame...	[-4.3242318e-003,2.0298825e-003,-7.3829181e-003,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.26000...	2026-03-02 12:41:45.26000...
3	3	Kinetic Sand, hand therapy, 5 kg box. Kinetic Sand. Sa...	[1.3602118e-002,1.5698463e-002,-5.3274455e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.27666...	2026-03-02 12:41:45.27666...
4	4	Snack Factory Pretzel Crisps Original Minis, Single-Ser...	[8.2855700e-003,-1.6639078e-002,-1.1962424e-003,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.28666...	2026-03-02 12:41:45.28666...
5	5	Acrylic Ball And Half Ball Twister, 12 Ga,3/8" (10Mm),Or...	[-6.2988782e-003,-1.9315742e-002,6.3013588e-003,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.30333...	2026-03-02 12:41:45.30333...
6	6	7mm NeoSport Hi-Top Booties w/ Zipper. NeoSport. M...	[-1.7430274e-002,-2.2373316e-003,-2.5526227e-00...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.31666...	2026-03-02 12:41:45.31666...
7	7	Beard Comb Two-Sided Green Sandalwood - Brown. A...	[-1.4237930e-002,-1.8639700e-002,-6.3427053e-00...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.33333...	2026-03-02 12:41:45.33333...
8	8	KidCo Foam Edge Protector, Black. KidCo. Edge and C...	[1.3002222e-003,-7.6550399e-003,2.2528848e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.34666...	2026-03-02 12:41:45.34666...
9	9	Coke Coca Cola 27.5x19.25 Delicious and Refreshing ...	[-6.6781021e-003,1.2822798e-002,1.4240593e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.36000...	2026-03-02 12:41:45.36000...
10	10	Sunlite Bicycle MX II Deluxe Foam Handlebar Bike Gri...	[2.8007995e-002,1.4137205e-002,-6.1108097e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.37000...	2026-03-02 12:41:45.37000...
11	11	Sevin Ready to Use Spray Garden Insect Killer, 32 fl oz....	[-2.8863377e-003,-2.7965685e-002,2.1775663e-002,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.38000...	2026-03-02 12:41:45.38000...
12	12	RingStop Tinnitus Treatment Liquid Drops - 1/2oz Bottl...	[-3.2694548e-002,-1.3109672e-002,2.7068932e-002,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.39333...	2026-03-02 12:41:45.39333...
13	13	Motorcycle Kickstand Silver Tone 28cm Length Univers...	[3.5099421e-002,1.2799066e-002,3.4860303e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.40666...	2026-03-02 12:41:45.40666...
14	14	The Bishops Decoy. Harding, H. J.. Contemporary Fant...	[-1.5853059e-003,-2.4946053e-002,1.3978641e-002,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.42000...	2026-03-02 12:41:45.42000...
15	15	Paw Patrol Feeding Bottle with Handles (8 oz.). PAW P...	[-7.5676669e-003,-1.2803688e-002,3.7678391e-002,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.43333...	2026-03-02 12:41:45.43333...
16	16	iPhone 8 Plus Screen Protector, 9H Hardness Temper...	[1.5370823e-002,-8.5132932e-003,-1.0245332e-002,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.44666...	2026-03-02 12:41:45.44666...
17	17	ULTNICE 3-12 Years Children Waterproof Smock Appr...	[3.3067323e-003,4.7660638e-003,5.6565545e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.46000...	2026-03-02 12:41:45.46000...
18	18	Elliptical Trainer Magnetic Control Smooth Quiet Ellipti...	[1.5558712e-002,1.1989940e-002,3.5462349e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.46666...	2026-03-02 12:41:45.46666...
19	19	The Amazing Quality Flitz Metal Pre-Clean - All Metals I...	[-9.9392785e-003,-1.2672594e-002,2.7782940e-003,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.48000...	2026-03-02 12:41:45.48000...
20	20	Suki\s 2001 Honda S2000 Pink \Fast & Furious\" Movie...	[-1.5859442e-002,-1.1712575e-002,7.4614156e-003,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.49000...	2026-03-02 12:41:45.49000...
21	21	Braun LF40US01 ThermoScan Lens Filters for Ear The...	[7.3108827e-003,6.1936900e-003,1.7816681e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.50333...	2026-03-02 12:41:45.50333...
22	22	Saint Eve Girls Unicorn Plush Robe. Saint Eve. Unicorn...	[1.9647865e-002,8.5001756e-003,1.4723883e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.51666...	2026-03-02 12:41:45.51666...
23	23	Transitional AWLT3052-23 Door Mat Rug Rectangle 2 ...	[1.1264286e-002,-1.6828995e-002,5.8127241e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.53000...	2026-03-02 12:41:45.53000...
24	24	Ableware 745260002 Food Bumper by Maddak-White ...	[-8.0044325e-003,2.7870678e-004,2.6086142e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.54000...	2026-03-02 12:41:45.54000...
25	25	Big Sky Thunder Blast Towable Tube for 1-2 People. Bi...	[8.6633004e-003,-2.4720352e-002,1.3161383e-002,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.55333...	2026-03-02 12:41:45.55333...
26	26	Avon SSS SPF 30 Bug Guard Cool N Fabulous 4 ounce...	[2.1879137e-003,-1.9058784e-002,-7.2908234e-003,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.56666...	2026-03-02 12:41:45.56666...
27	27	Aluminum Foil Candy Chocolate Wrappers Tinfoil Pape...	[-2.1970460e-004,2.2487000e-003,6.3543138e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.58000...	2026-03-02 12:41:45.58000...
28	28	(4 Pack) Norcom 80 Sheets 4x4 Graph Filler Paper, 10....	[-6.6534807e-003,-1.6111318e-002,2.3888571e-002,...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.59000...	2026-03-02 12:41:45.59000...
29	29	Alaffia Body Lotion, Passion Fruit, 32 Oz. Alaffia. Body L...	[3.3018140e-003,-9.5227184e-003,1.4282120e-003,-...	gemini-embedding-...	WALMARTITEMS	NULL	NULL	b7346421	4	2026-03-02 12:41:45.60333...	2026-03-02 12:41:45.60333...

