



# Searching for similar music tracks



David Pilato | [@dadoonet](#)





# Elasticsearch

You Know, for **Vector** Search

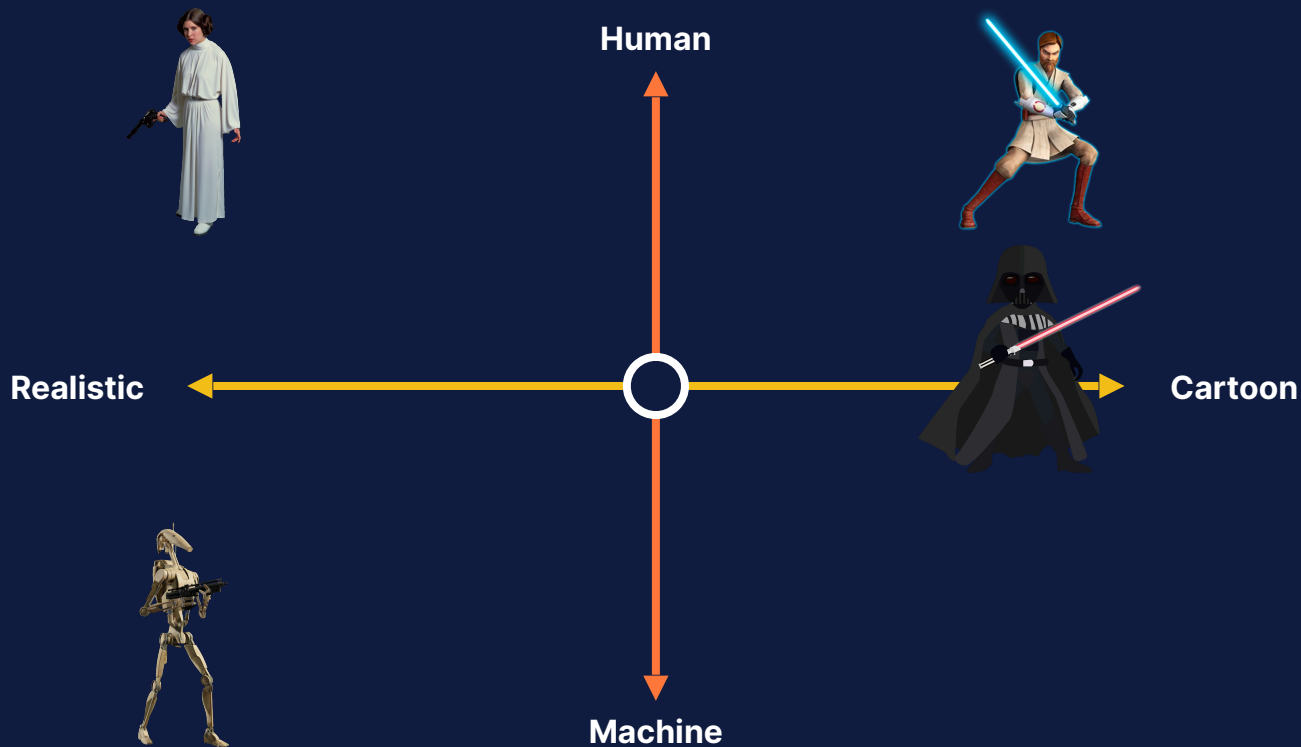
