PGConf NYC

ADVANCING DRUG SEARCH WITH POSTGRESQL AND AZURE AI

Microsoft

Taras Kloba
Senior Solutions Architect, Data & AI at Microsoft

- 15+ years in data engineering
- 9 years working deeply with PostgreSQL
- Co-Leader of PostgreSQL Ukraine Community
- Winner of Ukrainian IT Awards 2019 – Best Software Architect in Ukraine
- Two-time Winner of TIDE NATO Hackathon
- Ex-Microsoft MVP specialized in Azure Databases for PostgreSQL
- Microsoft Certified Trainer (MCT)
- Certified Cloud Architect & Data Engineer on Microsoft Azure, Google Cloud Platform, and Amazon Web Services
- Ph.D.
- Father of three daughters

the challenge.

understanding our challenge

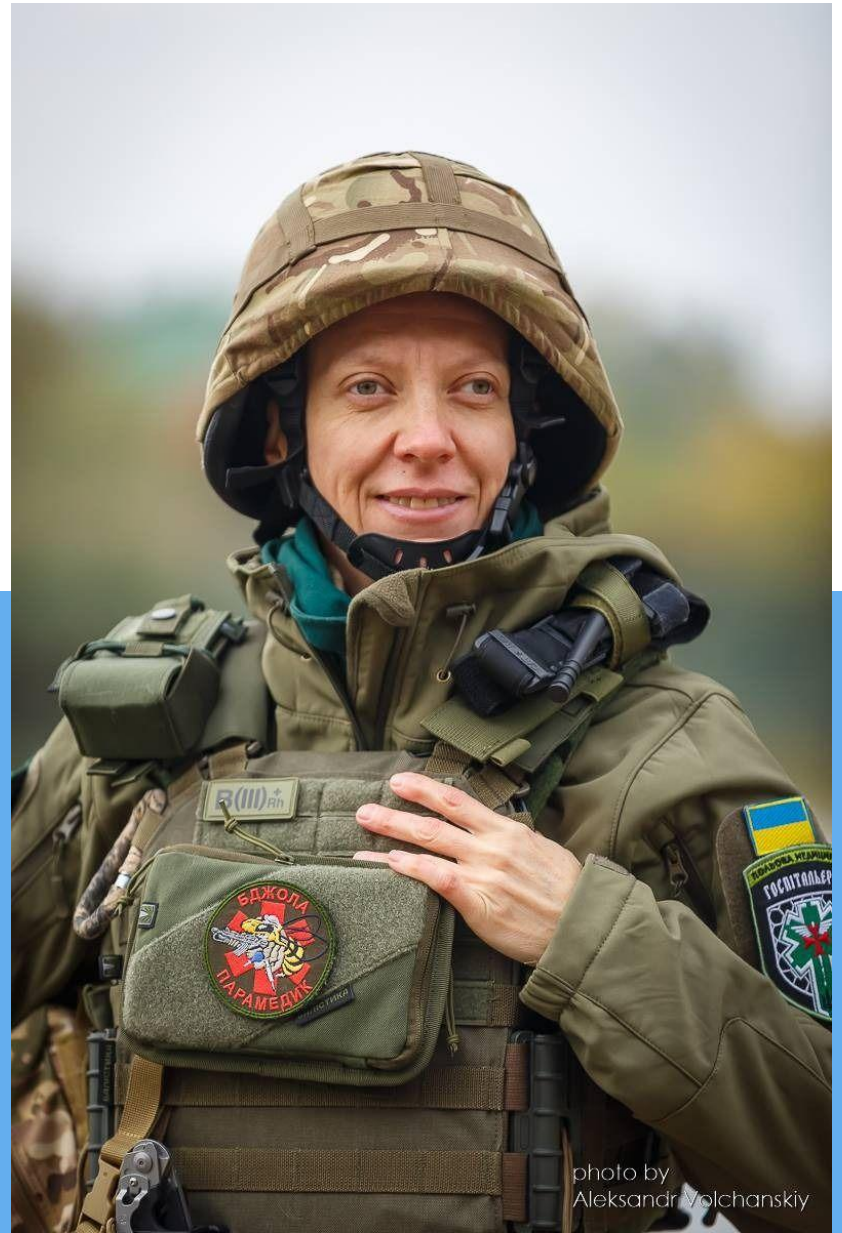
Medication Match-Up.

When a new war starts, NATO troops from different countries unite, each with their medical supplies. Our challenge was to make an app that helps **find the same medicine across different country brands**.

The Suwalki Gap: A Hypothetical Flashpoint

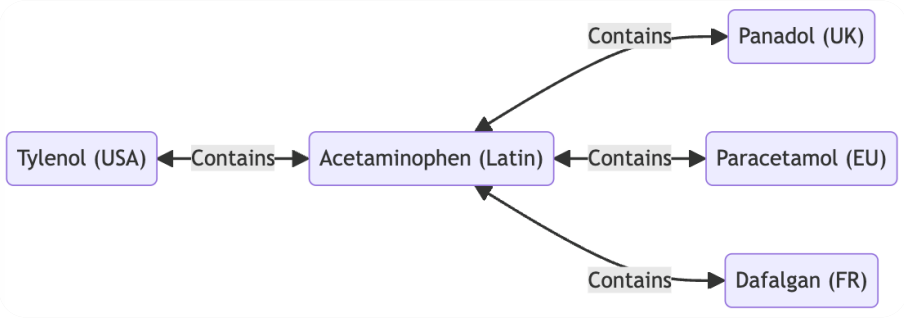
the solution.