5. Vector Serch On SQL Server

```
1  exec searchSimilar_1536
2  @inputText=N'I am looking for a coffee which makes me feel premium'
3  @modelName=N'gemini-embedding-001',
4  @topK=10,
5  @experimentId='b7346421',
6  @datasourceId=4
```

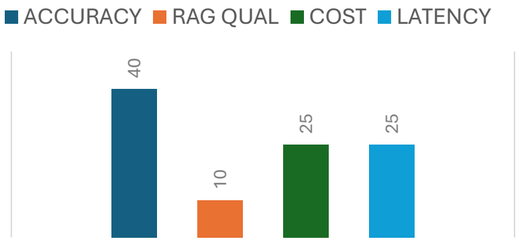
181 % | No issues found | Ln: 4, Ch: 9 | TABS | CRLF | Windows

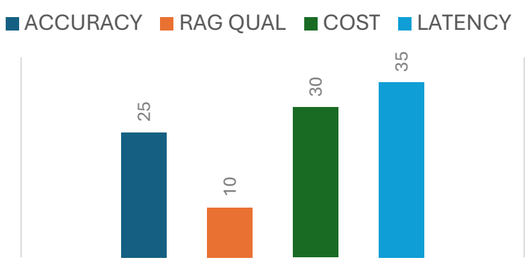
Results | Messages

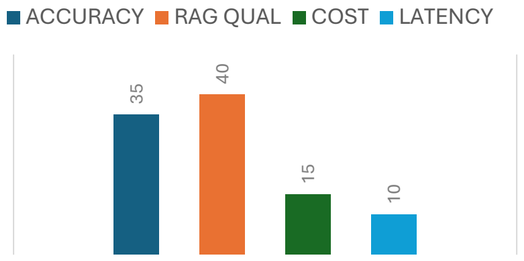
	id	content	similarity	model_name	source	experiment_id	datasource_id
1	612	Toasted Island Coconut Cold Brew Iced Coffee Hot C...	0,578050911426544	gemini-embedding-001	WALMARTITEMS	b7346421	4
2	1157	Folgers Coffee Filter Packs, Special Roast, 40/Carton. ...	0,562315165996552	gemini-embedding-001	WALMARTITEMS	b7346421	4
3	358	illy Ready To Drink Iced Coffee, Caffe Unsweetened, ...	0,555744171142578	gemini-embedding-001	WALMARTITEMS	b7346421	4
4	1585	Nestle Coffee-mate Liquid Creamer Pump, French Va...	0,553637146949768	gemini-embedding-001	WALMARTITEMS	b7346421	4
5	745	Cafe Du Monde French Roast Coffee 13Oz (3 Pack) +...	0,545998156070709	gemini-embedding-001	WALMARTITEMS	b7346421	4

