<epam>

RAG deep dive

JUNE 2025

DIALX COMMUNITY
POWERED BY <epam>

There are no magic! It is simple REST API!

Before we start:

- Raise your hand and ask questions if you have any
- It is better to ask questions when you have
- Also, type them in chat
- We will need DIAL API key for this session

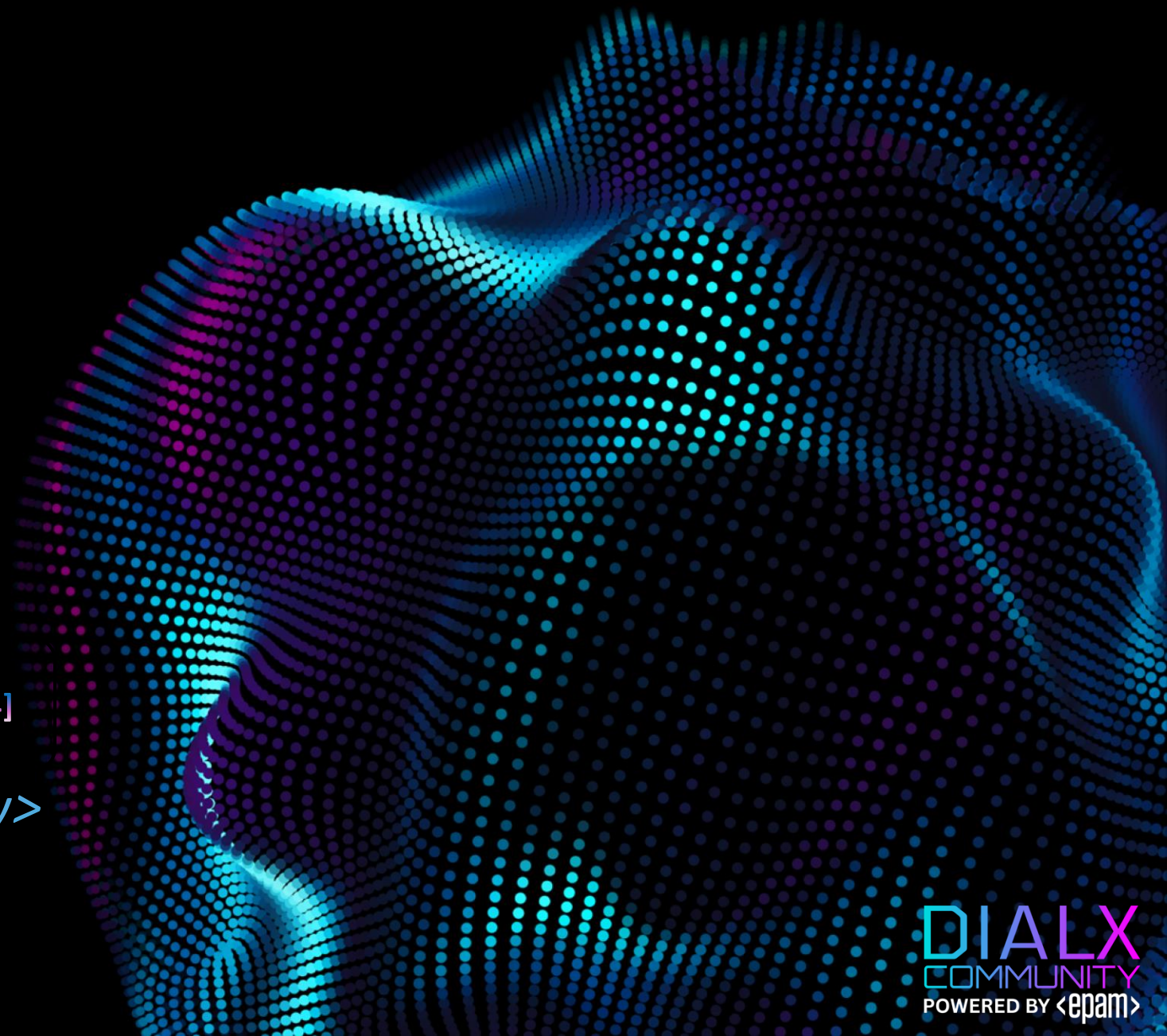
Agenda:

- **Presentation:**
  - **Embeddings**
  - **Search in Vector DB**
  - ***top_k* and *score* paramters for search**
- **Workshop:**
  - **Create MicrowaveRAG application**

Embeddings

[0.026394594,0.0014399963,-0.023680586,...,-0.019032994]

< Application components overview>



DIALX COMMUNITY POWERED BY <epam>

Embeddings:

text	embedding
isk of electric shock.⏎Note:⏎1. If you have any questio…	[0.007929899,-0.0103833815,-0.06096753,-0.0041557755,-0.042187788,0.01960363,-0.048827…
be adjusted or repaired by anyone except qualified serv…	[0.019625312,0.07142368,-0.0130586205,-0.018416643,-0.053729743,0.0055418145,-0.053580…
dangerous to repair or maintain the appliance by no oth…	[0.007216289,0.049263548,-0.020450495,-0.012869484,-0.06486845,0.012693635,-0.01578726…
r.⏎10. Do not fry food in the oven. Hot oil can damage …	[0.022751395,0.030335193,0.016967906,-0.007043706,-0.013974892,0.002911436,-0.01603399…
over, and then press START/QUICK⏎START button to resume…	[0.00082089123,0.0022738606,-0.031506665,-0.014780904,-0.033373725,-0.008920405,0.0358…
ation, the oven must have sufficient airflow. Allow minim…	[0.009636983,0.023800073,-0.0062219356,-0.027741412,-0.02087623,0.02209255,-0.01481803…
ectly connected to a low voltage power supply network w…	[-0.007313127,0.037871324,-0.03555582,-0.018665519,-0.02153417,0.045306657,-0.03712521…
r.⏎⏎AUTO COOK⏎For the following food or cooking mode, i…	[-0.0016354206,-0.013394599,-0.042835053,-0.040621832,0.008080563,-0.018178385,0.01598…
and 15% time⏎the oven will stop working in one cycle. U…	[-0.01445577,0.02532314,-0.022590334,-0.055830944,-0.026038265,0.018963622,-0.00489733…
trim⏎should not be used.⏎3. Do not use recycled paper p…	[0.025805177,0.016437544,0.0084770955,-0.011957668,0.008708227,0.017008575,6.410294e-0…
e dial to select others food code.⏎2. Press START/QUICK…	[-0.0048348736,-0.026497386,-0.012877011,-0.07045187,-0.045177445,-0.017325213,0.01469…
and the door or allow soil or⏎cleaner residue to accumu…	[0.011242517,0.041140627,-0.027245961,0.016288064,-0.012025224,0.002053797,0.009373076…
l sound every two minutes⏎until user presses any button…	[-0.012621777,0.002297385,0.0051053,0.005726547,-0.026842805,0.025588008,0.043130554,0…
dd more as⏎needed. Foods severely overcooked can smoke …	[0.03621854,-0.023570212,-0.010781076,0.004781386,-0.046626166,-0.0075730784,0.0137625…
)⏎1. Input the first microwave cooking program. Do not p…	[-0.008847023,-0.021002412,-0.029569935,-0.030482586,-0.064433254,-0.009645594,0.01033…
d heating of warming pads, slippers, sponges, damp clot…	[-0.007398685,0.051085625,-0.0047710882,-0.016767368,-0.023013156,-0.020438727,-0.0279…
select 24- hour clock.⏎2. Turn MENU/TIME dial to set ho…	[-0.0017236428,0.03179804,-0.002225171,-0.03276651,-0.064979605,-0.0032282274,0.016648…
. The microwave oven shall not be placed in a cabinet.⏎…	[-0.01100224,0.0060034357,-0.01566008,0.008335399,-0.02978581,0.036069326,-0.018156435…
m(D)⏎25Litres⏎Approx.14kg⏎⏎Rated Microwave Power Output…	[0.026943313,0.017219318,-0.012886474,-0.027789962,-0.059763357,0.021788724,0.01533926…
fore using the appliance and keep for future reference.…	[-0.0037010042,0.04981585,-0.024775151,-0.04551405,-0.011770189,-0.00916123,-0.0332725…
auto cooking menus.⏎START/QUICK START (the dial)⏎Press…	[-0.018388081,-0.0135191595,-0.024863113,-0.06292922,-0.02891001,-0.023560518,0.031161…
re that they are suitable for use in microwave⏎oven.⏎16…	[0.039369497,0.04368466,-0.03406439,-0.02761703,-0.06604735,0.020116266,-0.027363198,0…
or defrosting, cooking and steaming of food only.⏎35. U…	[0.013555558,0.064180404,-0.0076864623,-0.023551071,0.0023152744,-0.0037841045,-0.0294…
)⏎40% (40)⏎30% (30)⏎20% (20)⏎10% (10)⏎0% (00)⏎⏎NOTE: Yo…	[-0.01830597,-0.012414614,0.0066955937,-0.030407824,0.0057764654,0.0033318396,0.018025…
mer or⏎separate remote-control system.⏎30. The microwav…	[0.015503083,0.023018092,-0.028411057,-0.011204931,-0.054389197,0.010772413,-0.0117928…

Embeddings are numerical representations of data (like text, images, or audio) in a high-dimensional vector space where semantic similarities are preserved as geometric relationships.

Embeddings :

Input Text

Start with the raw text you want to convert.

Example: "I love machine learning."

Embeddings :

Input Text

Text Preprocessing (Optional but common)

Depending on the model, some preprocessing may be required:
- Lowercasing
- Removing punctuation

Example after preprocessing: ["I", "love", "machine", "learning", "."]

Note: Most modern embedding models like OpenAI's or SentenceTransformers handle this internally.

Embeddings :

Input Text

Text Preprocessing (Optional but common)

Tokenization

The text is split into tokens, usually using a tokenizer specific to the embedding model (e.g., BPE for OpenAI models).

Each token is mapped to an integer ID via a vocabulary.

Example:

"I" → 374
"love" → 1438
"machine" → 5826
"learning" → 8372
"." → 13

https://platform.openai.com/tokenizer

Embeddings :

Input Text

Text Preprocessing (Optional but common)

Tokenization

Model Encoding

The token IDs are fed into a pre-trained embedding model, such as:
- OpenAI's embedding models (e.g., text-embedding-3-small)
- Sentence-BERT (SBERT)
- HuggingFace transformer models

The model computes embedding vectors for the text.

Output: a vector of real numbers (floats), typically of size 1536, 768, 384, etc.

