

*Bachelor or Master Thesis*

# Efficient Pixelwise Access of Geospatial Raster Data

Are you looking for a Bachelor or Master thesis in advanced data science? Are you interested in data storage technologies? Do you want to get experience in the emerging field of geospatial databases and are you interested in defining new standards? Do you want to be part of the group of leading experts in this field? Then apply now!

We look forward to you joining us as a Bachelor or Master Thesis student (d/f/m) within the Big Geospatial Data Management Group at the Department for Aerospace and Geodesy. A Supervision in the School of Computation, Information and Technology is also possible.

**Location:** Ottobrunn/Munich/Remote

**Duration:** 3 to 6 months depending on your study program

**Your topic:**

Rising amounts of generated geospatial data, either trajectory-like tracking data, raster-like imagery, or vector-like mappings as in OpenStreetMap (OSM), grow the need for multimodal algorithmic investigation of this data. Data amounts thereby increase a lot in recent years with improved sensory technology and societal interest in creating and storing such data.

Anyway, an efficient cross domain access of data is currently computationally complex as data is often stored in compressed format and does not necessarily allow direct time- and location-based query of randomly distributed subsets of the data. Current methods still assume proximity of query objects. For example, compressed raster images in their standard format have to be loaded as a whole spatial patch into main memory to access single pixels for evaluation. Especially for complex queries like getting the image pixels close to independent trajectories in space and time, traditional approaches for indexing and query processing based on space partitioning methods have problems, as trajectories might span several branches of such trees. Therefore we want to investigate:

- How to minimize loading images into main memory by an efficient ordering of queries?
- How to improve indexing for randomized spatial queries for more efficient access?

This may include:

- Literature Review of existing multi-modal geospatial data access policies
- Development of heuristics to avoid reloading of single image patches
- Ordering pixel queries based on their probabilistic representations, e.g. in data structures like [GloBiMaps](#)
- Development of a new spatial index for random spatial queries

**Related Work:**

- Large Scale Analytics of Vector+Raster Big Spatial Data ([Eldawy et al, 2017](#))
- Scalable and queryable compressed storage structure for raster data ([Ladra et al, 2017](#))
- Efficient processing of raster and vector data ([Silva-Coira et al, 2020](#))
- GloBiMapsAI: An AI-Enhanced Probabilistic Data Structure for Global Raster Datasets ([Werner, 2021](#))

**Qualifications:**

- Interest in emerging field of Big (Geospatial) Data
- Advanced programming skills (preferably C / C++, alternatively Python, Rust)
- Experience with Databases and Data Management
- Interest and experience in literature-based work with a good scientific practice
- Enrolled full time student within Computer Science, Electrical Engineering, Geo Informatics or similar field of study
- Fluent English is mandatory; German would be an asset

**Applications via Mail with CV and transcript to:**

Advisor: M.Sc. Paul Walther  
Raum: 9377.01.113  
Telefon: 089/289-555 59  
Email: paul.walther@tum.de