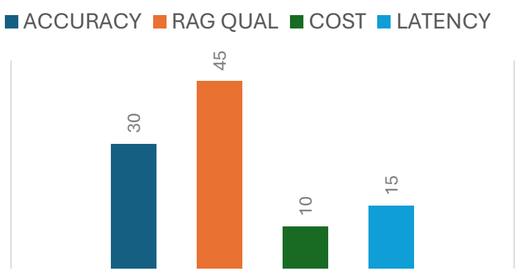
What about other datasets?

What about other datasets?









Strategy

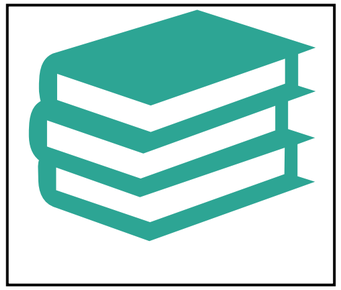
Reviews

Amazon Camera Reviews

E Commerce

Walmart Products

Book



SQL Server 2025 Unveiled Book PDF

Knowledge Base

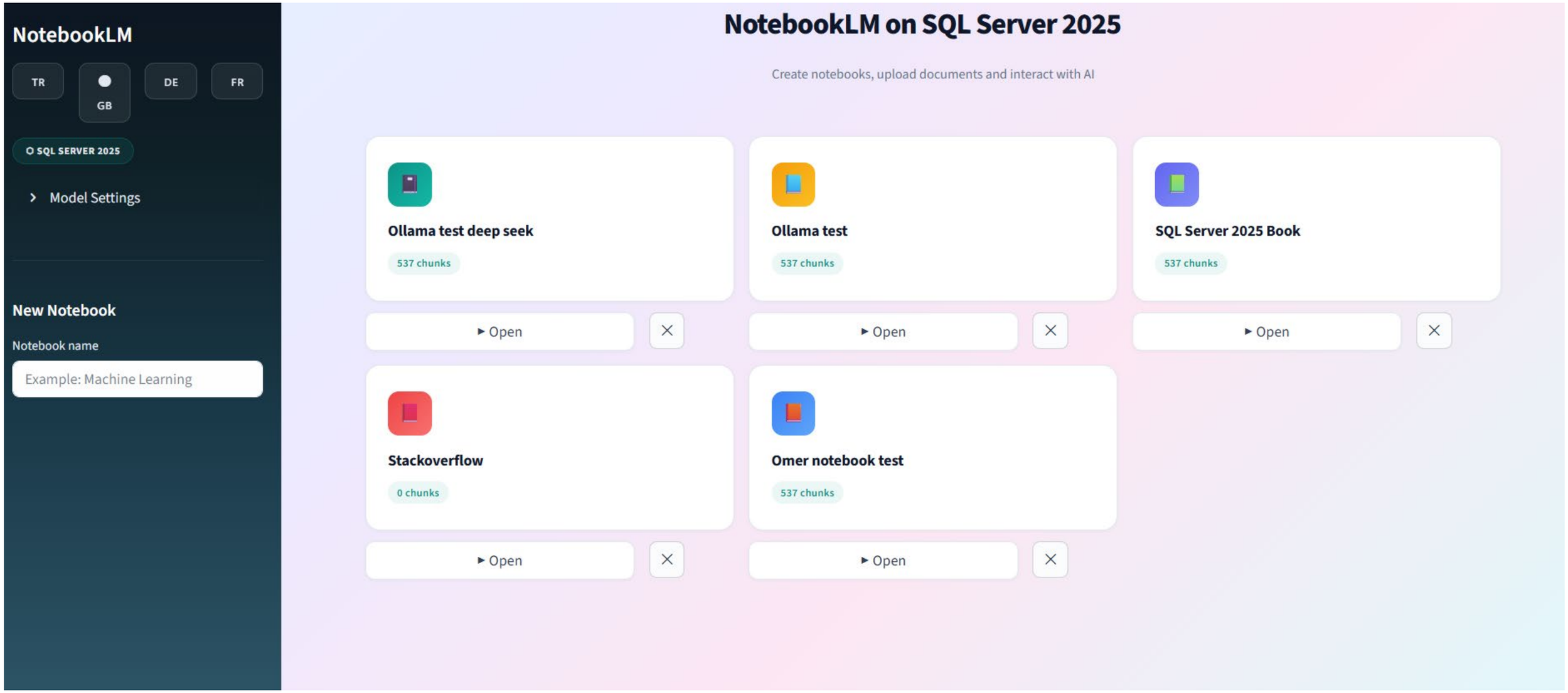
StackOverflow Q&A

Honestly, I have more than you wonder.

