Avian influenza

The *Wageningen Common Data Solutions* (WDCS) programme is striving to connect use cases from Wageningen University & Research (WUR) research institutes to the research data management infrastructure for FAIR Data. Open-source tools such as iRODS and Yoda are used for this purpose. WCDS is funded by the Ministry of LVNV for two years with a budget of 2 million euro.

In this fact sheet, we outline the progress with the project Avian influenza. The aim of this project is to create a data-handling workflow for epidemiological research data concerning outbreaks of avian influenza

Technical results

- We have introduced a general email address to which researchers can send their epidemiological data. Data on poultry sizes, wild bird mortality and avian influenza outbreaks has been loaded into Adagio. In addition, a connection has been created between the Adagio data platform and iRods.
- When publishing the avian flu data pipeline from within Adagio, an iRods collection is automatically created. This includes all metadata and authorisations as defined in Adagio. Every time data is loaded into the platform, the raw data is also stored in the iRods collection. Both raw data and metadata can be accessed from iRods by authorised users.
- Background documents and other reference metadata can also be added to Adagio. These are stored in iRods and made available through both channels.

- The main weight of cleaning, organising and linking data was pulled by Adagio. The connection with Yoda provides a convenient feature that makes all data and metadata available for storage in Yoda.

Current limitations for further improvement
Right now, it is not possible for a user to upload common concepts in bulk. This must be done by an administrator. Our aim is to make batch uploads easy for users.

Additional developments and implementations planned
- We are currently loading geographic point data into Adagio. Our aim is to be able to create geographical grids from point data in Adagio itself. In addition, we also want to make it possible to store grids, line data and polygon data. These features are also useful for ecological research in the terrestial and maritime real.
- It would be great to have good tools for automation of reading hand-written data.

Key outcomes

- A list with all avian influenza-related data sources.
- A common email address for epidemiological data.
- A connection between Adagio and Yoda. Through Adagio, avian influenza case data, poultry farm details and wild bird mortality data can be accessed.

Thanks to this project, other researchers within WUR can get access to the models and data in Adagio on request. Scientifically, the data pipeline makes it much easier to assure that all research projects access the same data

source. In addition, data handling operations are stored in Adagio. This makes it much easier to have consistent data use between projects and facilitates comparison of the outcomes.

Reflections

In this project, it was challenging to prioritise the data to include in the pipeline. There was not enough budget to include all the potential data sources. At times, it was also a challenge to get activities from WBVR and WEcR aligned.

At organisational level, the MT still needs to see the literal value of well-stored data and why it is worthwhile to invest in good data management now.

Project takeaway
Making a pipeline for epidemiological data is highly feasible in Adagio. To be able to work more independent, we would like to do more of the work in Adagio at WBVR.

Recommendations

- Create a unified toolbox for data management at WUR, with guidelines for which type of data fits best with which solution. Think of Yoda for archiving, Adagio as data warehouse for relatively small data and FAIR by design for large genomic datasets.
- Invest in the technical infrastructure. From the perspective of WUR, it is also important to invest in better policy & support to make all researchers within WUR data literate.
- Keep on organising interesting meetings and make these better available for people in the satellite

institutes outside Wageningen.
- Ensure clear project management and communication from the outset.

Contact
Wageningen Bioveterinary Research
Dr. R. Petie
E ronald.petie@wur.nl