Example: [0.0213, -0.5732, 0.9981, ..., 0.0724]

Embeddings :

Input Text

Text Preprocessing (Optional but common)

Tokenization

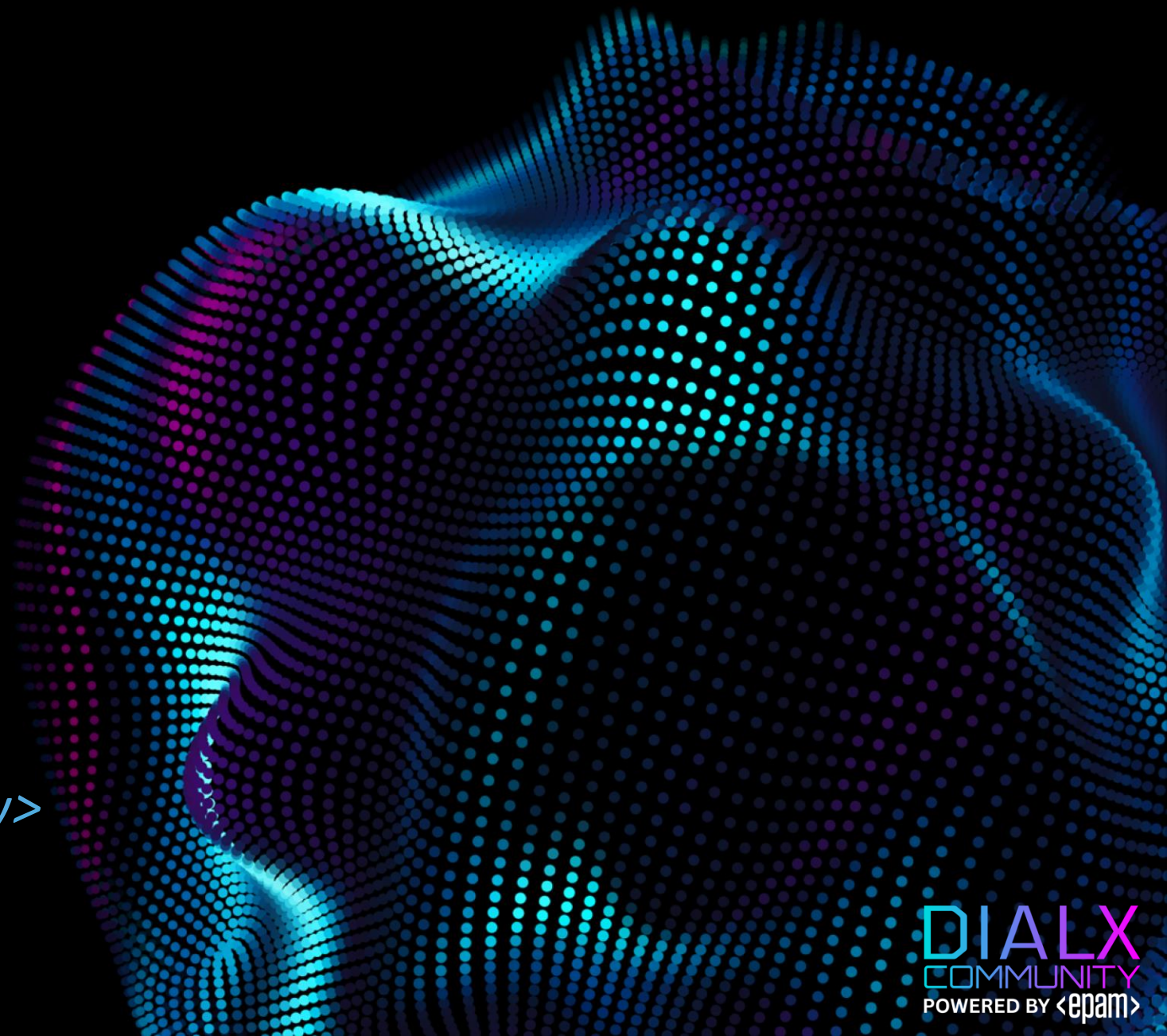
Model Encoding

Post-processing (optional)

Depending on the use case, you might:
- Normalize the embedding (e.g., L2 normalization)
- Average word embeddings for sentence-level tasks (if not already handled)
- Store the embedding in a vector database (e.g., PGVector, Pinecone, FAISS)

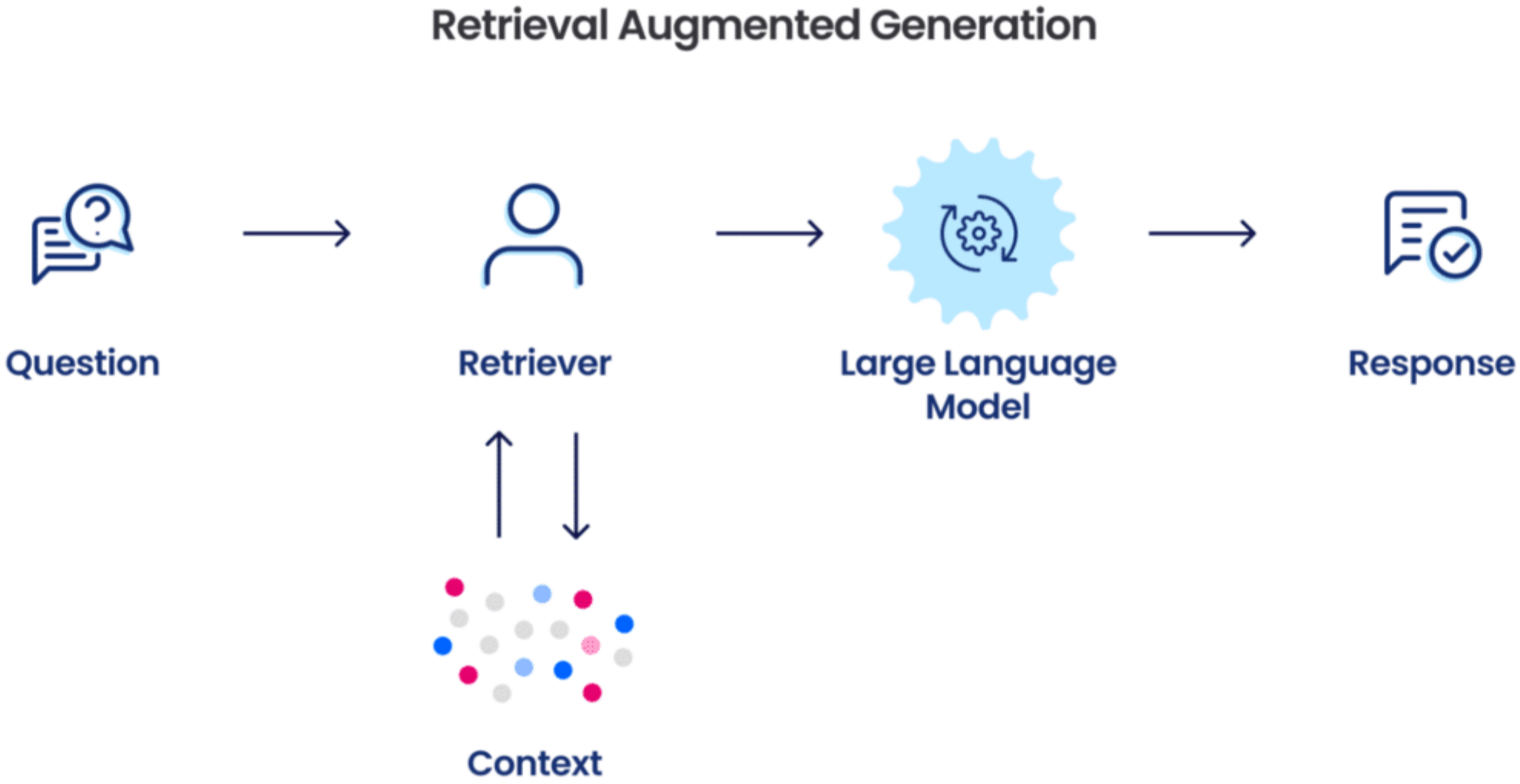
Application

< Application components overview>

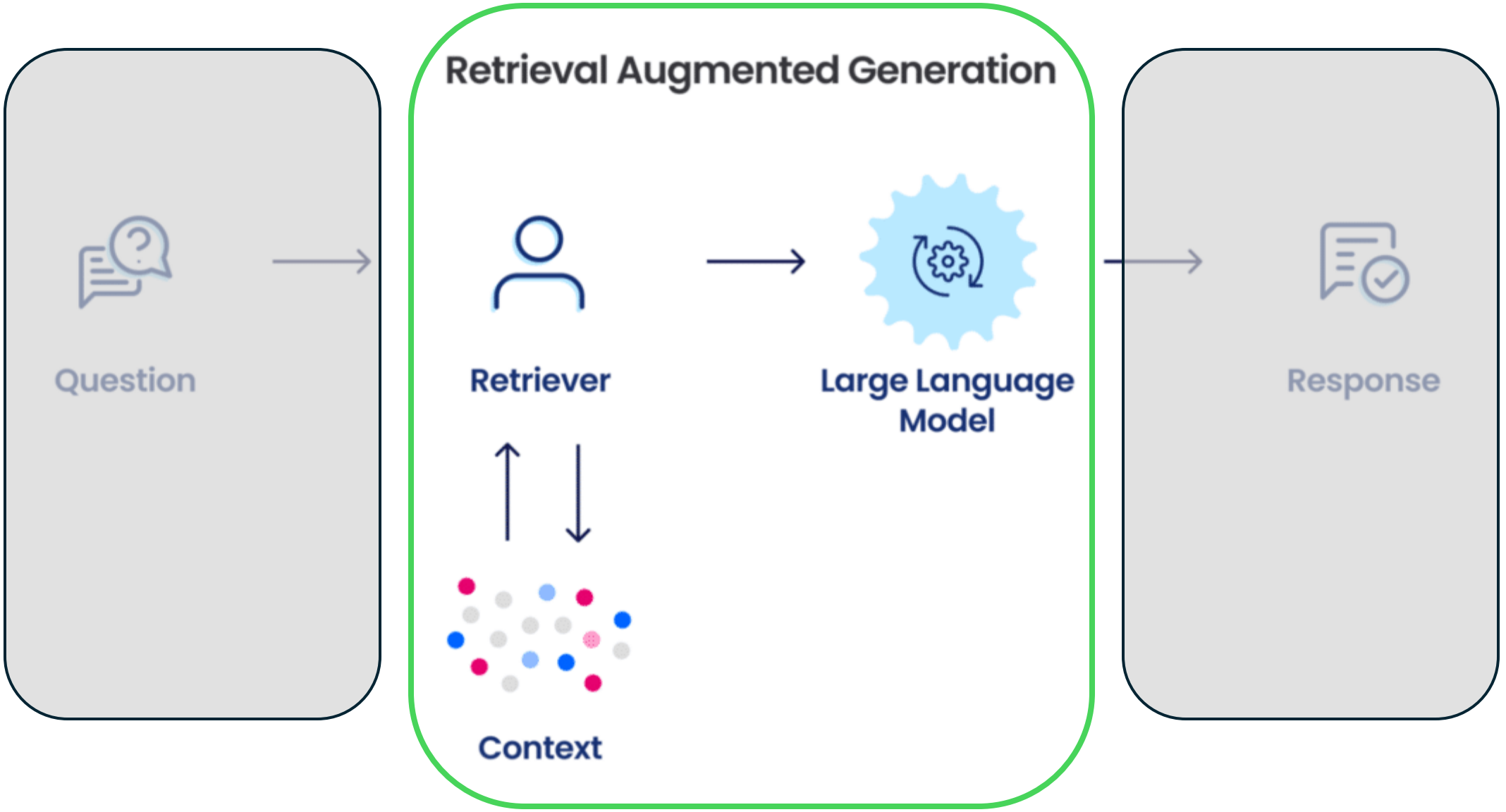


DIALX COMMUNITY POWERED BY <epam>

Chat with RAG under the hood:

