



SeMI Technologies

Weaviate - Vector Search Engine

Knowledge Graph Conference 2021

By Laura Ham, Community Solution Engineer at SeMI Technologies

Agenda



- Introduction to Weaviate: Vector Search Engine
- Demo
- Weaviate's core features
- Use cases



Weaviate

The AI-based vector search engine

Weaviate - why a vector search engine?

```
{ "data": [{  
  "Wine": "Covey Run 2005 Chardonnay",  
  "Description": "... good with fish ..."  
}]}
```

"Wine for seafood"



No products found ...

Traditional search engine



"Wine for seafood"

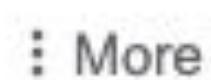


Covey Run 2005
Chardonnay

Vector search engine



What color of wine is Chardonnay?



Settings

Tools

About 12.800.000 results (0,81 seconds)

Chardonnay / Wine color

White Wine



"Chardonnay is the most compelling and popular **white wine** in the world, because it is the red wine of whites," Ramey said. "It's so complex, so interesting. And it's the red wine of whites for two reasons: barrel fermentation and malolactic." 26 Jul 2019

And how so extremely **fast**?


The question is very **abstract**.

How did Google find exactly this **data node**?
How can we predict this **relation**?

How to do this on your **own data**?


What's the problem?

- 80% to 90% of data is **unstructured**, that's about **34 trillion gigabytes!**
- We keep storing more and more data but businesses can't get **valuable insights** from it.
- 95% of businesses have issues **getting value** from their unstructured data.




How to achieve **search** and automatically
make **relations** within your **own data**?


In a **easy, fast, secure** and **scalable** way?




Weaviate is a cloud-native, modular,
real-time vector search engine
built to scale your machine learning models




Weaviate is a cloud-native, modular,
real-time vector search engine
built to scale your machine learning models



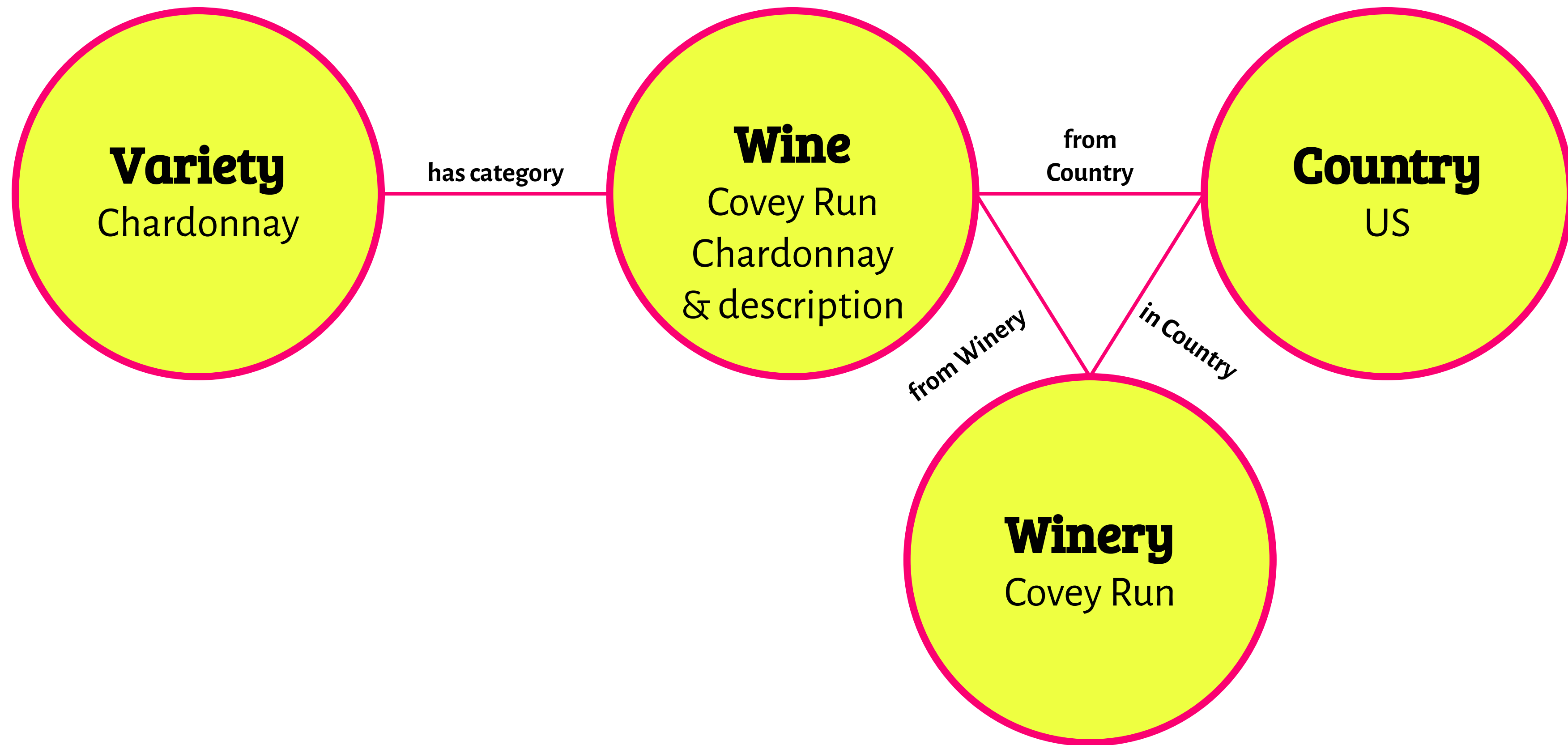
Weaviate is a cloud-native, modular,
real-time vector search engine
built to scale your machine learning models