# Embeddings represent your data



## Example: 1-dimensional vector



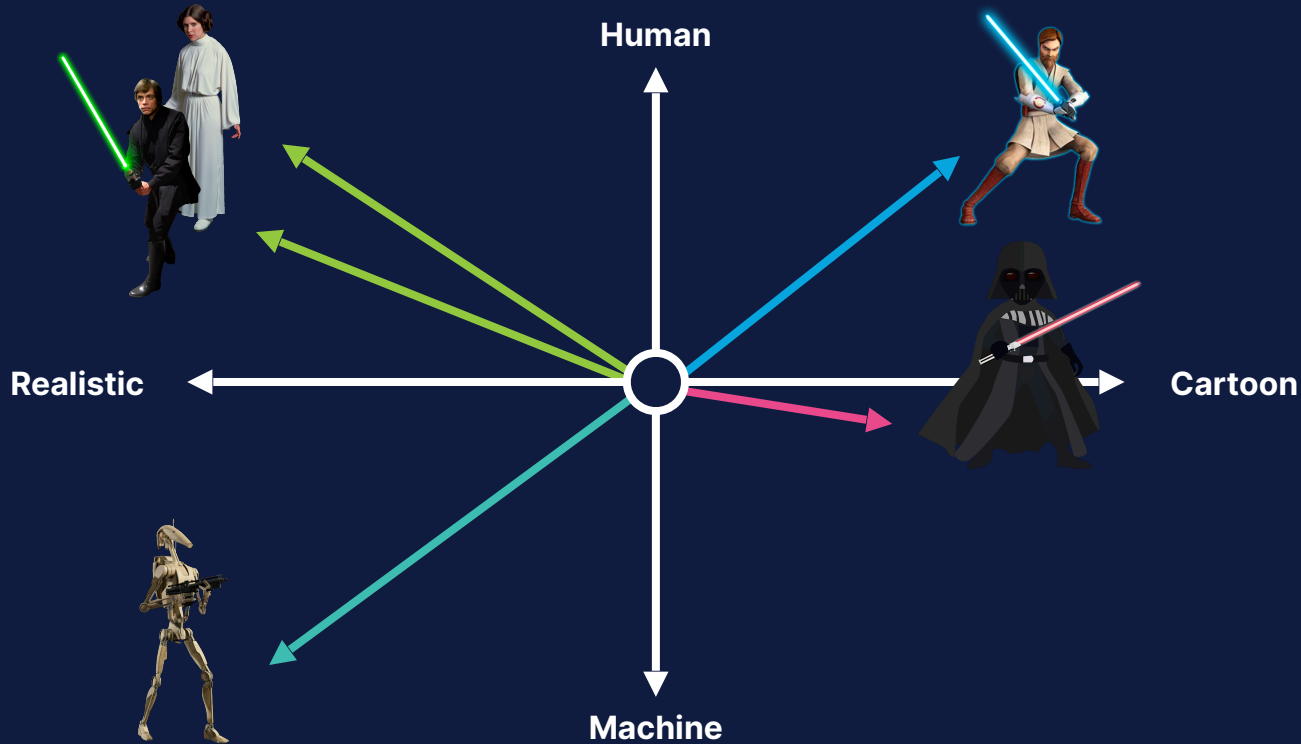
Character	Vector
	$[-1]$
	$[1]$




# Multiple dimensions represent different data aspects



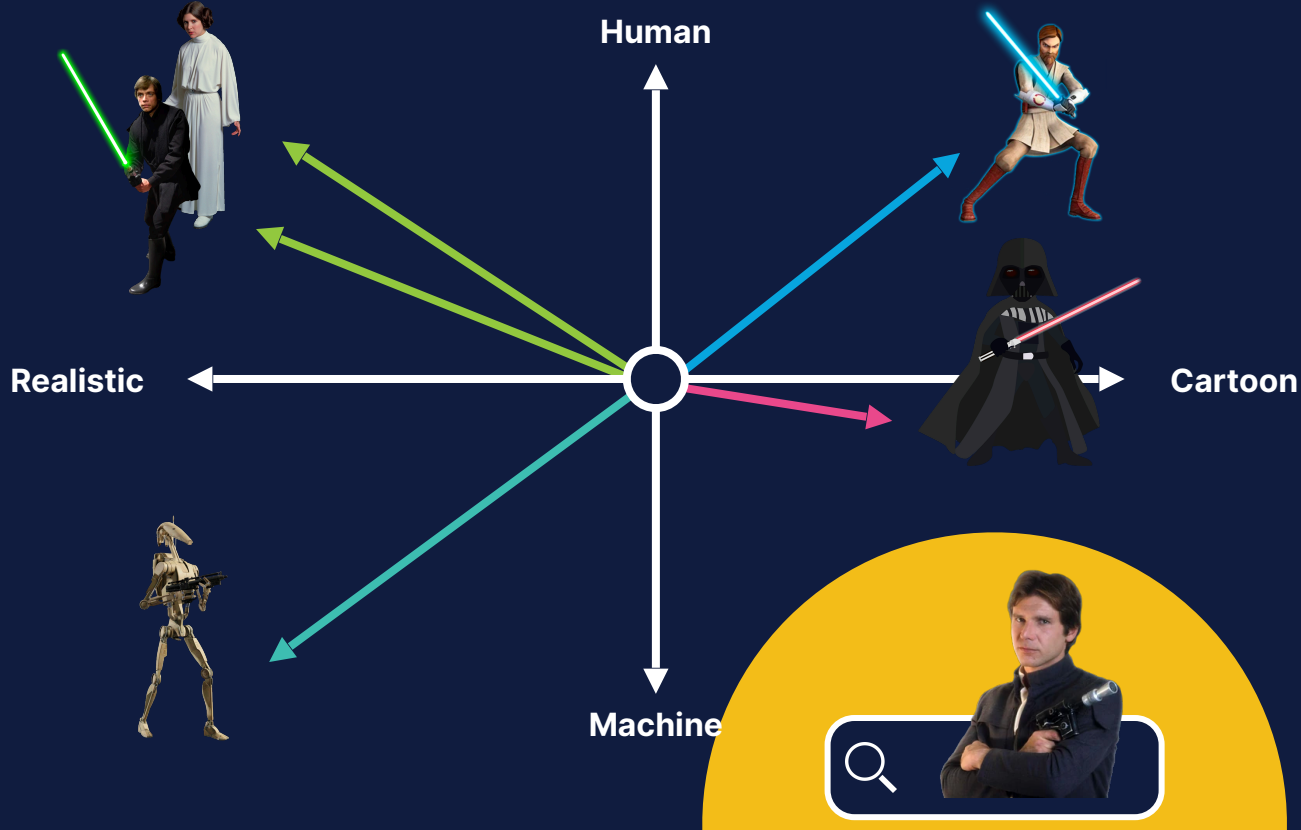
