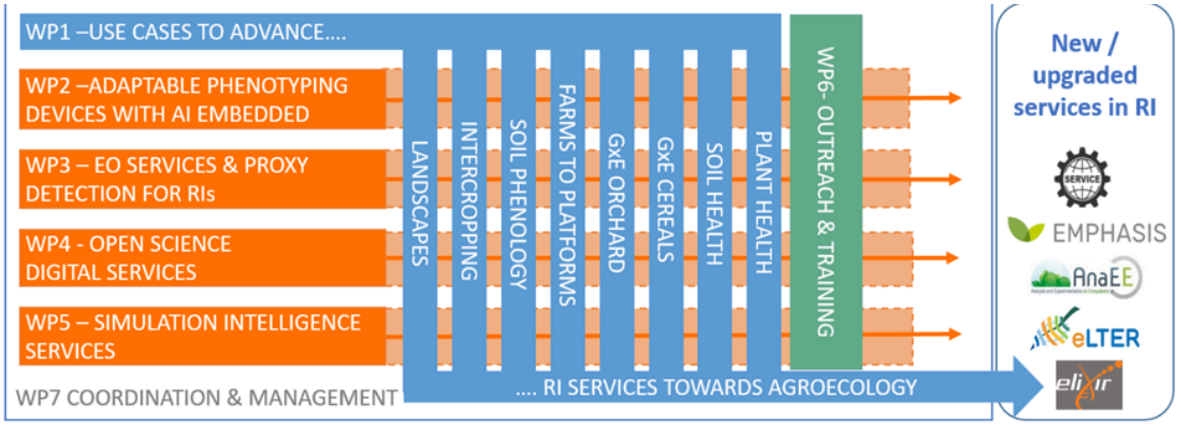
Phd offer funded on PHENET EU project

Embedded AI on connected Stick for plant phenotyping

Context: Europe urgently needs to find pathways towards agroecological transition of agroecosystems in support to food security, climate change resilience, biodiversity and soil carbon stocks restoration. In PHENET (https://www.phenet.eu/en/about-phenet), the European Research Infrastructures (RI) on plant phenotyping (EMPHASIS), ecosystems experimentation (AnaEE), long-term observation (eLTER) and data management and bioinformatics (ELIXIR) will join their forces to co-develop, with a diversity of innovative companies, new tools and methods – meant to contribute to new RI services - for the identification of future-proofed combinations of species, genotypes and management practices in front of the most likely climatic scenarios across Europe.



The Phd stands in the framework of the Work Package 2 which aims to co-develop a series of adaptable phenotyping devices with AI embedded targeted to agroecological traits. In particular, it will provide fully operational devices with multi-optical head, provide the associated series of computer vision models based AI, and provide embedded AI to directly extract traits.

Applied objective : For this Phd, we target the imaging systems dedicated for the use cases in soil phenology, GxE cereals and Farms to Platforms. We envision a connected stick (vector, autonomous battery, telecommunication ressources and imaging and non sensors). The systems themselves have been designed or selected by the partners of the project. The objective of the Phd is to extract information from these imaging sensors (segmentation of vegetation in mixture of crops, detection of diseases, prediction of stress from time series, …). The challenge is to perform these processes as much as possible on the connected stick processing capabilities themselves.

Methodological aspects

To achieve the targeted applied objective we envision the investigate the following methodological aspect

- **Tiny AI:** to take into account the power limitation, we need light models. We will revisit the so-called *distillation process* [3] with the current state-of-the-art AI models available for plant imaging. We will also take into account the current literature on the encoding of the weights in order to maximize the robustness of the process when some reduction in the voltage is encountered.
- **Automatic quality control:** a strategy to reduce the load of data to be transmitted would be to process all the data that are expected to be conformed to the training set of the model and to transfer to the server the data that are found not to be conformed as these can constitute new types of data or trouble shooting on the stick. To this purpose we will revisit the recent algorithms introduced for *conformal prediction* [1].
- **Few shot learning :** the recent introduction of foundation models [2], i.e. LLM-based models trained on huge data sets, opens the possibilities of generalization to unseen data with few shot or simple prompts. We will test such an approach on time-series of data from plants in development to limit the time of retraining of the models.
- **Adaptive sample :** Bayesian approaches will be developed to adjust automatically the sampling of the connected stick in order to synchronize with the plant growth and not to the clock of the CPU used in the stick.

Supervision: David Rousseau david.rousseau@univ-angers.fr ; Pejman Rasti ; pejman.rasti@univ-angers.fr

Collaboration network : UCL (Université Catholique de Louvain UC F2P), INRAe (Avignon-Clermont UC GxE), CNRS and Vienna (UC soil phenology). An engineer in charge of the deployment of the stick, and all the hardware aspects will be hired during the Phd.

[1] Angelopoulos, A. N., & Bates, S. (2023). Conformal prediction: A gentle introduction. *Foundations and Trends® in Machine Learning*, *16*(4), 494-591.

[2] Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., ... & Sun, L. (2023). A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. *arXiv preprint arXiv:2302.09419*.

[3] Yu, R., Liu, S., & Wang, X. (2023). Dataset distillation: A comprehensive review. *arXiv preprint arXiv:2301.07014*.