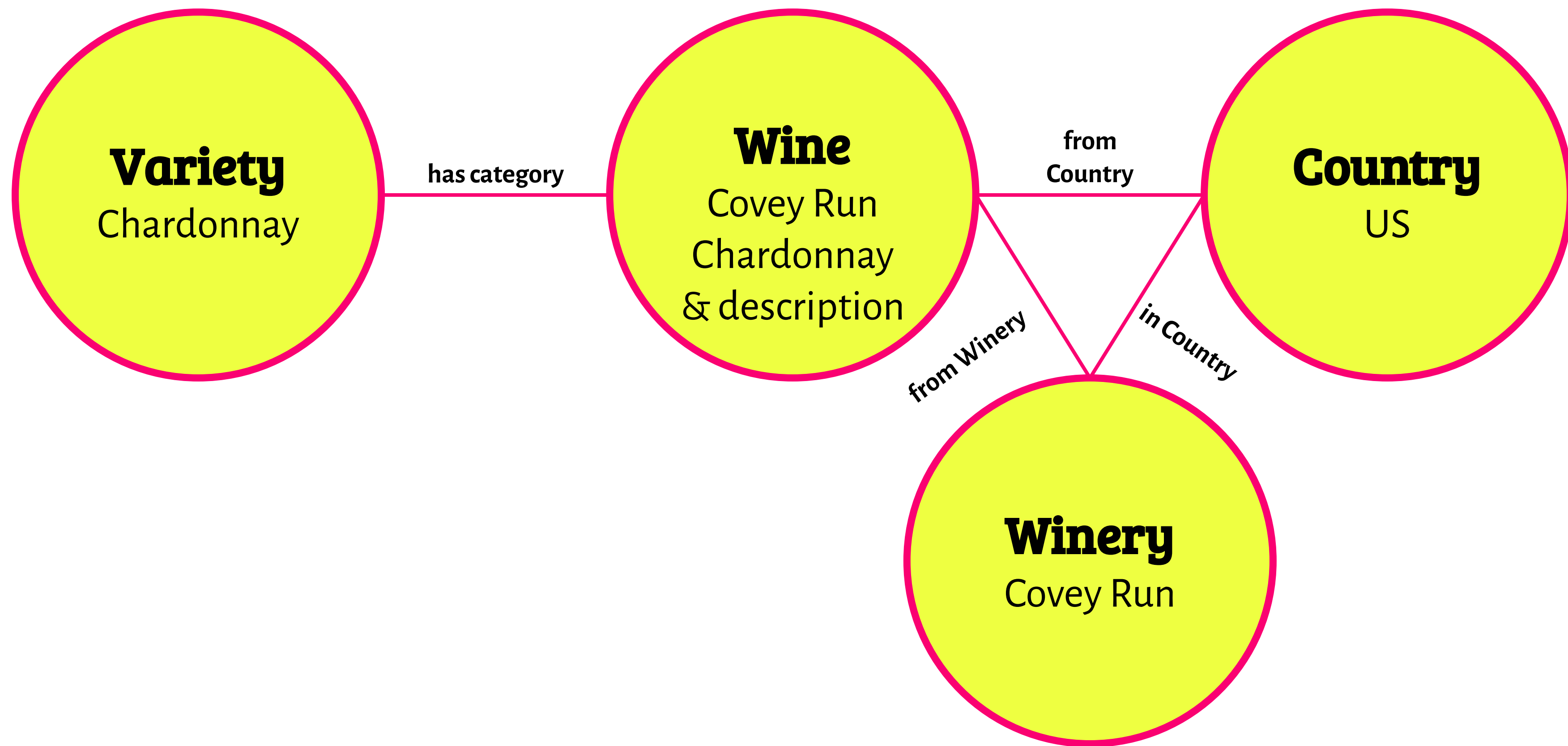


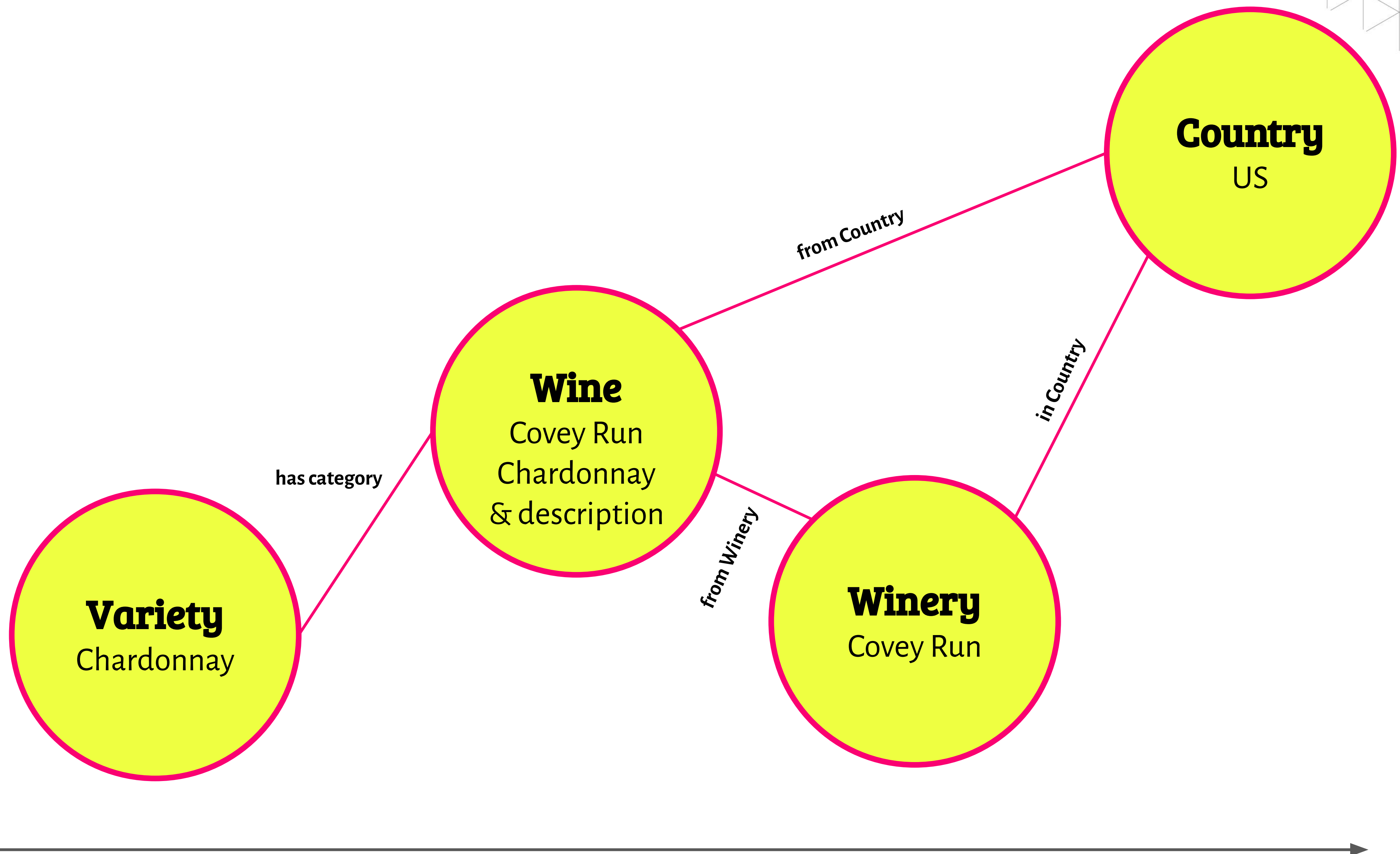
Weaviate is a cloud-native, modular,
real-time vector search engine
built to scale your machine learning models

