

# txtai – the all-in-one embeddings database

txtai is an all-in-one embeddings database for semantic search, LLM orchestration and language model workflows.

Summary of txtai features:

- *Vector search* with SQL, object storage, topic modeling
- Create *embeddings* for text, documents, audio, images and video
- *Pipelines* powered by language models that run LLM prompts
- *Workflows* to join pipelines together and aggregate business logic
- Build with *Python* or *YAML*. API bindings available for JavaScript, Java, Rust and Go.
- *Run local or scale out with container orchestration*

## Examples

List of example notebooks.

Notebook	Description
Introducing txtai	Overview of the functionality provided by txtai
Similarity search with images	Embed images and text into the same space for search
Build a QA database	Question matching with semantic search
Semantic Graphs	Explore topics, data connectivity and run network analysis

## Install

The easiest way to install is via pip and PyPI

```
pip install txtai
```

Python 3.8+ is supported. Using a Python virtual environment is **recommended**.

See the detailed install instructions for more information covering optional dependencies, environment specific prerequisites, installing from source, conda support and how to run with containers.

## Model guide

The following shows a list of suggested models.

Component	Model(s)
Embeddings	all-MiniLM-L6-v2 E5-base-v2
Image Captions	BLIP
Labels - Zero Shot	BART-Large-MNLI
Labels - Fixed	Fine-tune with training pipeline
Large Language Model (LLM)	Flan T5 XL Mistral 7B OpenOrca
Summarization	DistilBART
Text-to-Speech	ESPnet JETS
Transcription	Whisper
Translation	OPUS Model Series