# **AI AND DATACUBES:** TOWARDS A HAPPY MARRIAGE

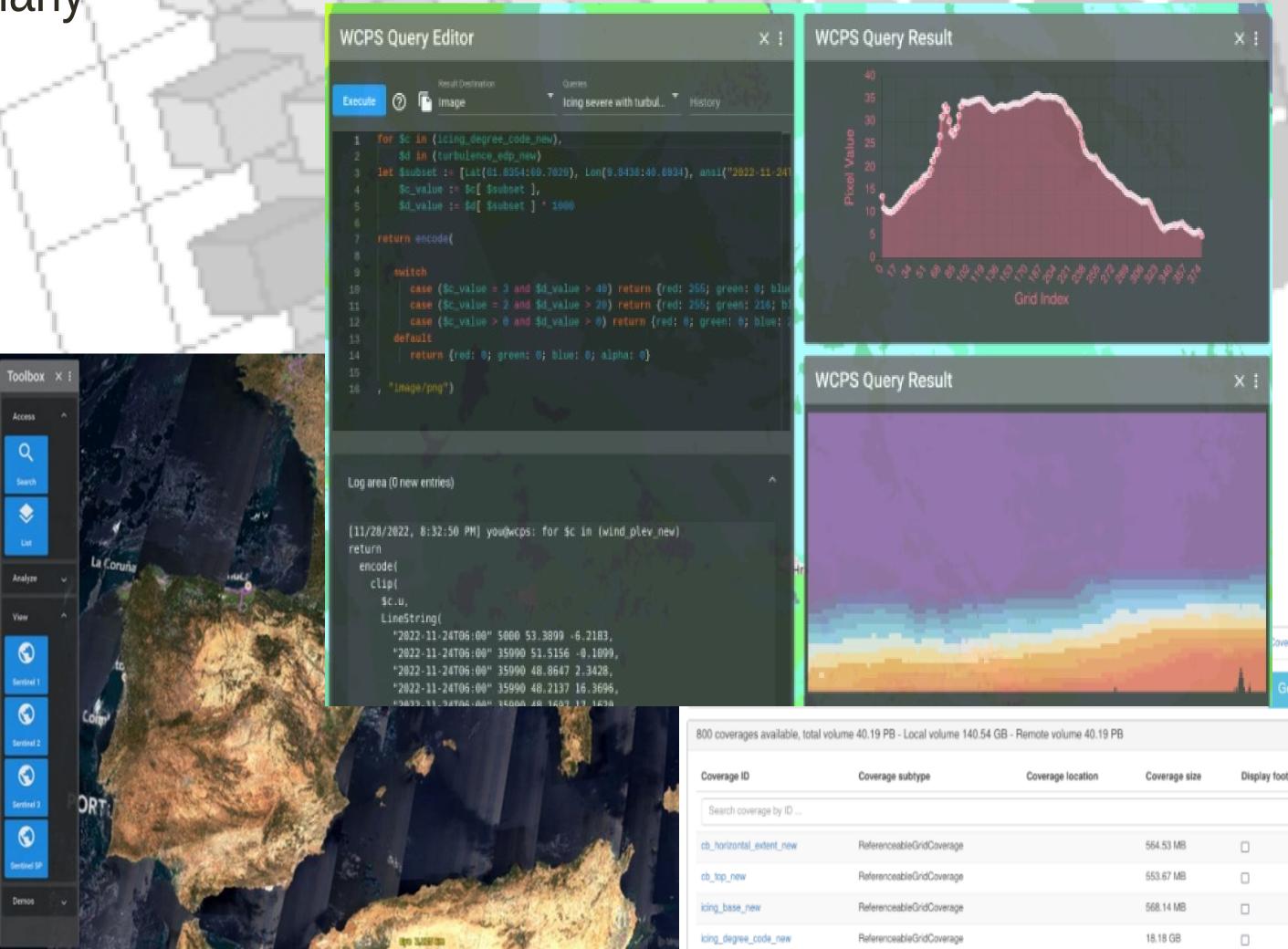
Peter Baumann, Dimitar Misev Constructor University, Bremen, Germany dmisev@constructor.university

# **1. Introduction**

- Datacubes are an accepted cornerstone towards Analysis-Ready Data, enabling high-level services on Big Earth Data (Figure 1)
- Al methods provide enhanced insight on vast amounts of data
- We integrate AI model application on datacubes with focus on
  - user-friendliness declarative datacube query language based on Tensor Algebra