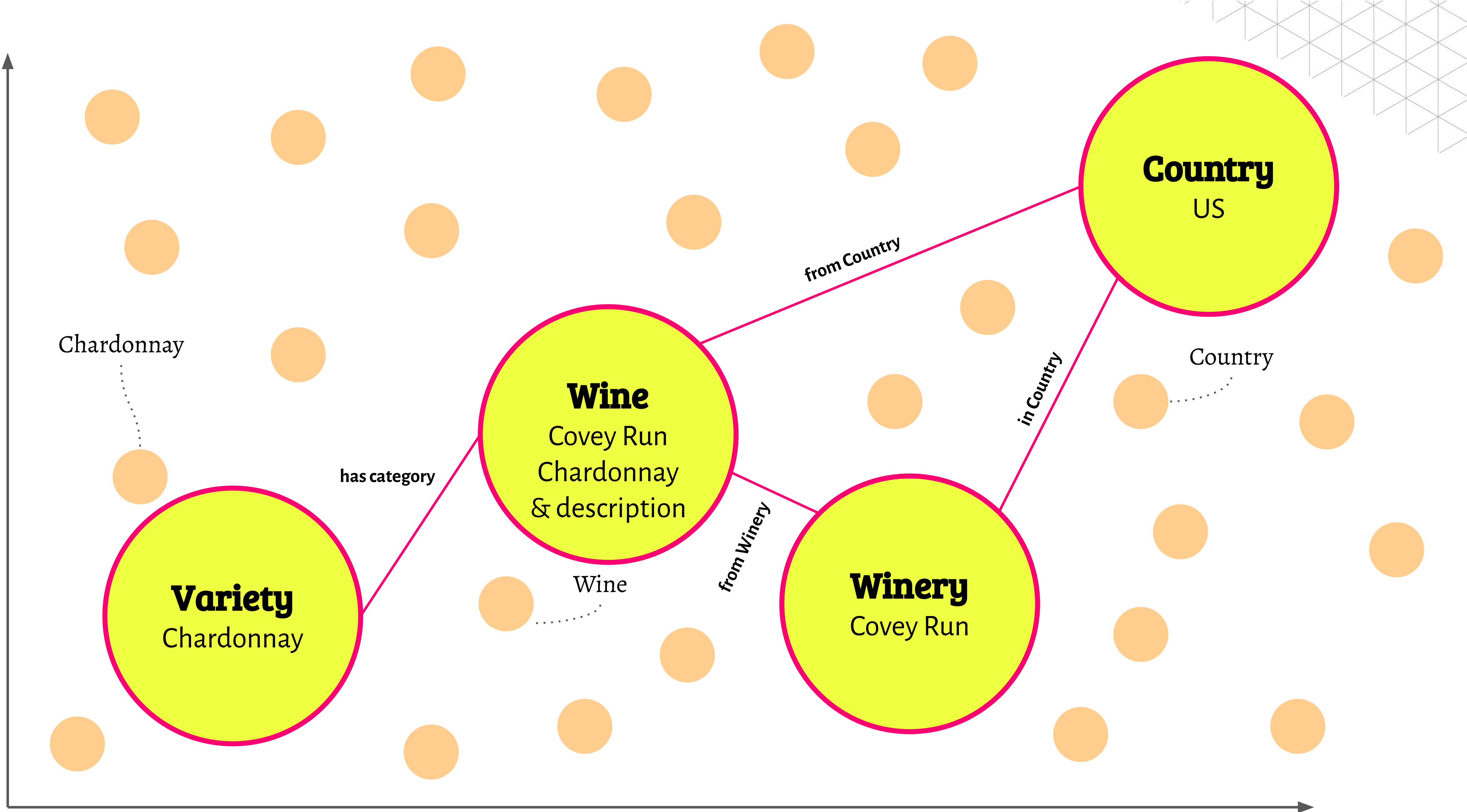


Weaviate is a cloud-native, modular,
real-time vector search engine
built to scale your machine learning models