innovation born of necessity

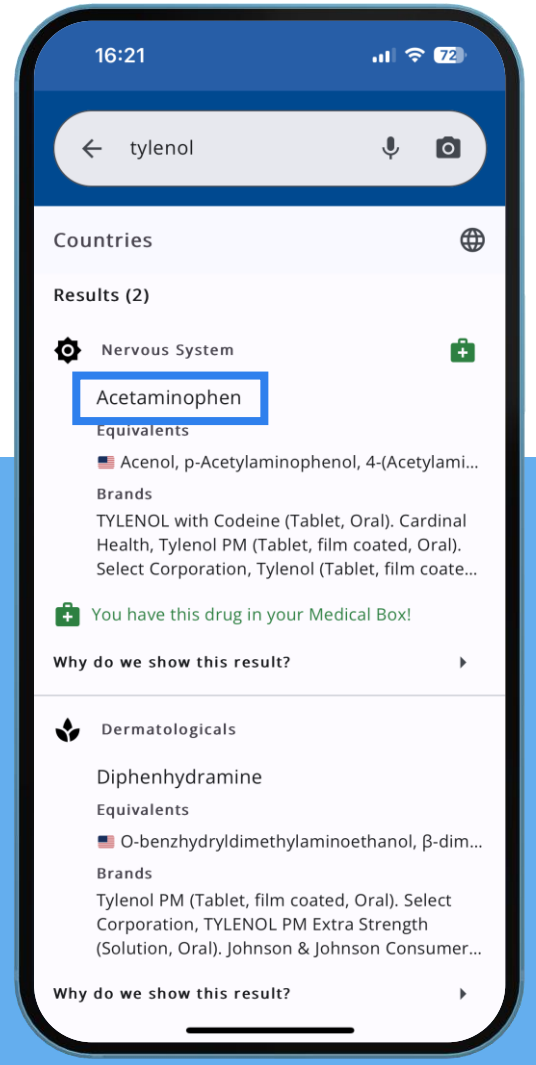


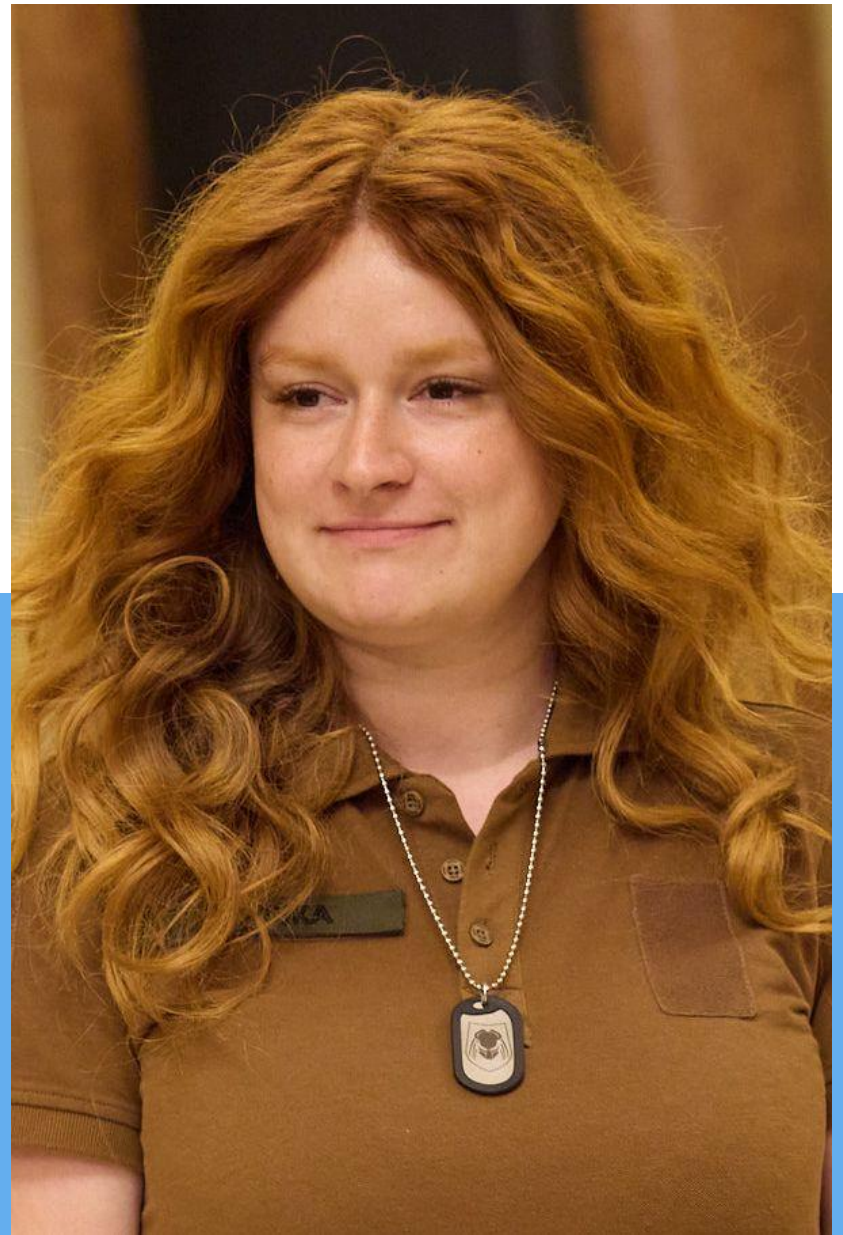
known as the "Bee Paramedic"

Inna Olkhova.



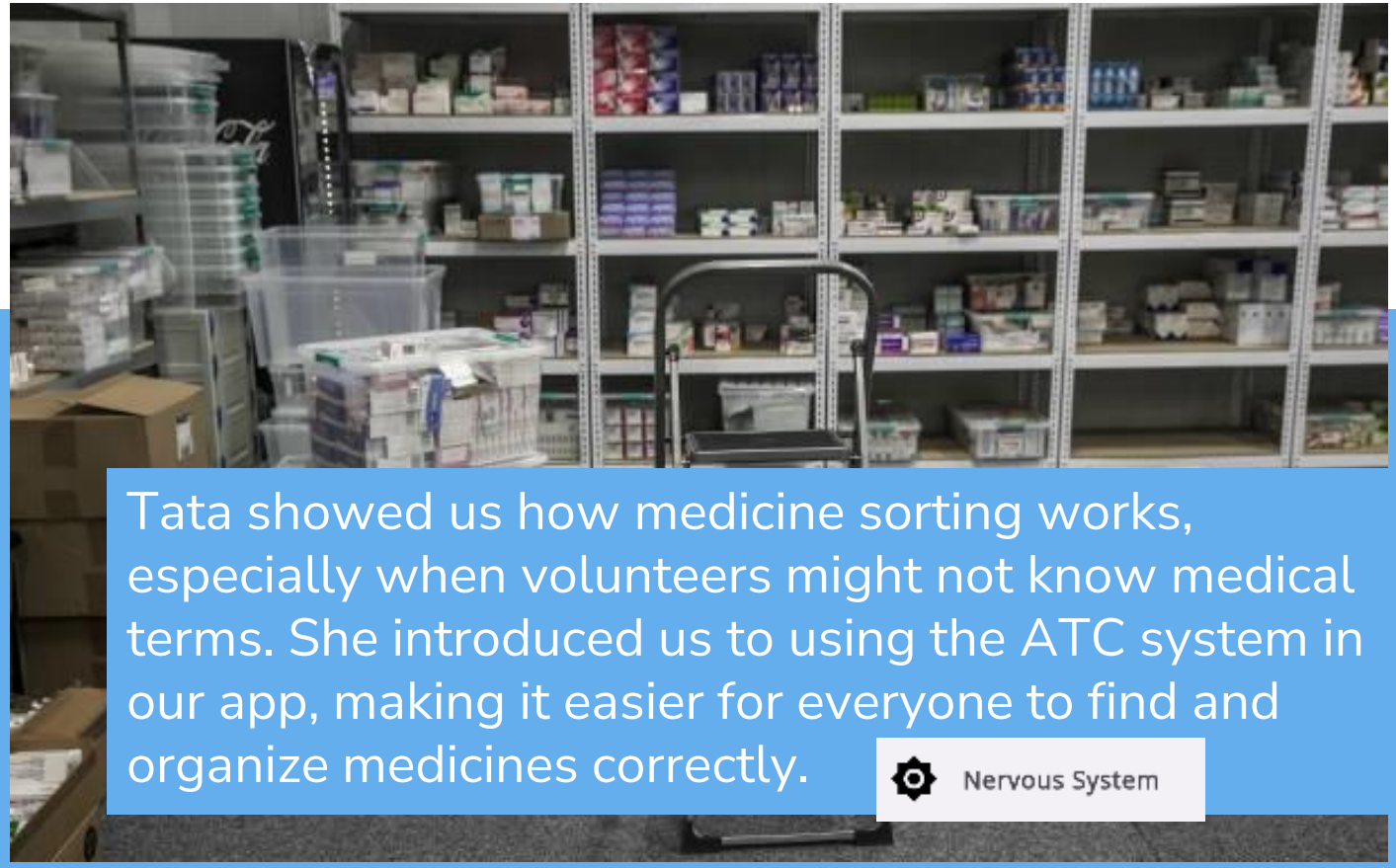
The diagram shows our app's ability to match different medicine brands through their active ingredient, Acetaminophen. Inna's idea helps find the same medicine, whether it's Tylenol in the US, Panadol in the UK, Paracetamol in the EU, or Dafalgan in France.





