National Research Data Infrastructure (NFDI)



December 2nd, 2019, Bonn

Xenia Specka

Matthias Lange

Jan-Henrik Haunert

Claus Weiland



Ulrike Stahl









Consortium & Community

Xenia Specka

Research Data Management

Jan-Henrik Haunert

Scientific Data & Infrastructure Services

Ulrike Stahl

ENCYCLOPÆDIA BRITANNICA

"Agricultural sciences, sciences dealing with food and fibre production and processing. They include the technologies of soil cultivation, crop cultivation and harvesting, animal production, and the processing of plant and animal products for human consumption and use"









Agricultural research disciplines



- Plant production
- Livestock farming
- Applied genetics & physiology
- Ecology of agricultural landscapes
- Soil sciences
- Forestry
- Agricultural economics & sociology





Agricultural research data domains









SUSTAINABLE GEALS





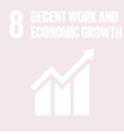
































(Co-)Applicant Institutions and Core Infrastructures



16 (co-)applicant institutions

Non-university research institutes and universities





























Examples of continuously funded core infrastructures

22 repositories & information systems 8 data infrastructures 3 tools











™TSN-ONLINE



















Supported by





www.nfdi4agri.de

5

Status quo

Heterogeneous landscape of data infrastructures

User requirements

24 user stories →

4 consolidated use cases

Need for action

6 task areas

25 measures

4 consolidated use cases



Use case 1: Findability of agricultural research data

e.g. site, type, date, content/value specific



Use case 2: Support agricultural research data management e.g. planning, compilation, publication



Use case 3: Quality of data, services & infrastructures

e.g. completeness, plausibility, machine readable



Use case 4: Applications & integrated analysis

e.g. statistics, machine learning

- 1. Set up an interoperable and scalable infrastructure by connecting available agricultural disciplinary repositories
- 2. Implement the NFDI4Agri portal as a central access point for searching agricultural data to enable semantic queries within the federated infrastructures
- 3. Develop a platform for knowledge exchange and technology transfer to engage in dialogue with the agricultural community

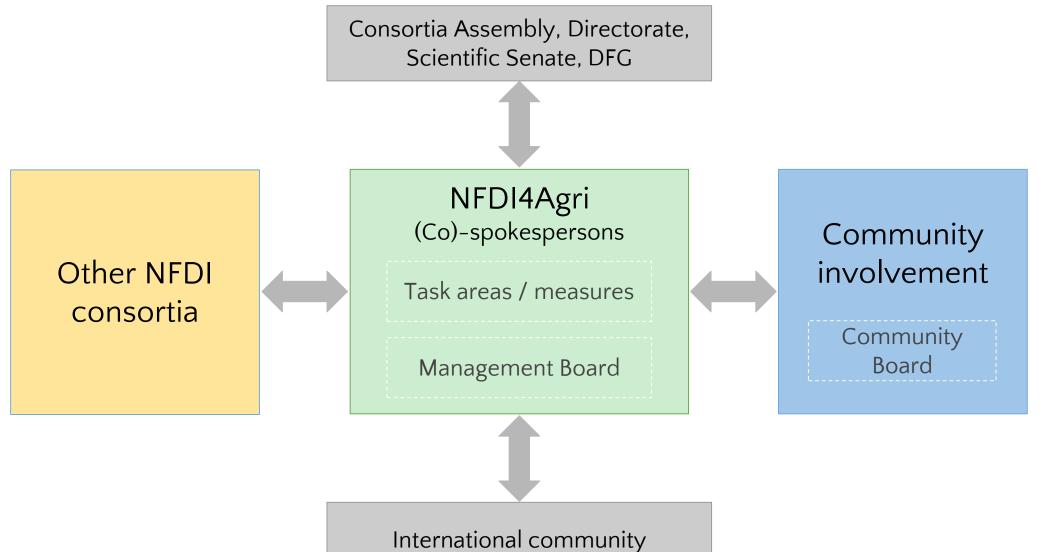
NFDI4Agri

TA 1: Coordination and Communication TA 2: Agricultural Community TA 3: Data Fitness for Use in Agricultural Research

TA 4:
Standards and
Interoperability –
Making Agricultural
Science Data FAIR

TA 5:
Agricultural Scientific
Data and
Infrastructure
Services

TA 6: Data Protection and Data Policies





Starting with 17 Participants (13 societies, 3 universities, 1 library network)





