

ID04:

Large-Scale Foundation Models for 3D Radiology

Data Science supervisor: Prof. Dr. Klaus Maier-Hein

Life Science supervisor: Dr. Hannes Kenngott

Large-scale foundation models have transformed natural-image analysis, but multimodal pretraining for 3D radiology remains limited by dataset scale and the difficulty of learning from full volumetric imaging data. This project develops multimodal foundation models for 3D radiology by extending The Human Radiome Project (THRP), a corpus of 2.1 million CT, MRI, and PET volumes, from image-only self-supervised pretraining toward joint learning from imaging, radiology reports, and structured clinical metadata.

The central scientific question is how multimodal supervision can be integrated effectively into large-scale 3D radiology pretraining. The project investigates scalable vision-language pretraining strategies for volumetric imaging, combining contrastive image-text alignment with reconstruction-based objectives and systematically studying supervision granularity from finding-level entities to full reports and volume-level representations.

The resulting models are evaluated across the THRP benchmark on multi-finding classification, retrieval, report generation, visual question answering, regression, segmentation, and zero-shot transfer, with particular focus on robustness under domain shift and low-data settings. Unlike prior work in this area, the required imaging corpus, paired report data, and pretraining infrastructure are already available within the participating groups.

By combining large-scale 3D imaging data with multimodal clinical supervision, the project aims to establish a scalable methodological basis for multimodal radiology foundation models and clinically relevant generative AI applications in medical imaging.



[Prof. Dr. Klaus Maier-Hein](#)



[Dr. Hannes Kenngott](#)

Special requirements for applicants:

- Master's degree in Computer Science, Data Science, or a related field
- Strong programming skills in Python
- Solid background in deep learning, ideally with experience in self-supervised learning, contrastive training, or vision-language models
- Experience with medical imaging data or large-scale model training is beneficial
- Excellent English communication skills, German is beneficial
- Ability to work independently and as part of a team
- Strong motivation and commitment to research