

Introducing the VISQAM Dataset: Toward Automated Map Interpretation

Eftychia Koukouraki, Ajay Ajay, Ahmad Abubakar, Yomna Eid
GeoAI 2026, Universiteit Ghent



Maps and Information Communication

- Thematic maps are key tools in the Geosciences for information communication.
- Reading maps involves complex spatial reasoning: multiple visual variables simultaneously processed.



<https://www.canva.com/photos/MAGZKqrw8BM-close-up-of-a-map/>

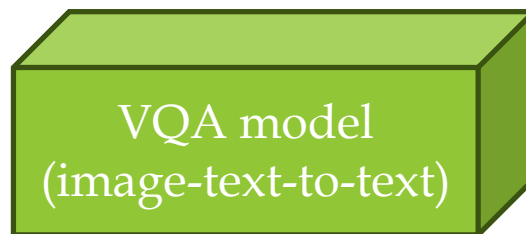
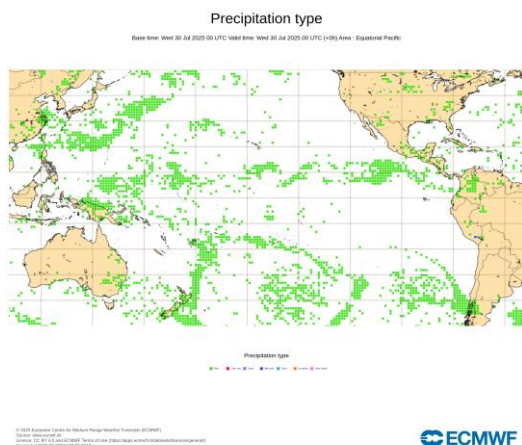
Visual Question-Answering offers new ways to access map content!



Visual Question-Answering (VQA)

VQA can help:

- Visually impaired individuals access content in figures
- Overcome language barriers
- With assessing content similarity, etc.