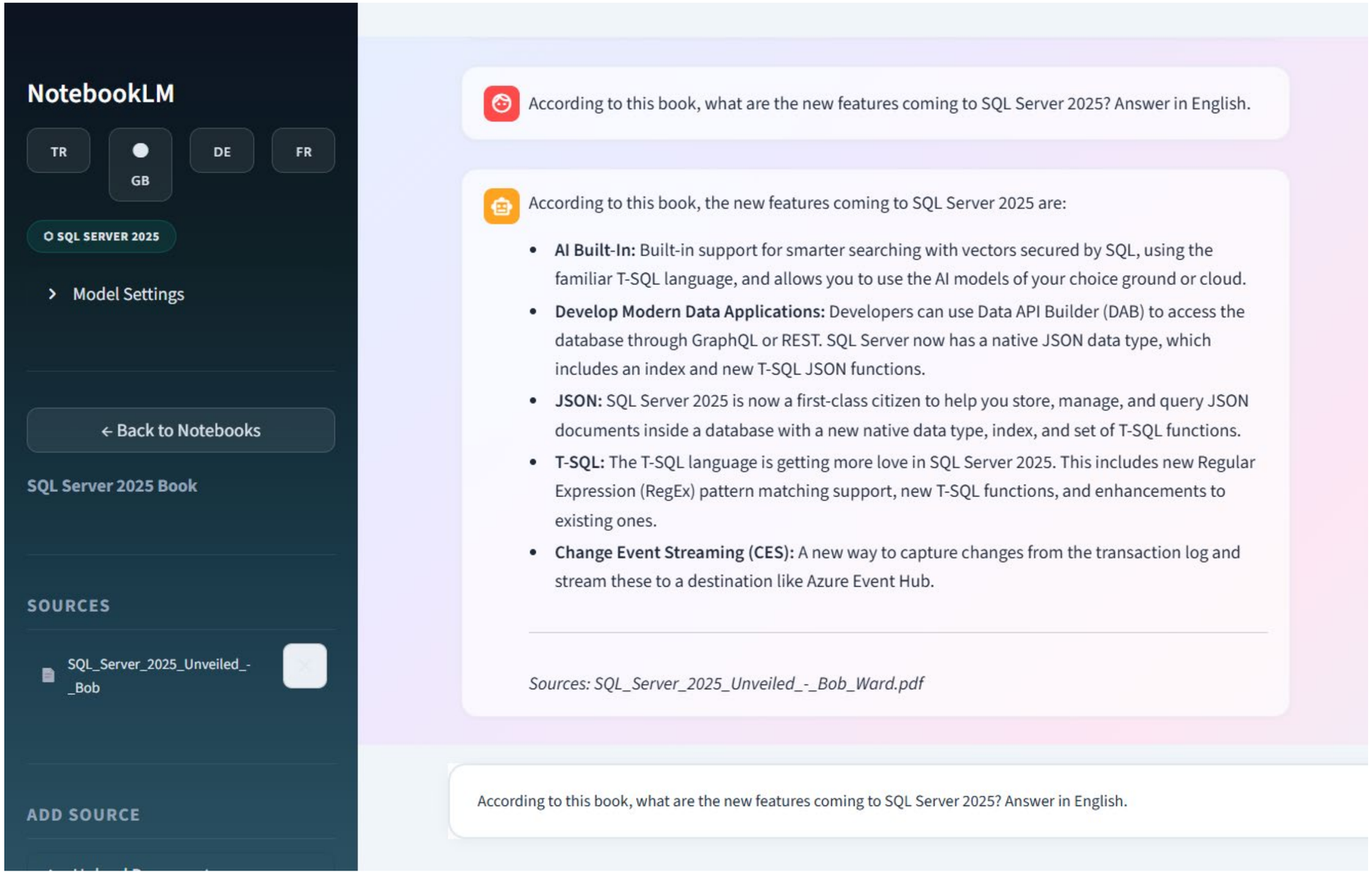
Ladies and gentelmen! Here is....

