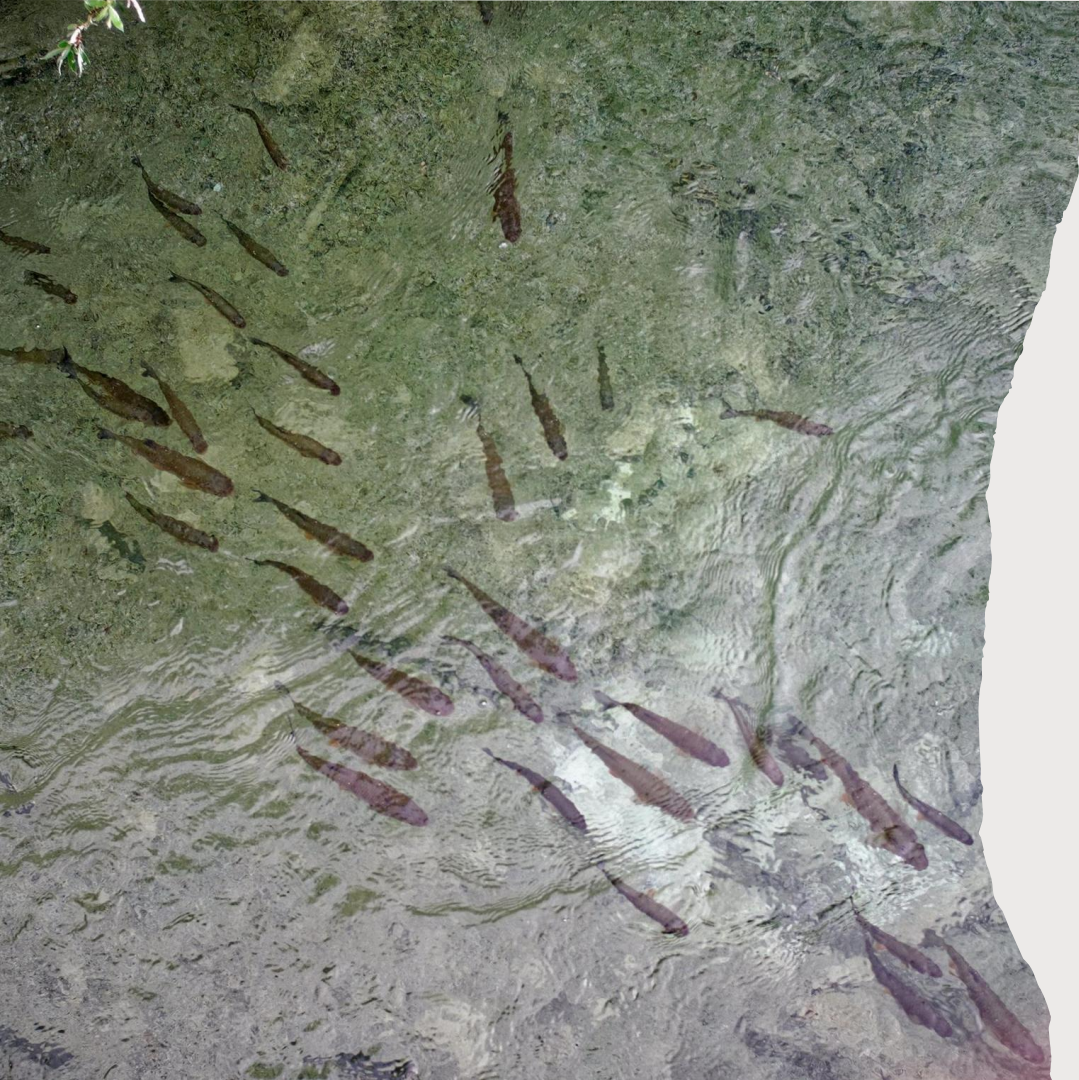


## Connecting data and people: CODATA and the cross domain interoperability framework

Matti Heikkurinen, CODATA  
Project Portfolio Manager



# People issues

- Academic research relies on short-term staff (PhD students, short-term research grants). **Organisational memory?**
- Individual incentives to support FAIR data: often **sticks instead of carrots**
- Cognitive limits vs open sets of:
  - Data
  - Tools
  - Use cases
  - Resource providers
- **Data longa, vita brevis** – the usefulness and relevance of data has never been limited to the researcher's professional lifetime.
- From **data deluge** to **metadata deluge?**

# The market

- **Market sizes**
  - Big data analytics: from 75 to over 300b\$/year (world)
  - Data professionals (EU + UK): ~10M (Statista)
  - Over 2 million researchers in the EU (2023)
- Focus on a single discipline and/or organisation
- **Cross-domain** activities:
  - **EOSC** annual budget 300M€ (~3000 FTEs)
  - **RDA** network of ~15,000 volunteers





# Shipping data across domains

- Global cross-domain interoperability is a formidable challenge.
- Remains relatively neglected in data stewardship community.
- Moving **communities, standards and policies** (and mountains) – **Levers** and **Fulcrums** needed!

