Seminar in the winter term 2023/2024:

Data Science in Urban Research and Remote Sensing

M.Sc. Elena Vollmer
M.Sc. Elias Naber
M.Sc. Mihir Rambhia
Dr.-Ing. Rebekka Volk

Kick-off-meeting:
27.10.2023, 10:00 - 12:00
Final meeting with presentations:
Mid-January: TBD
Submission deadline for the seminar theses:
February: TBD

Attendance is required on all dates.

Content of the seminar:
The seminar deals with topics related to the analysis of urban areas and structures. The topics are worked on individually with the possibility of receiving feedback during the working process from the supervisors.

The topics are designed as literature research projects or small coding challenges. The goal of each project is to summarize and clearly present current approaches and research trends for analyzing urban structures. The preparation of the seminar thesis serves as preparation for the bachelor/master thesis.

IMPORTANT: Please check your e-mails regularly after your application, especially from 13.10.2023, to respond to a seminar place offer as quickly as possible. If you do not respond in time, the seminar places will be passed on in the backlog procedure.
Topics:

1. Remote thermography and AI: What effect does data preprocessing have on performance? (Supervisor: Elena Vollmer)

2. Generating digital surface models (DSMs) from thermal images (Supervisor: Elena Vollmer)

3. What is “enough” data in the context of AI? Finding a saturation curve (Supervisor: Elena Vollmer)

4. How does heat spread underground? Identifying the influence of pipeline, ground, and surface composition on thermal conductivity (Supervisor: Elena Vollmer)

5. Graphical user interface (GUI) for the Namares model (Supervisor: Elias Naber)

6. Nature-based solutions and urban development (Supervisor: Elias Naber)

7. Data Science in Sustainable Urban Development (Supervisor: Elias Naber)

8. Assessment of methods used to quantify the quality of data in public datasets. (Supervisor: Mihir Rambhia)

9. Analysing the influence of various forms of green spaces on the urban heat island effect. (Supervisor: Mihir Rambhia)

---

**Topic 1:** Remote thermography and AI: What effect does data preprocessing have on performance? (Type: Review; supervisor Elena Vollmer)

In recent years, artificial intelligence has gained prominence for its uses in analysing data – in particular, images. While scientific publications show an ongoing competition in benchmarking ever more efficient models, the foundation for high performance resides with the quality of data used. The way in which the images are preprocessed and filtered may therefore have a significant impact on results. This study aims at generating a comprehensive overview of existing implementations and their effects on performance, with a focus on neural networks and thermal data in the urban context.

**Recommended knowledge and skills:** English, systematic and scientific way of working

**Topic 2:** Generating digital surface models (DSMs) from thermal images (Type: Review, Coding; supervisor Elena Vollmer)

Digital surface models (DSMs, e.g. 3D height maps) have many uses in visualisation, modeling, and as supplementary data for analyses in the urban context. Various software exists which can generate these kinds of models by mosaicking high-resolution RGB (red green blue) images. However, their ability of handling aerial, thermal imagery is less chartered. After a preliminary survey of existing analyses, this study focuses on systematically implementing a choice of open-source software (e.g. WebODM, NicNac, etc.) and assessing their proficiency in generating DSMs using select, real-world datasets.

**Recommended knowledge and skills:** General / Python coding experience, systematic and scientific way of working, English
Topic 3: What is “enough” data in the context of AI? Finding a saturation curve (Type: Coding; supervisor Elena Vollmer)

Artificial intelligence models are defined by the data they have been trained on, and it is often said that the amount of accessible, labeled information significantly impacts a model’s performance. However, the annotation process is often cumbersome and time-consuming in real-world applications. The question of when “enough” data has been generated to achieve an acceptable performance inevitably arises. This study will try to answer that question in the context of multi-class semantic image segmentation through systematic experimentation with an existing convolutional neural network and a real-world dataset.