On Prem Notebook LM
Based on SQL Server 2025
and AutoML Embedding Pipeline

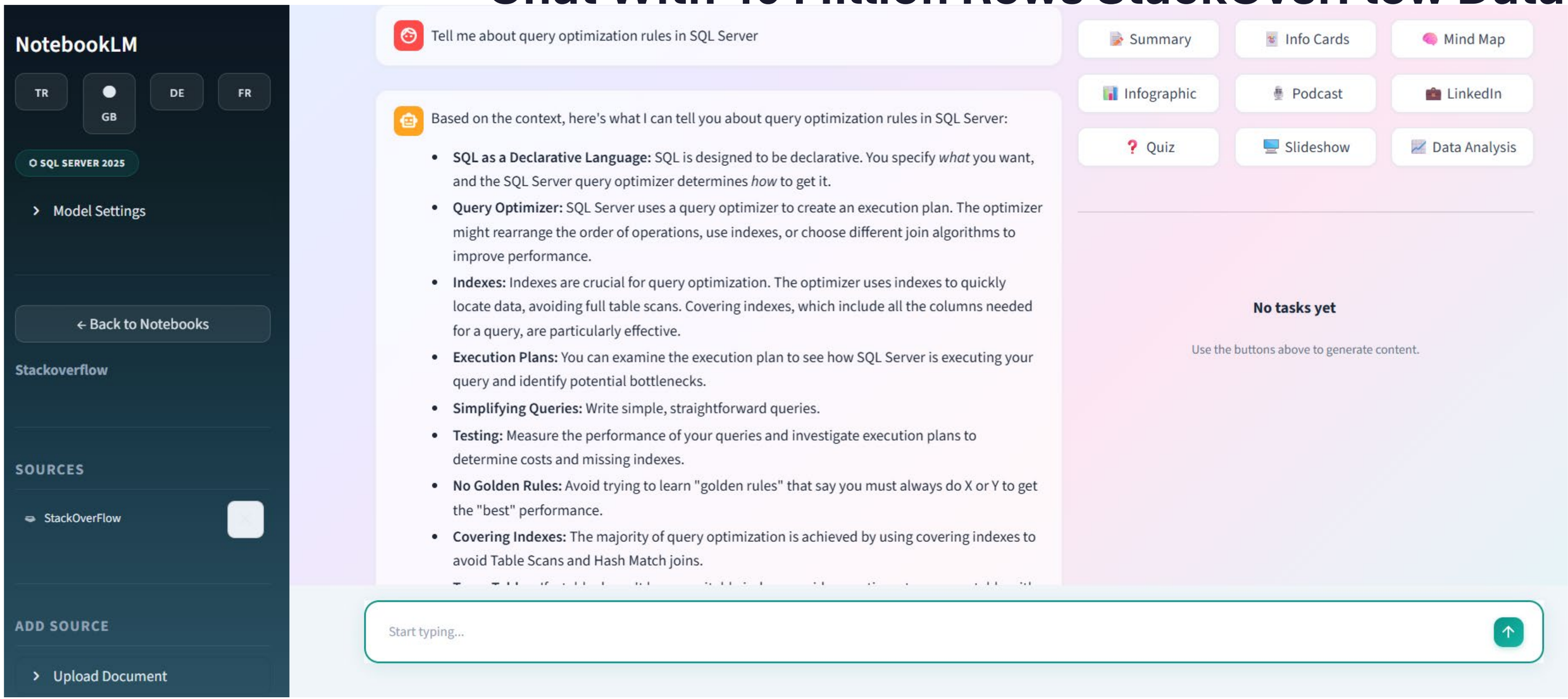
On Prem Notebook LM Based on SQL Server 2025



