PIDs for physical objects



Contribution ID: 19

Type: not specified

Application of PIDs and digital twins of plant genetic resources at IPK Gatersleben

Thursday 20 February 2025 11:35 (25 minutes)

The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) is a leading international plant science institute with a research focus on biodiversity and crop performance. In order to implement a sustainable data and material management infrastructure three pillars has been built in the last 15 years. An institutional policy for research data management, defined processes and a technical infrastructure. The technical backbone is a Research and Laboratory Information Management System (RALIMS) which was established 2011 and is operated as general-purpose data management system across all research groups and departments.

This RALIMS based ecosystem of databases, file storage, desktop clients, web applications and APIs serves two major classes of data management processes. The first class are service processes for centrally managed instruments, facilities and service units. They follow institutional agreed processes and operated by permanent staff. Examples are the high-throughput sequencing, chemicals management or phenotyping service processes. Service processes comprises (a) defined personnel and organizational responsibilities including defined transition points between the laboratories, the scientist and the LIMS project team, (b) defined standard-compliant and machine-processable data formats, (c) mandatory metadata standards and d) defined data publication processes, i.e. the minting of PUIDs, like DOIs, and data upload into international data repositories.

The second class of process are data flows in research projects. Here a more agile and are less rigidly structured processes are in place that reflecting the nature of innovation-driven science. Nevertheless, they are dovetailed with the core service processes and support immersive analytics driven knowledge generation in research projects. For example, research project for the genotypic and phenotypic characterization handle of thousands of plant samples and connect them with millions of data points. Scientist and technician work hand in and to interweave scientific data analysis and visualization pipelines and tools. This data servant approach, which is operated over more than 15 years, enabled the preservation of more than 6 million samples and terabytes of data files in a FAIR manner. The interplay of policies, processes and IT is a central backbone to support research data and material management at IPK and contributes data services to networks such as the European life-sciences infrastructure for biological information (ELIXIR), the German Bioinformatics Network (de.NBI) or the National Research Data Infrastructure (NFDI) in the consortia, FAIRAgro, DataPLANT and NFDI4Biodiversity.

This talk provide an overview to the policies, technology and processes at the IPK to implement FAIR data and material management and show case the application of digital twins in recent research projects.

Presenter: LANGE, Matthias (IPK Gatersleben)

Session Classification: Session 2