Character	Vector
	$[-1, 1]$
	$[1, 0]$

# Similar data is grouped together



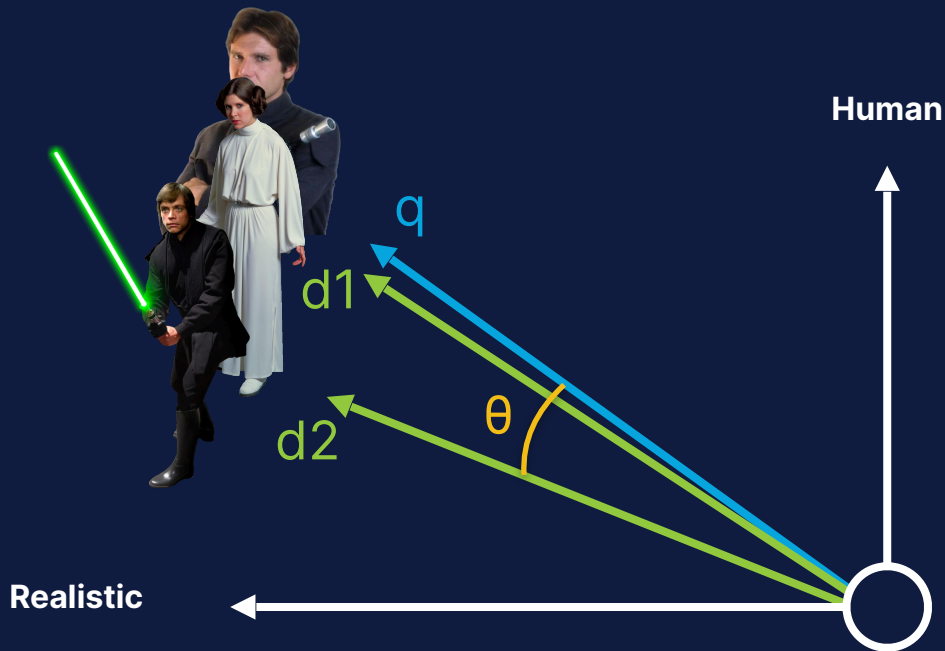
Character	Vector
	$[-1.0, 1.0]$
	$[1.0, 0.0]$
	$[-1.0, 0.8]$

