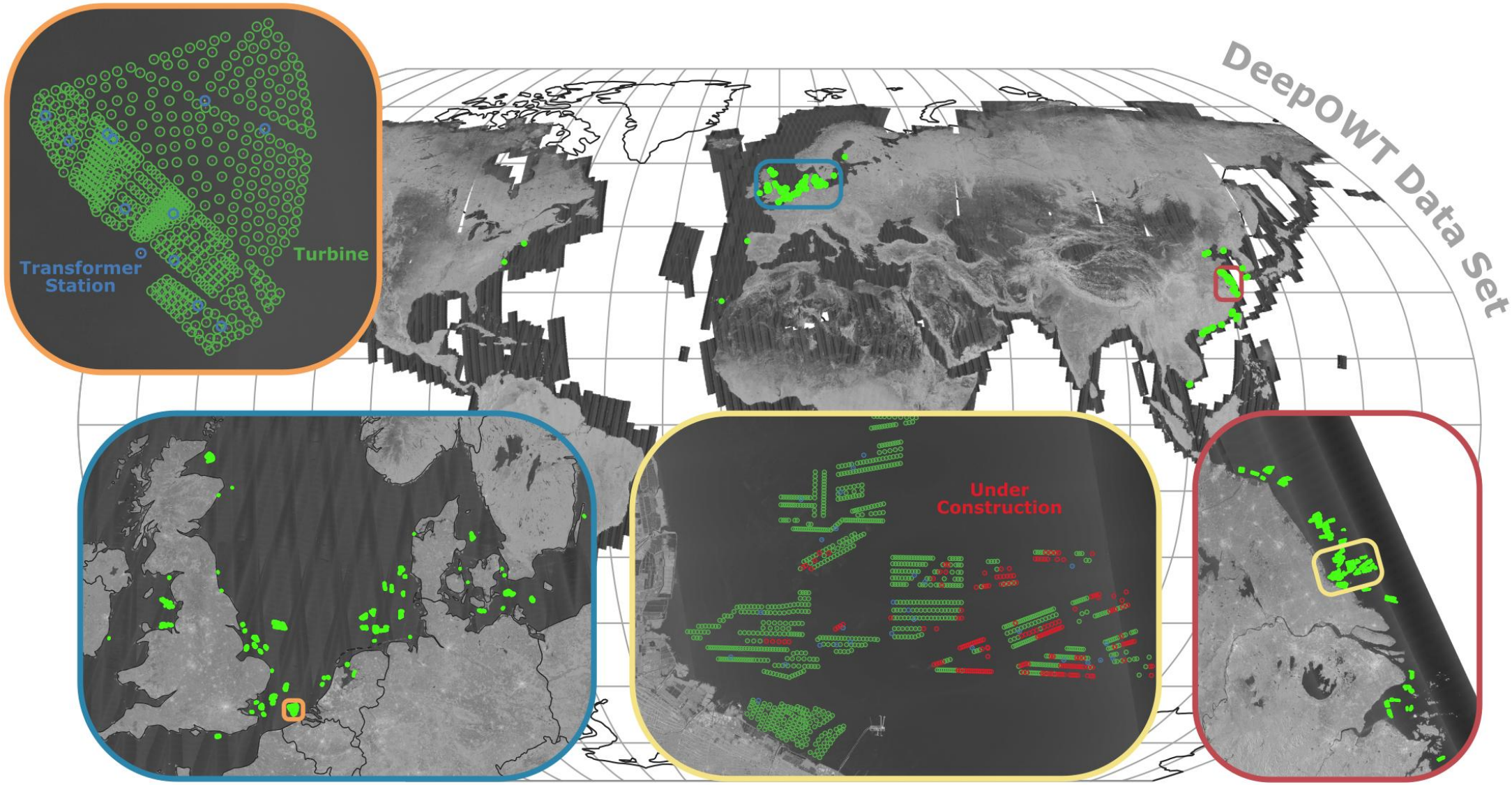
SEMANTIC ENRICHMENT OF SYNTHETIC EARTH OBSERVATION DATA: CONCEPT AND TECHNICAL IMPLEMENTATION

Dr. Thorsten Höser, Dr. Marco Ottinger, Robin Spanier | German Aerospace Center |
Living Planet Symposium 2025