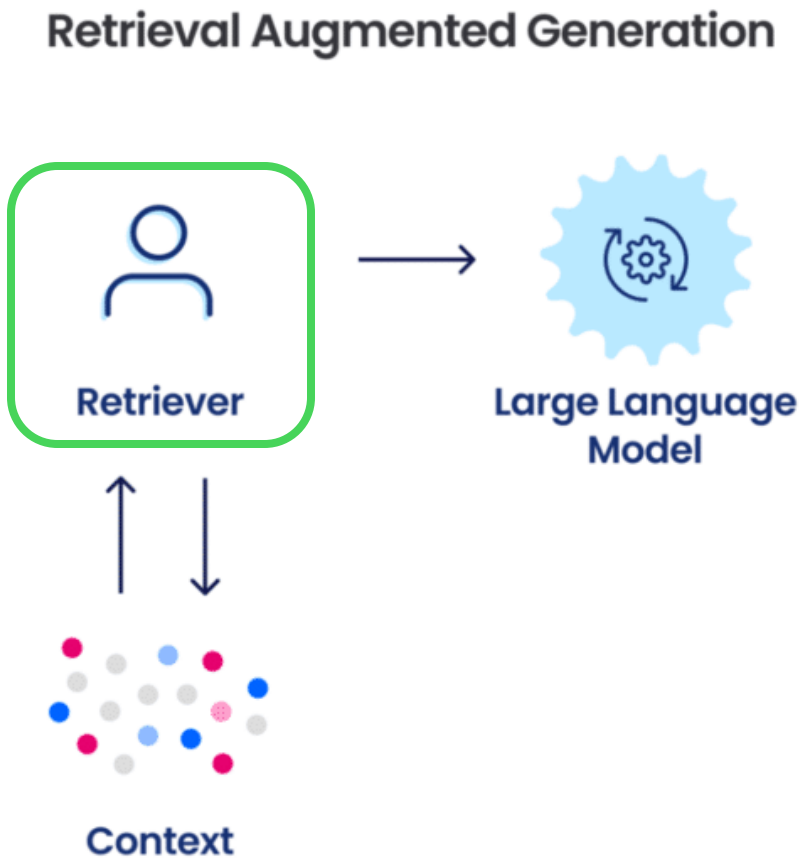


Chat with RAG under the hood:



Retriever:

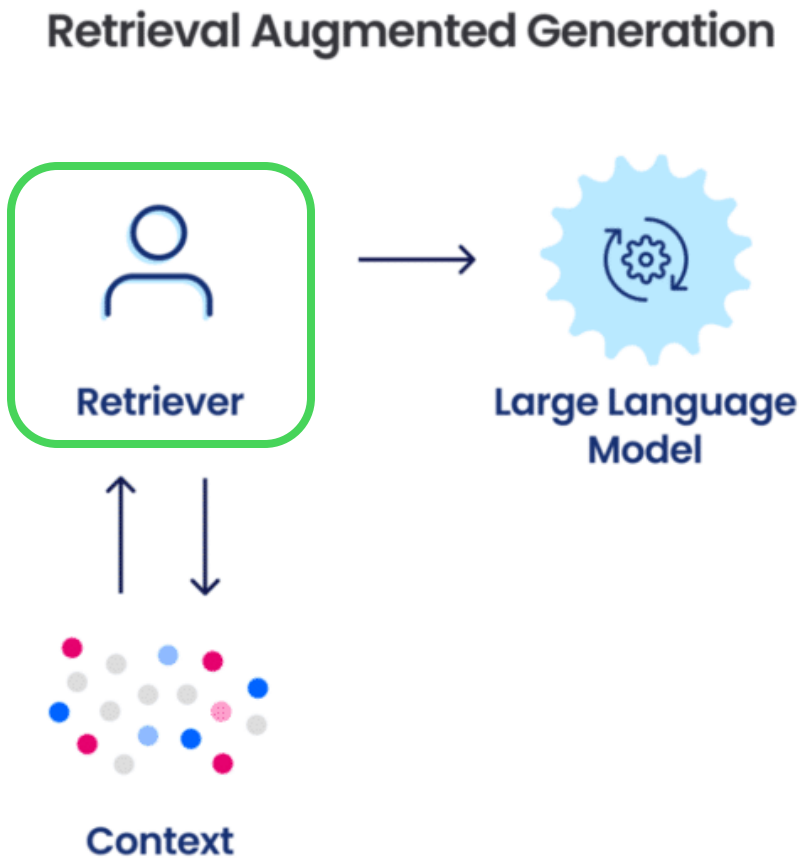
Convert user request into embeddings



Retriever:

Convert user request into embeddings

Embeddings can be generated via neural models

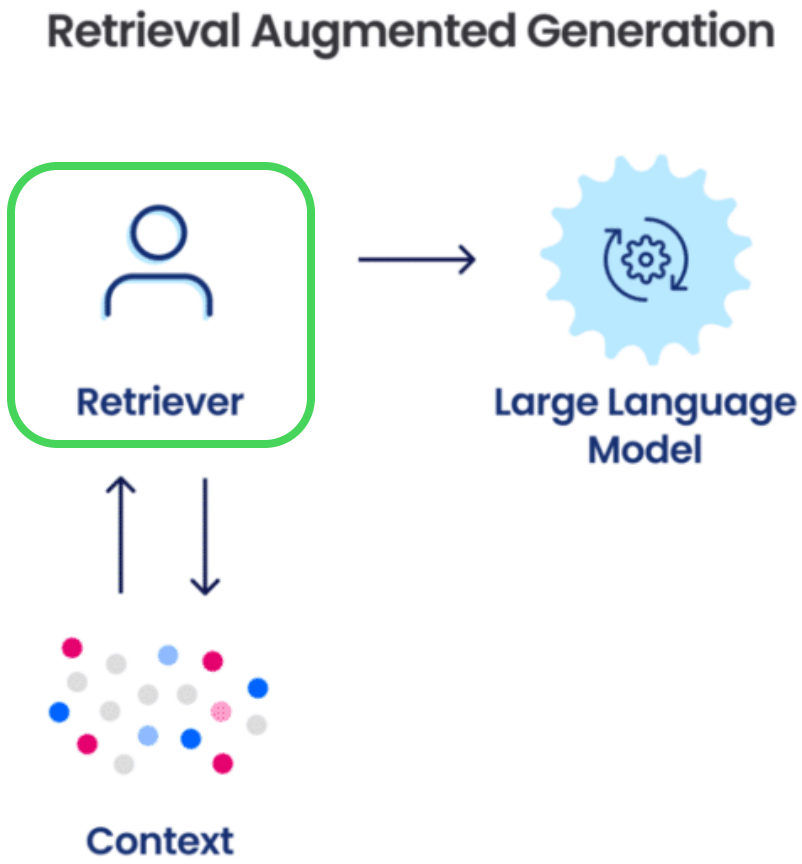


Retriever:

Convert user request into embeddings

Embeddings can be generated via neural models

Open AI models:
text-embedding-3-small (1536)
text-embedding-3-large (3072)



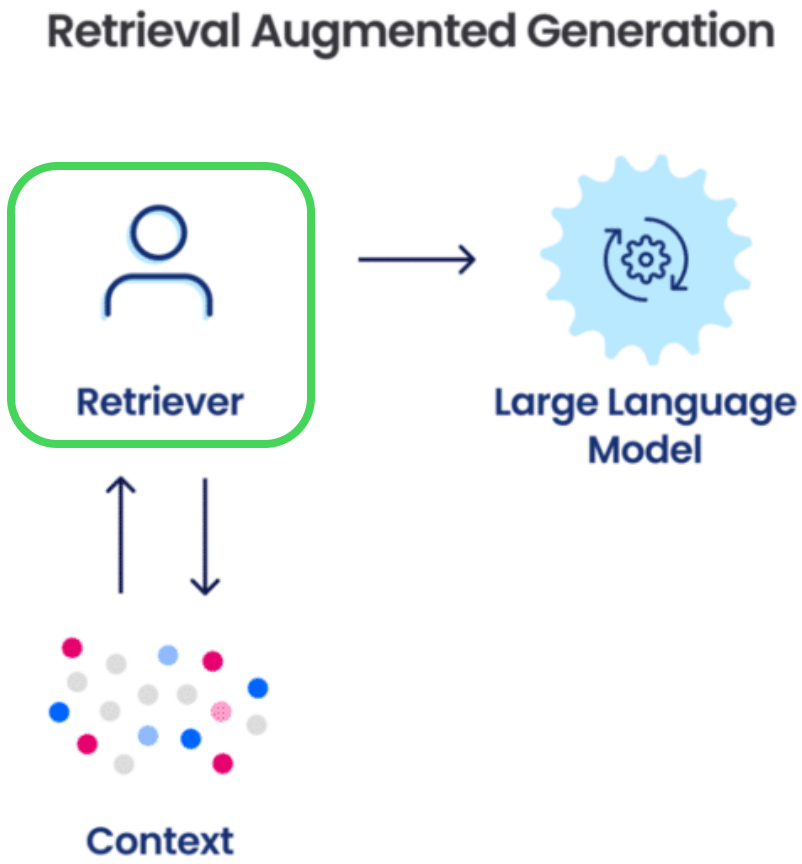
Retriever:

Convert user request into embeddings

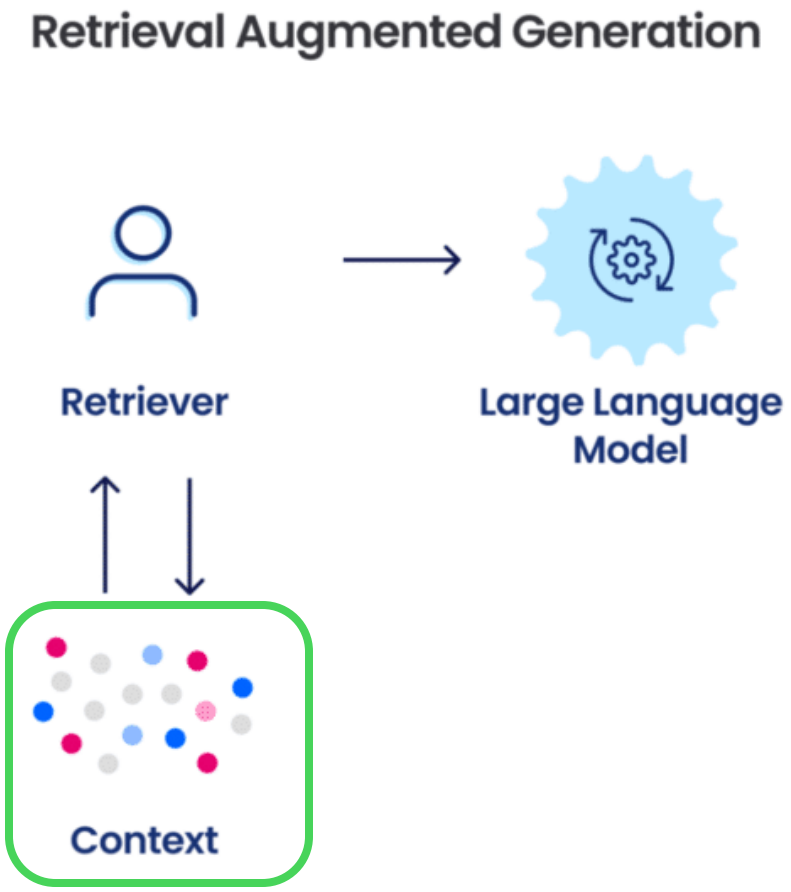
Embeddings can be generated via neural models

Open AI models:
text-embedding-3-small (1536)
text-embedding-3-large (3072)

Hugging Face models:
sentence-transformers/all-MiniLM-L6-v2
...



Context:



‘Context’ is some Vector DB with data and its embeddings and some API to communicate with such DB

Context:

As a Source for context generation can be taken:

.TXT, .PDF, .XLSX, .XML, .HTML DOCUMENTS

CONFLUENCE

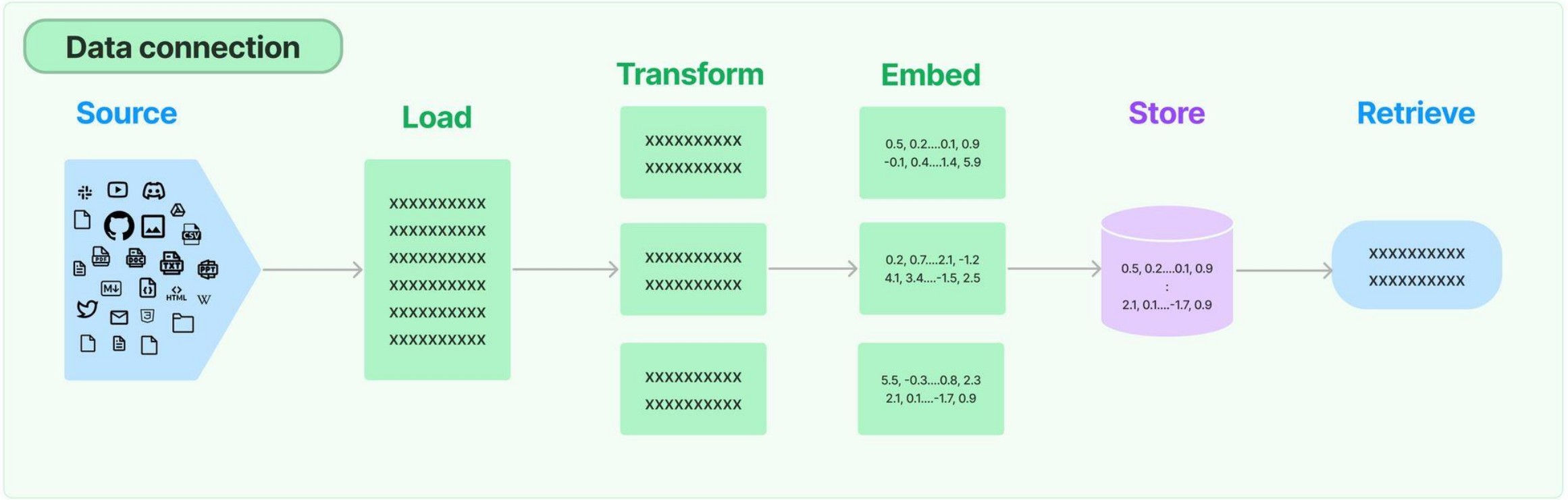
GITHUB

DISCORD

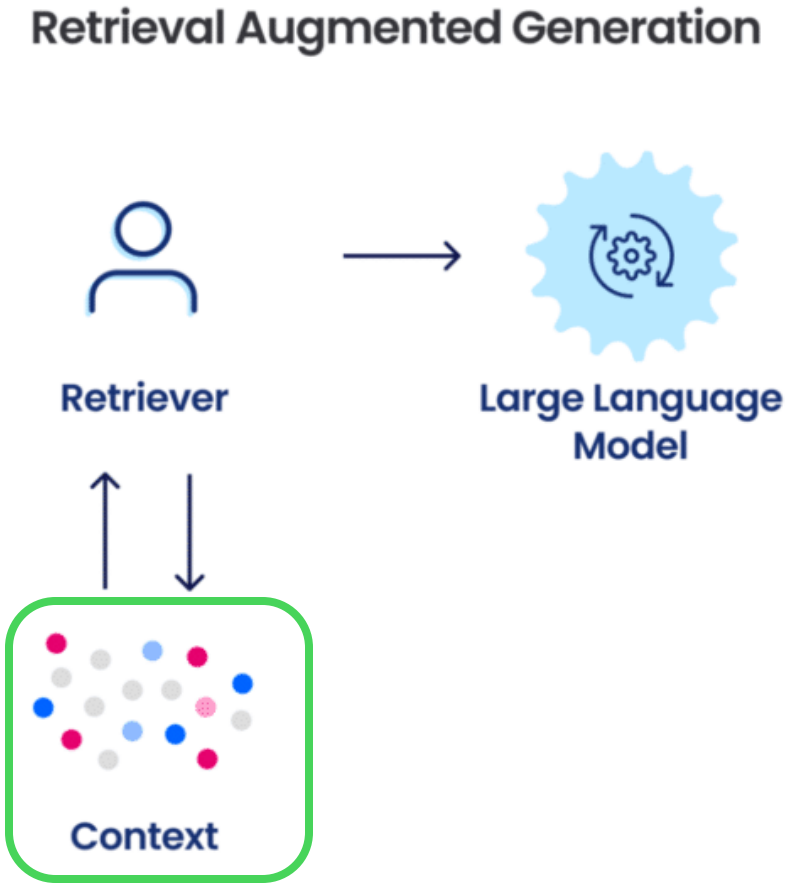
WEB RESOURCE

Context:

Pipeline of data transformation:



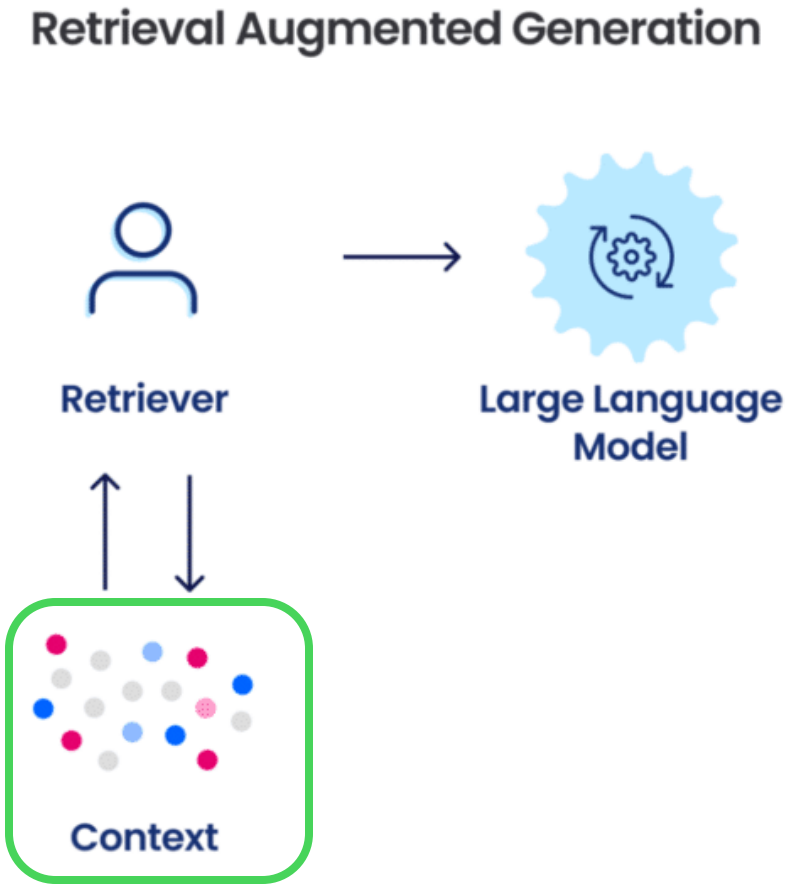
Context:



‘document_name’ contains some metadata related to document

document_name	text	embedding
files/microwave_manual.txt	isk of electric shock.↵Note:↵1. If you have any questio…	[0.007929899,-0.0103833815,-0.06096753,-0.0041557755,-0.042187788,0.01960363,-0.048827…
files/microwave_manual.txt	be adjusted or repaired by anyone except qualified serv…	[0.019625312,0.07142368,-0.0130586205,-0.018416643,-0.053729743,0.0055418145,-0.053580…
files/microwave_manual.txt	dangerous to repair or maintain the appliance by no oth…	[0.007216289,0.049263548,-0.020450495,-0.012869484,-0.06486845,0.012693635,-0.01578726…
files/microwave_manual.txt	r.↵10. Do not fry food in the oven. Hot oil can damage …	[0.022751395,0.030335193,0.016967906,-0.007043706,-0.013974892,0.002911436,-0.01603399…
files/microwave_manual.txt	over, and then press START/QUICK↵START button to resume…	[0.00082089123,0.0022738606,-0.031506665,-0.014780904,-0.033373725,-0.008920405,0.0358…
files/microwave_manual.txt	ation, the oven must have sufficient airflow. Allow minim…	[0.009636983,0.023800073,-0.0062219356,-0.027741412,-0.02087623,0.02209255,-0.01481803…
files/microwave_manual.txt	ectly connected to a low voltage power supply network w…	[-0.007313127,0.037871324,-0.03555582,-0.018665519,-0.02153417,0.045306657,-0.03712521…
files/microwave_manual.txt	r.↵↵AUTO COOK↵For the following food or cooking mode, i…	[-0.0016354206,-0.013394599,-0.042835053,-0.040621832,0.008080563,-0.018178385,0.01598…
files/microwave_manual.txt	and 15% time↵the oven will stop working in one cycle. U…	[-0.01445577,0.02532314,-0.022590334,-0.055830944,-0.026038265,0.018963622,-0.00489733…
files/microwave_manual.txt	trim↵should not be used.↵3. Do not use recycled paper p…	[0.025805177,0.016437544,0.0084770955,-0.011957668,0.008708227,0.017008575,6.410294e-0…
files/microwave_manual.txt	e dial to select others food code.↵2. Press START/QUICK…	[-0.0048348736,-0.026497386,-0.012877011,-0.07045187,-0.045177445,-0.017325213,0.01469…
files/microwave_manual.txt	and the door or allow soil or↵cleaner residue to accumu…	[0.011242517,0.041140627,-0.027245961,0.016288064,-0.012025224,0.002053797,0.009373076…
files/microwave_manual.txt	l sound every two minutes↵until user presses any button…	[-0.012621777,0.002297385,0.0051053,0.005726547,-0.026842805,0.025588008,0.043130554,0…
files/microwave_manual.txt	dd more as↵needed. Foods severely overcooked can smoke …	[0.03621854,-0.023570212,-0.010781076,0.004781386,-0.046626166,-0.0075730784,0.0137625…
files/microwave_manual.txt	)↵1. Input the first microwave cooking program. Do not p…	[-0.008847023,-0.021002412,-0.029569935,-0.030482586,-0.064433254,-0.009645594,0.01033…
files/microwave_manual.txt	d heating of warming pads, slippers, sponges, damp clot…	[-0.007398685,0.051085625,-0.0047710882,-0.016767368,-0.023013156,-0.020438727,-0.0279…
files/microwave_manual.txt	select 24- hour clock.↵2. Turn MENU/TIME dial to set ho…	[-0.0017236428,0.03179804,-0.002225171,-0.03276651,-0.064979605,-0.0032282274,0.016648…
files/microwave_manual.txt	. The microwave oven shall not be placed in a cabinet.↵…	[-0.01100224,0.0060034357,-0.01566008,0.008335399,-0.02978581,0.036069326,-0.018156435…
files/microwave_manual.txt	m(D)↵25Litres↵Approx.14kg↵↵Rated Microwave Power Output…	[0.026943313,0.017219318,-0.012886474,-0.027789962,-0.059763357,0.021788724,0.01533926…
files/microwave_manual.txt	fore using the appliance and keep for future reference.…	[-0.0037010042,0.04981585,-0.024775151,-0.04551405,-0.011770189,-0.00916123,-0.0332725…
files/microwave_manual.txt	auto cooking menus.↵START/QUICK START (the dial)↵Press…	[-0.018388081,-0.0135191595,-0.024863113,-0.06292922,-0.02891001,-0.023560518,0.031161…
files/microwave_manual.txt	re that they are suitable for use in microwave↵oven.↵16…	[0.039369497,0.04368466,-0.03406439,-0.02761703,-0.06604735,0.020116266,-0.027363198,0…
files/microwave_manual.txt	or defrosting, cooking and steaming of food only.↵35. U…	[0.013555558,0.064180404,-0.0076864623,-0.023551071,0.0023152744,-0.0037841045,-0.0294…
files/microwave_manual.txt	)↵40% (40)↵30% (30)↵20% (20)↵10% (10)↵0% (00)↵↵NOTE: Yo…	[-0.01830597,-0.012414614,0.0066955937,-0.030407824,0.0057764654,0.0033318396,0.018025…
files/microwave_manual.txt	mer or↵separate remote-control system.↵30. The microwav…	[0.015503083,0.023018092,-0.028411057,-0.011204931,-0.054389197,0.010772413,-0.0117928…

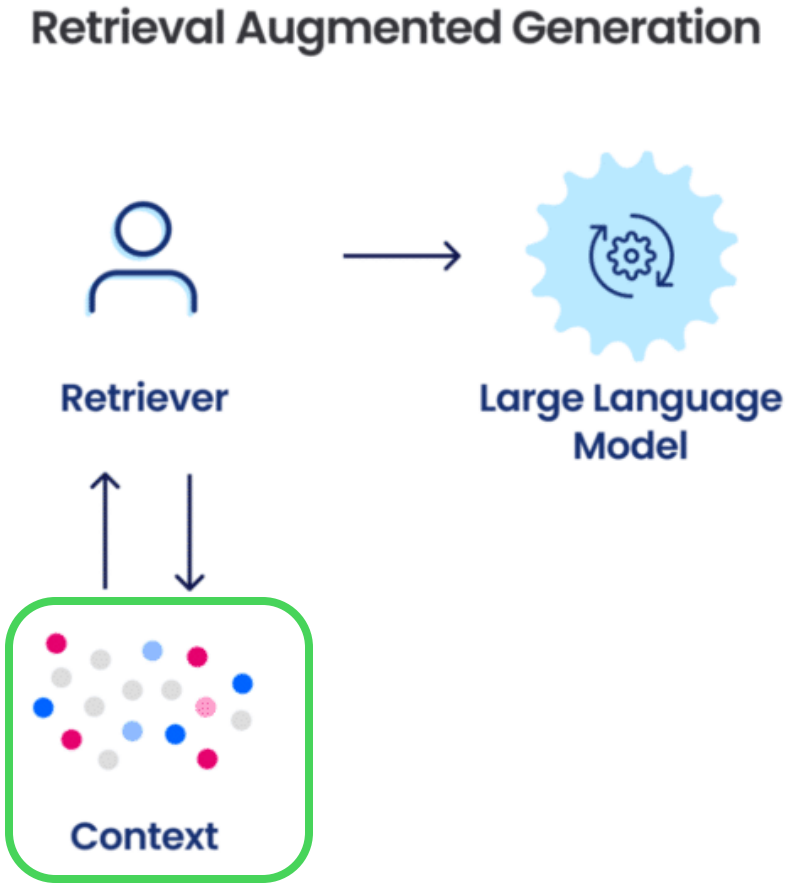
Context:



‘text’ is chunk of document data

document_name	text	embedding
files/microwave_manual.txt	isk of electric shock.↵Note:↵1. If you have any questio…	[0.007929899,-0.0103833815,-0.06096753,-0.0041557755,-0.042187788,0.01960363,-0.048827…
files/microwave_manual.txt	be adjusted or repaired by anyone except qualified serv…	[0.019625312,0.07142368,-0.0130586205,-0.018416643,-0.053729743,0.0055418145,-0.053580…
files/microwave_manual.txt	dangerous to repair or maintain the appliance by no oth…	[0.007216289,0.049263548,-0.020450495,-0.012869484,-0.06486845,0.012693635,-0.01578726…
files/microwave_manual.txt	r.↵10. Do not fry food in the oven. Hot oil can damage …	[0.022751395,0.030335193,0.016967906,-0.007043706,-0.013974892,0.002911436,-0.01603399…
files/microwave_manual.txt	over, and then press START/QUICK↵START button to resume…	[0.00082089123,0.0022738606,-0.031506665,-0.014780904,-0.033373725,-0.008920405,0.0358…
files/microwave_manual.txt	ation, the oven must have sufficient airflow. Allow minim…	[0.009636983,0.023800073,-0.0062219356,-0.027741412,-0.02087623,0.02209255,-0.01481803…
files/microwave_manual.txt	ectly connected to a low voltage power supply network w…	[-0.007313127,0.037871324,-0.03555582,-0.018665519,-0.02153417,0.045306657,-0.03712521…
files/microwave_manual.txt	r.↵↵AUTO COOK↵For the following food or cooking mode, i…	[-0.0016354206,-0.013394599,-0.042835053,-0.040621832,0.008080563,-0.018178385,0.01598…
files/microwave_manual.txt	and 15% time↵the oven will stop working in one cycle. U…	[-0.01445577,0.02532314,-0.022590334,-0.055830944,-0.026038265,0.018963622,-0.00489733…
files/microwave_manual.txt	trim↵should not be used.↵3. Do not use recycled paper p…	[0.025805177,0.016437544,0.0084770955,-0.011957668,0.008708227,0.017008575,6.410294e-0…
files/microwave_manual.txt	e dial to select others food code.↵2. Press START/QUICK…	[-0.0048348736,-0.026497386,-0.012877011,-0.07045187,-0.045177445,-0.017325213,0.01469…
files/microwave_manual.txt	and the door or allow soil or↵cleaner residue to accumu…	[0.011242517,0.041140627,-0.027245961,0.016288064,-0.012025224,0.002053797,0.009373076…
files/microwave_manual.txt	l sound every two minutes↵until user presses any button…	[-0.012621777,0.002297385,0.0051053,0.005726547,-0.026842805,0.025588008,0.043130554,0…
files/microwave_manual.txt	dd more as↵needed. Foods severely overcooked can smoke …	[0.03621854,-0.023570212,-0.010781076,0.004781386,-0.046626166,-0.0075730784,0.0137625…
files/microwave_manual.txt	)↵1. Input the first microwave cooking program. Do not p…	[-0.008847023,-0.021002412,-0.029569935,-0.030482586,-0.064433254,-0.009645594,0.01033…
files/microwave_manual.txt	d heating of warming pads, slippers, sponges, damp clot…	[-0.007398685,0.051085625,-0.0047710882,-0.016767368,-0.023013156,-0.020438727,-0.0279…
files/microwave_manual.txt	select 24- hour clock.↵2. Turn MENU/TIME dial to set ho…	[-0.0017236428,0.03179804,-0.002225171,-0.03276651,-0.064979605,-0.0032282274,0.016648…
files/microwave_manual.txt	. The microwave oven shall not be placed in a cabinet.↵…	[-0.01100224,0.0060034357,-0.01566008,0.008335399,-0.02978581,0.036069326,-0.018156435…
files/microwave_manual.txt	m(D)↵25Litres↵Approx.14kg↵↵Rated Microwave Power Output…	[0.026943313,0.017219318,-0.012886474,-0.027789962,-0.059763357,0.021788724,0.01533926…
files/microwave_manual.txt	fore using the appliance and keep for future reference.…	[-0.0037010042,0.04981585,-0.024775151,-0.04551405,-0.011770189,-0.00916123,-0.0332725…
files/microwave_manual.txt	auto cooking menus.↵START/QUICK START (the dial)↵Press…	[-0.018388081,-0.0135191595,-0.024863113,-0.06292922,-0.02891001,-0.023560518,0.031161…
files/microwave_manual.txt	re that they are suitable for use in microwave↵oven.↵16…	[0.039369497,0.04368466,-0.03406439,-0.02761703,-0.06604735,0.020116266,-0.027363198,0…
files/microwave_manual.txt	or defrosting, cooking and steaming of food only.↵35. U…	[0.013555558,0.064180404,-0.0076864623,-0.023551071,0.0023152744,-0.0037841045,-0.0294…
files/microwave_manual.txt	)↵40% (40)↵30% (30)↵20% (20)↵10% (10)↵0% (00)↵↵NOTE: Yo…	[-0.01830597,-0.012414614,0.0066955937,-0.030407824,0.0057764654,0.0033318396,0.018025…
files/microwave_manual.txt	mer or↵separate remote-control system.↵30. The microwav…	[0.015503083,0.023018092,-0.028411057,-0.011204931,-0.054389197,0.010772413,-0.0117928…