Task Area 2: Agricultural Community

NFDI4Agri

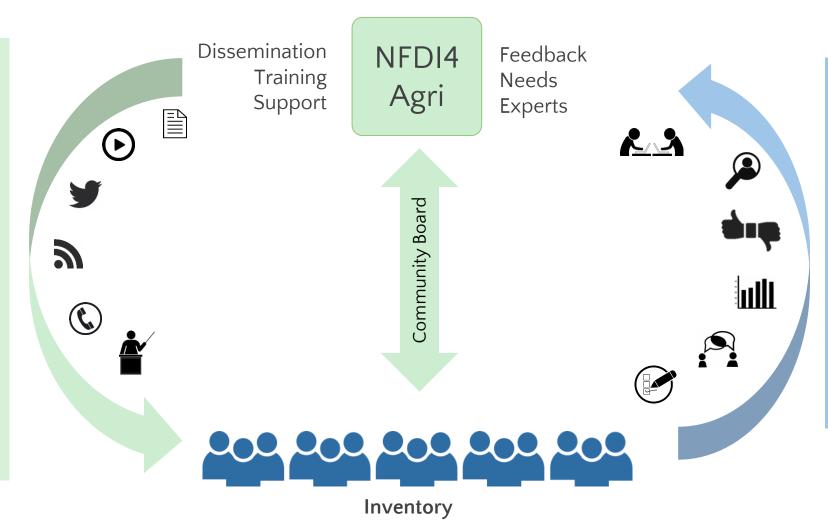
Content preparation

Web portal, mailing lists, webinars, newsletter, social media, scientific papers, press releases

Train the trainer Summer schools Conferences Workshops

Data stewardship

Helpdesk and hotline



Structured feedback system, Hackathons

Indirect feedback by key performance indicators

Surveys & interviews

Recruiting of experts

Participation on community events

National Research Data Infrastructure NFDI4Life NFDI4Agri NFDI4Health NFDI4Earth DataPLANT Umbrella Area of cooperation and networking NFDI4Bio MaRDI Diversity Berlin Declaration Special interest groups Legal & ethical aspects Cross-cutting topics Research Data Ontologies NFDI4 KonsortSWD Use cases Commons Microbiota NFDI4 NFDI4MSE NFDI4Culture NFDI4Chem MobileTech

Consortium & Community

Xenia Specka

Research Data Management

Jan-Henrik Haunert

Scientific Data & Infrastructure Services

Ulrike Stahl

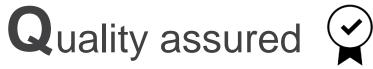


Findable Q

Accessible **

nteroperable **







Scalable

Long-term preserved









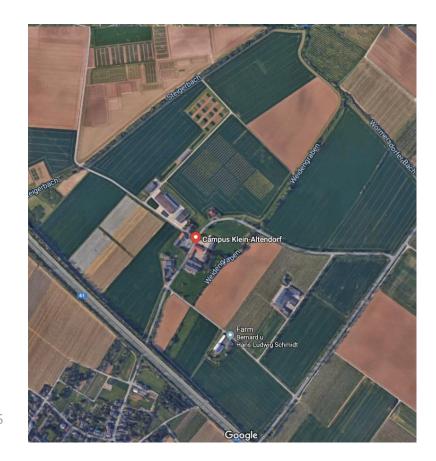


User Story: Robotics & Phenotyping for Sustainable Crop Production

Cluster of Excellence PhenoRob of University of Bonn & Forschungszentrum Jülich

- Field experiments at Campus Klein-Altendorf of University of Bonn
- Large variety of sensors and sensor platforms





4D model of the shoot

Hyperspectral images

Ground penetration radar

Earthworm density

3D geometry of canopy by laser or RGB images

Magnetic resonance

imaging

Electromagnetic induction

Root-tunnel-rhizotron observations

Good documentation of data quality needed, e.g. for integrated data analysis www.nfdi4agri.de

Task Area 3: Data Fitness for Use in Agricultural Research

• Even FAIR data may have low quality & thus be of little use e.g., see RfII 2019: Herausforderung Datenqualität



Quality requirements depend on application

Standards of data quality (e.g. ISO 19157) practices of data acquisition (e.g. KA5 for soil data)

Quality requirements of agricultural use cases & previous uses of data sets (application-data-matrix)

- Implement online platform for discussion & data reviewing
- Develop framework for algorithm-aided data curation









- Plant breeding scientist is looking for a regionally adapted winter barley genotype
- Relevant information:

phenology

stress physiology

interactions with other organisms



chemical ingredients

genotype

morphology

environment

• Currently, this needs to be collected from different unconnected repositories



- Making Agricultural Science Data FAIR

- Design and propose fundamental data models
- Review of established metadata & identifier schemas
- Take up key results of initiatives like RDA IGAD and Agrisemantics WG
- Define interfaces to Data Aggregators like the AgroLD
- Enable semantic search across terminologies & services building on FAO's AGROVOC Thesaurus