# Vector search ranks objects by similarity (~relevance) to the query



Rank	Result
Query	
1	
2	
3	
4	
5	

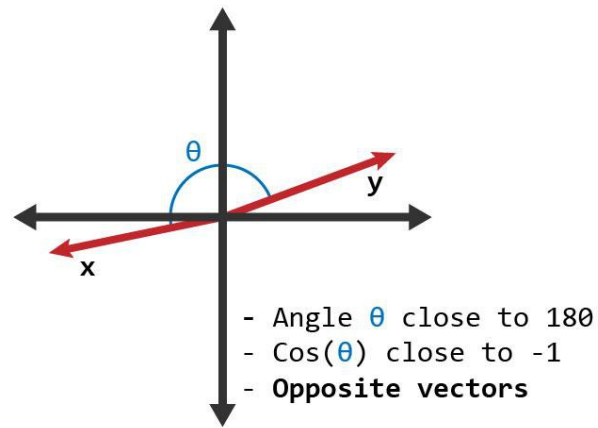
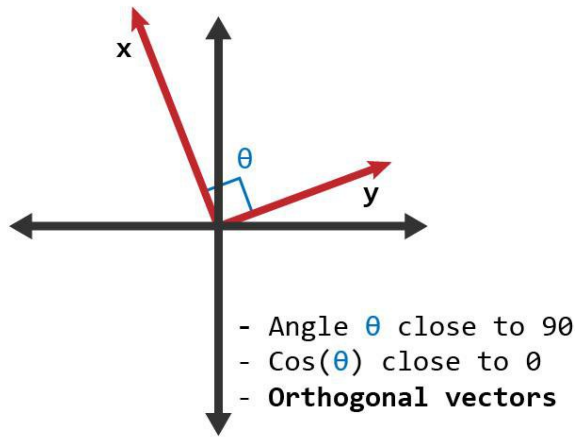
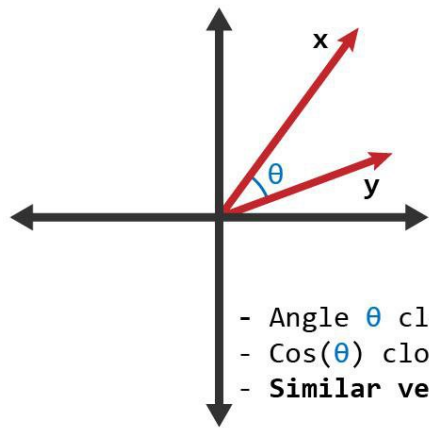
# Similarity: cosine (cosine)