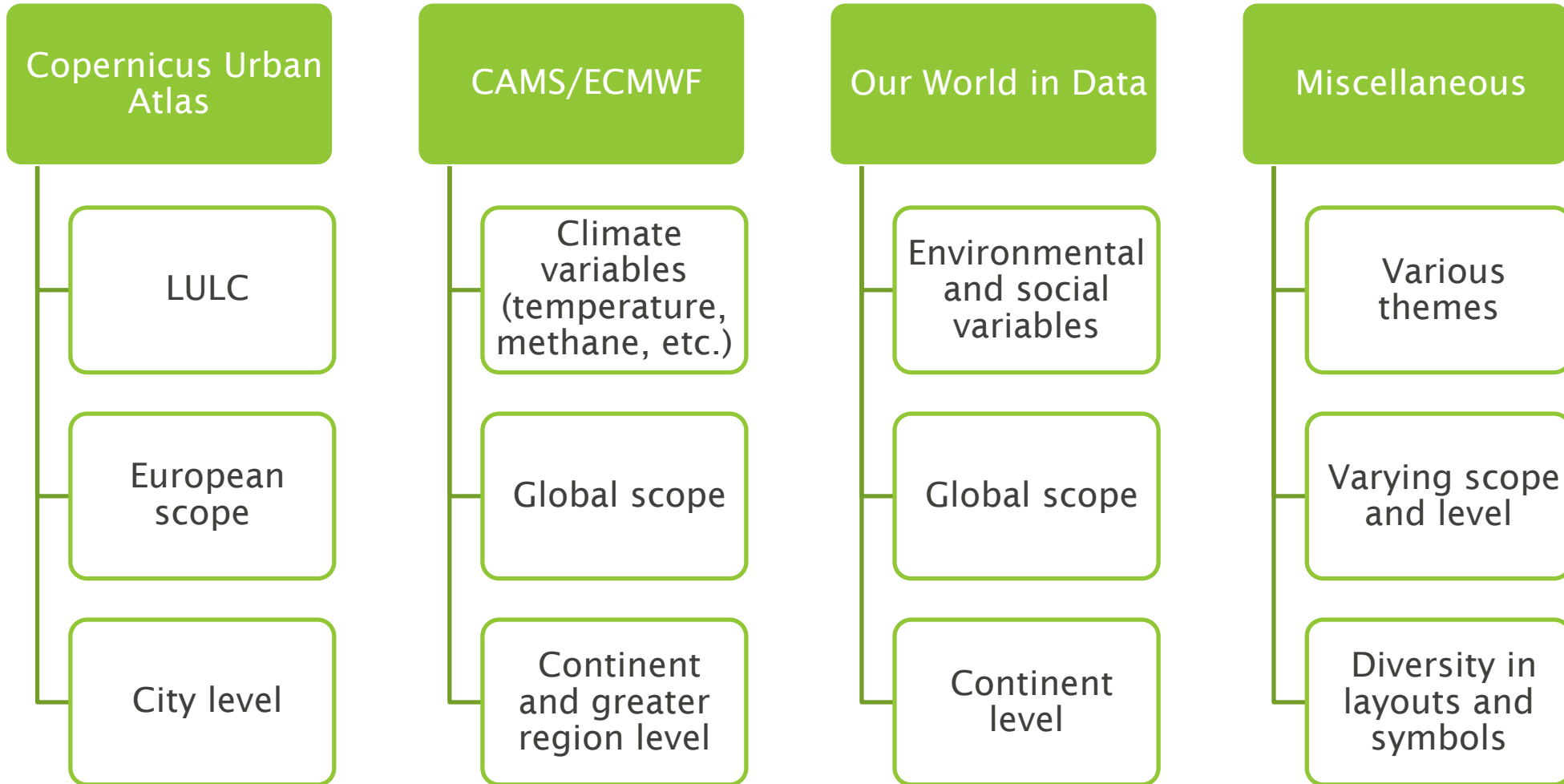


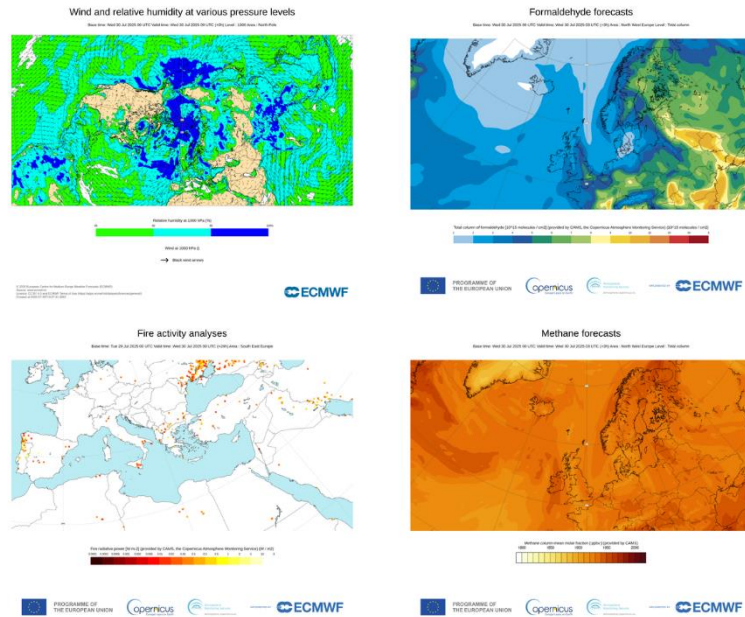
There is more rain in Mexico.

Is there more rain or snow in Mexico?