# CODATA Fulcrums (high-level view)

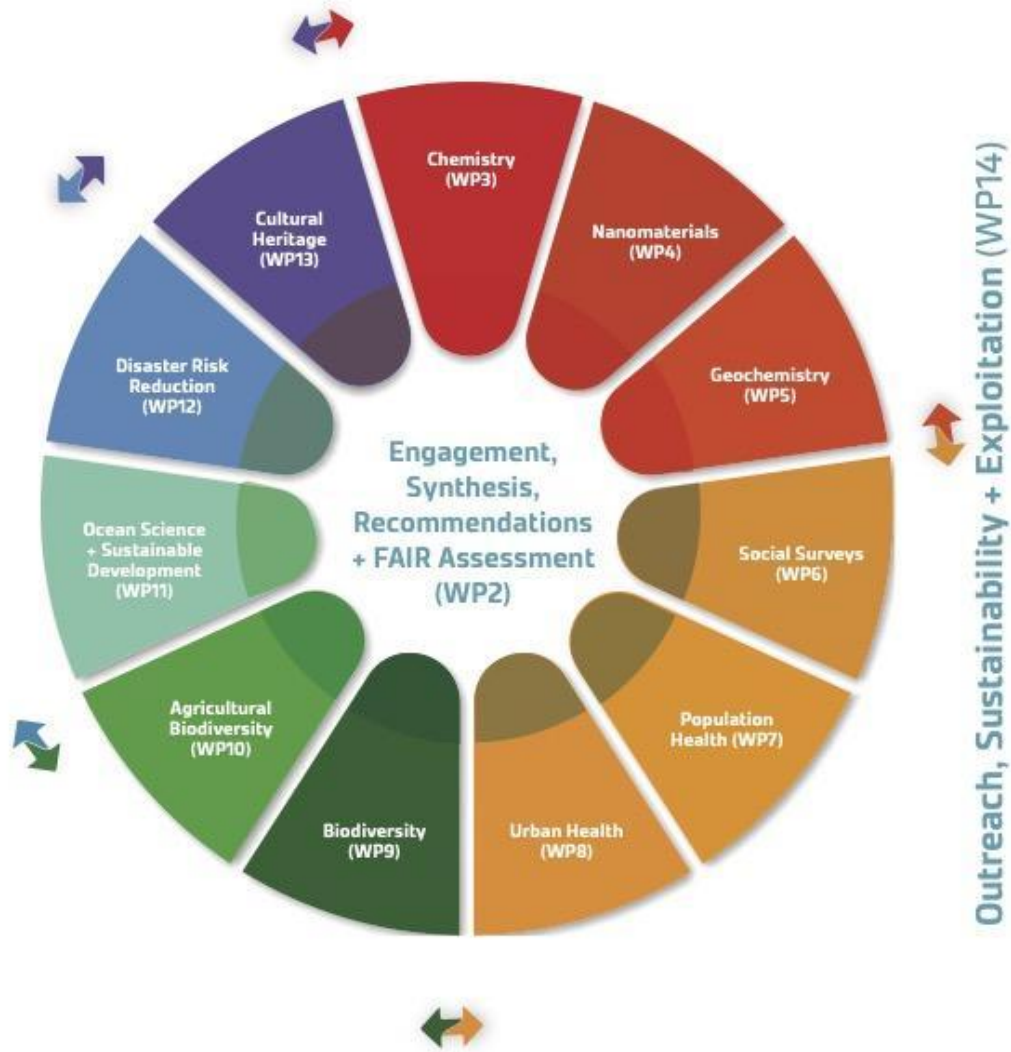
- CODATA serves a **membership** that includes national data committees, scientific academies, International Scientific Unions (ISC membership) and other organisations.
- CODATA works with many **partners**, including
  - ISC and ISC members – **Policies and technical recommendations.**
  - UN and intergovernmental agencies (UNESCO, UN Stats, UNEP, UNDRR; OECD; GBIF, BIPM) – **Standardisation**
  - RDA, GO FAIR, WDS – **Community engagement**
- **CODATA's vision is of a world in which science is empowered to address universal challenges through the transparent, trustworthy and equitable use of data and information.**