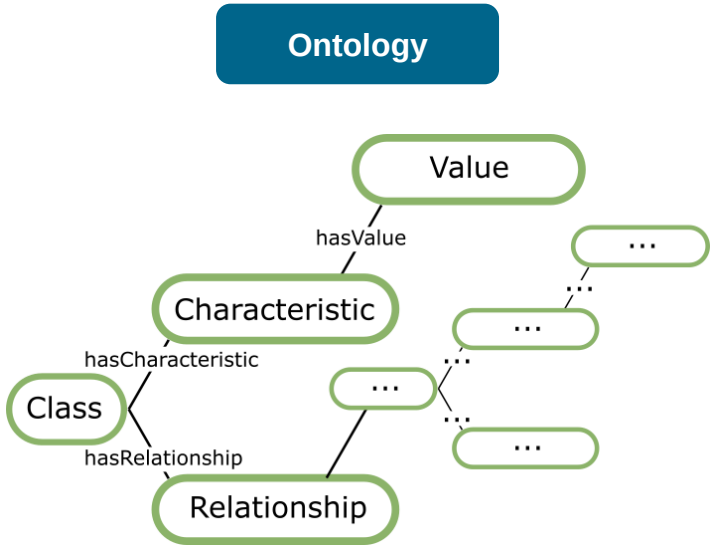
DLR

Global Offshore Wind Energy Infrastructure Derived from Sentinel-1



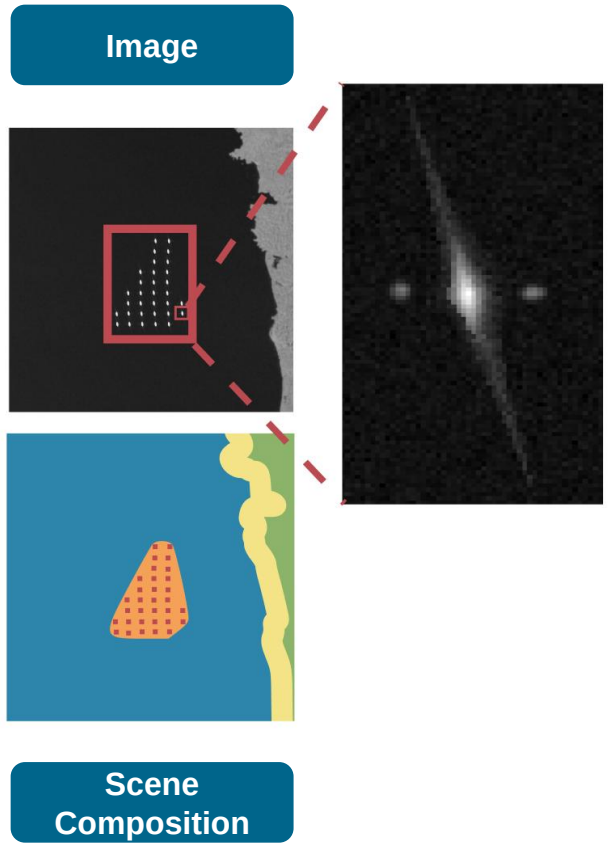
Hoeser et al. 2022a

SyntEO Conceptual Overview



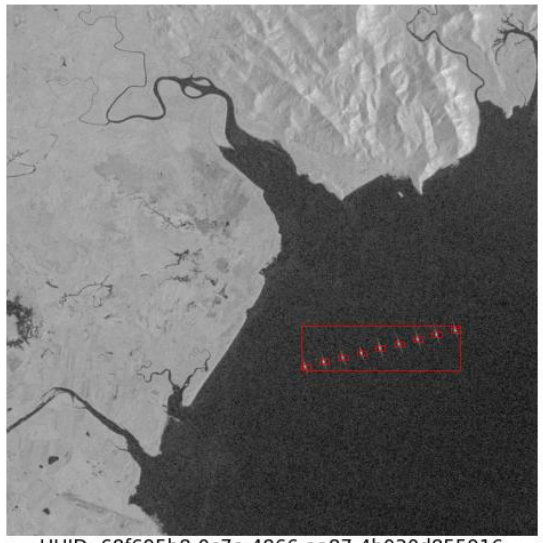
Hoeser and Kuenzer 2022

Create an individual with a constraint-based ontology instantiation

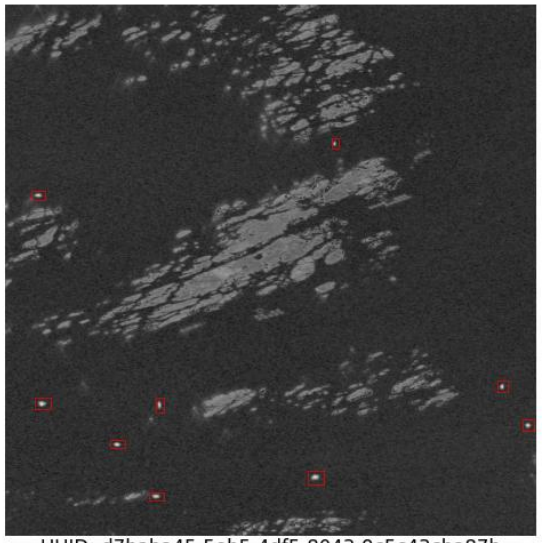


Generated Images and Derived Annotations

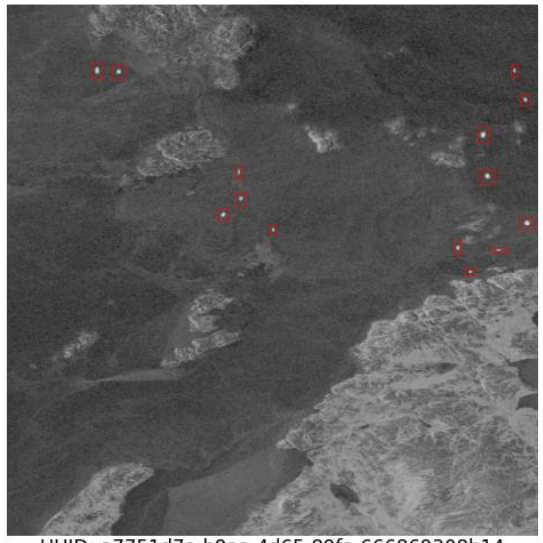
UUID: 23d1abba-526b-414a-b204-a82ea6f85c75



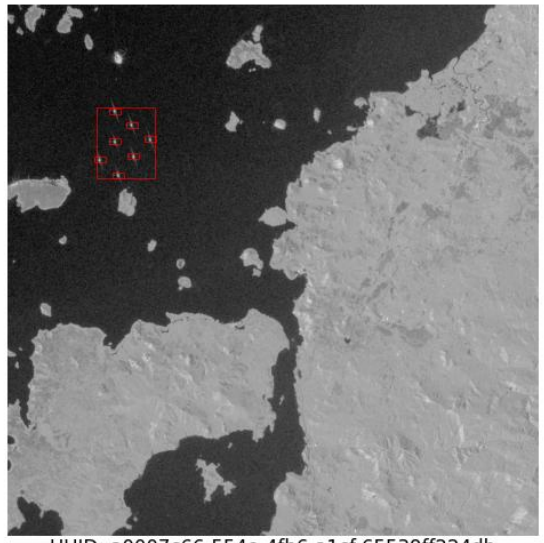
UUID: f2e07092-f8df-4276-aafc-54c83e063758