$$\cos(\theta) = \frac{\vec{q} \times \vec{d}}{|\vec{q}| \times |\vec{d}|}$$

$$\text{\_score} = \frac{1 + \cos(\theta)}{2}$$

# Similarity: cosine (cosine)



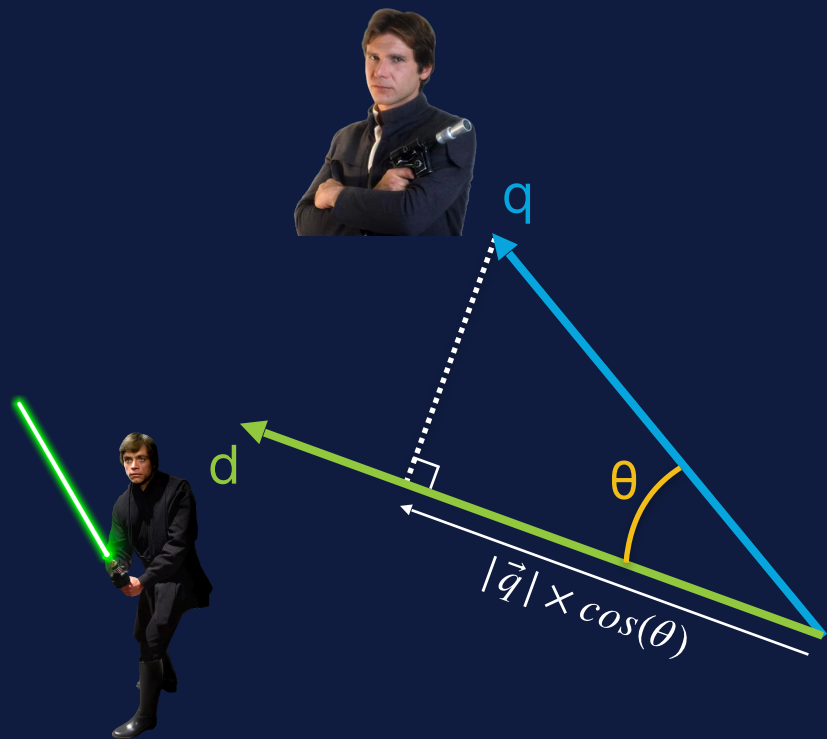
$$\text{\_score} = \frac{1 + 1}{2} = 1$$