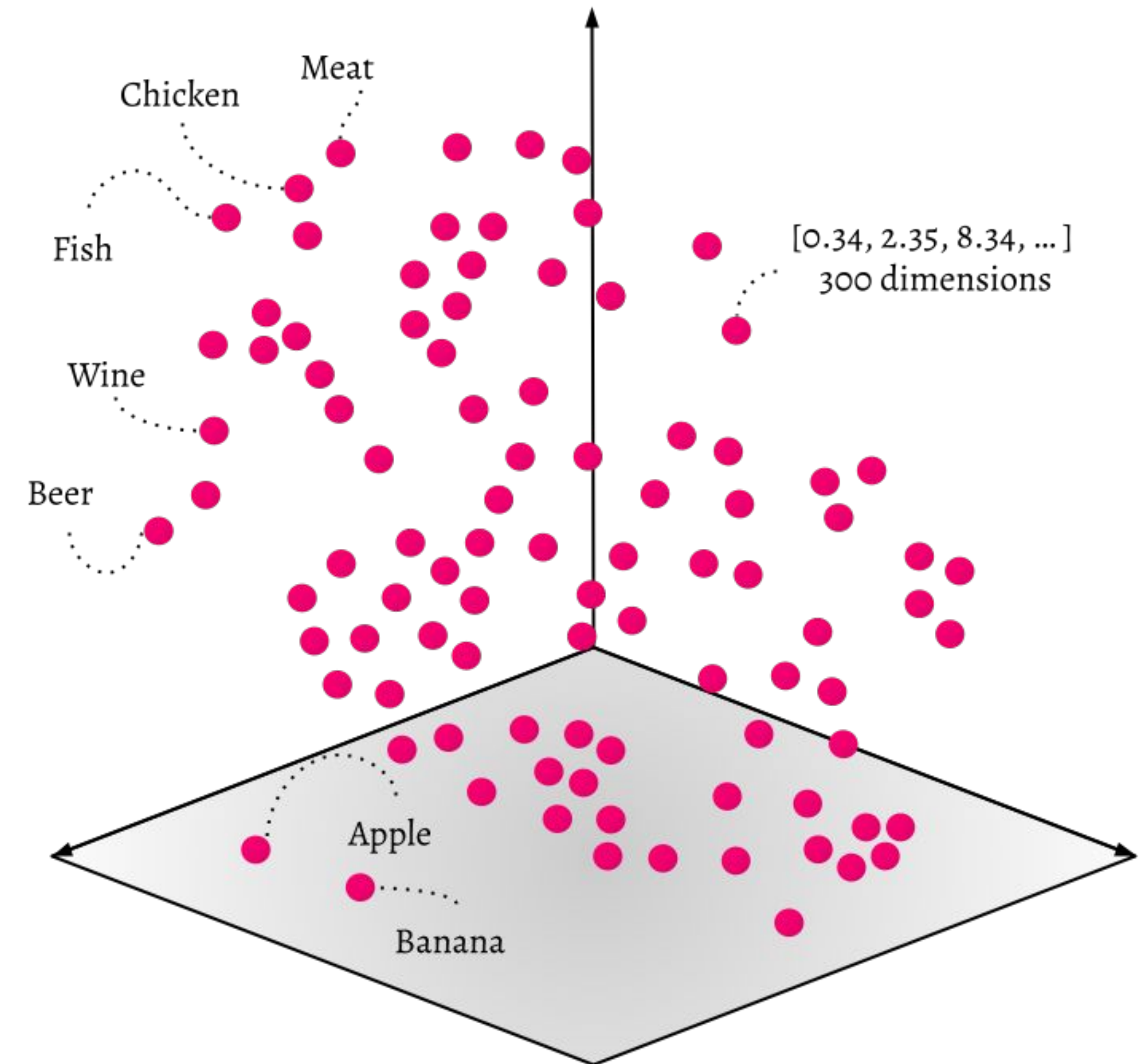




Weaviate - How does it work?

- Data is stored as high dimensional vectors
- Vector positions capture data meaning and context
- Pre-trained NLP module for automatic
 - Vectorization
 - Classification
 - Nearest neighbor search