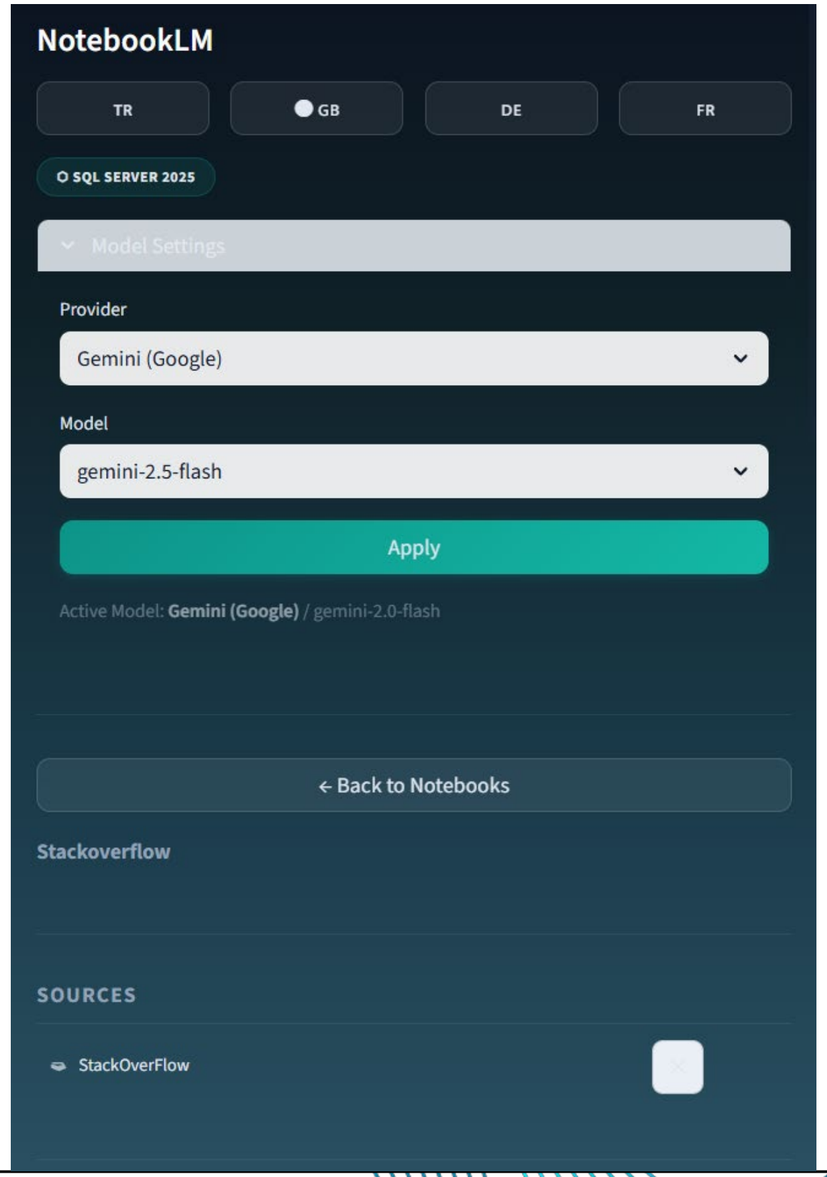
Chat With Your Embedded PDF



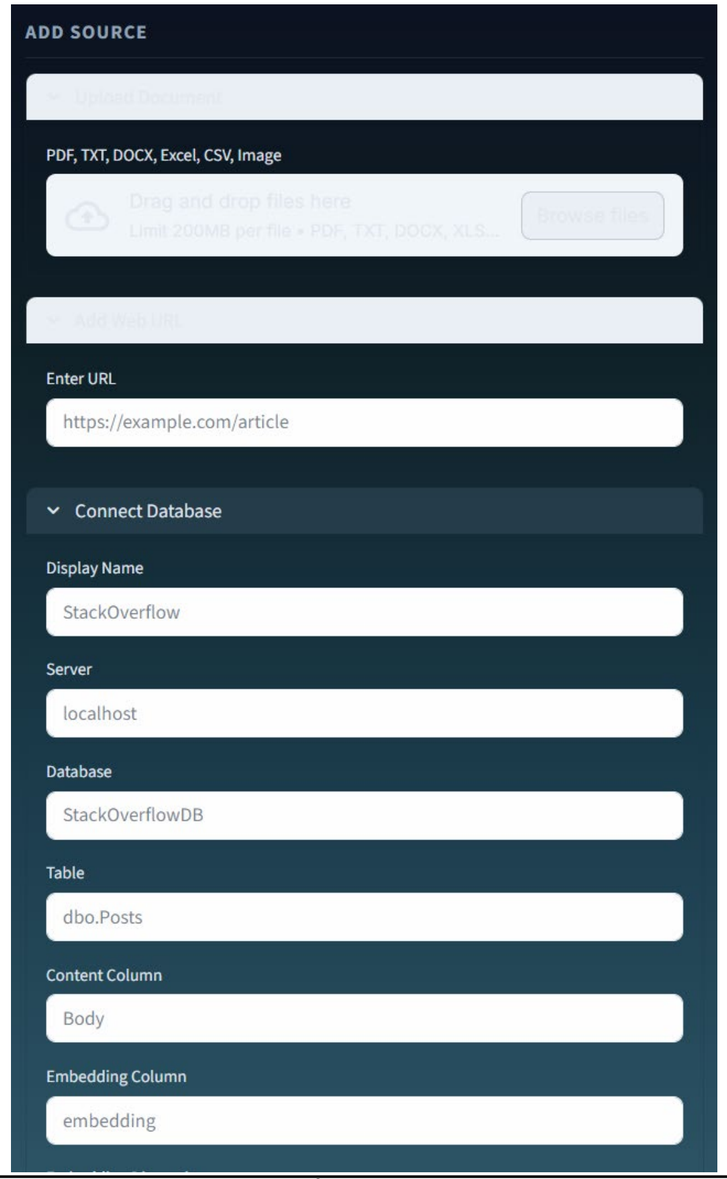
Chat With 40 Million Rows StackOverFlow Data