From Artistry to Aid: Shaping Accessible Care in the Field

Tata Kepler.

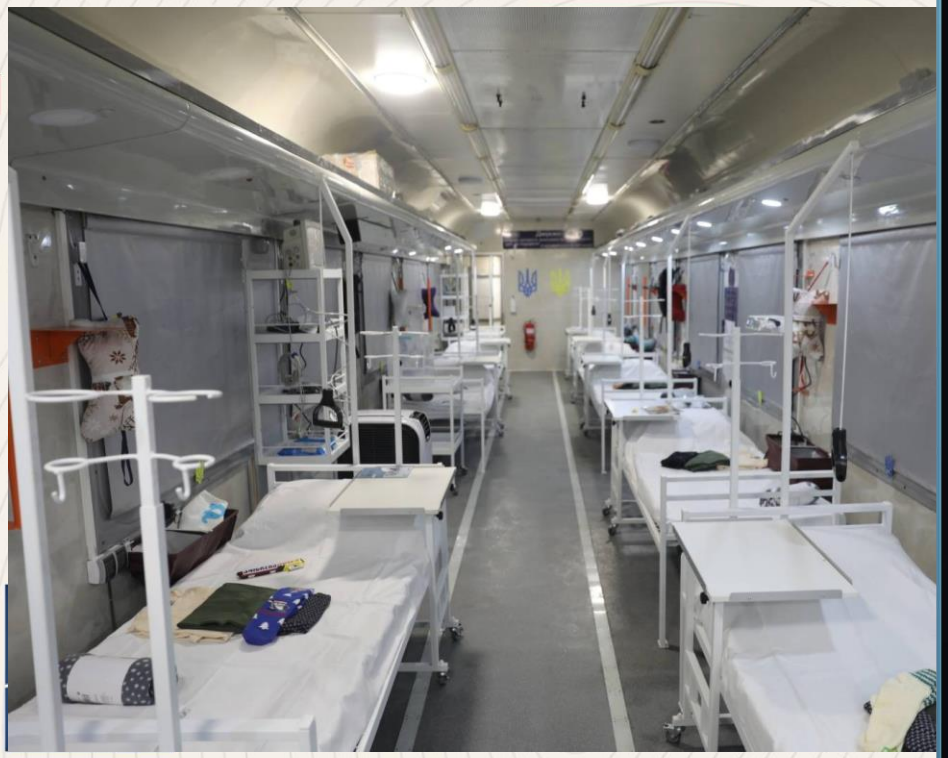


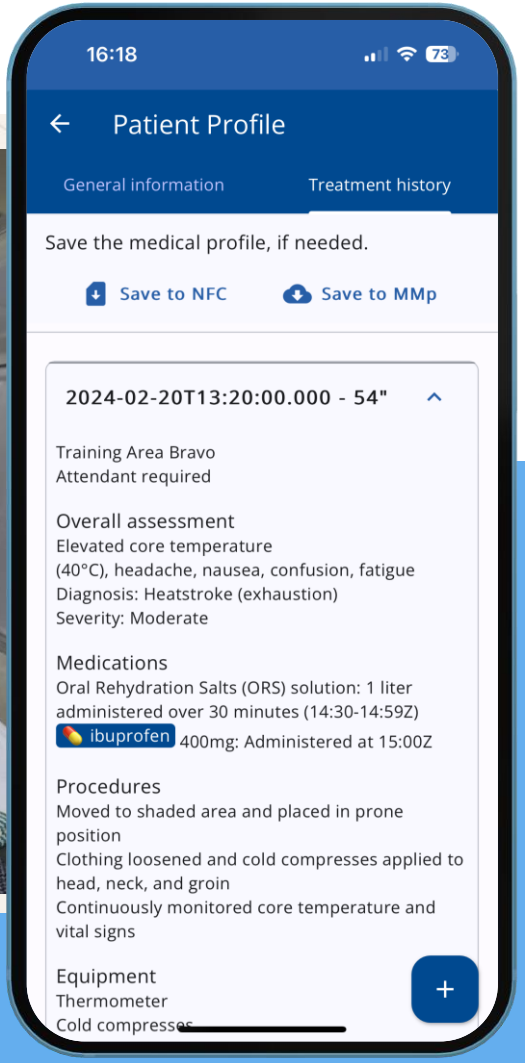
Tata showed us how medicine sorting works, especially when volunteers might not know medical terms. She introduced us to using the ATC system in our app, making it easier for everyone to find and organize medicines correctly.

Nervous System

Train Evacuations in Ukraine: Pioneering Lifesaving Journeys

Iryna Soloshenko.

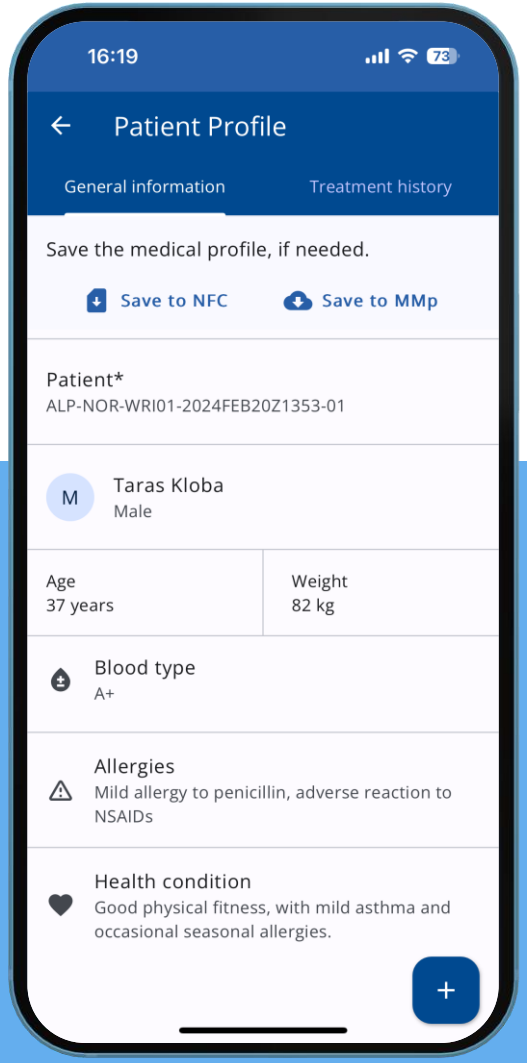




Immediate Action: Key to Battlefield Survival

Daryna Smolnikova.