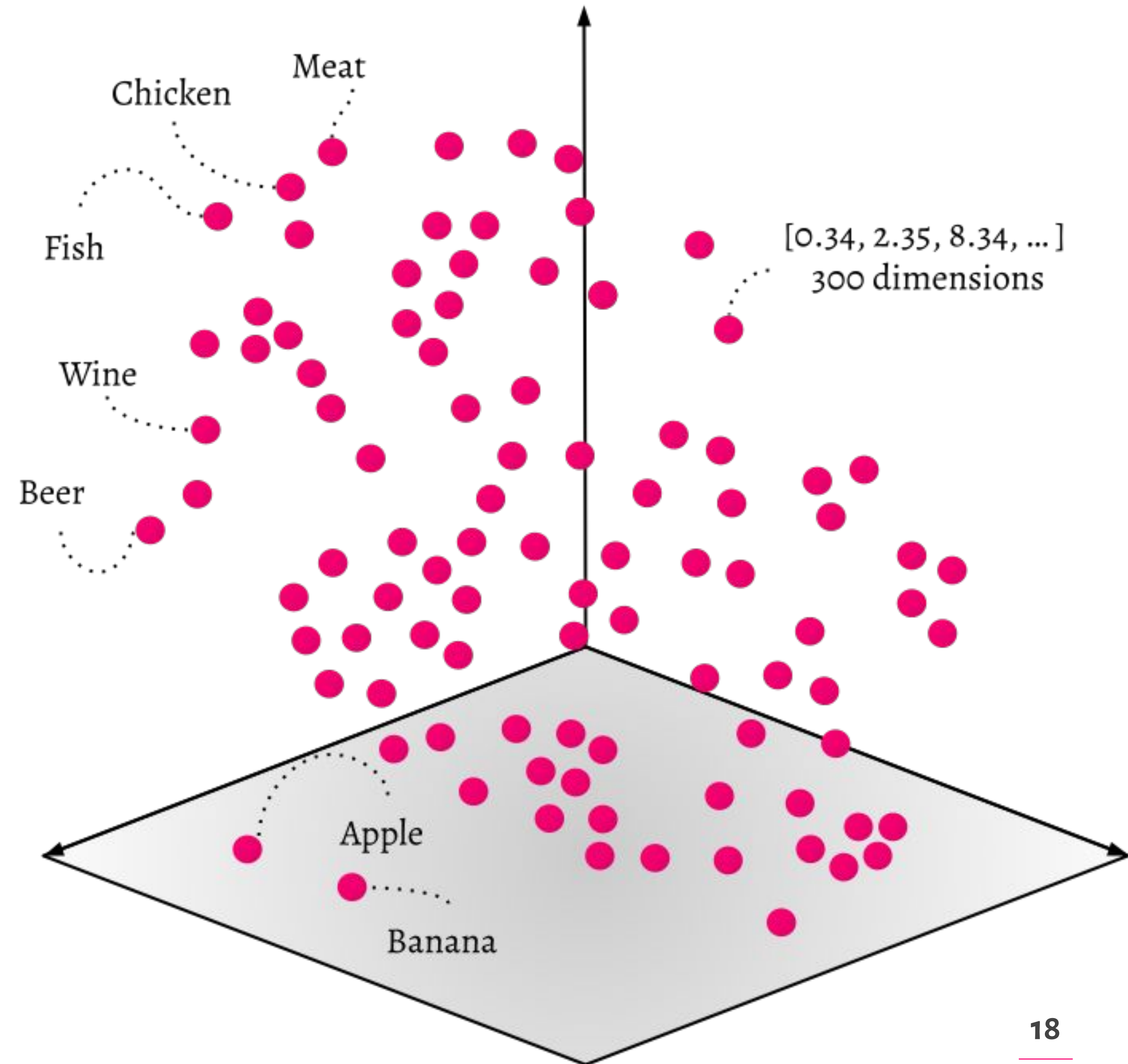
⇒ Weaviate *understands* the data



Weaviate - How does it work?

1. Weaviate Machine Learning Modules

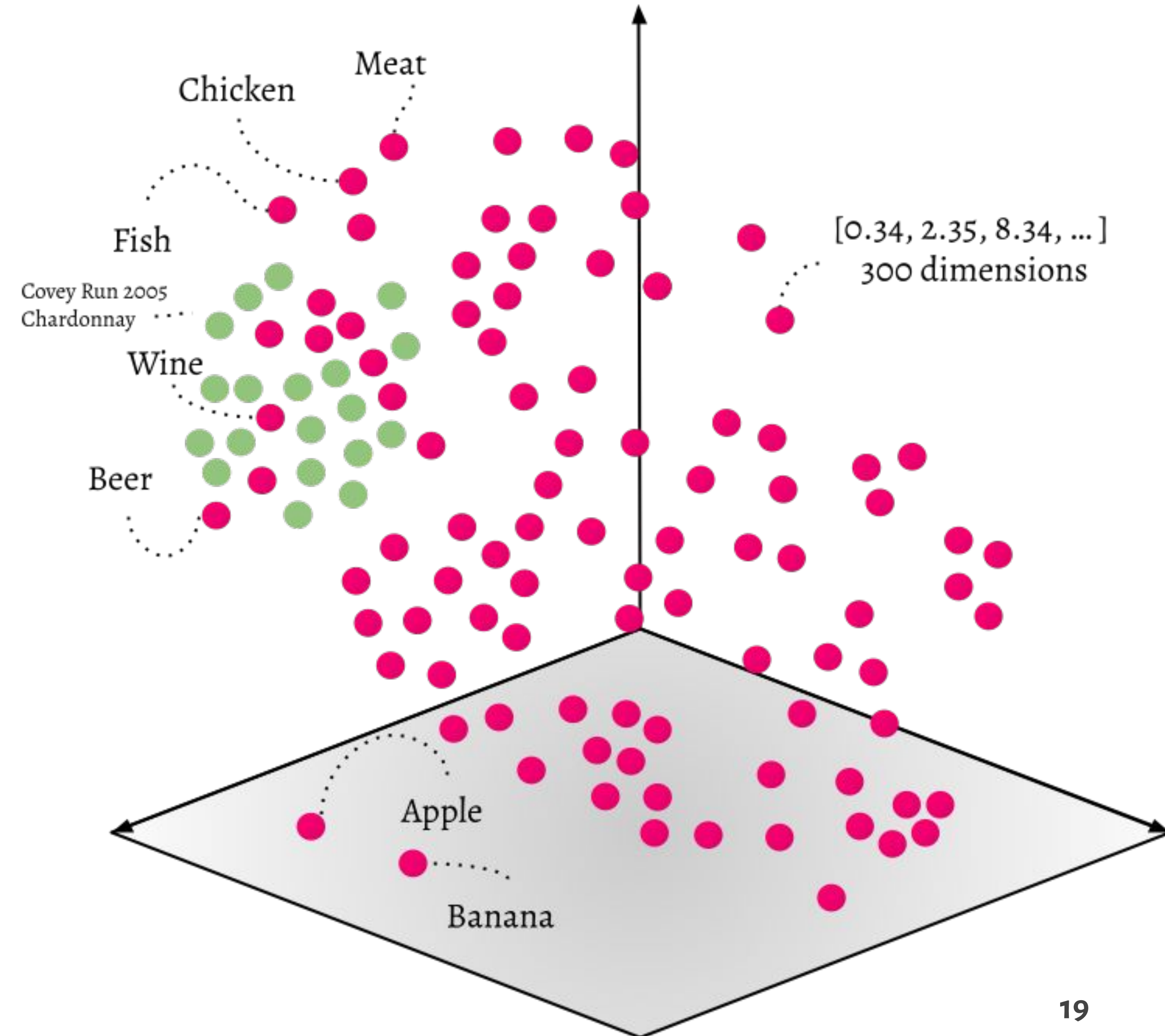
- E.g. NLP module trained with *fasttext*
- This language model represents all words and concepts in the hyperspace.