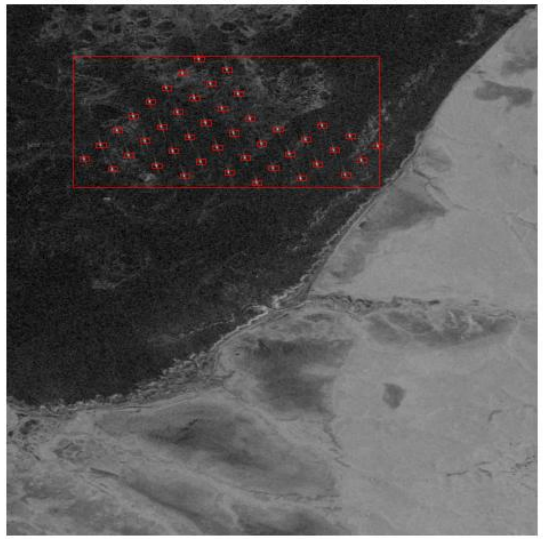
UUID: f3ece385-774b-4b22-af64-a7635262d296



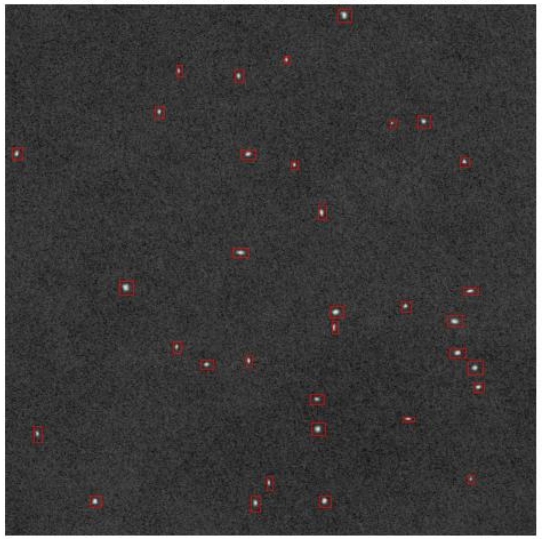
UUID: 7db0be02-81b4-4b5a-b89c-2d805f6095ae



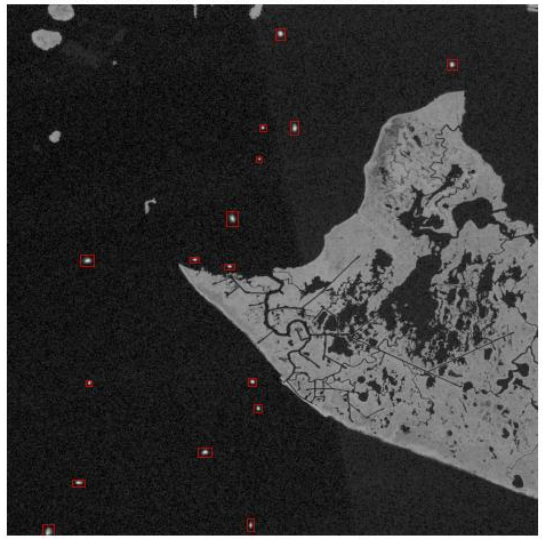
UUID: 68f695b8-0c7e-4866-aa87-4b030d855916



UUID: d7baba45-5eb5-4df5-8043-9c5c43cba87b



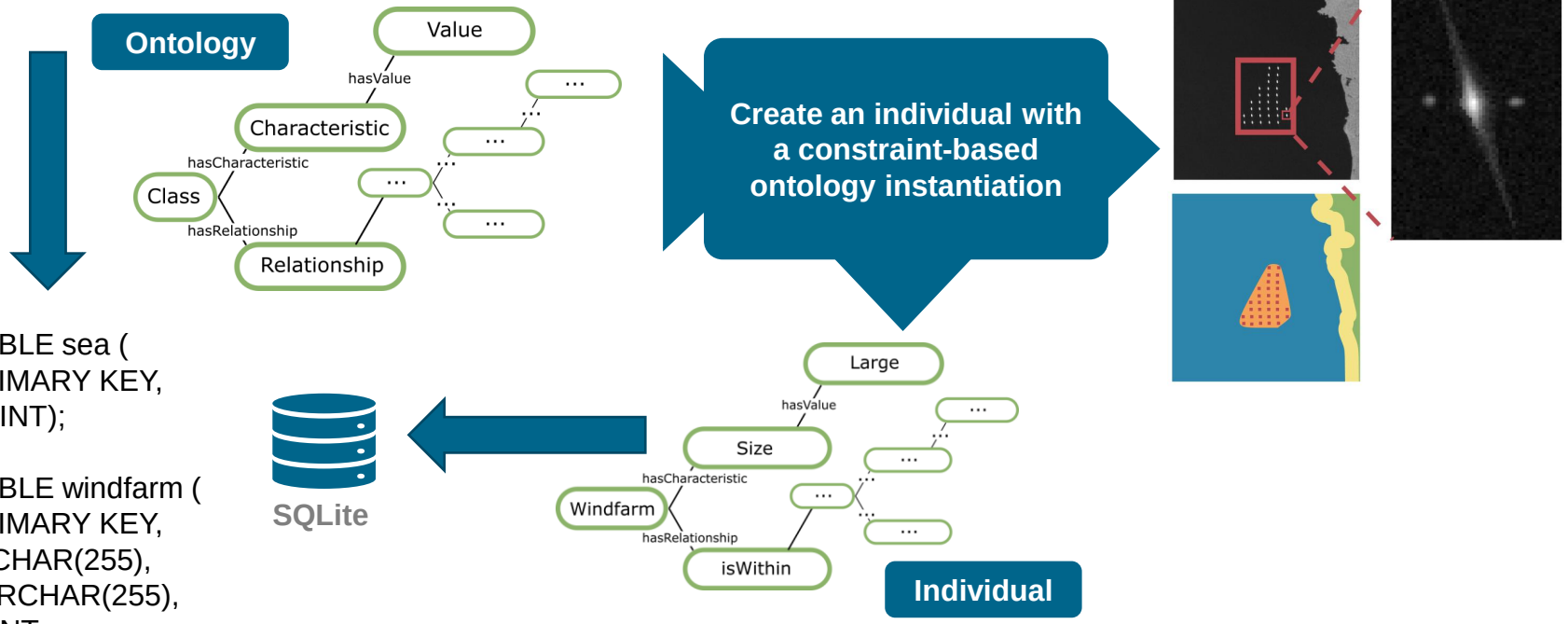
UUID: a7751d7a-b9aa-4d65-89fa-666869308b14