$$\text{\_score} = \frac{1 + 0}{2} = 0.5$$

$$\text{\_score} = \frac{1 - 1}{2} = 0$$



# Similarity: Dot Product (`dot_product`)

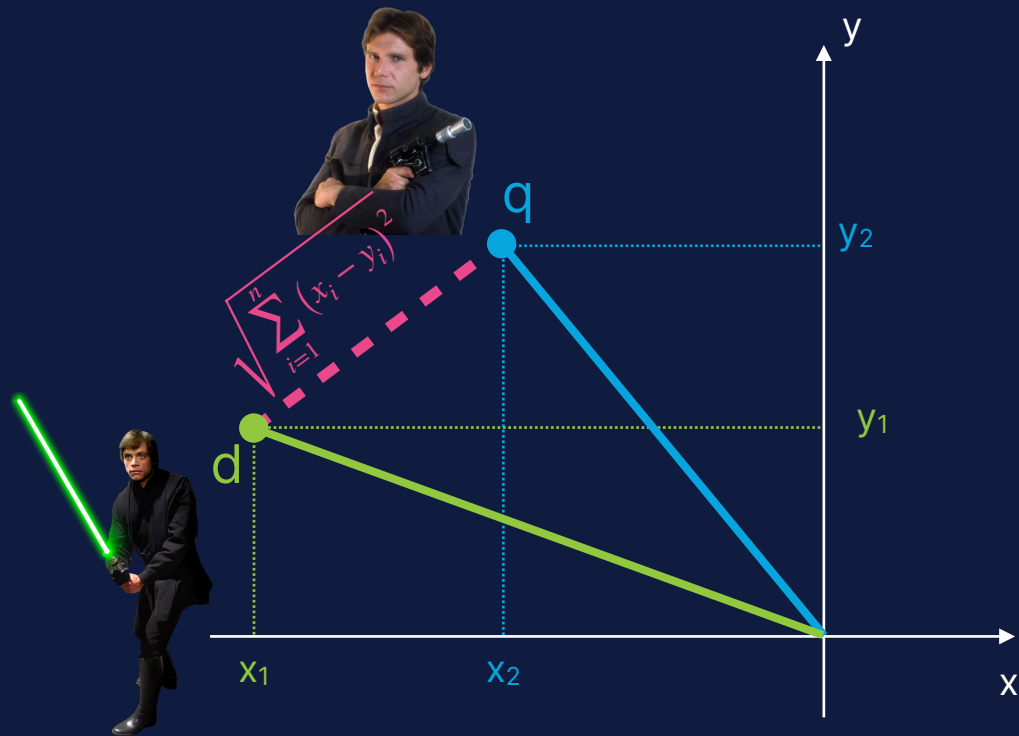


$$\vec{q} \times \vec{d} = |\vec{q}| \times \cos(\theta) \times |\vec{d}|$$

$$\_score_{float} = \frac{1 + dot\_product(q, d)}{2}$$

$$\_score_{byte} = \frac{0.5 + dot\_product(q, d)}{32768 \times dims}$$

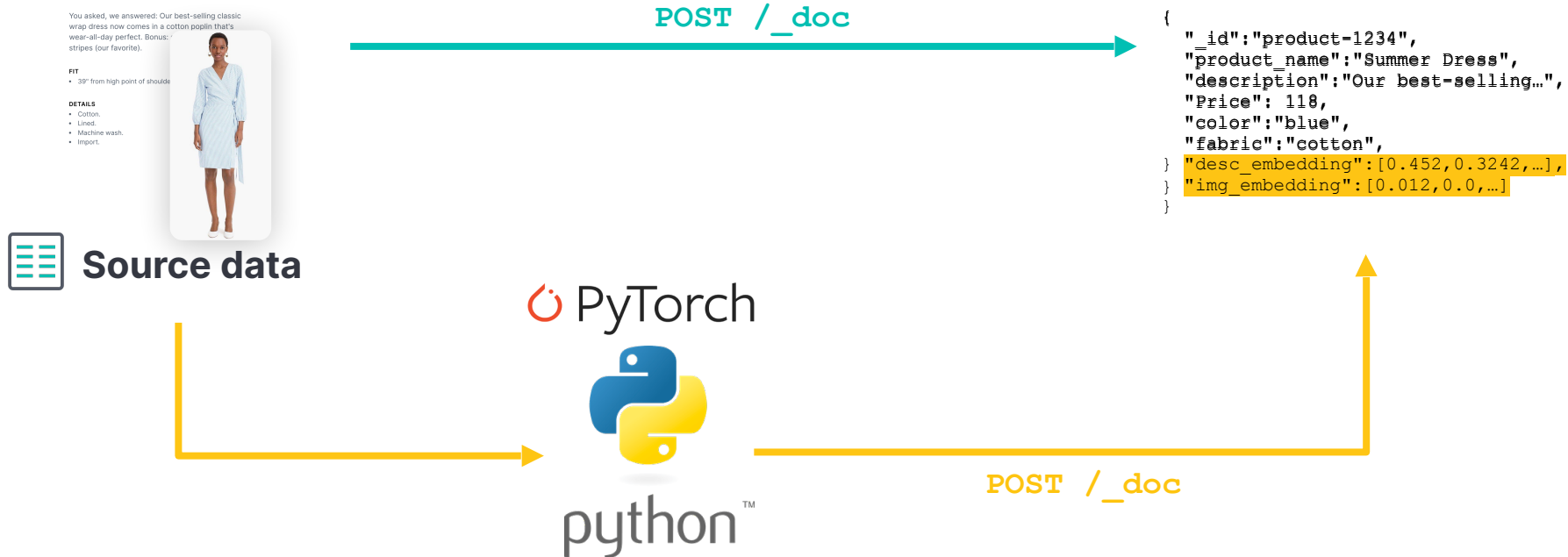
# Similarity: Euclidean distance (l2\_norm)



$$l2\_norm_{q,d} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$
$$\_score = \frac{1}{1 + (l2\_norm_{q,d})^2}$$