Weaviate - How does it work?

2. Automatically vectorize and index your data

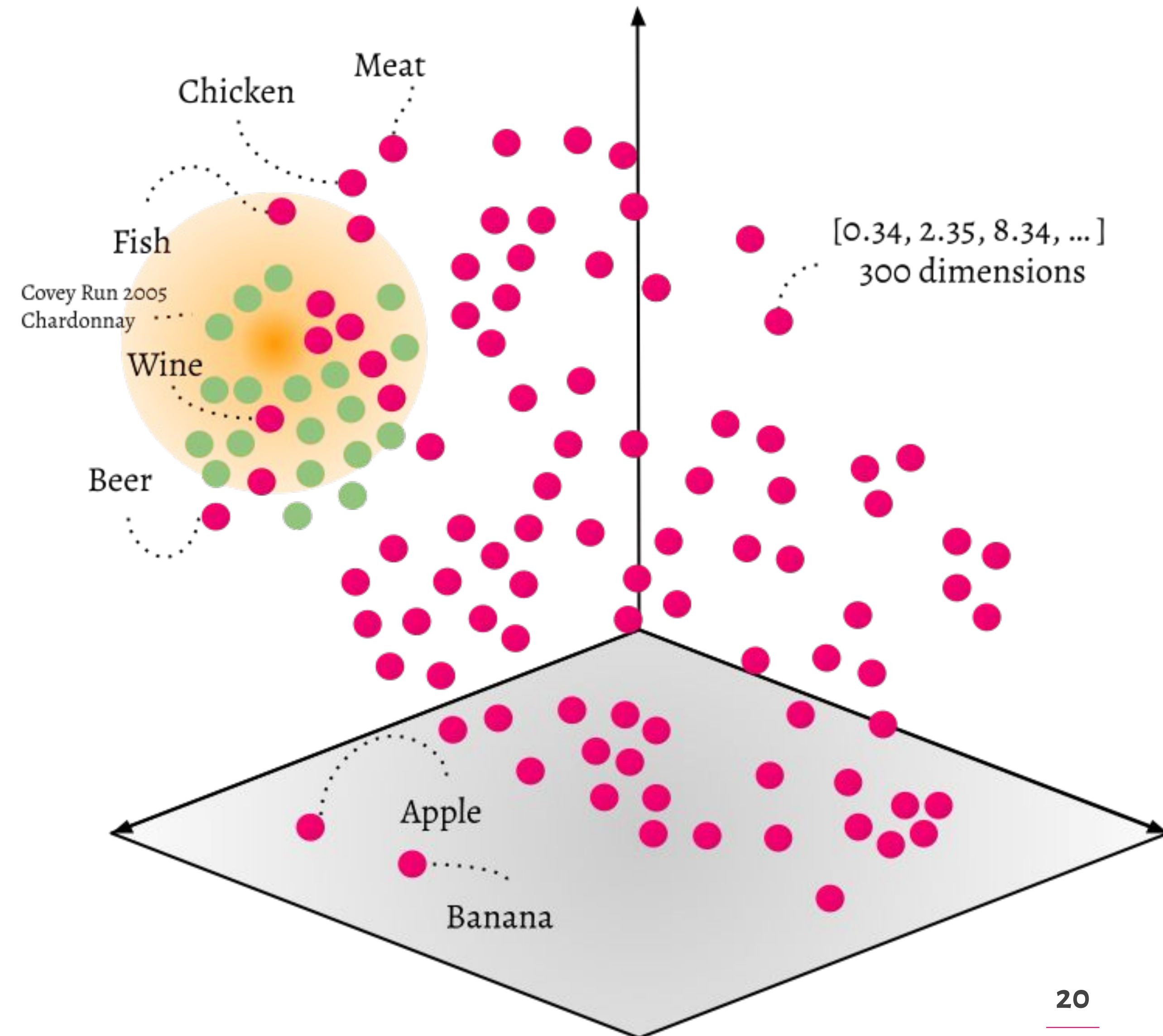
- Weaviate *understands* your data
- When you import data: Weaviate looks at the language in your data object
- E.g. a *Chardonnay* is closely related to *Wine*, *White* and the *food* it fits with



Weaviate - How does it work?