# Levers & fulcrums: WorldFAIR

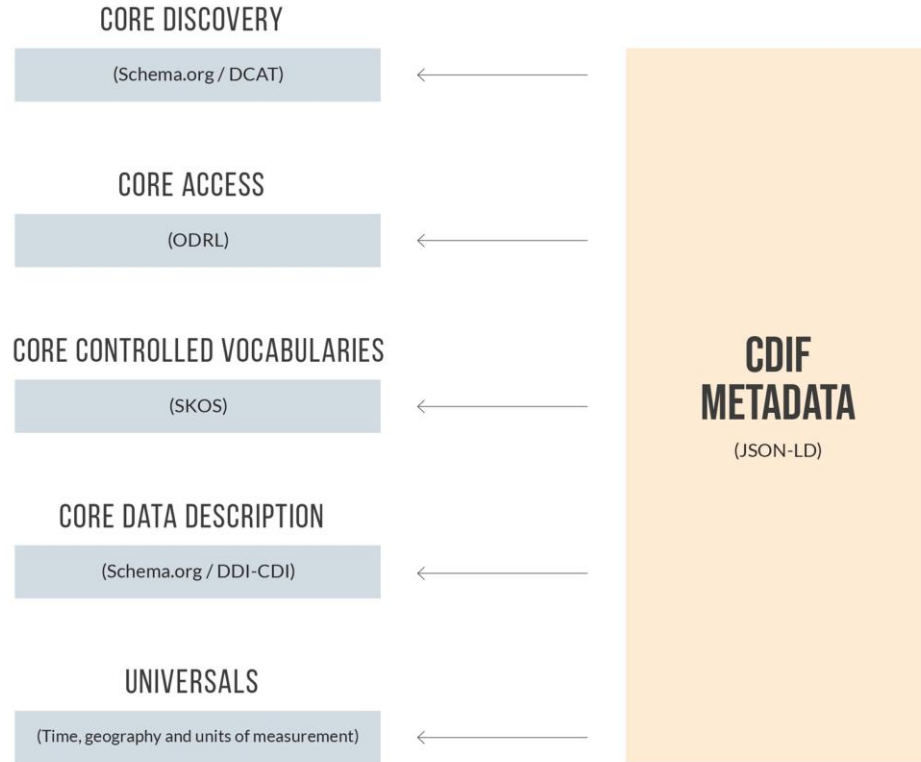
- **Chemistry** – making IUPAC assets FAIR
- **Nanomaterials** – applying NanoInchi and FAIR recommendations in Nanosafety.
- **Geochemistry** – recommendations for FAIR in geochemistry, particularly vocabularies.
- **Social Surveys Data** – data harmonisation between ESS and AussiESS.
- **Population Health** – INSPIRE - Integration of population surveys with clinical and genomics data for COVID-19 research in eastern and southern Africa.
- **Urban Health** – terminologies and making urban health data FAIR
- **Biodiversity** – improving GBIF data model in collaboration with TDWG - GBIF (Global Biodiversity Information Facility)
- **Agricultural Biodiversity** – pollinator data (KALRO, Embrapa, Meise, HiveTracks)
- **Ocean Science** – Implementing FAIR in the ODIS (Ocean Data and Information System) for the UNESCO Oceans' decade.
- **Disaster Risk Reduction** – recommendations on making DRR data and terminologies FAIR, case studies in Africa and Pacific Islands
- **Cultural Heritage** – recommendations on making cultural heritage data FAIR (particularly digital representation of heritage artefacts)

Project Management + Coordination (WP1)



# The Primary Lever: CDIF

- The Cross Domain Interoperability Framework (CDIF) is a set of practical, implementation-level principles designed to improve data management practices within any community and lower the barriers to cross-domain data reuse. CDIF offers standards and methodologies for achieving different levels of interoperability necessary for reusing data across diverse domains. It is built around five core profiles that address the essential functions for implementing cross-domain FAIR principles.
- CDIF was first released in May 2024 as an output of the WorldFAIR project:  
<https://doi.org/10.5281/zenodo.11236871>
- The point of reference for CDIF and its component profiles is now the CDIF Book: <https://cross-domain-interoperability-framework.github.io/cdifbook/introduction.html>



