LiveStack: An open-source framework for creating livestock data pipelines

Fieguth Braeden*[1], Alcantara Lucas[1], Frewin Andrew[1], Cant John[1]

[1] University of Guelph

The increased installment of data systems across livestock operations creates a critical need for standardized data orchestration. While ICAR's ADE guidelines establish the foundational architecture and protocols for global interoperability, a practical gap remains in orchestrating diverse on-farm data sources into unified, compliant pipelines.

LiveStack addresses this gap as an open-source, Python-based framework leveraging Apache Airflow and Docker to automate complex data flows. Designed for deployment by technical consultants, it utilizes a declarative approach in which the users describe the data processing and infrastructure they need, while the system handles container orchestration, ETL tasks, and robust monitoring. By leveraging open-source tools for container orchestration, LiveStack can provide centralized user authentication and coordinated access to distributed container environments. The architecture is modular, featuring sub-packages for specialized data cleaning and a plugin system for extensibility, such as integrating Agrifood Data Canada's Semantic Engine for automated data verification. This modular design allows the core functionality to remain agnostic to livestock species. To ensure resilience in rural environments and support data sovereignty, LiveStack is designed to allow for edge, cloud, or hybrid deployments, utilizing local caching to mitigate connectivity issues. A central feature is a standardized API module that outputs data in compliance with ICAR ADE standards. By serving a provider-agnostic data stream, LiveStack lowers the barrier for third-party developers to deploy precision tools by removing the burden of implementing custom ingestion overhead.

LiveStack aims to provide a robust, open-source foundation for an interoperable and scalable livestock data ecosystem. By remaining open-source, LiveStack allows the community to develop and maintain integrations that the entire livestock industry can benefit from.