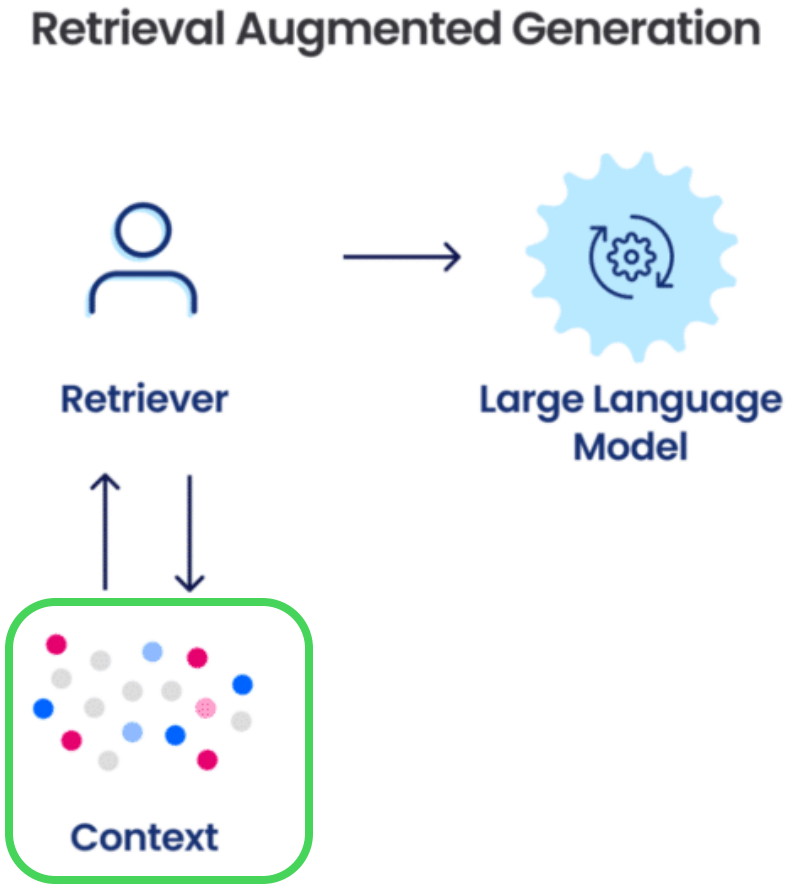
Context:



‘embedding’ is is numerical representation of text chunk

document_name	text	embedding
files/microwave_manual.txt	isk of electric shock.⏎Note:⏎1. If you have any questio…	[0.007929899,-0.0103833815,-0.06096753,-0.0041557755,-0.042187788,0.01960363,-0.048827…
files/microwave_manual.txt	be adjusted or repaired by anyone except qualified serv…	[0.019625312,0.07142368,-0.0130586205,-0.018416643,-0.053729743,0.0055418145,-0.053580…
files/microwave_manual.txt	dangerous to repair or maintain the appliance by no oth…	[0.007216289,0.049263548,-0.020450495,-0.012869484,-0.06486845,0.012693635,-0.01578726…
files/microwave_manual.txt	r.⏎10. Do not fry food in the oven. Hot oil can damage …	[0.022751395,0.030335193,0.016967906,-0.007043706,-0.013974892,0.002911436,-0.01603399…
files/microwave_manual.txt	over, and then press START/QUICK⏎START button to resume…	[0.00082089123,0.0022738606,-0.031506665,-0.014780904,-0.033373725,-0.008920405,0.0358…
files/microwave_manual.txt	ation, the oven must have sufficient airflow. Allow minim…	[0.009636983,0.023800073,-0.0062219356,-0.027741412,-0.02087623,0.02209255,-0.01481803…
files/microwave_manual.txt	ectly connected to a low voltage power supply network w…	[-0.007313127,0.037871324,-0.03555582,-0.018665519,-0.02153417,0.045306657,-0.03712521…
files/microwave_manual.txt	r.⏎⏎AUTO COOK⏎For the following food or cooking mode, i…	[-0.0016354206,-0.013394599,-0.042835053,-0.040621832,0.008080563,-0.018178385,0.01598…
files/microwave_manual.txt	and 15% time⏎the oven will stop working in one cycle. U…	[-0.01445577,0.02532314,-0.022590334,-0.055830944,-0.026038265,0.018963622,-0.00489733…
files/microwave_manual.txt	trim⏎should not be used.⏎3. Do not use recycled paper p…	[0.025805177,0.016437544,0.0084770955,-0.011957668,0.008708227,0.017008575,6.410294e-0…
files/microwave_manual.txt	e dial to select others food code.⏎2. Press START/QUICK…	[-0.0048348736,-0.026497386,-0.012877011,-0.07045187,-0.045177445,-0.017325213,0.01469…
files/microwave_manual.txt	and the door or allow soil or⏎cleaner residue to accumu…	[0.011242517,0.041140627,-0.027245961,0.016288064,-0.012025224,0.002053797,0.009373076…
files/microwave_manual.txt	l sound every two minutes⏎until user presses any button…	[-0.012621777,0.002297385,0.0051053,0.005726547,-0.026842805,0.025588008,0.043130554,0…
files/microwave_manual.txt	dd more as⏎needed. Foods severely overcooked can smoke …	[0.03621854,-0.023570212,-0.010781076,0.004781386,-0.046626166,-0.0075730784,0.0137625…
files/microwave_manual.txt	)⏎1. Input the first microwave cooking program. Do not p…	[-0.008847023,-0.021002412,-0.029569935,-0.030482586,-0.064433254,-0.009645594,0.01033…
files/microwave_manual.txt	d heating of warming pads, slippers, sponges, damp clot…	[-0.007398685,0.051085625,-0.0047710882,-0.016767368,-0.023013156,-0.020438727,-0.0279…
files/microwave_manual.txt	select 24- hour clock.⏎2. Turn MENU/TIME dial to set ho…	[-0.0017236428,0.03179804,-0.002225171,-0.03276651,-0.064979605,-0.0032282274,0.016648…
files/microwave_manual.txt	. The microwave oven shall not be placed in a cabinet.⏎…	[-0.01100224,0.0060034357,-0.01566008,0.008335399,-0.02978581,0.036069326,-0.018156435…
files/microwave_manual.txt	m(D)⏎25Litres⏎Approx.14kg⏎⏎Rated Microwave Power Output…	[0.026943313,0.017219318,-0.012886474,-0.027789962,-0.059763357,0.021788724,0.01533926…
files/microwave_manual.txt	fore using the appliance and keep for future reference.…	[-0.0037010042,0.04981585,-0.024775151,-0.04551405,-0.011770189,-0.00916123,-0.0332725…
files/microwave_manual.txt	auto cooking menus.⏎START/QUICK START (the dial)⏎Press…	[-0.018388081,-0.0135191595,-0.024863113,-0.06292922,-0.02891001,-0.023560518,0.031161…
files/microwave_manual.txt	re that they are suitable for use in microwave⏎oven.⏎16…	[0.039369497,0.04368466,-0.03406439,-0.02761703,-0.06604735,0.020116266,-0.027363198,0…
files/microwave_manual.txt	or defrosting, cooking and steaming of food only.⏎35. U…	[0.013555558,0.064180404,-0.0076864623,-0.023551071,0.0023152744,-0.0037841045,-0.0294…
files/microwave_manual.txt	)⏎40% (40)⏎30% (30)⏎20% (20)⏎10% (10)⏎0% (00)⏎⏎NOTE: Yo…	[-0.01830597,-0.012414614,0.0066955937,-0.030407824,0.0057764654,0.0033318396,0.018025…
files/microwave_manual.txt	mer or⏎separate remote-control system.⏎30. The microwav…	[0.015503083,0.023018092,-0.028411057,-0.011204931,-0.054389197,0.010772413,-0.0117928…