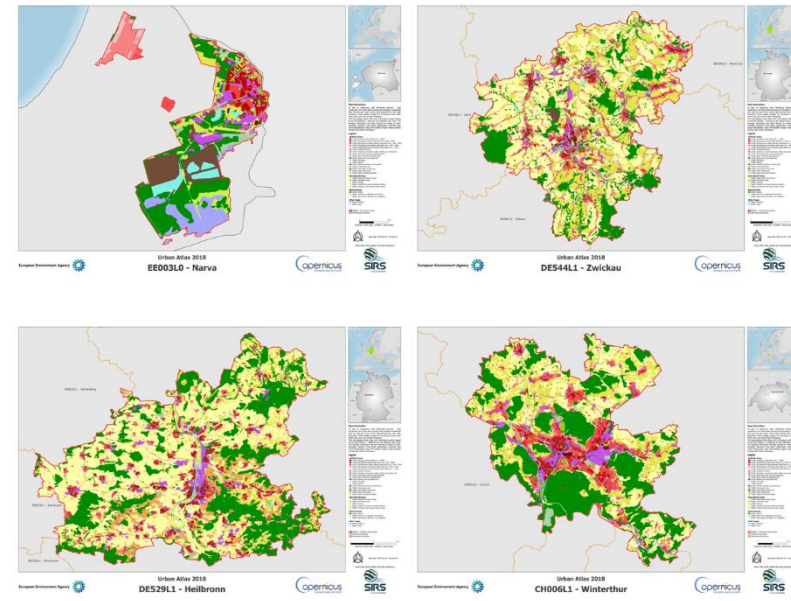


Dataset – Images Sources

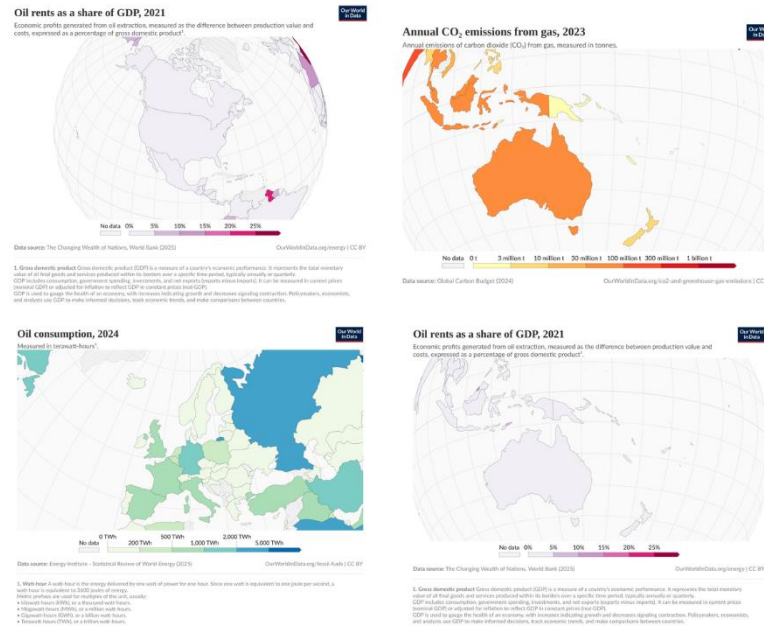




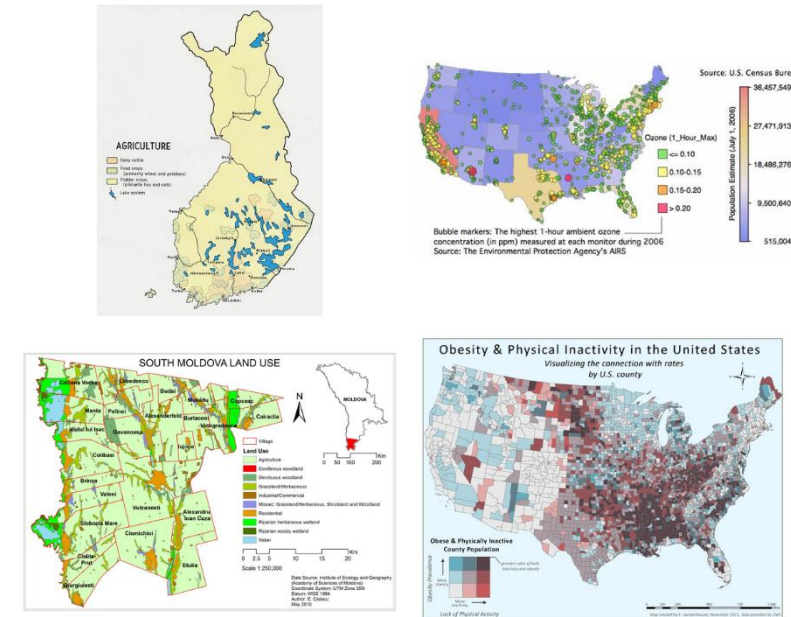
(A) Climate Variables



(B) Land Use Land Cover



(C) Our World In Data



(D) Ubimap-L



QA Annotations

- GQA Taxonomy¹: structural, semantic.
- 34 template questions in total.
- QA pairs per image: 3-5.
- 4594 verified QA pairs for 1200 images in total.