UUID: a0007c66-554e-4fb6-a1cf-65539ff224db



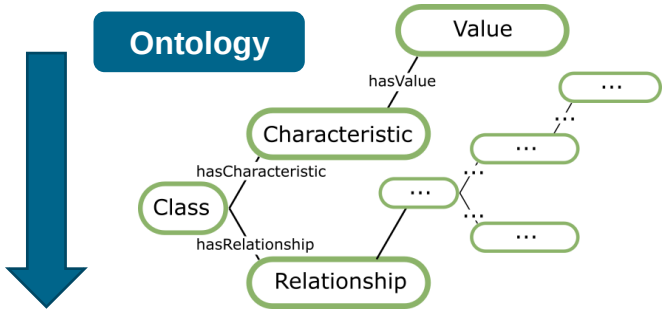
Select Samples to Compile a Training Dataset



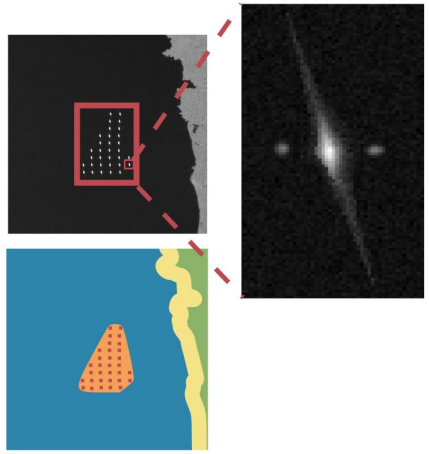

```
CREATE TABLE sea (
    id INT PRIMARY KEY,
    coverage INT);

CREATE TABLE windfarm (
    id INT PRIMARY KEY,
    size VARCHAR(255),
    shape VARCHAR(255),
    is_within INT,
    FOREIGN KEY (is_within) REFERENCES sea(id));
```

Semantic Enrichment – High Level Implementation



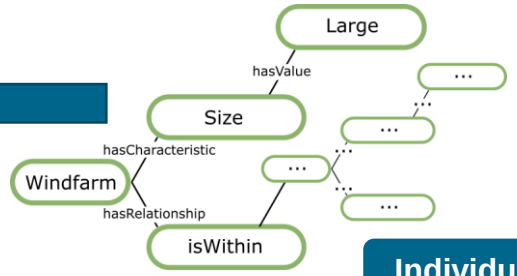
Create an individual with a constraint-based ontology instantiation



```
CREATE TABLE sea (
    id INT PRIMARY KEY,
    coverage INT);

CREATE TABLE windfarm (
    id INT PRIMARY KEY,
    size VARCHAR(255),
    shape VARCHAR(255),
    is_within INT,
    FOREIGN KEY (is_within) REFERENCES sea(id));
```

SQLite



Individual

Serialize to JSON → Serve LLM → Inference with structured output, Temperature fixed 0.2 → Scene Description

Pydantic

Bringing schema and sanity to your data

vLLM powered by terrabyte

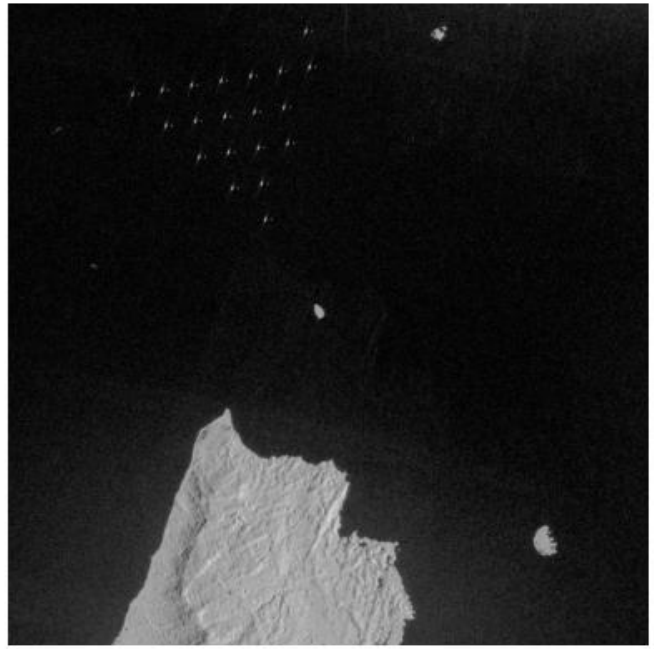
deepseek

deepseek-R1-distill-llama-70b
(DeepSeek-AI, Guo, D. et al. 2025)