### rasdaman raster data manager



- scalability transparently use PB of datacubes
- *flexibility* model-agnostic method via Python UDFs

(Web) client for whereUDF() returnUDF() ras	WCPS with UDF calls	<pre>def predict(B02, B03, B04, B08, B05, B06, B07, B11, B12,</pre>
selectUDF() from whereUDF()	Java UDF rasql with UDF calls C++/Python	<pre>model = torch.jit.load(model_path) question_tokens = tokenize_question(     question=question_text,     model='bert-base-uncased').unsqueeze(0) res = model([image, question_tokens]) answer_index = torch.argmax(res, dim=1).item() answer = get_answer_by_index(answer_index) return answer</pre>

Fig. 2: Architecture of the rasdaman datacube DBMS with UDFs

**3.** Applications

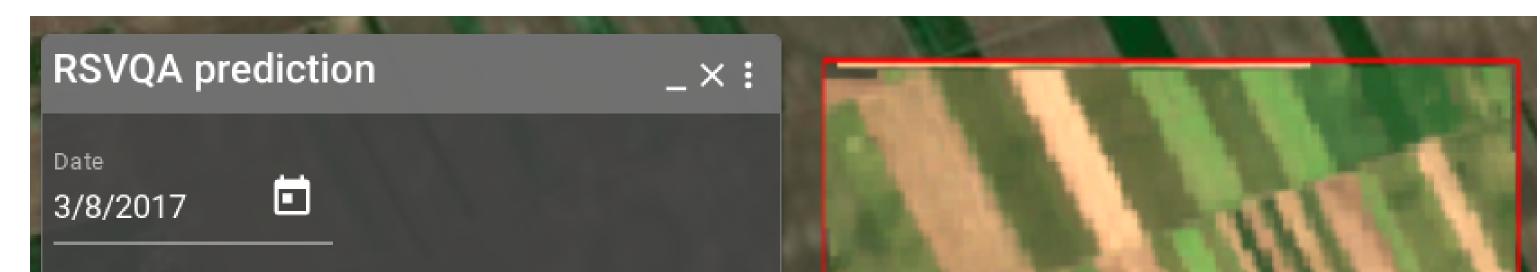
	Coverages
	Get Capabilities
ges available, total volume 40.19 PB - Local volume 140.54 GB - Remote volume 40.19 PB	
ges available, total volume 40.19 FB - Local volume 140.54 GB - Remote volume 40.19 FB	*

Coverage ID	Coverage subtype	Coverage location	Coverage size	Display footprints
Search coverage by ID				
cb_horizontal_extent_new	ReferenceableGridCoverage		564.53 MB	
cb_top_new	ReferenceableGridCoverage		553.67 MB	0
icing_base_new	ReferenceableGridCoverage		568.14 MB	
icing_degree_code_new	ReferenceableGridCoverage		18.18 GB	0

Fig. 1: High-level Datacube services: WMS/WMTS, WCPS, WCS

# **2. User-Defined Functions**

- Extend query language on datacubes with new functions
- High performance Java, C++ or Python extensions, directly embedded in the query processing pipeline
- WCPS (geo-datacubes) and rasql (domain-neutral arrays)



#### • Natural Language Processing (Fig. 3)

using Visual Question Answering for Remote Sensing Data (RSVQA)

for \$B02 in (S2\_L2A\_B02\_10m), \$B03 in (S2\_L2A\_B03\_10m), \$B04 in (S2\_L2A\_B04\_10m), \$B08 in (S2\_L2A\_B08\_10m), \$B05 in (S2\_L2A\_B05\_20m), \$B06 in (S2\_L2A\_B06\_20m), \$B07 in (S2\_L2A\_B07\_20m), \$B11 in (S2\_L2A\_B11\_20m), \$B12 in (S2\_L2A\_B12\_20m), \$B8A in (S2\_L2A\_B8A\_20m), \$VH in (S1\_GRDH\_IW\_DV\_VH), \$VV in (S1\_GRDH\_IW\_DV\_VV)

**let** subs2 := [ansi("2023-06-08"), Lat(53.0450:53.0570), Lon(8.7800:8.7920)], subs1 := [ansi("2023-06-04"), Lat(53.0450:53.0570), Lon(8.7800:8.7920)]

return rsvqa.predict(\$B02[subs2], \$B03[subs2], \$B04[subs2], \$B08[subs2], \$B05[subs2], \$B06[subs2], \$B07[subs2], \$B11[subs2], **\$B12[subs2], \$B8A[subs2], \$VH[subs1], \$VV[subs1],** "rsvqa\_trained\_model.pt", "Are agricultural areas or wetlands present?")