¹ <https://doi.org/10.48550/arXiv.1902.09506>



Question Taxonomy

Structural Types

Verify

Yes/no questions about map content

Query

Open-ended information retrieval

Choose

Select from given options

Logical

Inference and logical reasoning

Compare

Comparison between two features

Semantic Types

Object

Geographic entity existence

Attribute

Geographic features

Category

Feature classification

Relation

Geographic relationships between features

Global

General aspects of the map



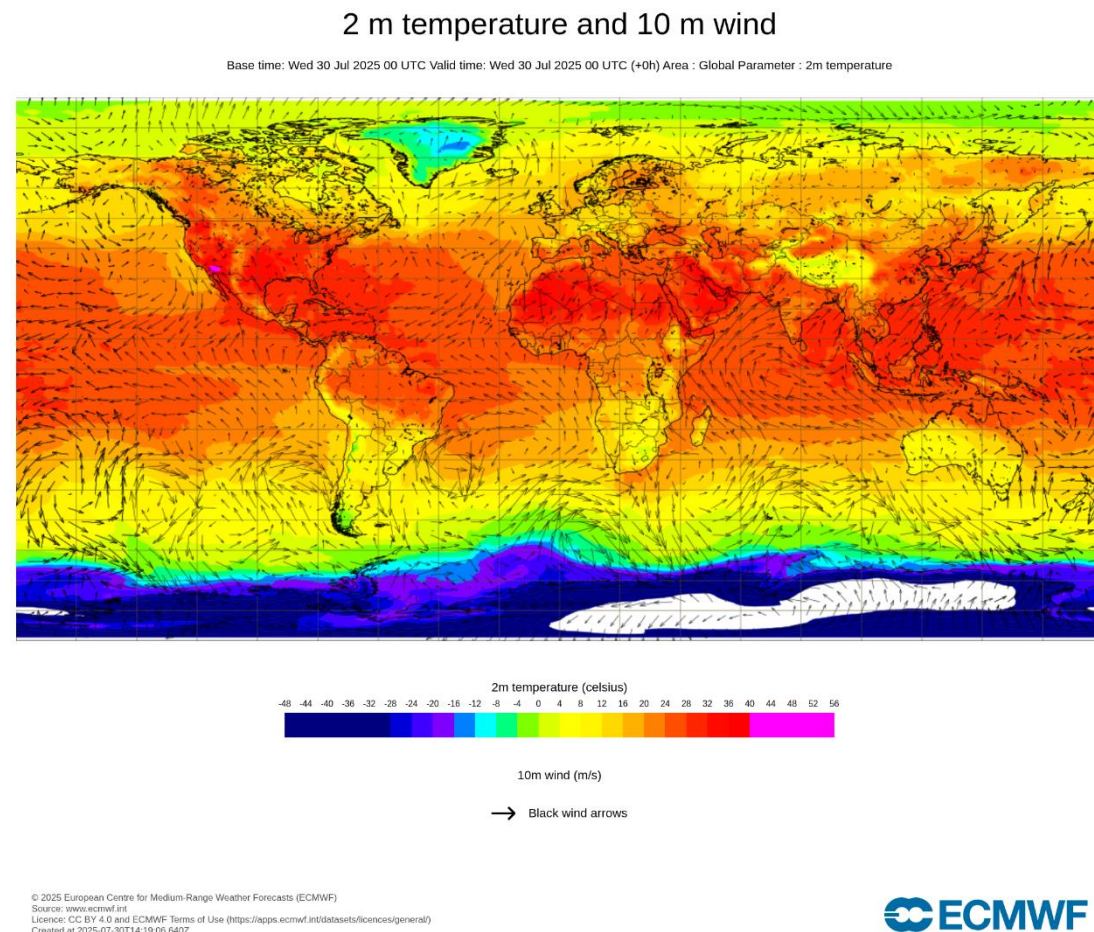
QA annotation example

- Which color represents 0-4 degrees Celsius of temperature?
- The green shade with the hexadecimal code #d9ff00.

structural: query, semantic: attribute

- Is the overall trend of temperature increasing across the map?
- Yes, it is increasing from the south and the north poles towards the equator.