3. Search query

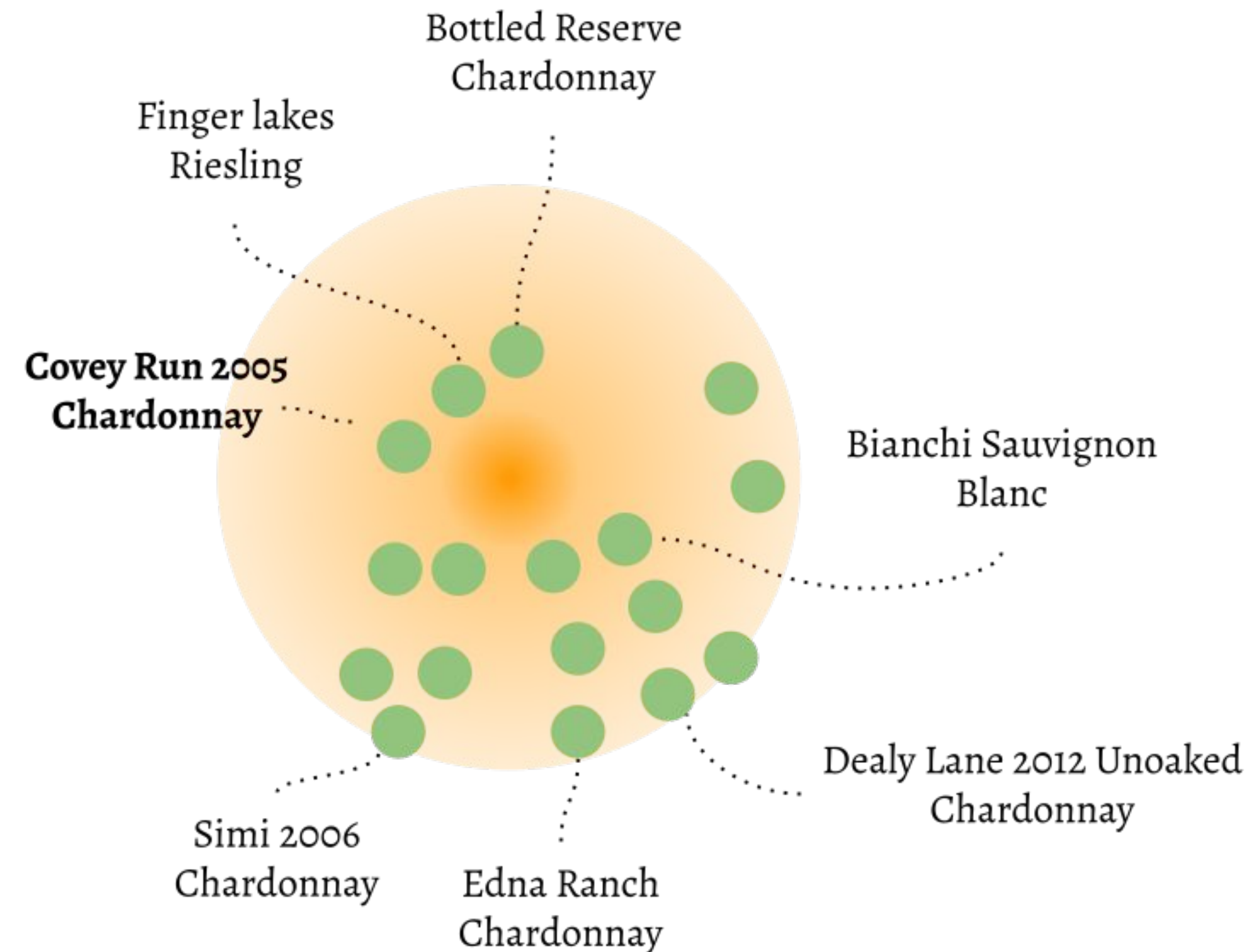
- Your search queries in natural language will also be vectorized and understood by the machine learning module of Weaviate
- It is places close to the words and data object that are semantically related to the query
- E.g. *"Wine that fits with a seafood dish"*



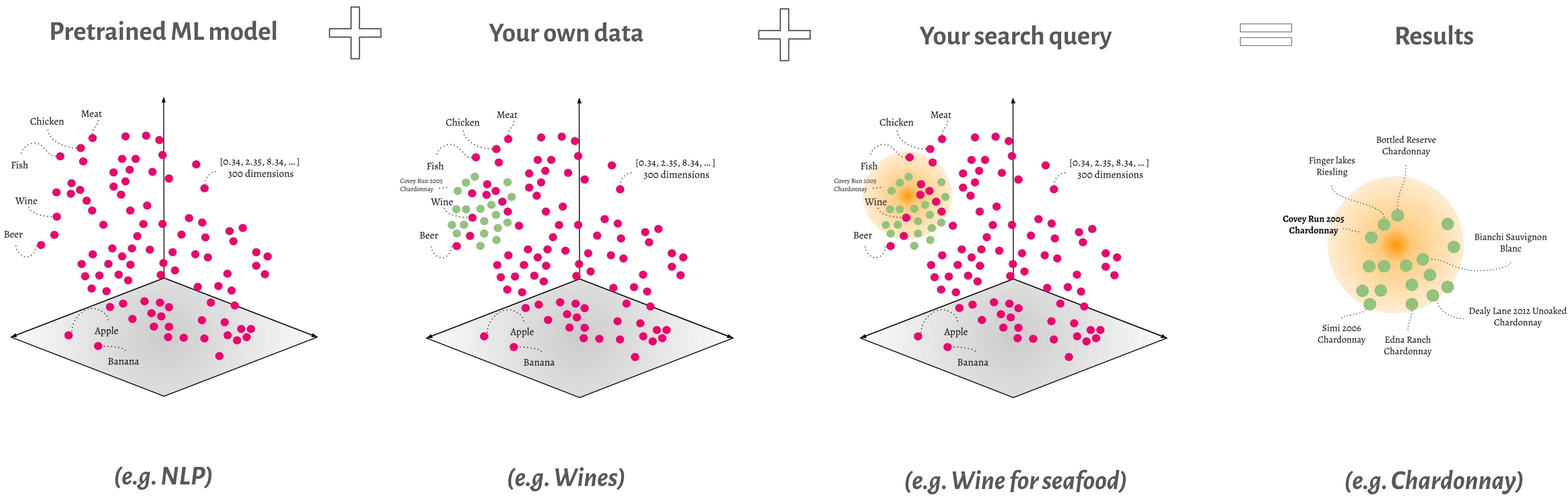
Weaviate - How does it work?

4. Results

- The data objects that are closest to the search query are retrieved from the dataset



Weaviate - How does it work?





