Use case 6: Automated data flows for crop simulation models

FAIRagro

Benjamin M. L. Leroy[1,*], David Gacktetter[1], Andreas Donaubauer[2], Joseph Gitahi[2], Marija Knezevic[2], Thomas H. Kolbe[2], Sebastian Burkhart[3], Patrick O. Noack[3], Adolf Kellermann[4], Lina Hörl[5], Giada Matheisen[5], Senthold Asseng[1]

[1] Technical University of Munich, Hans Eisenmann-Forum for Agricultural Sciences, Freising, Germany
[2] Technical University of Munich, Chair of Geoinformatics, Munich, Germany
[3] University of Applied Sciences Weihenstephan-Triesdorf, Competence Center for Digital Agriculture, Weidenbach, Germany
[4] Bavarian State Institute for Agriculture, Freising, Germany
[5] Bavarian State Archives, Munich, Germany
* Corresponding author: benjamin.leroy@tum.de

Background

Crop simulation models have become important tools in agricultural research and crop systems analysis. While they are typically used to forecast crop performance under different environmental conditions, they also possess largely unexploited potential as decision-support systems in crop management. Crop models require the input of various data types, including crop trial measurements, soil surveys, weather time series and forecast, or remote sensing data. The integration of these heterogeneous source materials requires experts in a range of disciplines for locating, accessing, and transferring data, as well as workflows for data integration, transformation and quality control. Additionally, crop models also generate large amounts of output data that must be quality checked, annotated, made available for deriving crop management decisions, and prepared for publication and long-term archiving. Currently, the heterogeneity of data quality and formats, the scattered nature of data sources and a lack of widely adopted domain-specific standards for research data management make data preparation for crop modeling applications a cumbersome and hardly replicable process.

Objectives

- Establish domain-specific research data management guidelines and make them available to agronomists via publications and workshops.
- Develop FAIR-compliant workflows to automate data processing steps for compiling input datasets for crop models.
- Conduct crop trials leveraging these workflows "from field to simulation" to showcase the potential of crop models as decision-support tools.
- Publish the generated datasets annotated with rich metadata by registering digital twins of all involved components in the SRADI catalogue.
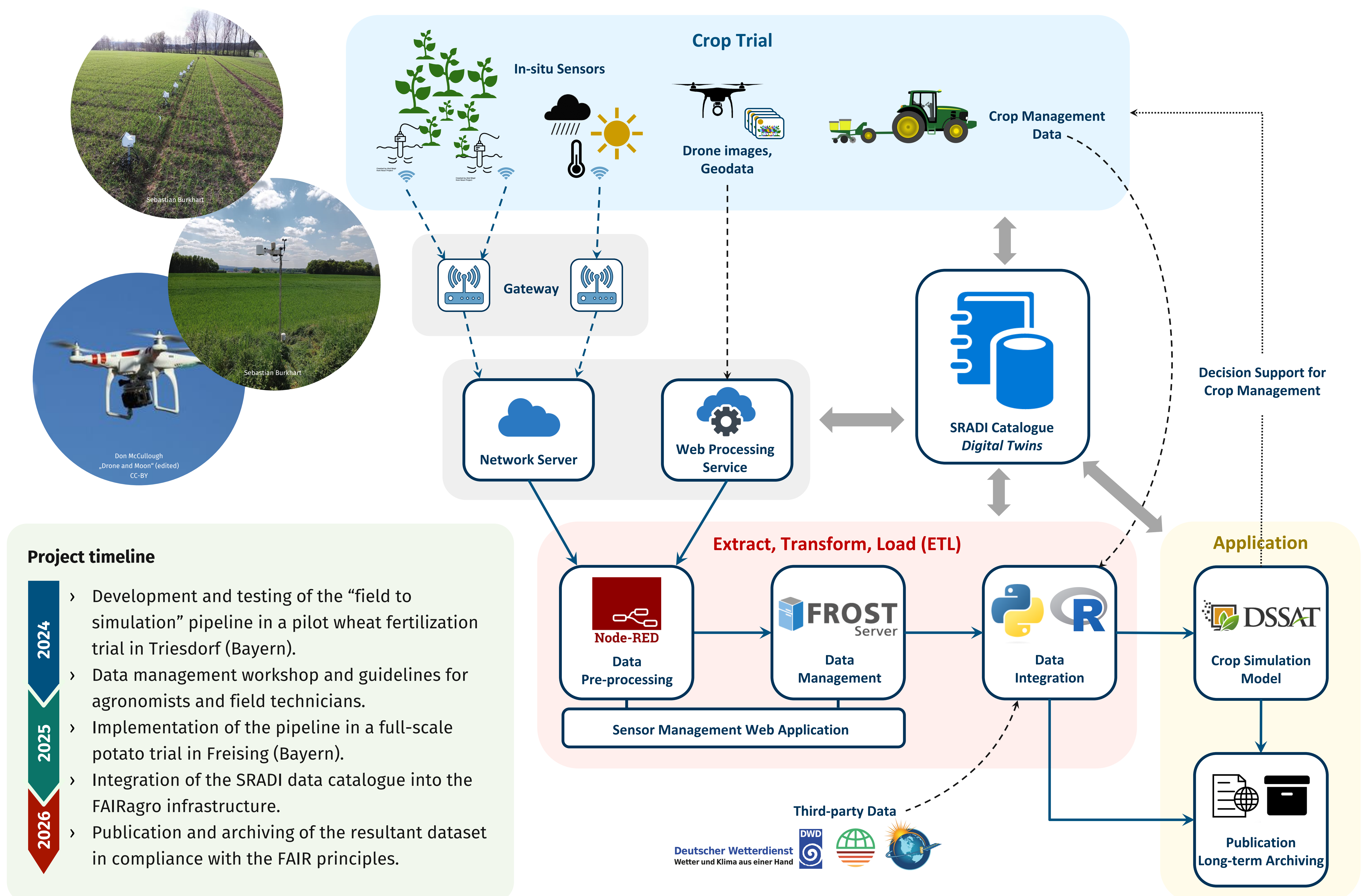
Project timeline

- **2024**
  - › Development and testing of the "field to simulation" pipeline in a pilot wheat fertilization trial in Triesdorf (Bayern).
  - › Data management workshop and guidelines for agronomists and field technicians.
- **2025**
  - › Implementation of the pipeline in a full-scale potato trial in Freising (Bayern).
  - › Integration of the SRADI data catalogue into the FAIRagro infrastructure.
- **2026**
  - › Publication and archiving of the resultant dataset in compliance with the FAIR principles.