# CDIF – second impressions

## CORE DISCOVERY

(Schema.org / DCAT)

## CORE ACCESS

(ODRL)

## CORE CONTROLLED VOCABULARIES

(SKOS)

## CORE DATA DESCRIPTION

(Schema.org / DDI-CDI)

## UNIVERSALS

(Time, geography and units of measurement)

## CDIF METADATA

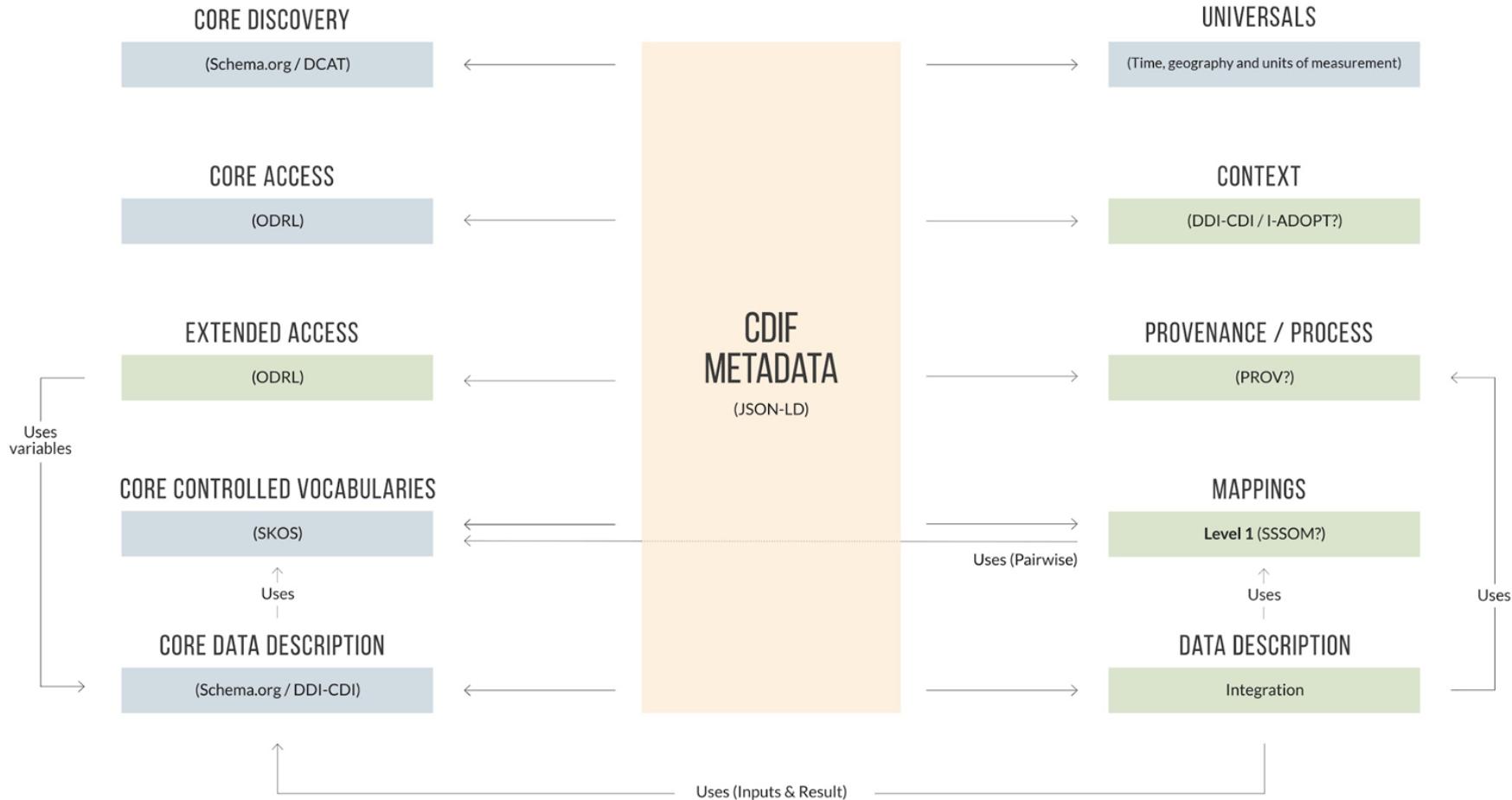
(JSON-LD)

### Feedback received from a colleague at a national data infrastructure

“I have a long reading list that I’m working through and initially wasn’t too excited to be sitting down to read another technical report, and a massive one at that, but as I started reading, it was like it stopped me in my tracks to ask **“Is it your job to try to work out the design of metadata for a cross-domain repository and would you like us to tell you how you might do that in the best, most FAIR way?”** to which I had to reply **“Yes, yes that’s exactly what I’m trying to do...”**”.

