Session 1: Drivers and Challenges for new uptakes of new sources and uses of data recording

S01.O-05

AGRI-FOOD DATA CANADA: A DATA ECOSYSTEM SERVING AGRI-FOOD SUSTAINABILITY

Lucas M. Alcantara, Carly Huitema, A. Michelle Edwards.

University of Guelph, Guelph, Canada.

Agri-food Data Canada (ADC) is creating a data ecosystem serving agri-food sustainability. Through investments in technology, infrastructure, and culture, we are helping researchers and the research community get more value from the data researchers are already collecting. Agri-food Data Canada's approach is guided by the FAIR data principles (that data should be Findable, Accessible, Interoperable and Reusable). To improve data FAIRness ADC is 1) Creating a semantic engine that will help researchers create and use better machine-actionable, reusable, and accessible descriptions and governance for their data, projects, algorithms, tools, workflows, and other digital research outputs; 2) Collaborating on projects supporting the federation of data silos, to ensure that data, metadata, and access rights can travel with the data from source to destination within the ADC federation; 3) Developing tools to help researchers with data provenance and traceability; and 4) Creating a culture of FAIR data by developing knowledge-sharing resources such as webinars, training, and teaching materials. ADC works with partners to align our approaches and contribute to the global research community, with the goal to ensure research data is FAIR. One collection of tools that are under development at ADC is the Semantic Engine. While there are many approaches to harmonizing data through the creation of data platforms, ADC sees the value in adding value to heterogeneous data through the creation of tools that improve data without the necessity of data platform infrastructure. Researchers can improve their data documentation workflows by adding context to their data through the creation of machine-actionable data schemas. At the heart of the Semantic Engine is the Overlays Capture Architecture (OCA), an international open standard created by the non-profit organization Human Colossus Foundation. OCA's layered architecture which is machine-actionable and easy to generate. OCA schemas allow multiple contributors to improve a schema independently and permits the bundling of schemas with appropriate task-specific schema overlays. Schemas can be internationalized through the creation of language-independent overlays, and their additions do not change the underlying structure of the schema which ensures interoperability and allows schemas to be continually improved throughout the dataset's lifecycle. OCA also permits the use of downstream data validation rules carried by schemas and enables the incorporation of ontological terms. For example, ontologies, terms, and data standards endorsed by ICAR can be added to schemas to improve data interoperability and harmonization, which are essential for advancing the international agri-food sector. Agri-food Data Canada is developing a powerful collection of tools and creating a data ecosystem that will reduce barriers to data documentation, ease data sharing, and support the international agri-food sector's data needs.