Single Scene Description Example


```
{"scene": {
    "sea_coverage_in_percent": 90.0,
    "land_coverage_in_percent": 10.0,
    "scene_objects": {
    "sea": {
        "child_objects": {
            "offshore_wind_farm": {
                "parent_relationship": "is within",
                "turbine_count": 20,
                "shape": "polygon",
                "size": "medium",
                "spatial_appearence": {
                "presence": {
                    "northwest": true,
                    "north": true
                },
                "coverage": {
                    "northwest": 15,
                    "north": 27
                }
    "land": {
        "spatial_appearence": {…
```


UUID: 6c6b042f-63d6-47ba-9bd6-a95ed4449f22



'The scene primarily captures a vast expanse of sea, covering most of the area, with a small portion of land visible. The land is concentrated in the southern regions with a smaller patch near the center. The sea dominates the northern parts of the image. Within the northern section, there is a medium-sized offshore wind farm. This wind farm contains 20 turbines arranged in a polygon shape, adding a notable feature to the otherwise open sea.'

Embedding visualization of synthetic scene descriptions

Descriptions

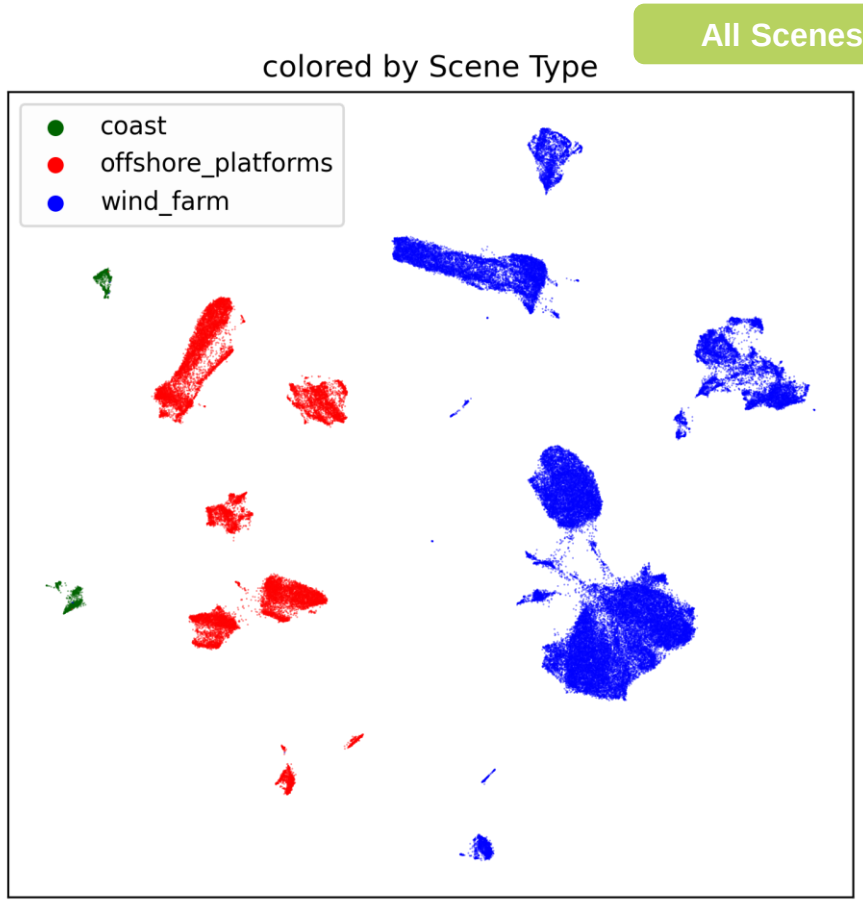
Embedding

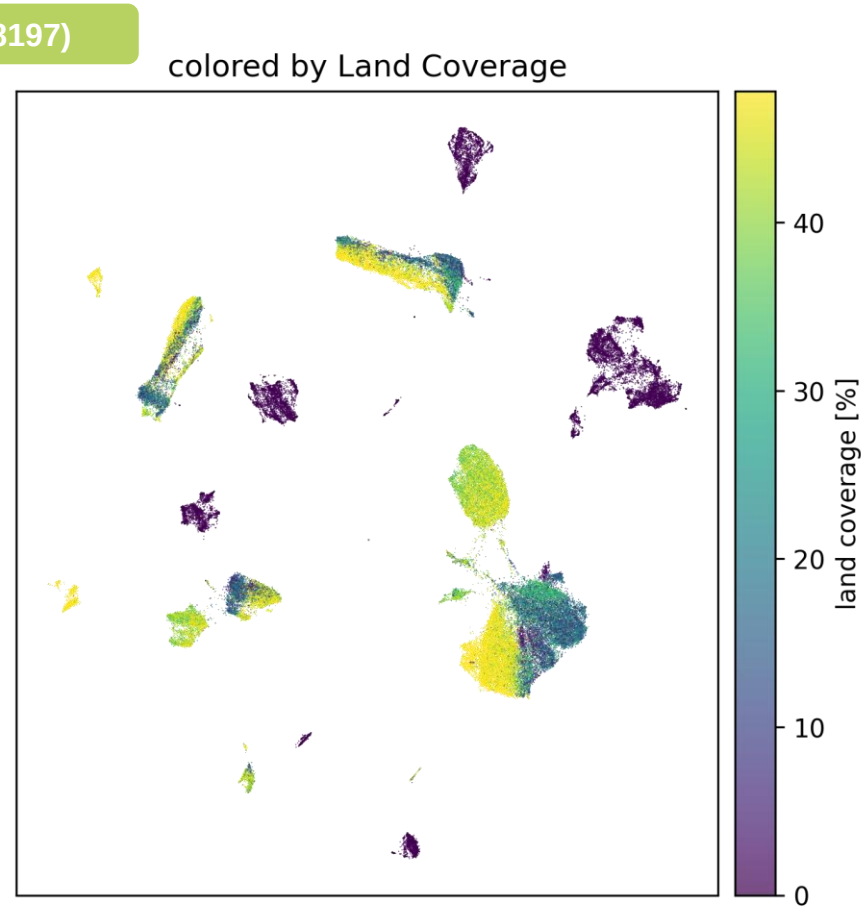
Jina-embed-v3
1024 Dimensions
(Sturua et al. 2024)