On Prem Notebook LM Based on SQL Server 2025



Add Source



Summary, Mindmap,Info Cards

Summary | Info Cards | Mind Map

Infographic | Podcast | LinkedIn

Quiz | Slideshow | Data Analysis

Summary

Here's a comprehensive summary of the provided document:

Main Topic:

The document is a book, "SQL Server 2025 Unveiled: The AI-Ready Enterprise Database with Microsoft Fabric Integration," authored by Bob Ward with a foreword by Shireesh Thota. It provides an in-depth look at the features, capabilities, and integration aspects of the new SQL Server 2025.

Key Points:

- **SQL Server 2025 as an AI-Ready Database:** A central theme is the integration of AI capabilities directly into the database engine. This includes features like vector data types, model definitions, embedding generation, and vector search, enabling smarter searching and AI-driven applications.
- **Microsoft Fabric Integration:** The book emphasizes the seamless integration of SQL Server 2025 with Microsoft Fabric, allowing users to mirror SQL data to Fabric for analytics and

Summary

Here's a comprehensive summary of the provided document:

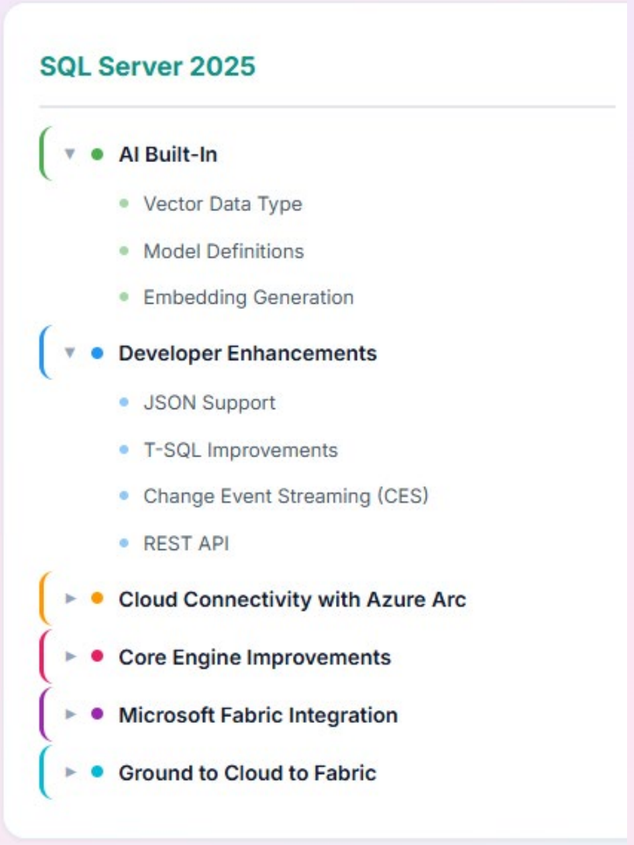
Main Topic:

The document is a book, "SQL Server 2025 Unveiled: The AI-Ready Enterprise Database with Microsoft Fabric Integration," authored by Bob Ward with a foreword by Shireesh Thota. It provides an in-depth look at the features, capabilities, and integration aspects of the new SQL Server 2025.