structural: verify, semantic: global



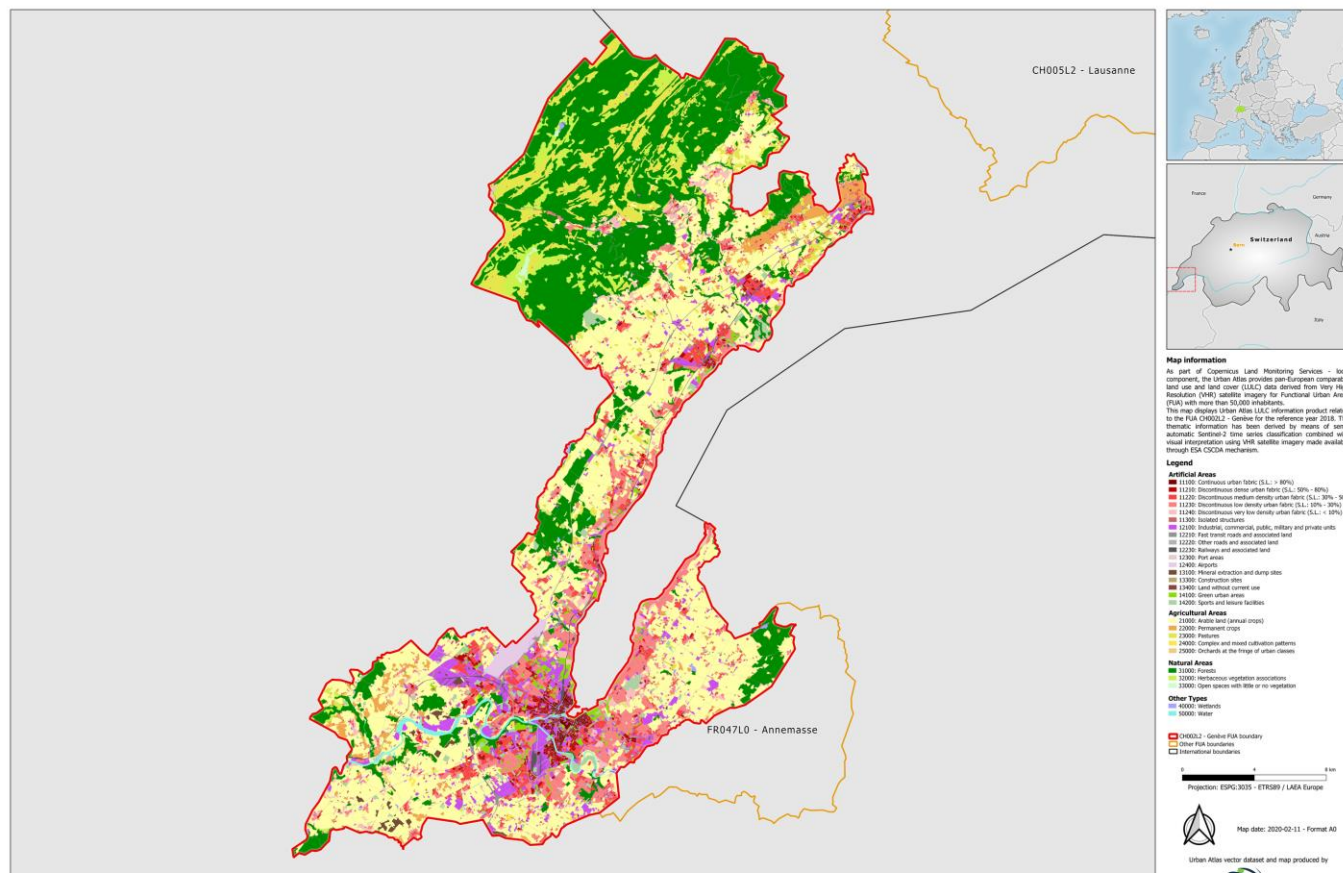


Baseline Model

- Reference for thematic map QA tasks.
- Fine-tuned [Qwen3-VL-2B-Instruct](#).
- Its big equivalent was the highest ranked open-weights model in the Vision Leaderboard in [arena.ai](#) as of Dec '25.
- LoRA fine-tuning on PALMA II HPC cluster (3 epochs).
- 95% train, 5% validation.



Example inference



USER: Where is the highest concentration of urban fabric?

VISQAM-FINETUNED: In the central area of the city.

BASE MODEL: The highest concentration of urban fabric is found in ****urban centers**** and ****dense metropolitan areas****. Urban fabric refers to the built environment —such as buildings, streets, parks, and other infrastructure — that constitutes the physical and social framework of a city.\n\nThe most densely populated and developed urban areas, such as:\n\n- ****New York City, New York**** (United States)\n- ****London, United Kingdom****\n- ****Tokyo, Japan****\n- ****Paris, France****\n- ****Cairo, Egypt****\n- ****Delhi, India****\n- ****São Paulo, Brazil****\n\n...typically exhibit the highest concentration of urban fabric.



Result Overview: Substantial Improvements

BERTScore-F1

Base: 0.43 → Fine-tuned: 0.72 (+29%)

Exact Match (EM)

Base: 0.00 → Fine-tuned: 0.24 (new!)

BERTScore-F1: Indicator of semantic similarity. Harmonic mean of BERTScore-Precision and BERTScore-Recall, calculated based on the cosine similarities of the contextual BERT embeddings.

EM: String match after lowercasing, punctuation and whitespace removal.



Contribution

- Multiple cartographic styles from real use-cases.
- Human quality annotations.
- Diverse projections.
- Geographically explicit questions (north/south/west/east rather than next to).
- Publicly available under CC BY 4.0.



Future Work

- Dataset expansion and retrained baselines.
- Develop architectures that explicitly use annotated legend bounding boxes to ground responses in map symbology.
- Different evaluation metrics for different structural question types.



VISQAM at a glance

1200
map images

4500+
QA pairs

Legend bbox
annotations

COCO format

Dataset



Code



Questions? Comments?



eftychia.koukouraki@uni-muenster.de