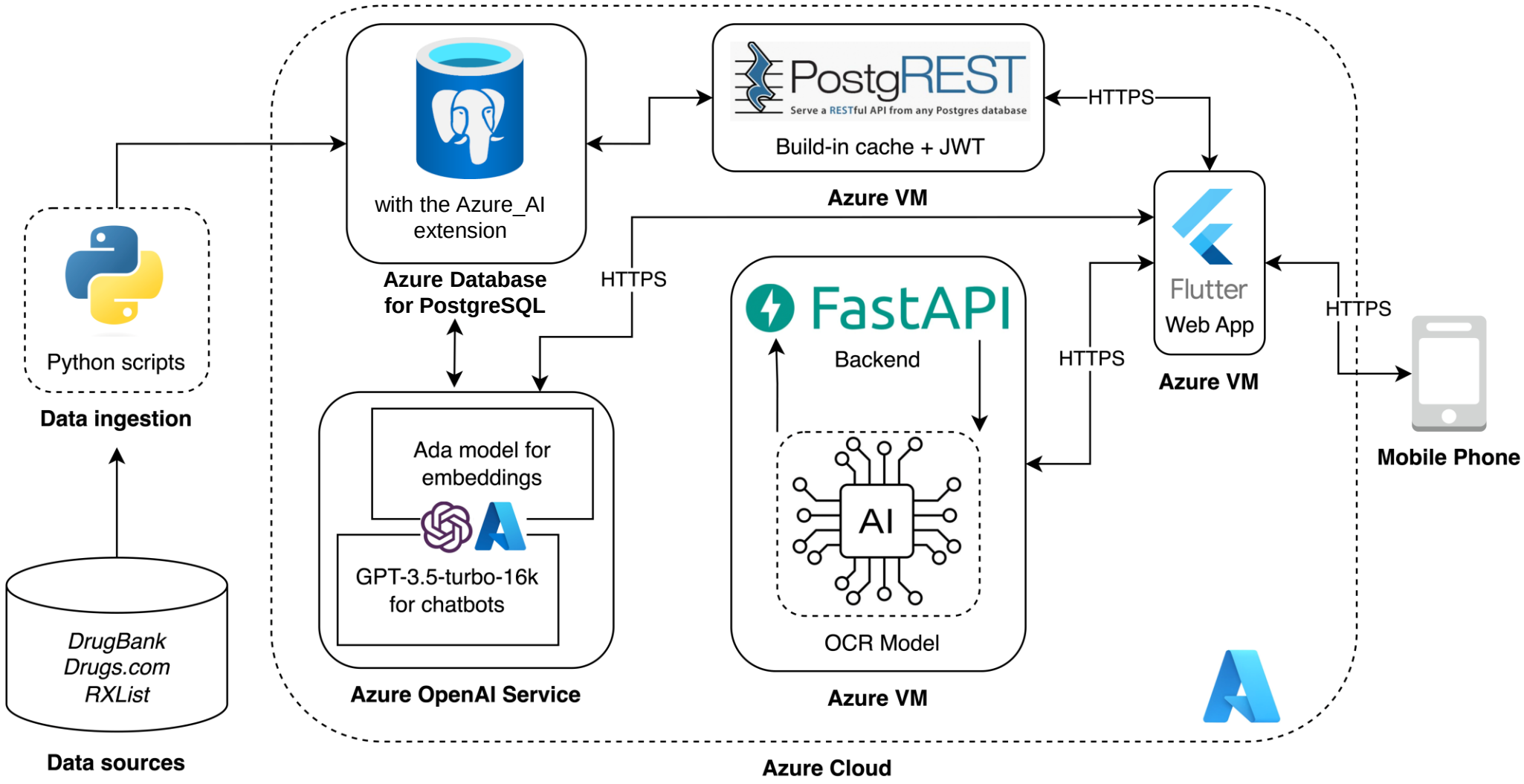
Daryna shared the idea of using NFC, QR codes, and short links to store and share information about the wounded. She emphasized that sometimes, sharing this information, even with the enemy, is crucial to saving lives.



tech specs.

engineering precision for lifesaving solutions

Architecture.



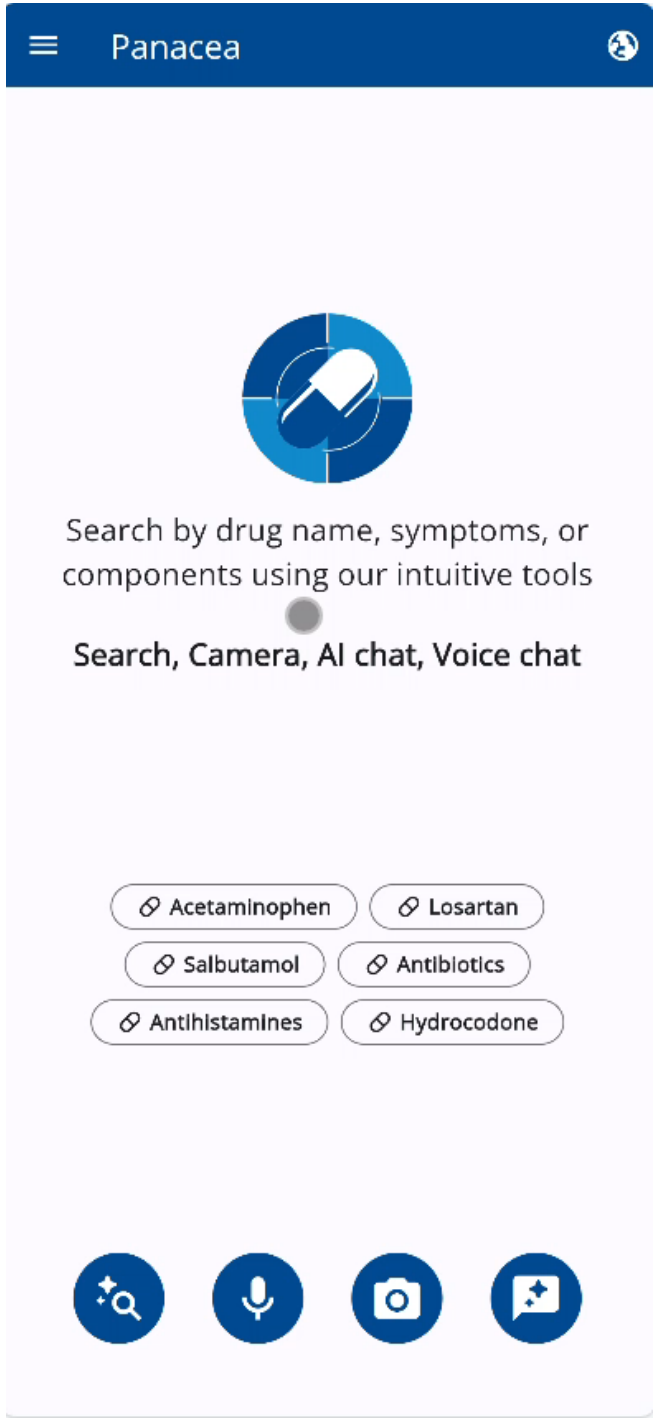
PANACEA

ILIKE.

The **ILIKE** operator in PostgreSQL is used for ***case-insensitive*** pattern matching. It functions similarly to the **LIKE** operator but ignores the case of the character

```
SELECT *
FROM drugs
WHERE drug_name ILIKE 'tylen%';
```

```
SELECT *
FROM drugs
WHERE drug_name ILIKE '%tylen%';
```