User Story: Animal Disease Reporting System (Tierseuchen-Nachrichtensystem TSN)



- Reporting of animal diseases for effective control and scientific assessment
- Data is highly sensitive
- Legal & ethical issues and anonymization are important

(6)

Targeted actors and measures:

ethical and practical aspects

Data Providers Repository Managers Data Users Tailored guidance on Tailored guidance on regulation, Good practices in terms of regulation, data ownership, legal, ethical and policy data ownership, ethical and

practical aspects

Sample materials such as checklists, exemplary consent forms, data policies and terms of use, e.g., in compliance with the Nagoya Protocol

aspects of data and metadata sharing

Consortium & Community

Xenia Specka

Research Data Management

Jan-Henrik Haunert

Scientific Data & Infrastructure Services

Ulrike Stahl

Task Area 5: Scientific Data & Infrastructure Services Operating Model

Agricultural Community Other Communities **Federated Service Topology** FAIR+: Quality assured, Scalable, Long-term preserved, Legally sound TA 1: Coordination & Communication TA 2: Agricultural Community TA 5: Scientific Data and Infrastructure Services TA 4: Standards & Interoperability TA 3: Data Fitness for Use TA 6: Data Protection & Policies **Services of Co-applicants** Thesaurus Repository Web Services & Tools **Community Use Cases NFDI Cross-Cutting Topics** Information System IT-Infrastructure **Education Services**



Inventory service – for continuous monitoring of infrastructures



Middleware of services - for harmonized access to data sources



Search service - for information retrieval



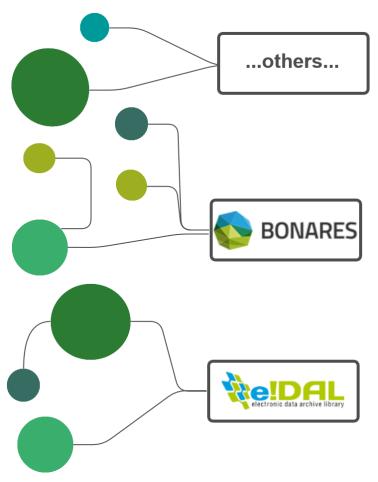
Open software components – for research data management



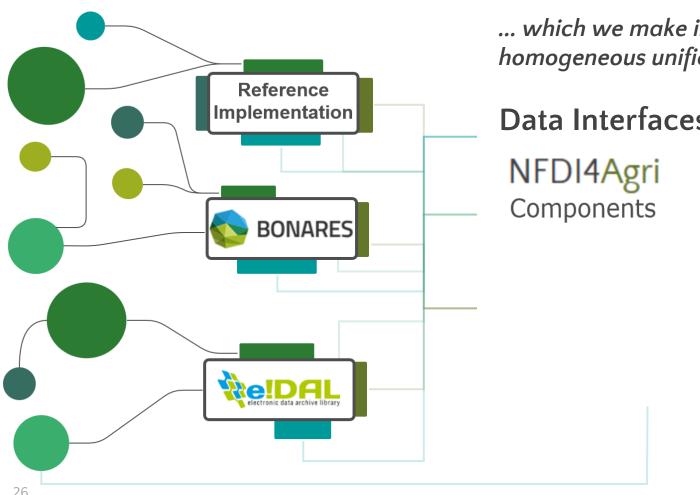
FAIR data containers - for warehousing harmonized data collections



(a)