## • Crop Classification (Fig. 4)

for \$s2 in ( sentinel2\_flevopolder\_10m\_7x4bands )

**let** aoi := [E(677400:679900), N(5834260:5836760)]

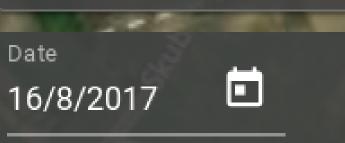
return encode(

fairicube.predictCropClass( \$s2[aoi], max(\$s2[aoi]) ), "png")

Prediction Ouestion \* What are all the land cover classes in the scene

#### $\Box$ **Run Prediction**

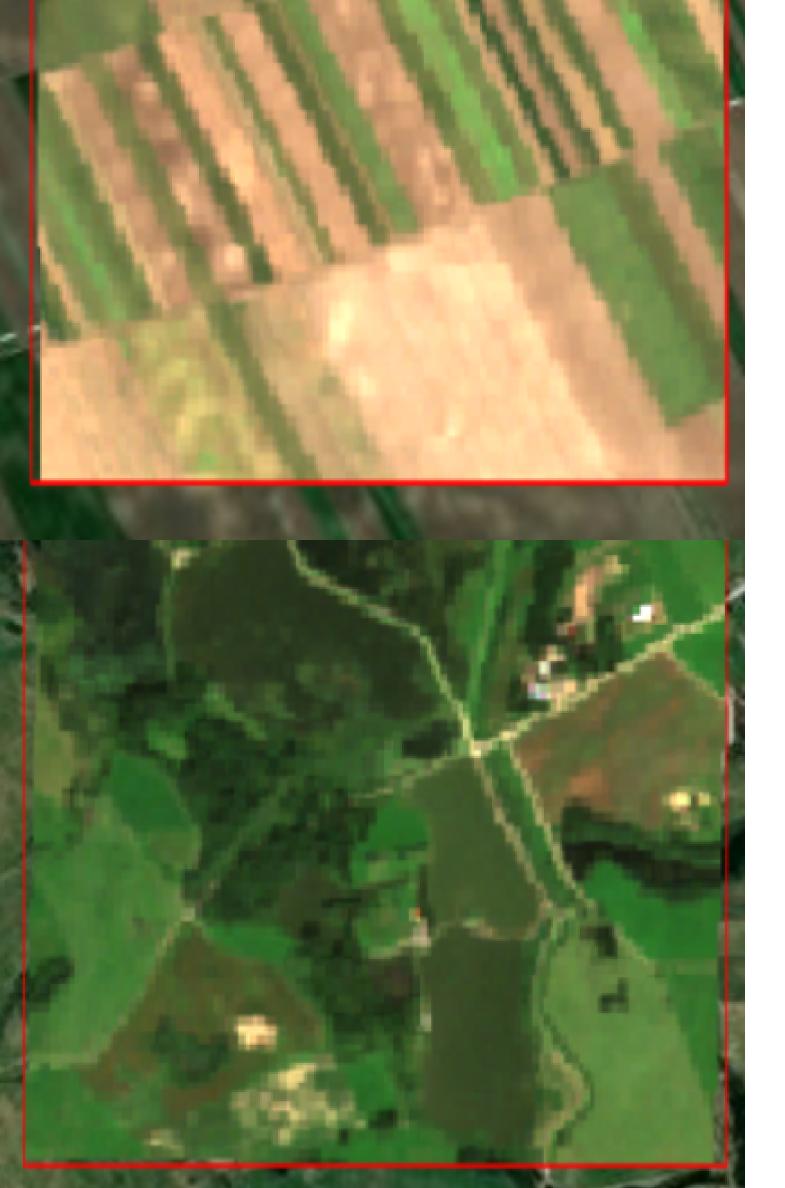
Q: What are all the land cover classes in the scene? A: agricultural areas, arable land and non-irrigated arable land