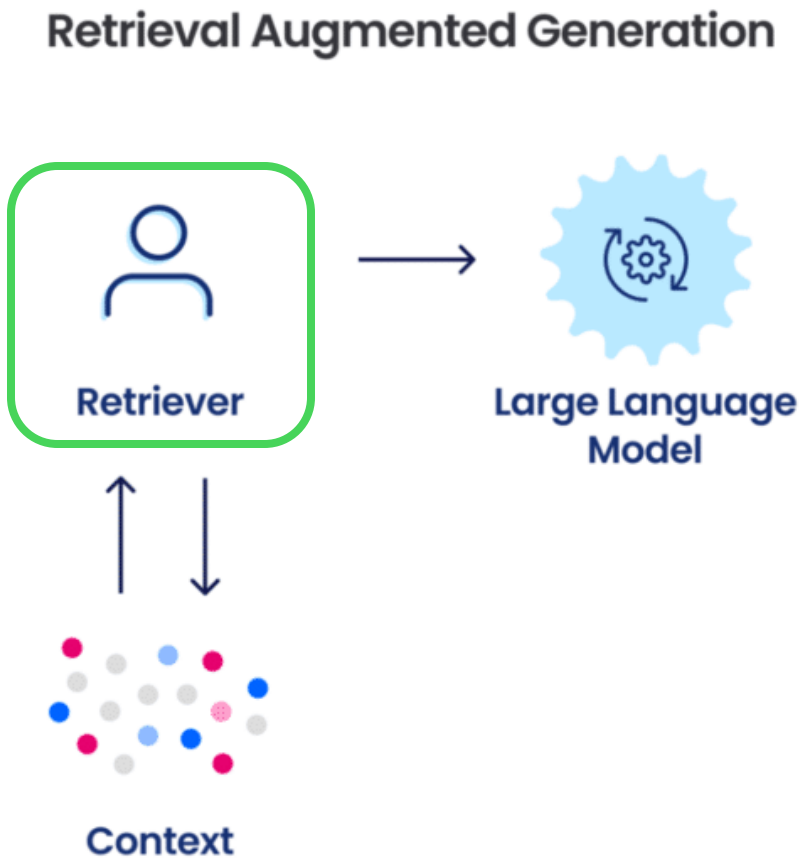
Context:



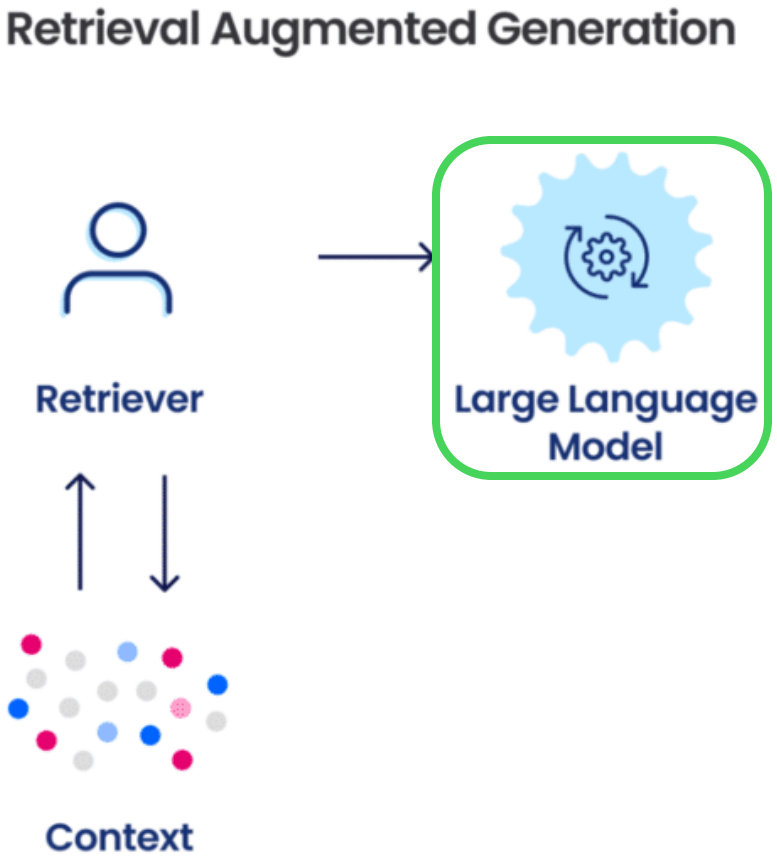
We retrieve here the chunks of information via search (similarity, semantic, etc...) by generated embedding from user request in DB

Augmentation:

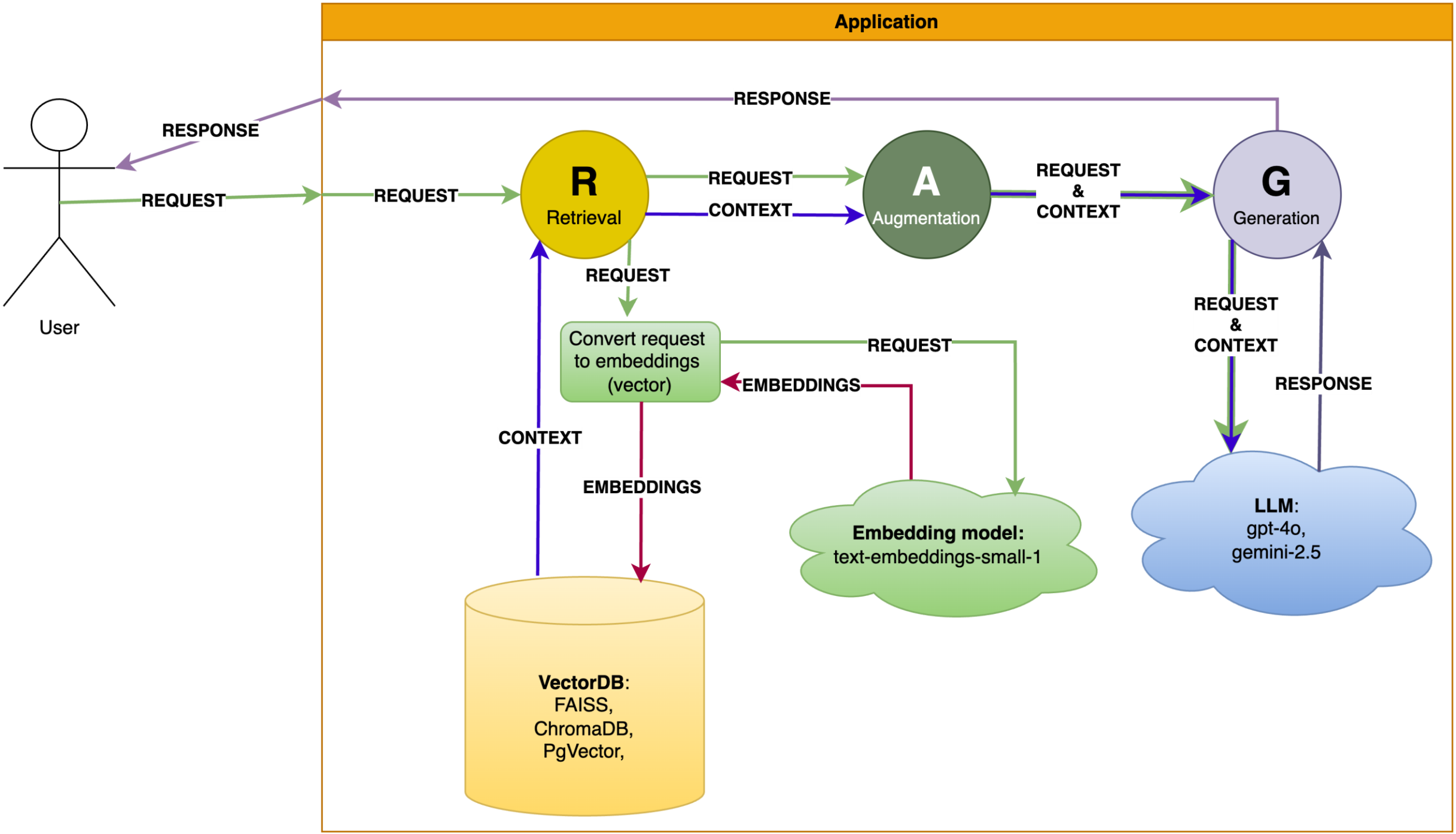
Now, we join user request and with collected data (context)



Generation:

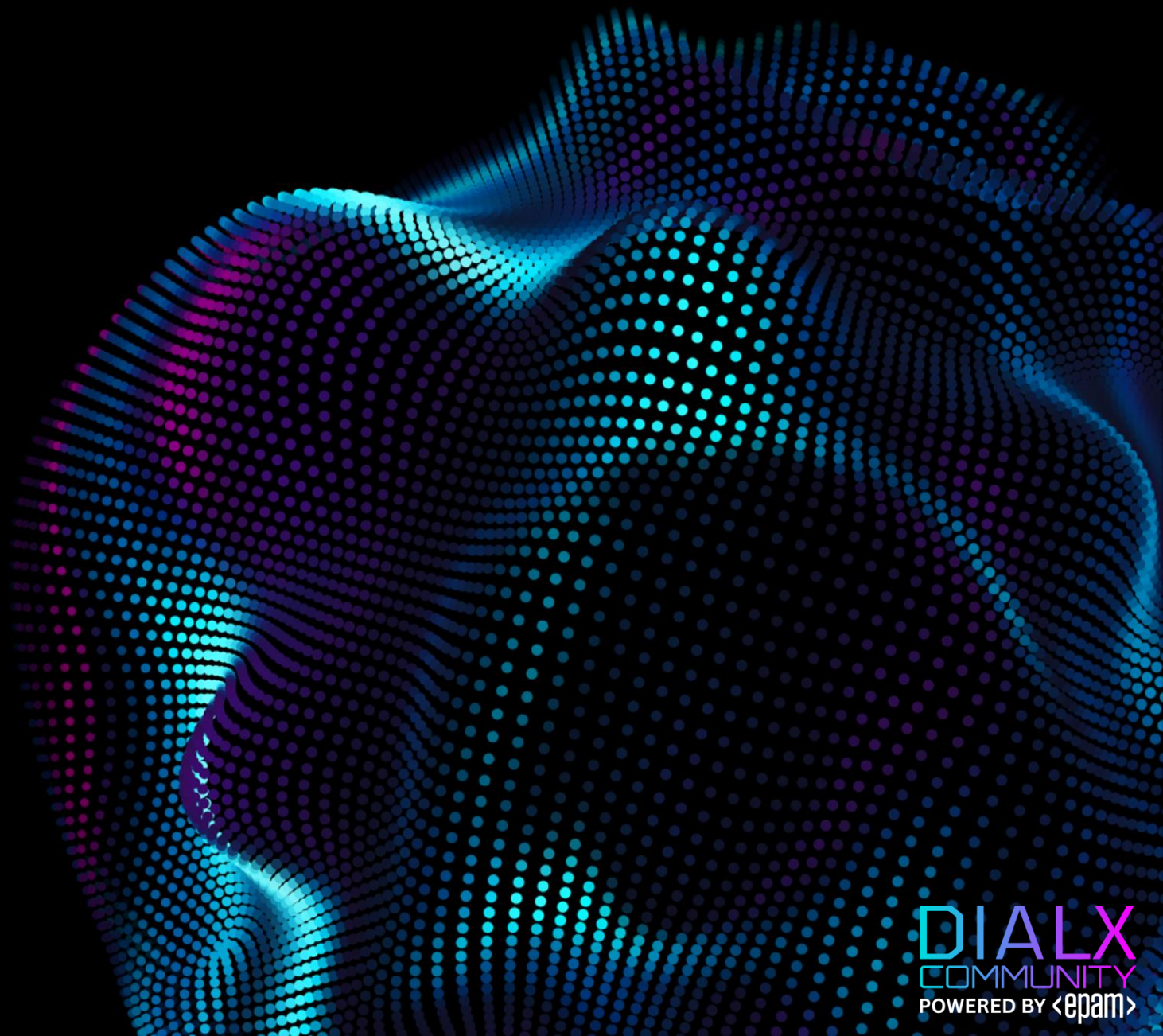


Now, with user request and some retrieved context we generate some content with LLM