... use long-term-committed heterogeneous but specialized **Data Repositories**



... which we make interoperable by homogeneous unified

Data Interfaces & Services

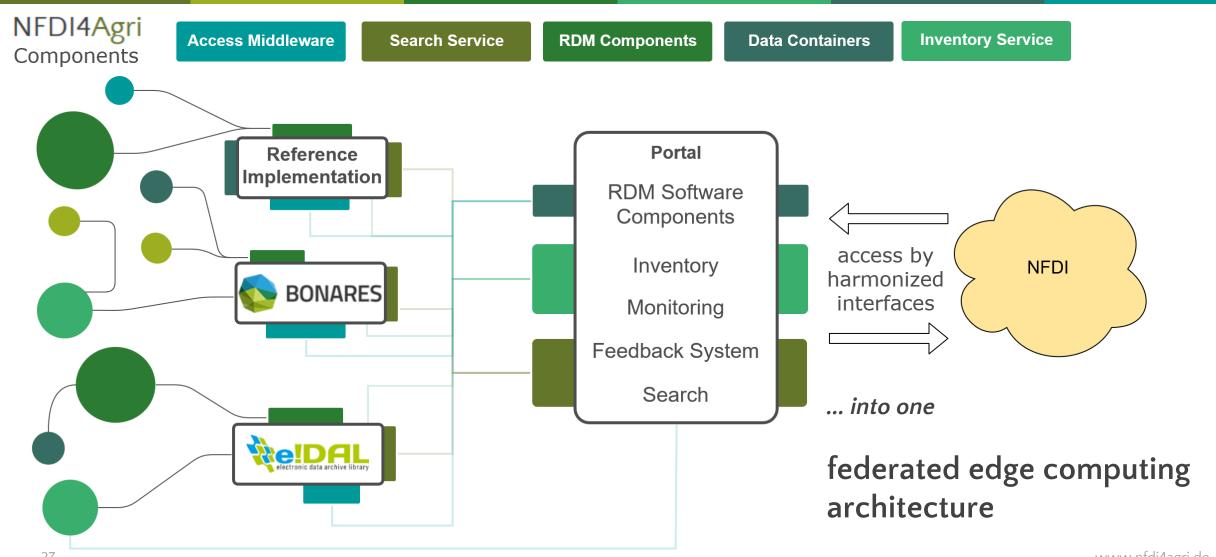
Access Middleware

Search Service

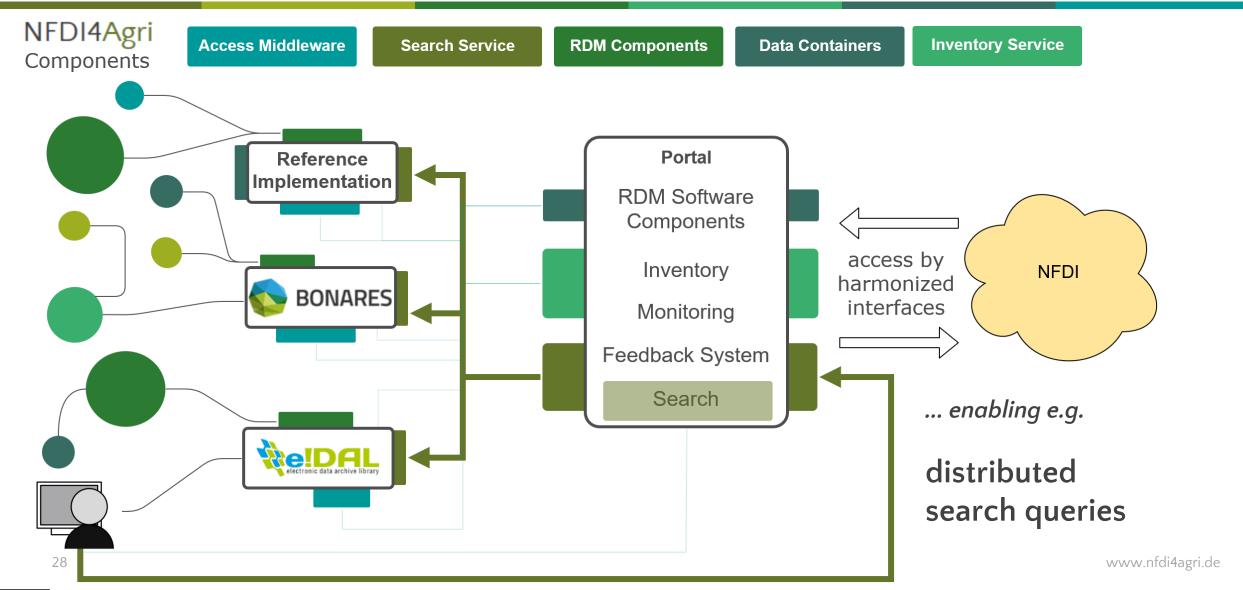
RDM Components

Data Containers

Inventory Service



Federated Service Topology



- Federated long-term funded and interoperable infrastructures
- Modular components for implementing or upgrading infrastructures
- Structure for actively involving the community (bottom-up-approach)
- FAIR+ (quality assured, legally sound, scalable, long-term preserved)

Enabling data owners to become data providers!





- Slide 1: Petair | Fotolia (Field)
- Slide 3: Nemanja Oticoticki | Fotolia (Tractor), Jarno Müller | ZALF (Glashaus, Cows)
- Slide 6: Eva Kufulium (Analytics graph)
- Slide 15: Google Earth (greater Bonn, Klein Altendorf region)
- Slide 19: Nemanja Oticoticki | Fotolia (Tractor), Jarno Müller | ZALF (Cows)
- Slide 23: Freepik from Flaticon (icons)
- Slide 29: Haunert (Roboter)

For pictures and graphics, which are not listed here, either CCO licenses are available or the rights are with one of the speakers.

31