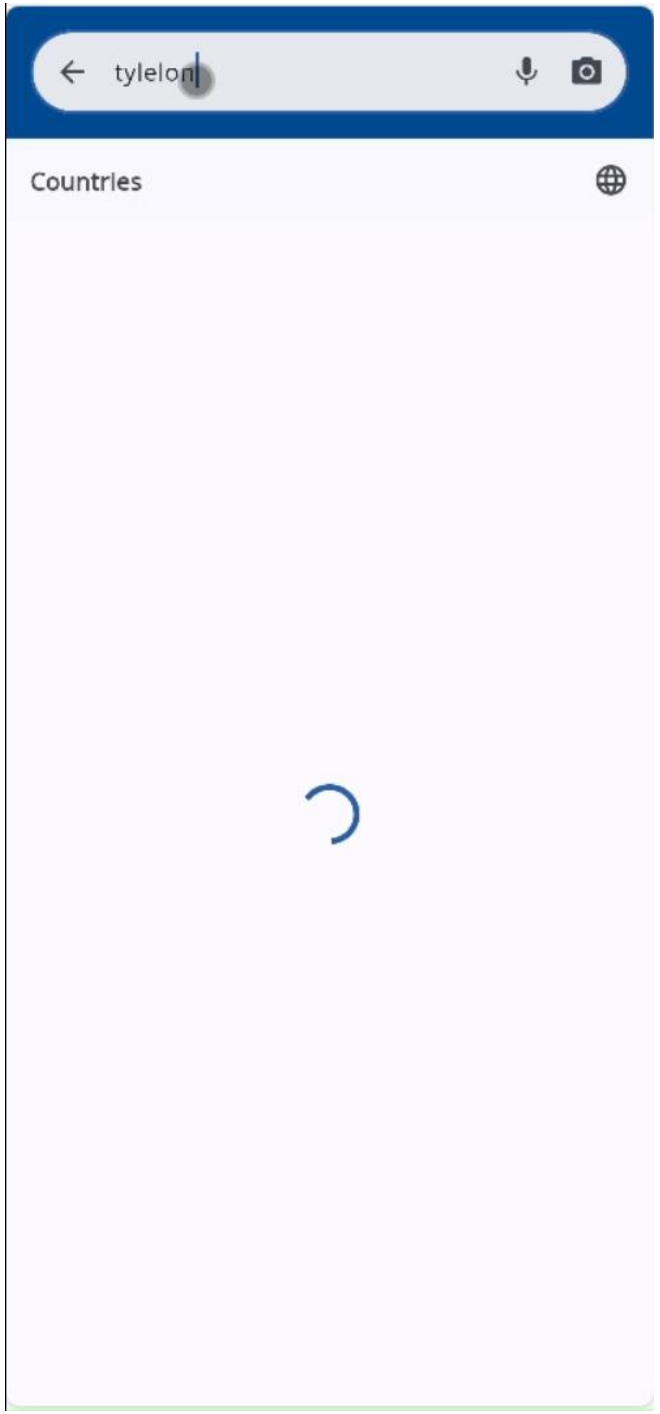


PANACEA

TRIGRAMS.

A trigram is a group of **three consecutive characters** in a string. By comparing these trigrams, PostgreSQL can measure the similarity between strings. Useful for finding similar strings despite minor spelling errors and variations.

```
CREATE EXTENSION pg_trgm;

SELECT drug_name
FROM drugs
WHERE drug_name % 'tylelon'
ORDER BY similarity(drug_name, 'tylelon') DESC;
```

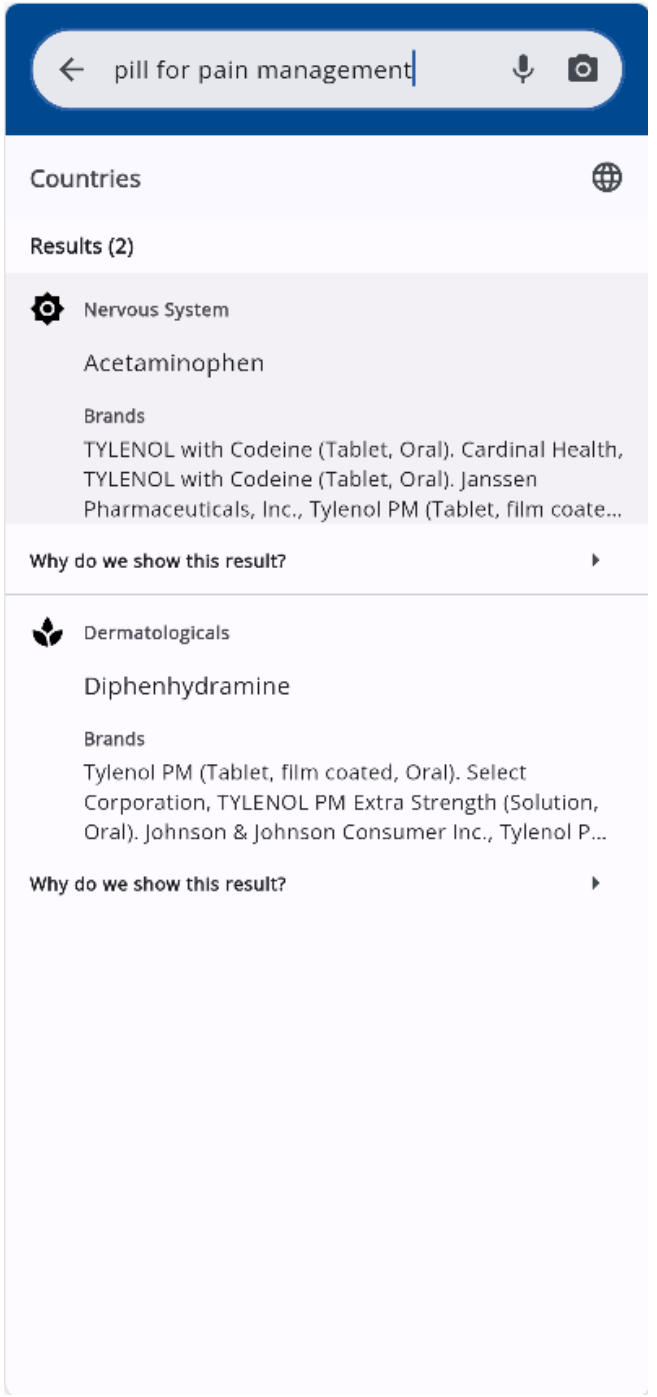


PANACEA

VECTOR SEARCH.

Vector search uses vector representations of text to find similar items. By comparing these vectors, PostgreSQL can measure the similarity between text descriptions.

```
CREATE EXTENSION azure_ai;

ALTER TABLE drugs
ADD COLUMN vector VECTOR GENERATED ALWAYS AS
(azure_ai.vectorize(drug_name)) STORED;

SELECT drug_name
FROM drugs
ORDER BY vector <=> azure_ai.vectorize('pill for pain
management')
LIMIT 5;
```