Key Points:

- **SQL Server 2025 as an AI-Ready Database:** A central theme is the integration of AI capabilities directly into the database engine. This includes features like vector data types, model definitions, embedding generation, and vector search, enabling smarter searching and AI-driven applications.
- **Microsoft Fabric Integration:** The book emphasizes the seamless integration of SQL Server 2025 with Microsoft Fabric, allowing users to mirror SQL data to Fabric for analytics and unified data management.
- **Developer Enhancements:** The book highlights significant improvements for developers, including enhanced JSON support, regular expressions in T-SQL, Change Event Streaming (CES), and REST API capabilities.
- **Core Engine Improvements:** The document details enhancements to the core database engine, focusing on security (Microsoft Entra integration, managed identities), performance (optimized locking, tempdb resource governance), and availability (Always On Availability Groups).

Mind Map



Info Cards

What is the main focus of SQL Server 2025?

SQL Server 2025 is primarily focused on being an AI-ready enterprise database. It achieves this through features like built-in AI capabilities, Microsoft Fabric integration, and enhanced tools for developers to build modern data applications.

How does SQL Server 2025 integrate with Microsoft Fabric?

SQL Server 2025 integrates with Microsoft Fabric through mirroring capabilities, allowing users to mirror SQL Server databases to Fabric. This enables unified data analytics and AI scenarios by leveraging Fabric's services like OneLake and Power BI.

What are some key security enhancements in SQL Server 2025?

SQL Server 2025 enhances security through Microsoft Entra integration, managed identities, and security caches. It also includes encryption and password enhancements, along with features like Row-Level Security (RLS), Transparent Data Encryption (TDE), and Dynamic Data Masking (DDM).

What are some of the core engine improvements in SQL Server 2025?

SQL Server 2025 includes core engine improvements such as optimized

Thank you.
It’s honor to be here and talk to you.

Sound off.
The mic is all yours.
Influence the product roadmap.

Join the Fabric User Panel

Share your feedback directly with our Fabric product group and researchers.

https://aka.ms/JoinFabricUserPanel

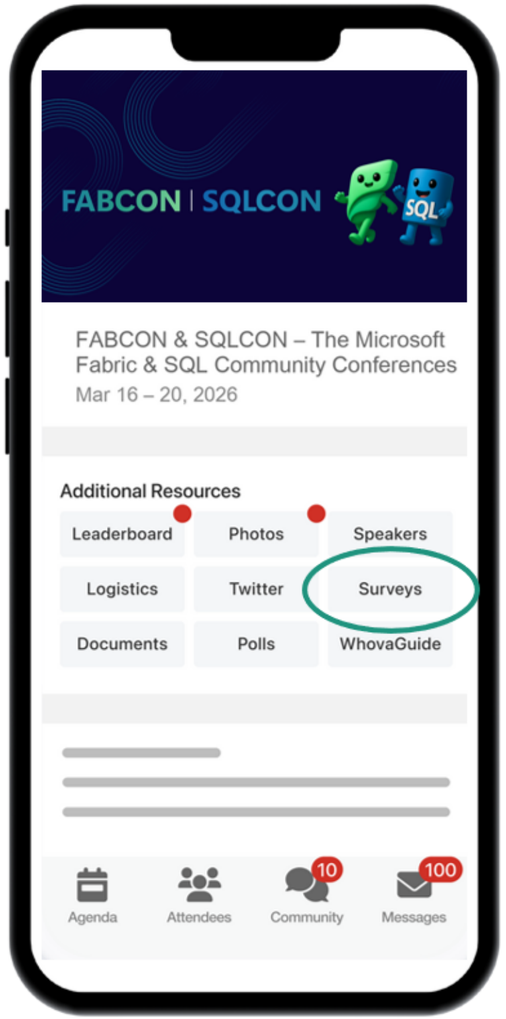
Join the SQL User Panel

Influence our SQL roadmap and ensure it meets your real-life needs

https://aka.ms/JoinSQLUserPanel

How was the session?

Complete Session Surveys in Whova for your chance to WIN PRIZES!



Two Fabric Certifications, One FREE Exam Included

Attendees can take the Fabric Analytics Engineer or Fabric Data Engineer exam for free. Be part of the 2 fastest growing role-based certifications in Microsoft history.

Request your voucher by March 31, 2026.

https://aka.ms/GetDataCertified