# How do you index **vectors**?

# Data Ingestion and Embedding Generation



# How do you search **vectors**?

# Vector Query

🔍 summer clothes | ✕ 🛒

PyTorch



python™

```
GET product-catalog/_search
{
  "knn": {
    "field": "desc_embedding",
    "k": 5,
    "num_candidates": 50,
    "query_vector": [0.123, 0.244, ...],
  },
  "filter": {
    "term": {
      "department": "women"
    }
  },
  "size": 10
}
```

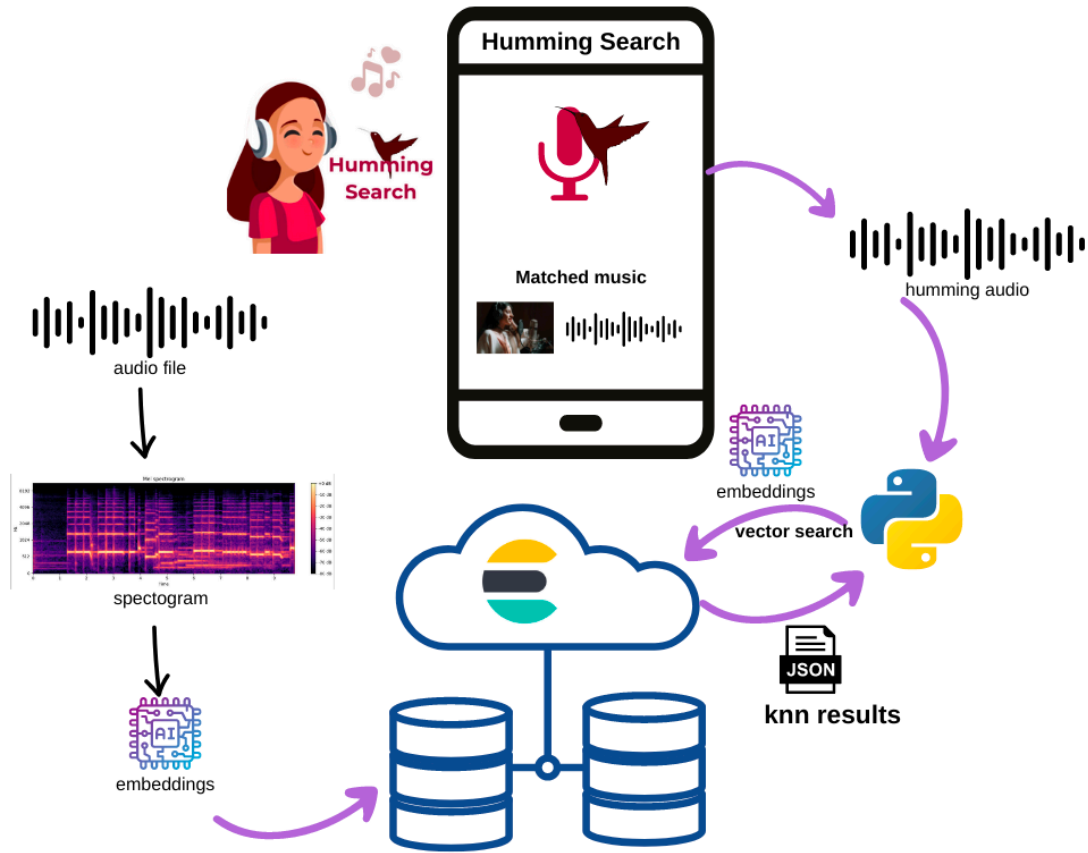
Anniversaire **Lucas** - 25 ans



**16/09/2023**



<https://djdadoo.pilato.fr/>



<https://github.com/dadoonet/music-search/>





# Searching for similar music tracks



David Pilato | [@dadoonet](#)



# Kahoot!