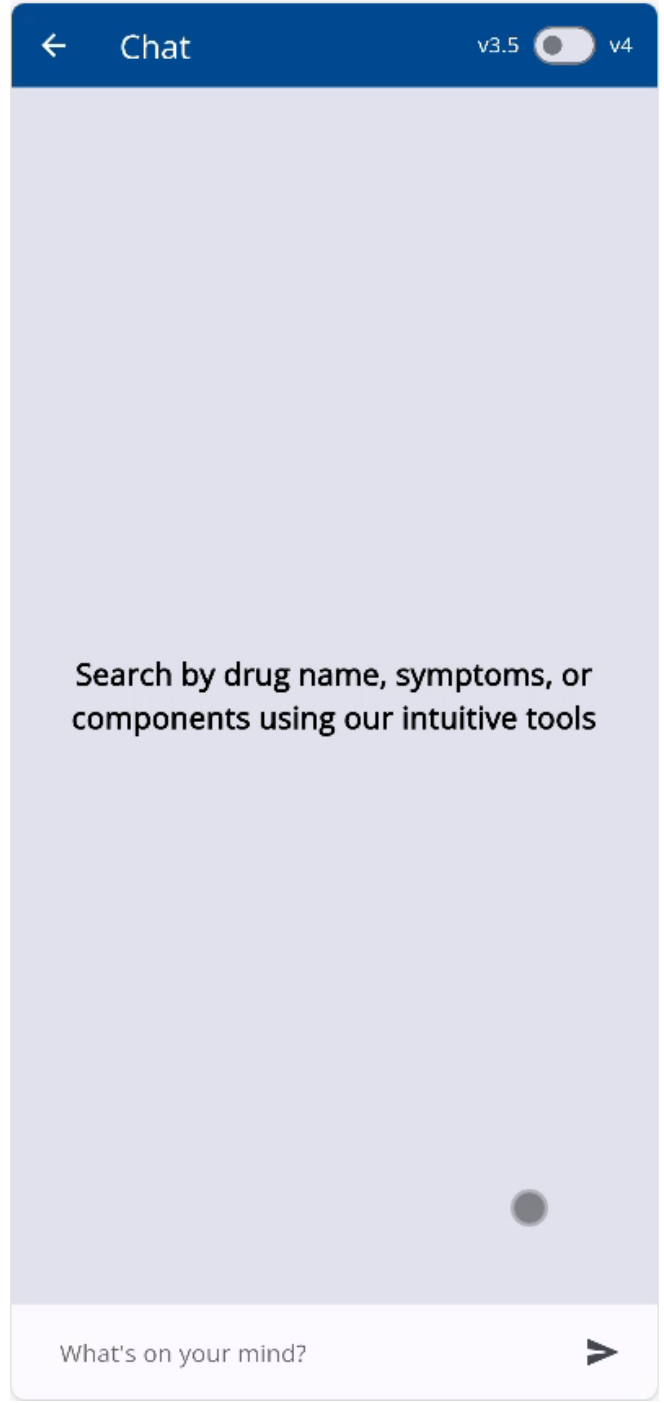


PANACEA

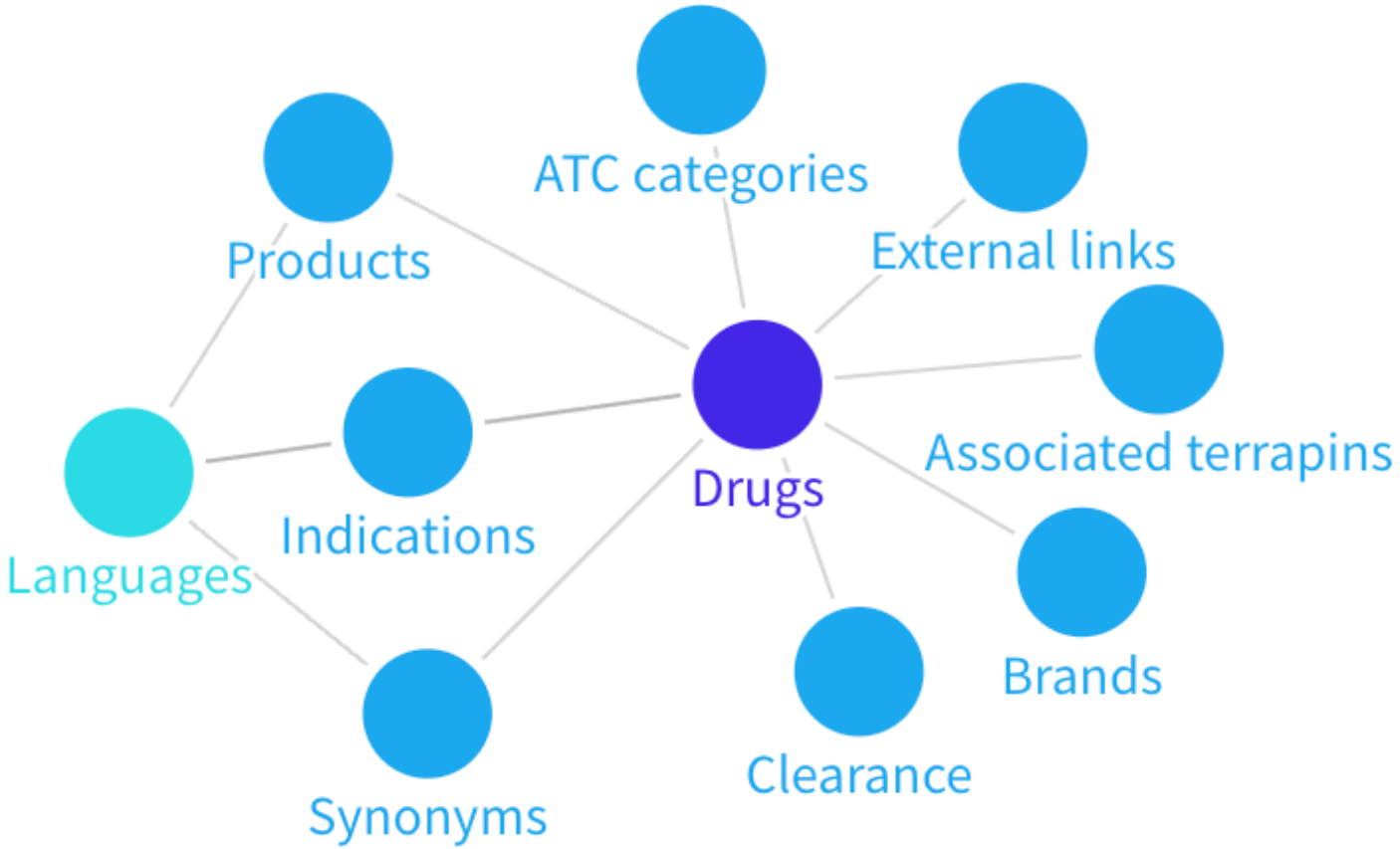
VECTOR SEARCH.

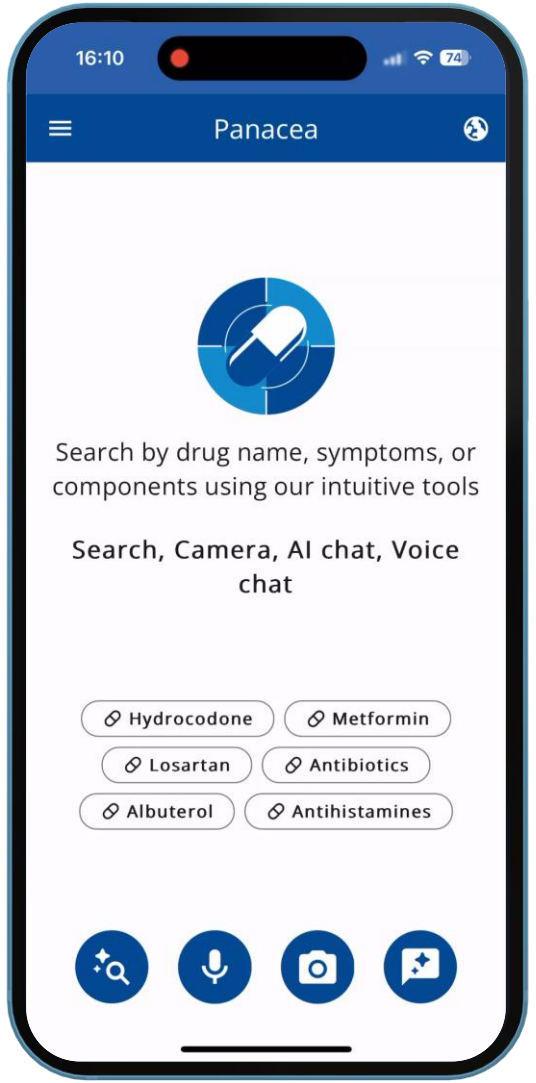
Vector search uses vector representations of text to find similar items. By comparing these vectors, PostgreSQL can measure the similarity between text descriptions.

```
CREATE EXTENSION azure_ai;

ALTER TABLE drugs
ADD COLUMN vector VECTOR GENERATED ALWAYS AS
(azure_ai.vectorize(drug_name)) STORED;

SELECT drug_name
FROM drugs
ORDER BY vector <=> azure_ai.vectorize('red pill for
headaches')
LIMIT 5;
```



Knowledge Graph.