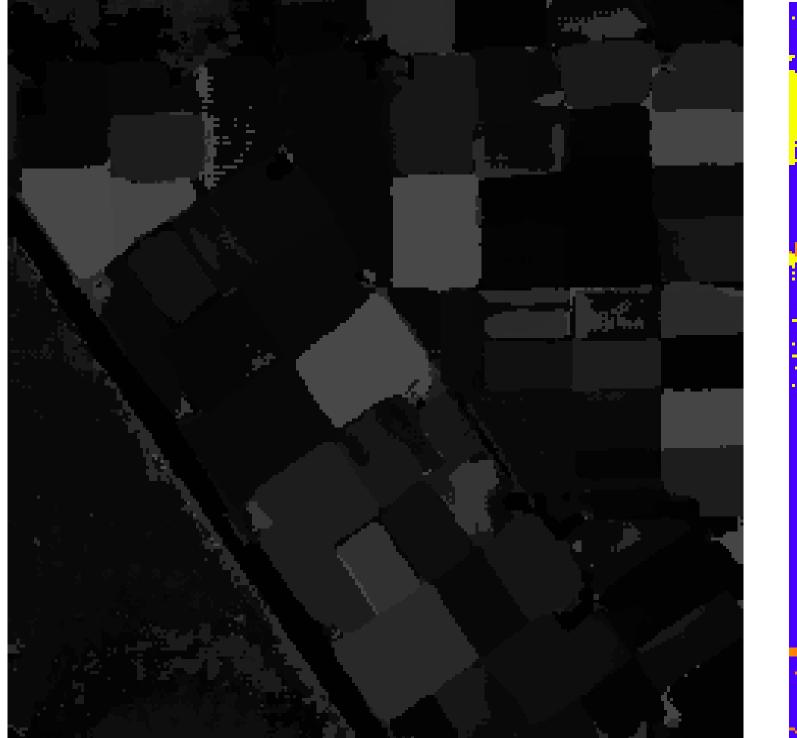


Prediction Question \* Are non-irrigated arable land or water courses r

**Run Prediction** D

Q: Are non-irrigated arable land or water courses present? A: yes





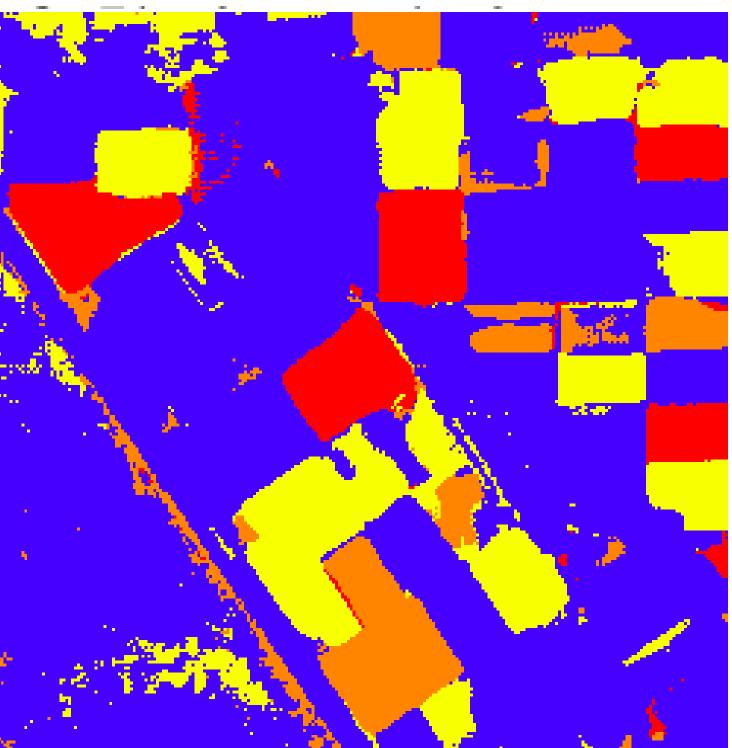


Fig. 4: Crop classification query results (FAIRiCUBE project [2])

#### Fig. 3: "Chatting" about remote sensing data (AI-Cube project [1])

# 4. Conclusion and Outlook

- Seamless integration of AI/ML prediction on datacubes through UDFs
- Open questions
- User-friendly model management
- Guidance on model applicability
- Integration of AI/ML training

- Live demo: https://aicube.rasdaman.com/rasdaman-dashboard/

#### Acknowledgments

1. AI-Cube has received funding from the German Federal Ministry of Food and Agriculture 2. FAIRiCUBE has received funding from EU Horizon 2020