Dimensionality Reduction

UMAP
2 Dimensions
(McInnes et al. 2018)

All Scenes (n= 88197)

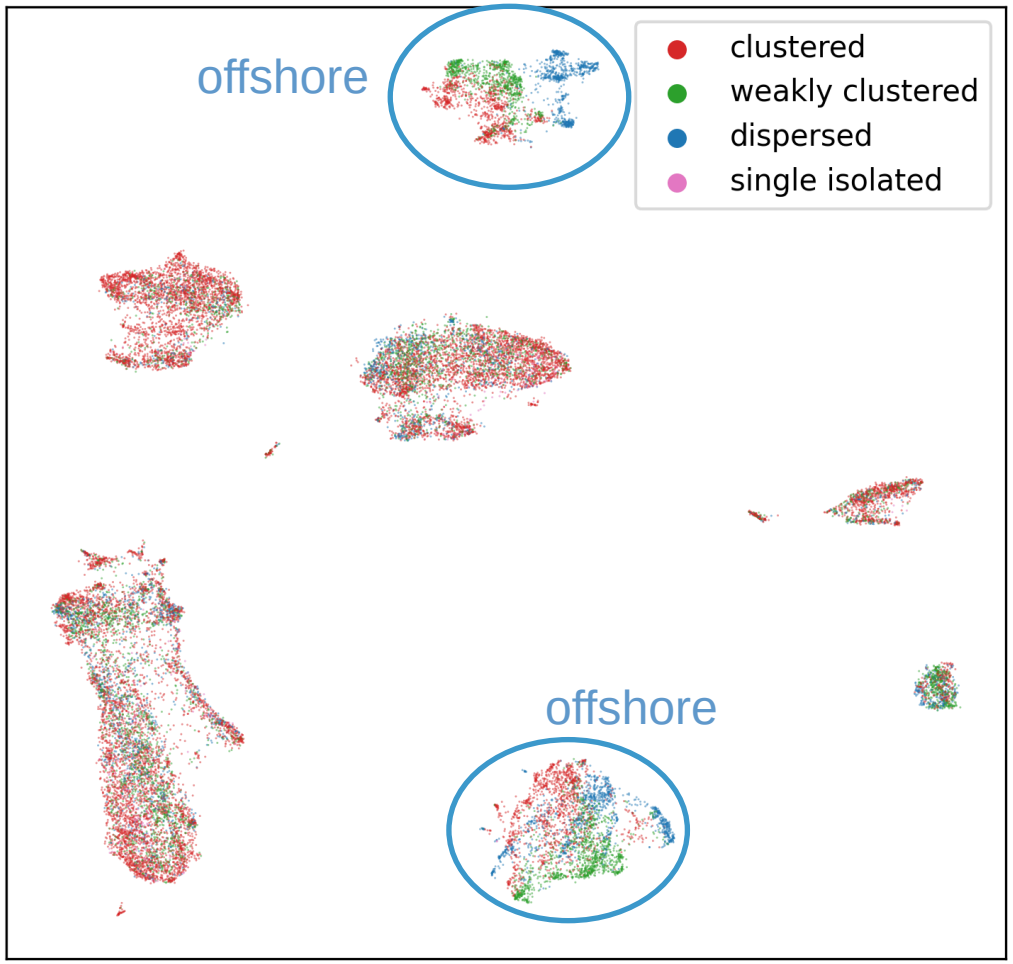




Embedding visualization of synthetic scene descriptions

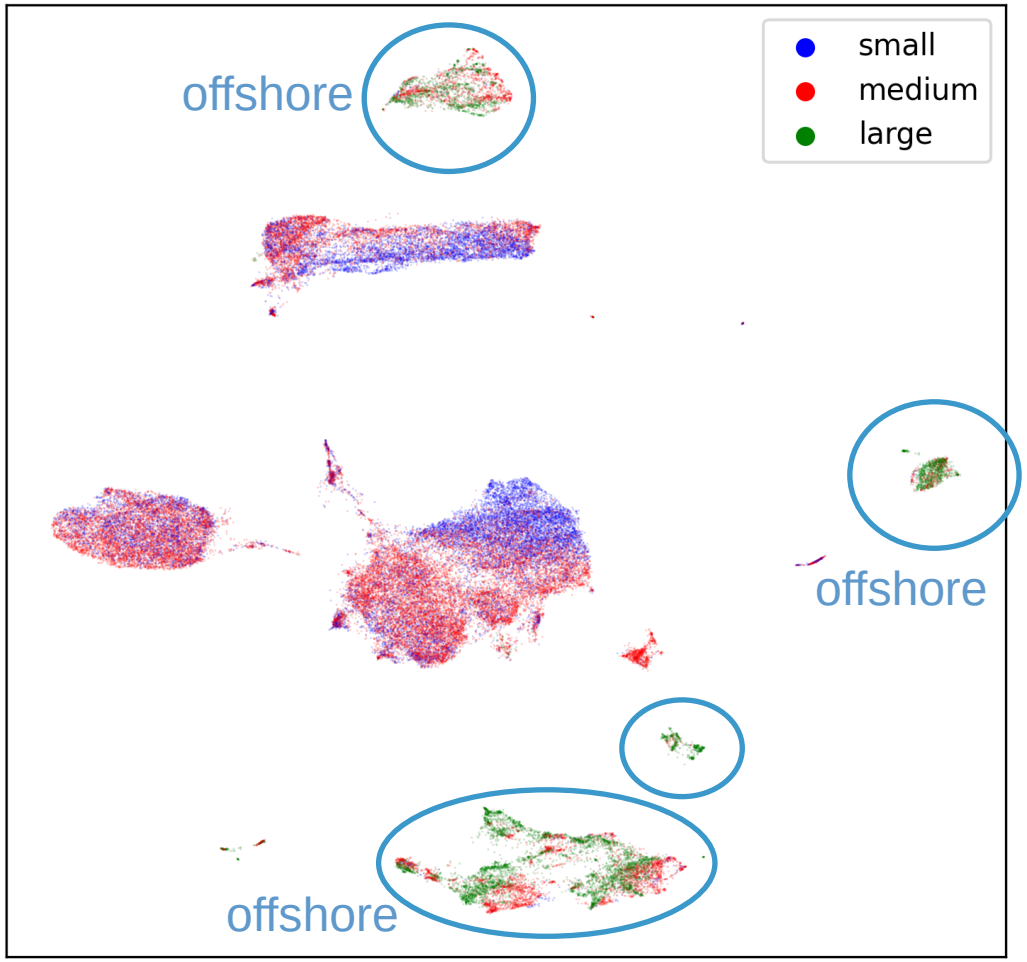
Only Offshore Platforms (n=24927)