Application
User
RESPONSE
RESPONSE
REQUEST
REQUEST
R
Retrieval
REQUEST
CONTEXT
A
Augmentation
REQUEST
&
CONTEXT
G
Generation
REQUEST
Convert request
to embeddings
(vector)
REQUEST
EMBEDDINGS
CONTEXT
EMBEDDINGS
Embedding model:
text-embeddings-small-1
REQUEST
&
CONTEXT
RESPONSE
LLM:
gpt-4o,
gemini-2.5
VectorDB:
FAISS,
ChromaDB,
PgVector,

Search

<Let's explore how search is working>



DIALX COMMUNITY POWERED BY <epam>

Search:

Search can be performed in different ways

The most common are Similarity and Semantic searches

We will practice with Similarity search

Search:

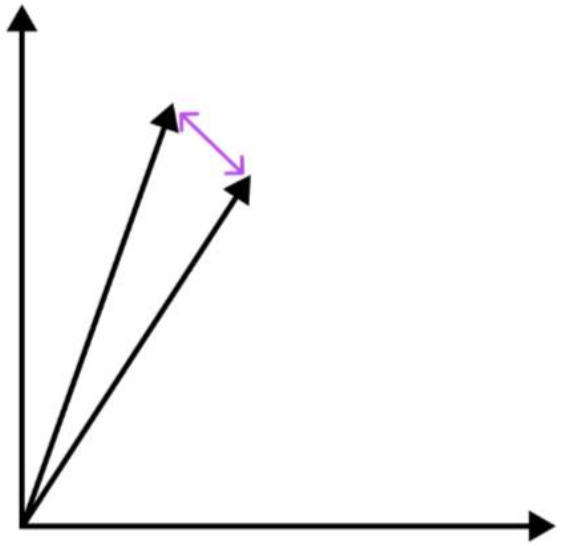
Measures the straight-line (geometric) distance between two points in vector space.

The smaller the distance, the more similar the vectors

In Postgres use <-> operator for similarity search with Euclidean distance

Euclidean Distance

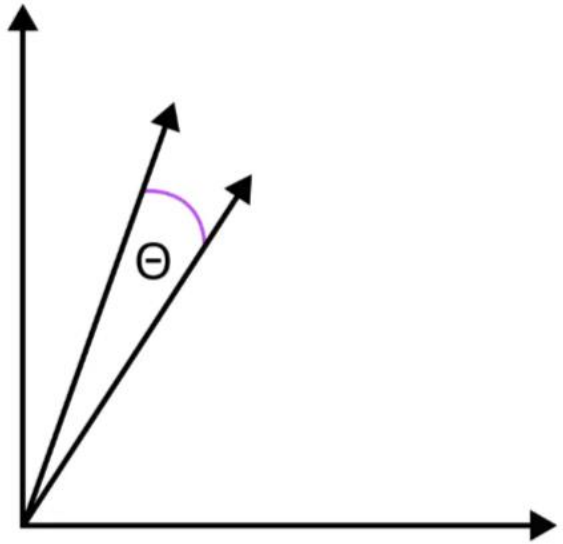
$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}$$



Search:

Cosine Similarity

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}$$



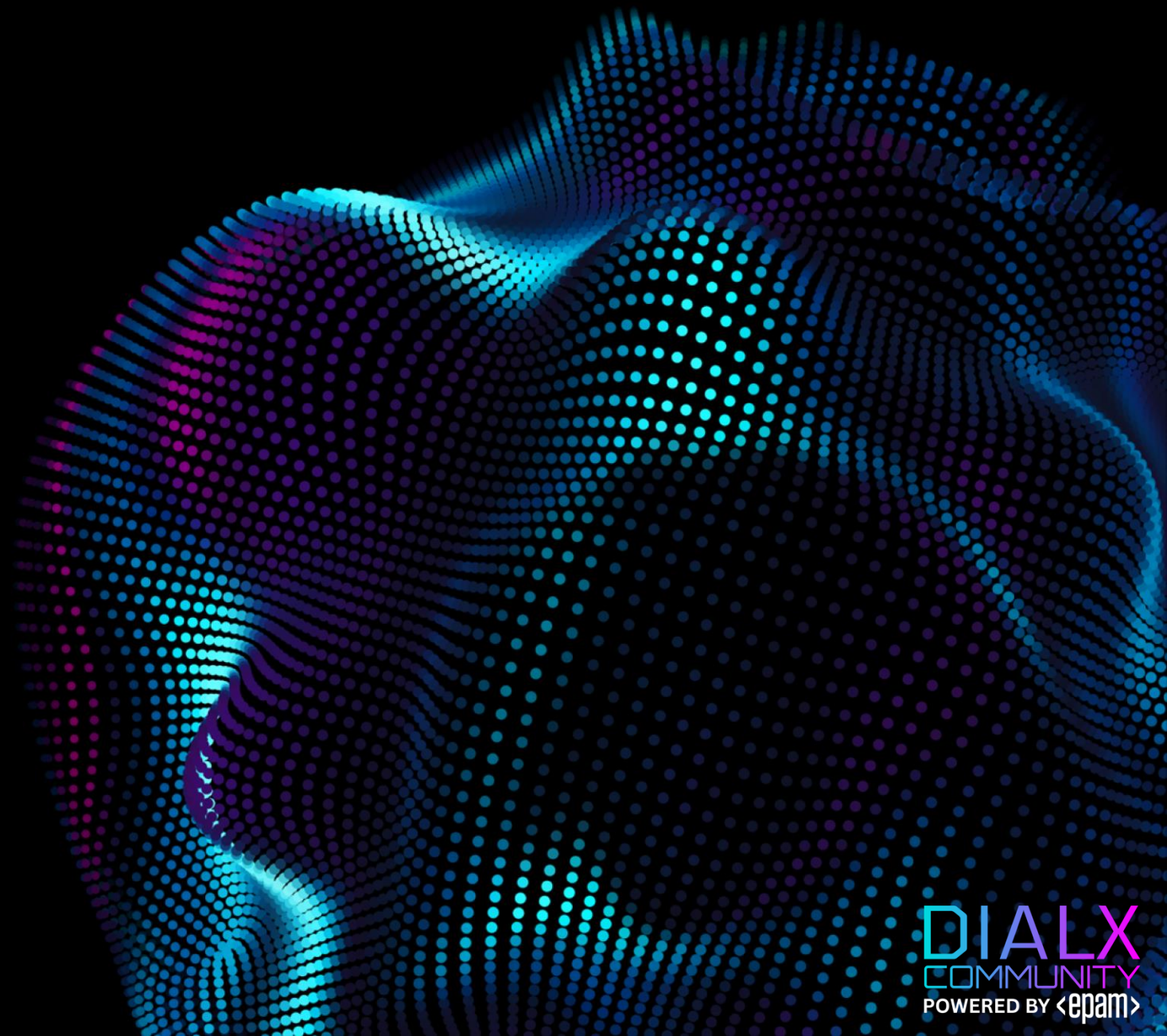
Measures the cosine of the angle between two vectors

It ranges from -1 (opposite), 0 (no similarity), to 1 (similar)

In Postgres use <=> operator for similarity search with Cosine similarity

Search with top_k and score

<Let's explore how search with the top_k and score filters>



DIALX COMMUNITY POWERED BY <epam>

User request

Wash all the red clothes

Retrieval

Got it. Now I'll retrieve all the red clothes from this bucket

Here are all the red clothes that I was able to find

Augmentation

'Wash all the red clothes' + the clothes

Generation

I've made the laundry with all the provided red stuff. There were a couple of rags and a pair of shoes. Any other instructions?

Wash all the red clothes

User request

Retrieval

Augmentation

Generation

Wash all the red clothes

Got it. Now I'll retrieve all the red clothes from this bucket.
top_k = 1

Here are all the red clothes that I was able to find

'Wash all the red clothes' + the clothes

I've made the laundry with all the provided red stuff. There were a couple of rags and a pair of shoes. Any other instructions?

Wash all the red clothes

User request

Retrieval

Augmentation

Generation

Wash all the red clothes

Got it. Now I'll retrieve all the red clothes from this bucket.
top_k = 10
score = 0.1

Here are all the red clothes that I was able to find

'Wash all the red clothes' + the clothes

I've made the laundry with all the provided red stuff. There were a couple of rags and a pair of shoes. Any other instructions?

Wash all the red clothes

User request

Retrieval

Augmentation

Generation

Wash all the red clothes

Got it. Now I'll retrieve all the red clothes from this bucket.
top_k = 10
score = 0.6

Here are all the red clothes that I was able to find

'Wash all the red clothes' + the clothes

I've made the laundry with all the provided red stuff. There were a couple of rags and a pair of shoes. Any other instructions?

Wash all the red clothes

User request

Retrieval

Augmentation

Generation

Wash all the red clothes

Got it. Now I'll retrieve all the red clothes from this bucket.
top_k = 10
score = 0.9

Here are all the red clothes that I was able to find

'Wash all the red clothes' + the clothes

I've made the laundry with all the provided red stuff. There were a couple of rags and a pair of shoes. Any other instructions?

Wash all the red clothes

DIALX COMMUNITY
POWERED BY epam

Join us:

Subscribe to WeAreCommuntiy

https://wearecommunity.io/communities/dial

Keep in touch with our latest updates. Here you find webinars, workshops and articles about DIALX features and products.

Subscribe to YouTube

https://www.youtube.com/@TeamDIALX

Here we publish videos about our newest products and features.

Join our Discord community

https://discord.gg/jvTCQv4E4q

✨AI DIALX Community✨ is the place where you can find help with your questions about DIALX, direct communication with DIALX team and contributors.

DIALX
COMMUNITY
POWERED BY epam
Thank you!