Recommended knowledge and skills: Python coding experience, systematic and scientific way of working, English

Topic 4: How does heat spread underground? Identifying the influence of pipeline, ground, and surface composition on thermal conductivity (Type: Review; supervisor Elena Vollmer)

District heating networks transport heat to end energy users via underground pipelines. When a leakage occurs through damage or fatigue, hot water is released into the environment and heat spreads to the surface, where it can be detected as a thermal hotspot. In what specific way the heat propagates depends on various factors, such as pipeline, ground, and surface type. Understanding the influencing factors and thus how the process can be simulated is key to aiding with the timely handling of leakages. This study focuses on creating a comprehensive overview of ground, surface, and pipeline construction types (particularly in Germany), distinguishing characteristics influencing thermal conductivity, and identifying methods for simulating the process.

Recommended knowledge and skills: German, English, systematic and scientific way of working

Topic 5: Graphical user interface (GUI) for the Namares model (Type: Coding; supervisor Elias Naber)

Many scientific models require detailed knowledge for a comprehensive operation. In the research project NaMaRes and namares 2.0, data and calculation models were developed to support urban development digitally. The very extensive analysis possibilities of the model outputs are challenging for users. Reducing complexity and designing an open, appealing, and practical user interface are the core challenges of this topic. Initially, a short overview of GUI structure and functionalities of related software products is generated. Then, a GUI for the Namares projects is designed. The expected results are a mock-up and a R Shiny (or similar) draft. The thesis documents the process and presents the findings of the overview.

Recommended knowledge and skills: Afinity to design, R or SQL coding experience, Systematic working

Topic 6: Nature-based solutions and urban development (Type: Review; supervisor Elias Naber)

Nature-based solutions (NBS, e.g., greening or ponds) can provide significant ecosystem services. In this topic you will research which and how NBS are modelled in academic assessment tools. The body of literature is extensive. Therefore, the thesis will systematically overview the respective NBS and their integration into assessment frameworks and models. The overview provides information on technical, economic, and ecological implementation.

Recommended knowledge and skills: English, Systematic working, Scientific working
Topic 7: Data Science in Sustainable Urban Development (Type: Review; supervisor Elias Naber)

Cities and municipalities collect and maintain incredible data treasures. Many of them still need to be recovered and processed. In this topic, you will research the innovative ways in which cities use data to mitigate climate and sustainability challenges. The focus is on spatially referenced data and how challenges and management concerning the local resources of land use, ecosystems, water, energy, and waste are tackled.

Recommended knowledge and skills: English, Systematic working, Scientific working

Topic 8: Assessment of methods used to quantify the quality of data in public datasets (Type: Review and Analysis; supervisor Mihir Rambhia)

Several cities have published a range of datasets as part of open data initiatives. When these datasets are of high quality and available in sufficient quantity, they can serve many urban applications and facilitate informed decision-making. However, to have a correct application, it is crucial first to establish a quantitative means of assessing whether a published dataset meets the criteria of being of good quality and usability. This subject comprises a systematic review of existing studies that analyse dataset quality. The study aims to address the following research questions:

- What is good data quality and quantity in the context of public datasets?
- Which are the available methodologies to quantify the data quality?
- How do different methodologies compare, and what are their strengths and weaknesses?
- Use the open tree-inventory datasets and quantify them on various quality parameters (a case study).

Recommended knowledge and skills: Systematic and scientific literature review, Qualitative and quantitative analysis, English

Topic 9: Analysing the influence of various forms of green spaces on the urban heat island effect. (Type: Review and Analysis; supervisor Mihir Rambhia)

The presence of green spaces in cities, regardless of their type or extent, offers numerous benefits to the city and its citizens. Green space interventions have been employed to enhance environmental conditions, protect and enhance biodiversity, promote active lifestyles, increase social interaction and exchange, and provide a healthy urban environment conducive to good physical and mental well-being. There is ample scientific literature recognizing the importance of green spaces in mitigating the urban heat island effect. However, the varied impact of different types of green spaces, such as street trees, parks, and forests, is not clearly understood. This topic constitutes a systematic literature review of studies analysing these specific benefits. The study aims to address the following research questions:

- What parameters determine the impact of green spaces in mitigating the urban heat island effect?
- What methodologies are available for quantifying these benefits?
- How do different methodologies compare, and what are their strengths and weaknesses?
- How has the heat island effect and the presence of green spaces changed over the last two decades, and what correlations can be identified (a case study).

Recommended knowledge and skills: Systematic and scientific literature review, Spatial data analytics, English