colored by Clusteredness



Only Wind Farms (n= 61068)

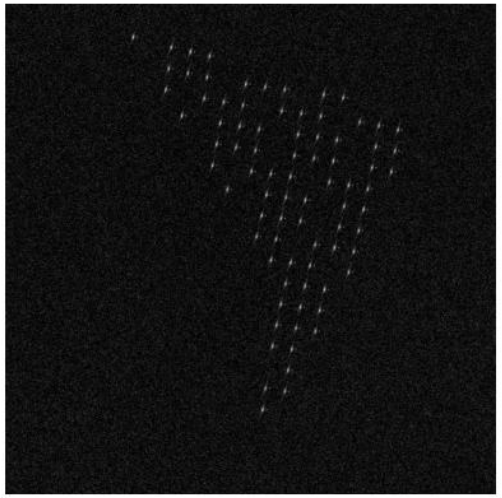
colored by Size



Synthetic Scene Retrieval with Semantic Queries

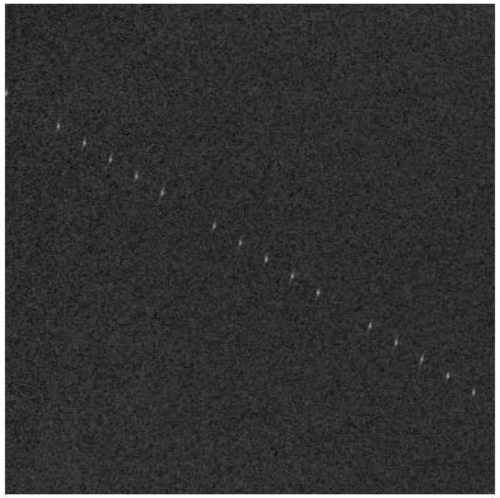
Marine energy infra-structure far offshore.

UUID: 9bacf7e5-991b-4133-ab1c-0a7c924481ae



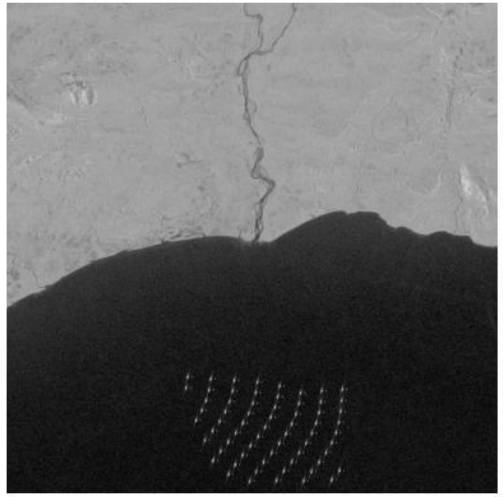
A large, linear offshore wind farm in open waters.

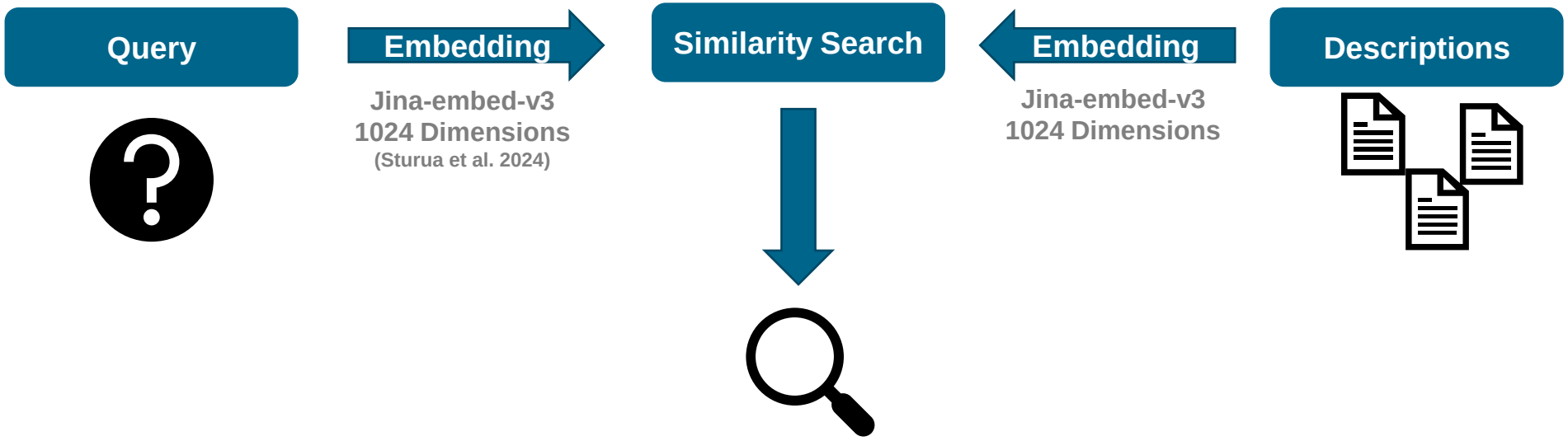
UUID: 6e8ebd39-84b0-4315-b5cb-0beac85dd793



A mixture of sea and land with coastal features and an offshore wind farm.

