

# Internship Report

**Intern Name:** Kshitij Raj Sharma

**Organization:** Christian Doppler Laboratory Geohum, Paris Lodron University (in collaboration with Doctors Without Borders – MSF, Vienna)

**Supervisor:** Dr. Lorenz Wendt

**Internship Duration:** June 30, 2025 – August 2025 (6 weeks full-time, equivalent to 3 months part-time)

## 1. Introduction

During my internship, I was involved in research for adopting standards while developing & serving AI models, infrastructure setup, and integration with MSF environments. My work focused on standardization of AI models, development of pipelines for geospatial analysis, and adapting workflows to different platforms. The main goal was to create models that can be interoperable and usable across different infrastructures, including Windows and ArcGIS environments.

## 2. Objective :

Research how to package and develop models in a way they can be shared, reused and explained in a better way !

## 3. Weekly Activities and Learnings

### Week 1:

- Onboarding , Supported the team on AGIT and attended my first academic conference (AGIT 2025, <https://agit.at/en/>).
- Met a variety of people from academia and industry, which helped me understand different approaches in the domain.
- Conducted preliminary research on AI standardization and performed a feasibility study to explore how AI model workflows could be structured.

### Week 2:

- Began research on which tools and platforms to use, and consulted with MSF and other experts about available models.
- Developed a pilot refugee camp building detection model using existing models and sample standards.
- Documented the training process and metadata to ensure the model could be tracked and shared efficiently.

### Week 3:

- Focused on cloud infrastructure: set up a server and S3 bucket to store model artifacts.
- Created a workflow to generate stack MLM items readable by others.

- Exported the model into different file formats and hosted the full workflow on my DigitalOcean server using a student grant.
- Learned how to connect the infrastructure and metadata effectively.

#### **Week 4:**

- Familiarized myself with MSF infrastructure, including Windows and ArcGIS, which were new for me.
- Developed a script to convert stack MLM items into ArcGIS-compatible formats (.emd and .blbk).
- Tested the refugee camp detection model within MSF's ArcGIS environment.
- Attended the FOSS4G Europe conference (<https://2025.europe.foss4g.org/>) received feedback from experts, including Kubernetes specialists.

#### **Week 5:**

- Worked on-site at MSF Vienna, collaborating with Leslie, Jan, and Ann.
- Adapted my workflow to MSF's solution-driven environment, focusing on practical usability.
- Managed Windows-based Python environments, set up Conda/Mamba, and fixed broken environments for ArcGIS compatibility.
- Learned about the limitations of ArcGIS and Windows servers, particularly the inability to deploy Docker, which affected parts of my model standardization plans.

#### **Week 6:**

- Developed a full pipeline for the Google Buildings model tailored to MSF infrastructure.
- Built preprocessing scripts to convert drone images into model-ready inputs.
- Implemented scripts to run the model, convert outputs into geospatial formats, and manage efficient tiling of drone images.
- Developed a UI for MSF to use and delivered a stable version that can be run using a few commands within MSF's environment.
- Explored packaging models into a .exe format for easier deployment, though this remains untested for now.

#### **Week 7:**

- Prepared reports and presentation materials summarizing my internship work and outcomes.

### **4. Key Learnings and Reflections**

- Practical environments, especially Windows and ArcGIS, require workflows to be simple, efficient, and compatible with existing infrastructure.

- Model standardization is important, but deployment constraints like Docker availability can limit full-scale implementation.
- Adaptation is crucial: I had to adjust my approach to meet MSF's infrastructure and operational needs.
- Working directly with MSF provided insights into real-world application of AI in humanitarian contexts.
- I learned a lot about metadata handling, ArcGIS file formats, and integrating models into diverse workflows.

## 5. Outcome

- Developed a sample refugee camp detection model with standard-compliant metadata.
- Implemented a full pipeline for Google Buildings model processing from drone images to geospatial outputs.
- Delivered a user interface and a functional deployment workflow for MSF environments.
- Contributed scripts to convert stack MLM items into ArcGIS-compatible formats.
- Gained experience with cloud infrastructure, Python environments, and practical deployment challenges. Specially with kubernetes
- It was a good experience while developing the complete pipeline for google buildings and package it in a container that can be executed directly ! Good experiment to package the model completely with its dependencies , scripts , versioning !

## 6. References and Links

- GitHub repository for model standardization: <https://github.com/kshitijrajsharma/opengeoaimodelshub>
- Blueprint for the model hub: <https://kshitijrajsharma.github.io/opengeoaimodelshub/#vision>
- Standarization detailed report : <https://github.com/kshitijrajsharma/opengeoaimodelshub/blob/master/MLOps%20for%20Humanitarian%20Geospatial%20Models.pdf>
- Tiler for drone images to input in ML models : <https://github.com/kshitijrajsharma/cog2tiles>
- Hosted mlflow : <https://mlflow.krschap.tech/>
- Google Buildings Pipeline : [https://github.com/kshitijrajsharma/google\\_buildings\\_msf](https://github.com/kshitijrajsharma/google_buildings_msf)
- AGIT Conference: <https://agit.at/en/>
- FOSS4G Europe: <https://2025.europe.foss4g.org/>
- GeoHum: <https://geohum.eu/>

## 7. Conclusion

This internship was a significant learning experience. It allowed me to combine research, development, and real-world deployment. I learned the importance of adapting AI workflows to practical environments, handling infrastructure constraints, and developing models that

can be easily shared and used. The work I completed lays the foundation for future research and deployment of standardized AI models in geospatial and humanitarian contexts.