

Data Stewardship @ WUR

Systems and Synthetic Biology (SSB) and Microbiology

Bart Nijssse

1st Data Steward @WUR network meeting, November 8, 2018



WAGENINGEN
UNIVERSITY & RESEARCH



100years
1918 — 2018

What kind of data at SSB and MIB

- Genomics
- Transcriptomics
- Proteomics
- NGS Marker/Amplicon data (MIB)
- ~5TB of (raw) data and growing rapidly

Current (meta) data control and storage

Control

- Excel sheet with metadata regarding the experiment
 - Samples, conditions, organism(s), people, sequencing center etc.
- When correct/complete → make data available for analysis

Storage

- Local NAS + offline/cold storage
- External backup (SURF SARA)

Current obstacles

- Tailored to raw data only
- Find ability
- Unaware researchers already having data
- Old data

(near) Future deployment

iRODS for new projects

- Open source, data management middle ware
- Integrated **user management**
- Browse and **search (meta)data**
- In collaboration with SURF SARA



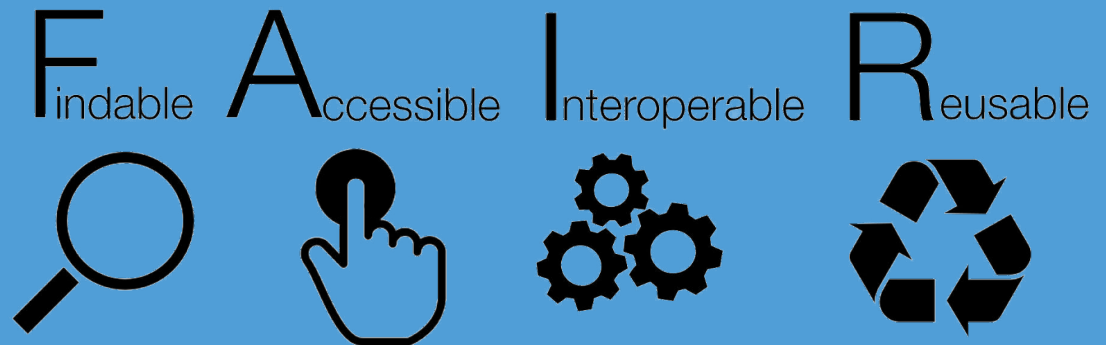
iRODS Rule Engine

- Create customized workflow and **trigger analyses**

Recommendations

- Actionable Data Management Plan (also group wide)
- Enforce policies, i.e. data embargo

- **Data steward**



- Work FAIR by Design