Demo

Weaviate - Scalable Vector Search Engine



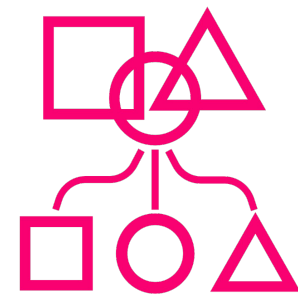
Search

how to search the vector space



Weaviate Modules

out-of-the-box modules you can use
e.g. Weaviate NLP modules,
Transformer modules
Custom modules



Classification

how to classify in the vector space

Weaviate Vectorization Modules

Machine learning models

- **Weaviate's NLP module: trained with *fasttext***
 - Contextionary, in multiple languages
- **Transformer modules: general-purpose NLP architectures**
 - e.g. BERT, GPT-2, RoBERTa, XLM, DistilBert, etc
- **Custom modules**

Weaviate - Scalable Vector Search Engine



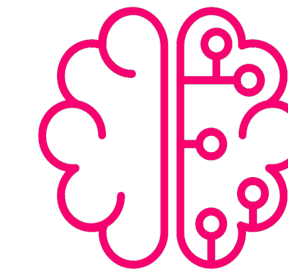
Search

how to search the vector space



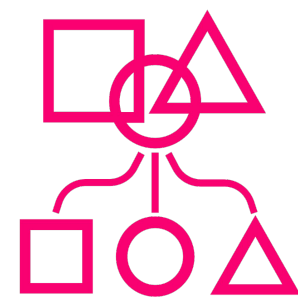
Weaviate Modules

out-of-the-box modules you can use
e.g. Weaviate NLP modules,
Transformer modules
Custom modules



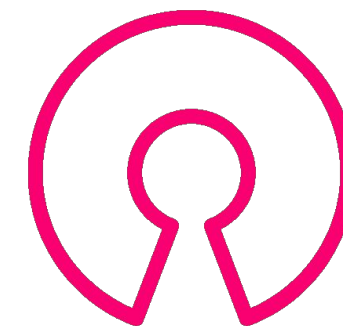
End-to-end

Complete solution for industry,
API driven
Cloud-native
Secure



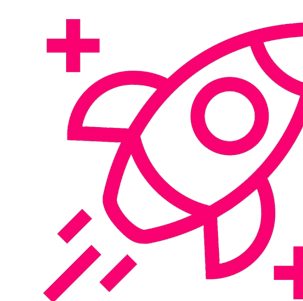
Classification

how to classify in the vector space



Open Source

all code and models are open



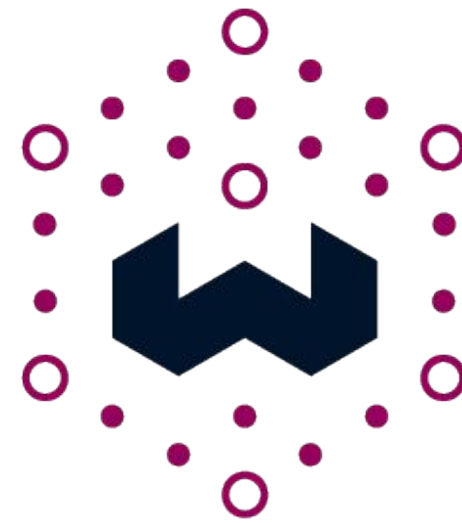
Scalable and fast

Vectorizing and querying big data

Example use cases



Weaviate
Enterprise Search



Weaviate
ERP classifications



Weaviate
Cybersecurity



Weaviate
Biology



Weaviate
Data Harmonization



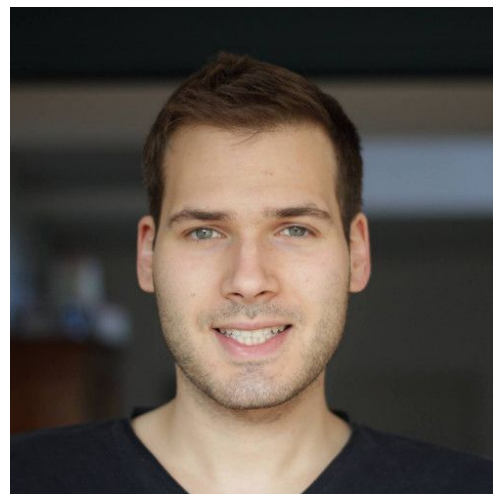
Weaviate
Social Media



Weaviate
Service Management

SeMI Technologies

Team



Etienne



Donna



Laura



Micha



Bob



Stefan



Marcin



Henrique



Amar



Alicja

Try out Weaviate yourself!

- Go to www.semi.technology/developers

The screenshot displays the Weaviate web interface. On the left is a pink sidebar with icons for a menu, a query editor, and a list. The top navigation bar includes the Weaviate logo, a RSS feed icon, a notification bell, a key icon, and a 'Log Out' button. Below the navigation bar is a toolbar with buttons for 'Prettify', 'Merge', 'Copy', 'History', 'Share this query', and a link icon. The main area is split into two panels. The left panel shows a query in a code editor with line numbers 1 through 12. The right panel shows the JSON response of the query.

```
1 {  
2   Get {  
3     Wine (  
4       nearText: {  
5         concepts: ["wine with fish"]  
6       }  
7     ) {  
8       title  
9       review  
10    }  
11  }  
12 }
```

```
{  
  "data": {  
    "Get": {  
      "Wine": [  
        {  
          "review": "Light and fruity with pineapple, melon  
and apple. This fruit forward, food-friendly wine has  
natural acids and a delicate balance. For the money it is a  
fine choice to accompany seafood, poultry and pasta  
salads.",  
          "title": "Covey Run 2005 Quail Series Chardonnay  
(Columbia Valley (WA))"  
        },  
        ...  
      ]  
    }  
  }  
}
```


Thank you!

www.semi.technology

Laura Ham

laura@semi.technology

