

**Title:** Auto-ML with synthetic training data and hardware in the Loop

**Team:** ATB: Cornelia Weltzien  
UOS: Mario Pormann  
HS: Stefan Stiene

**Abstract:**

Auto-ML is a highly researched field. The goal is to automatically determine an optimal model and its parameters for given data. The goal of this thesis is to extend this principle by not considering a static data set, but a dynamically extensible data set based on synthetic training data. Here, the work can make use of existing tools to generate this data. Furthermore, real edge devices will be integrated into the optimization loop on which the AI models will run and be evaluated in an automated way. This will provide the Auto-ML optimization problem with additional degrees of freedom and constraints. This scientific work is to be realized in an agricultural use case. Due to the frequent lack of connectivity, agriculture relies on edge applications to run AI procedures on agricultural machinery and agricultural robots. Furthermore, for many applications it is difficult to generate a sufficient amount of real data, so that, e.g., the detection of crops and weeds has shown that synthetic training data is a promising alternative here. The work is carried out in close cooperation with agricultural technology companies of the Agrotech-Valley-Forum, so that an evaluation of the results on real machines in real agricultural processes is possible.

**Desired skills of the applicant:**

- Scientific Background in AI and Auto-ML
- Strong programming skills in python
- Experiences in the application field agriculture is also a strong plus

**References:**

Naeem Iqbal , Justus Bracke , Anton Elmiger , Hunaid Hameed and Kai von Szadkowski: „Evaluating synthetic vs. real data generation for AI-based selective weeding” in C. Hoffmann et al.: Resiliente Agri-Food-Systeme, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2023