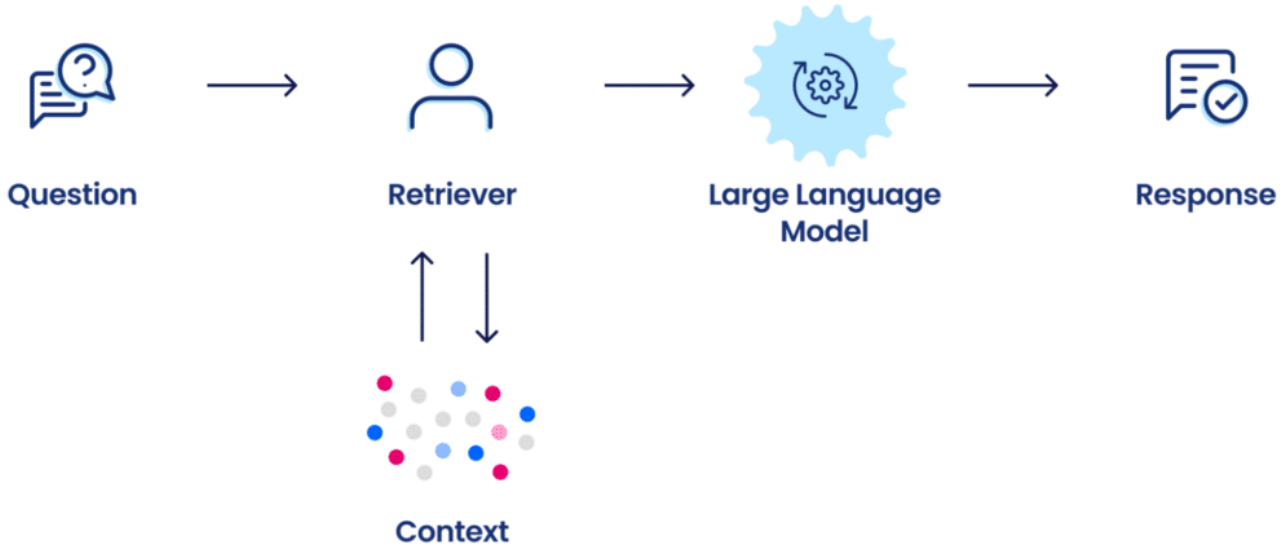


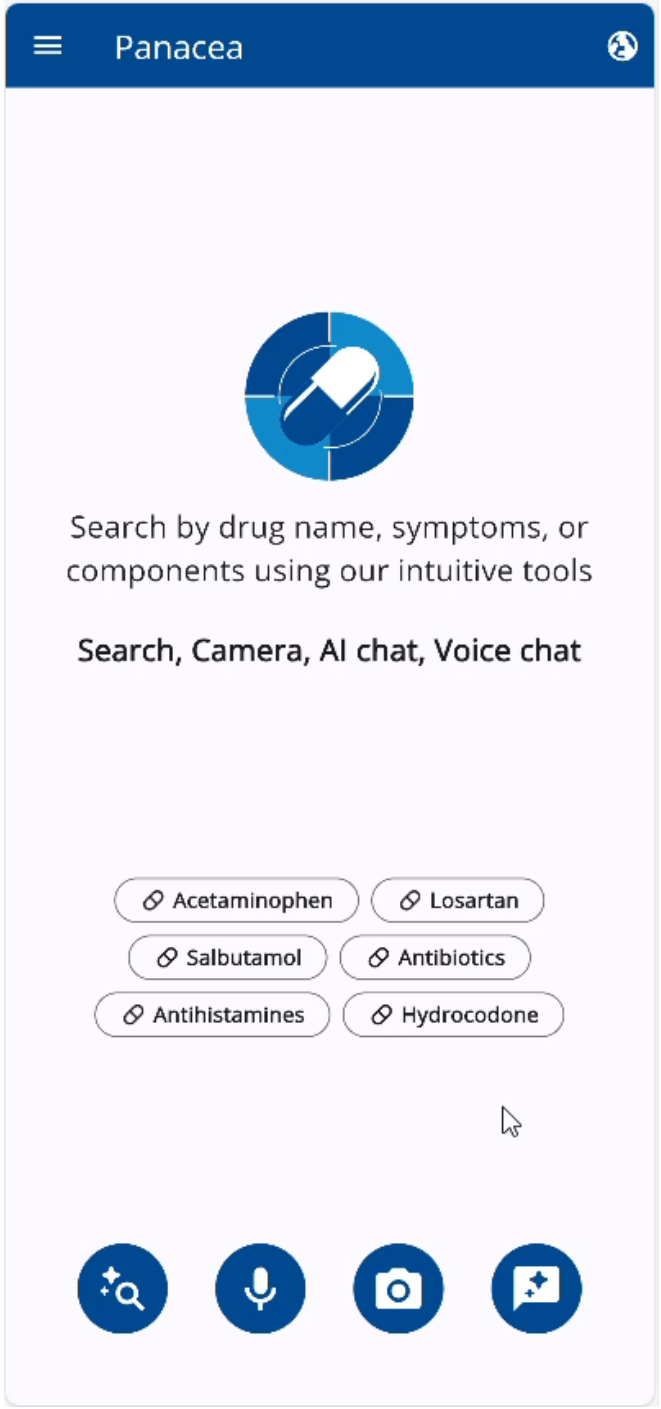
Powered by DrugBank.com database

PANACEA

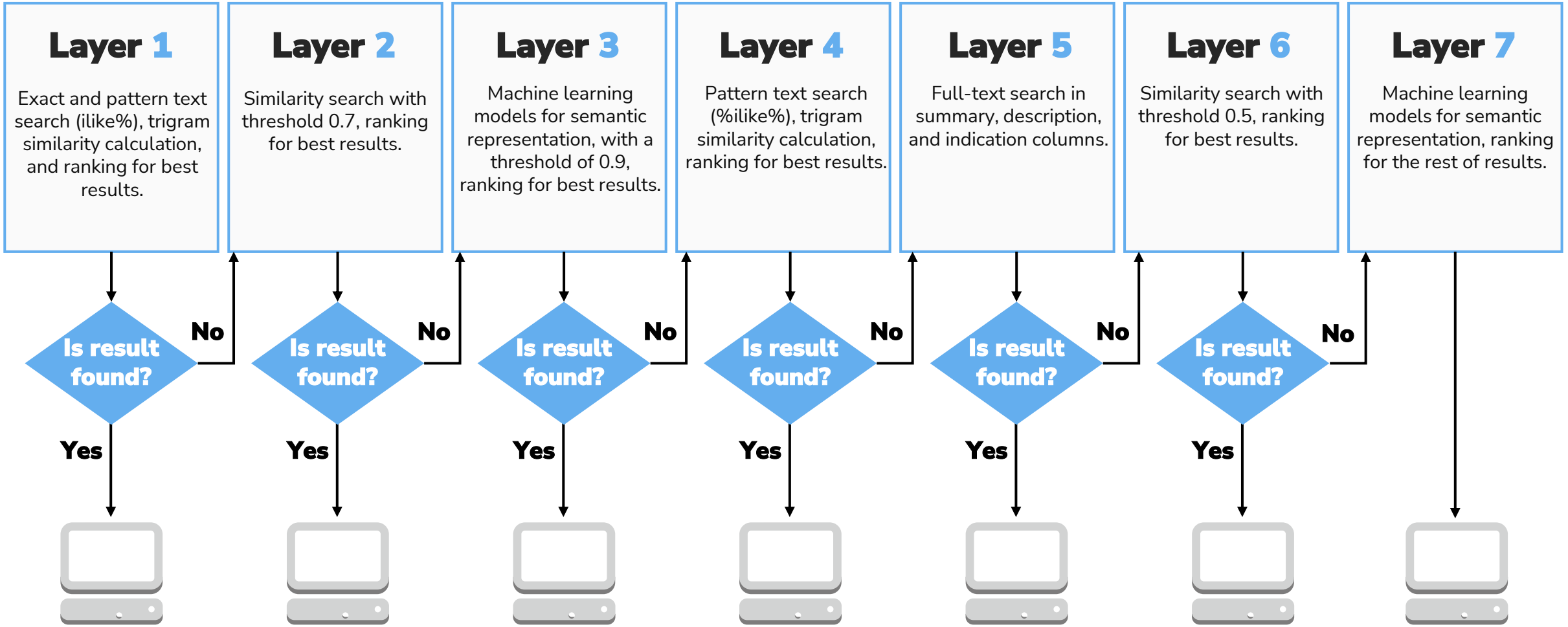
DRUGGPT (AI RAG ASSSISTANT).

RAG (Retrieval-Augmented Generation) is a hybrid approach that combines retrieval and generation techniques to enhance the quality of AI responses. It first retrieves relevant information from a knowledge base or database and then uses this information to generate more accurate and contextually relevant answers.

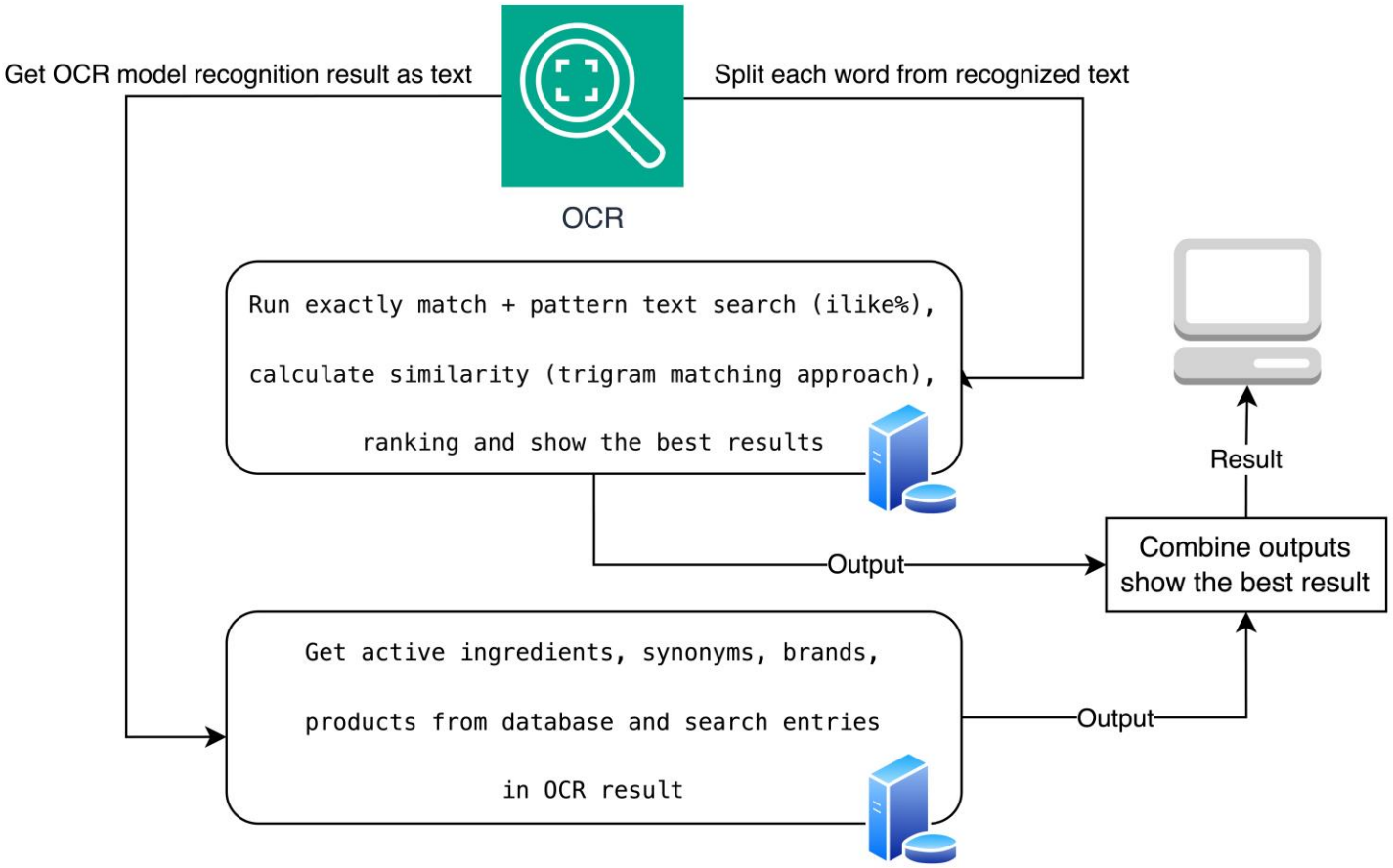




Search.


Layer 1
Exact and pattern text search (ilike%), trigram similarity calculation, and ranking for best results.
Layer 2
Similarity search with threshold 0.7, ranking for best results.
Layer 3
Machine learning models for semantic representation, with a threshold of 0.9, ranking for best results.
Layer 4
Pattern text search (%ilike%), trigram similarity calculation, ranking for best results.
Layer 5
Full-text search in summary, description, and indication columns.
Layer 6
Similarity search with threshold 0.5, ranking for best results.
Layer 7
Machine learning models for semantic representation, ranking for the rest of results.
Is result found?
No
Yes
Is result found?
No
Yes
Is result found?
No
Yes
Is result found?
No
Yes
Is result found?
No
Yes
Is result found?
No
Yes

OCR Search.



[Tylenol], [Acetaminophen], [Pain], [Reliever], [Fever], [Reducer], [Extra], [Strength], [FOR], [ADULTS], [500], [mg], [each]

Database

demo.

try it out.

get in touch.

Taras Kloba

aka.ms/taras

Microsoft

Got 3 minutes? We'd love your input on some of our Postgres work

Get your FREE socks
@ Microsoft booth

Microsoft

PGConf NYC

Meet our Postgres team at PGConf NYC 2024

Follow Azure Database for PostgreSQL on LinkedIn

https://aka.ms/azure-pg-linkedin

Have you listened to?

TalkingPostgres.com

Save the date
June 10-12, 2025

POSETTE: An Event for Postgres

2025 Now in its 4th year!

A free & virtual developer event

Subscribe to news → aka.ms/posette-subscribe