UUID: c4e6ac34-a628-4ed9-a86f-b8542edfeb04

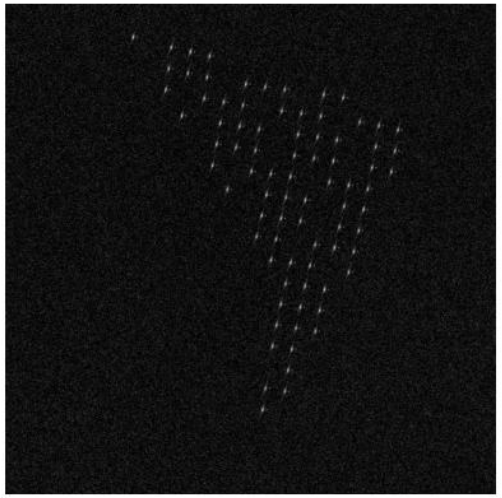




Synthetic Scene Retrieval with Semantic Queries

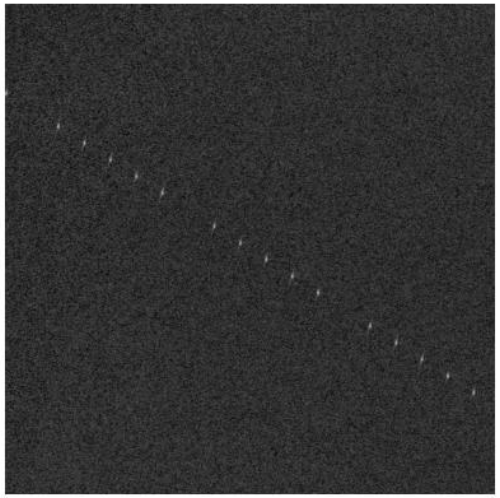
Marine energy infra-structure far offshore.

UUID: 9bacf7e5-991b-4133-ab1c-0a7c924481ae



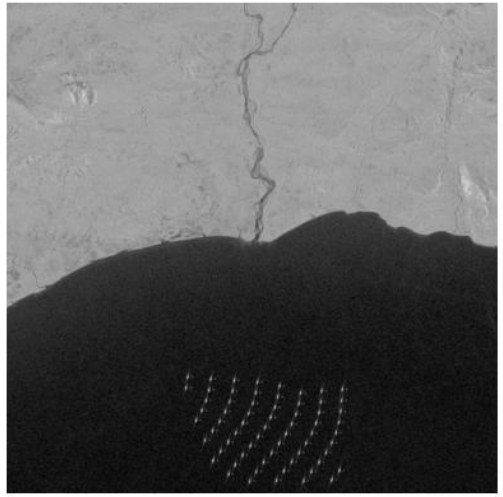
A large, linear offshore wind farm in open waters.

UUID: 6e8ebd39-84b0-4315-b5cb-0beac85dd793



A mixture of sea and land with coastal features and an offshore wind farm.

UUID: c4e6ac34-a628-4ed9-a86f-b8542edfeb04



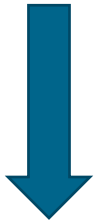
Query



Embedding

Jina-embed-v3
1024 Dimensions
(Sturua et al. 2024)

Similarity Search

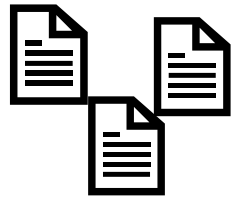




Embedding

Jina-embed-v3
1024 Dimensions

Descriptions



How many offshore turbines have been built in the North Sea last year?

Did the deployment of the Horn Sea offshore wind farm influence the migration rout of maritime mammals?

Quellen

DeepSeek-AI, Guo, D. et al., 2025: DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. arXiv:2501.12948 [cs.CL] https://doi.org/10.48550/arXiv.2501.12948

Hoeser, T., Feuerstein, S., Kuenzer, C., 2022. DeepOWT: A global offshore wind turbine data set derived with deep learning from Sentinel-1 data. Earth System Science Data 14, 4251-4270. doi: https://doi.org/10.5194/essd-14-4251-2022

Hoeser, T. and Kuenzer, C., 2022, SyntEO: Synthetic dataset generation for earth observation and deep learning – Demonstrated for offshore wind farm detection. ISPRS Journal of Photogrammetry and Remote Sensing 189, 163-184. https://doi.org/10.1016/j.isprsjprs.2022.04.029

McInnes et al., 2018. UMAP: Uniform Manifold Approximation and Projection. Journal of Open Source Software, 3(29), 861, https://doi.org/10.21105/joss.00861

Saba Sturua, Isabelle Mohr, Mohammad Kalim Akram, Michael Günther, Bo Wang, Markus Krimmel, Feng Wang, Georgios Mastrapas, Andreas Koukounas, Nan Wang, Han Xiao 2024. jina-embeddings-v3: Multilingual Embeddings With Task LoRA. arXiv:2409.10173 [cs.CL] https://doi.org/10.48550/arXiv.2409.10173.