This will make a real difference to guide and frame what we’re doing and save me much time by recommending best practices and summarising choices that we would be making along the way. It gives us an achievable first scope for our metadata but will allow us to grow this over time as CDIF develops beyond version 1. We were reassured by how well it aligns well with what we were thinking of doing — we had independently already decided to start at the top-level with a SKOS vocabulary and DCAT catalogue description. As such, we intend to use your report as a starting point for how to implement this and build it up as far as it fits with our project plan.”

# CDIF, Next Steps



# 'WorldFAIR+', CDIF Implementation Projects

1. **"Data Science Without Borders"**: Wellcome-funded project. Population health, building on WorldFAIR WP07. CDIF implementation. Combining population health / statistical data, clinical outcome data, phylogenetic data, environmental data. Also privacy management, ML/AI to enable federated analysis across four African health research centres (Kenya, Ethiopia, Senegal, Cameroon). Three years. Underway. Africa.
2. **"FAIR Data and Emergencies"**: ISC-funded. Applying the WorldFAIR methodology, implementing CDIF components in case studies on earthquake data (Turkey) and flooding / cholera (Malawi). 18 months. Started 1 September 2024. Africa and Turkey.
3. **"CDIF-4-XAS"**: OSCARS cascading grant (EC). Describing X-ray absorption spectroscopy data with CDIF to enhance interoperability and enable interdisciplinary reuse. Two years. Started 1 October 2024. Europe (Germany and UK, but with global relevance and partners).
4. **"CLIMATE-ADAPT4EOSC"**: Major EC funded project, FAIR data and innovative services for climate adaptation. CDIF implementation for legal, organisational, semantic and technical (LOST) interoperability. Three case studies: urban heat (Greece); oceans / coastal management (Portugal); clay soils / hydrology / built environment / insurance (France). Four years. Started 1 Jan 2025. Europe.
5. **"JUST SAFE"**: EU-funded, linked to CLIMATE-ADAPT. CDIF implementation, particular emphasis on citizen science data, legal and policy recommendations. Four years. Starts 1 May 2025. Europe.
6. **"FAIR Principles implementation for DDE"**: Implementation of FAIR principles, alignment of IUGS CGI standards with CDIF, for cross-domain research topics and data reuse in geology. Three years. Starts April 2025. Funding from IUGS. Global.
7. **"Citizen-Driven Living Labs for Urban Heat Island Mitigation"**: ISC Science Missions project selected for funding. Transdisciplinary approach to urban heat in India and SE Asia. CDIF core to data integration for multiple data sources. ISC brokered funding. Starts later 2025.



WorldFAIR





Next steps

# Save the Date!



## INTERNATIONAL DATA WEEK

13-16 October 2025  
Brisbane, Australia



billy

### HOSTED BY



Australian Research Data Commons



Australian Government



NCRIS  
National Research  
Infrastructure for Australia

ARDC is  
enabled  
by NCRIS



COMMITTEE FOR DATA  
CODATA  
INTERNATIONAL  
SCIENCE COUNCIL



RDA  
RESEARCH DATA ALLIANCE



WORLD  
DATA SYSTEM

### OUR PARTNERS

BUSINESS EVENTS AUSTRALIA



TOURISM  
& EVENTS  
Queensland

brisbane  
australia



BRISBANE  
CONVENTION  
& EXHIBITION  
CENTRE

[internationaldataweek.org/idw-2025/](https://internationaldataweek.org/idw-2025/)

# CODATA Secretariat – please get in touch!

## CODATA Secretariat

- **Radomir Buzek**, Executive assistant.
- **Nina Grau**, Project Officer, RDA Tiger.
- **Arofan Gregory**, Metadata and Technical Expert, ‘Making Data Work’, WorldFAIR+ initiative and DSWB project.
- **Matti Heikkurinen**, Project Portfolio Manager.
- **Simon Hodson**, Executive Director.
- **Asha Khandelwal**, Programme Assistant: communications, technical platforms and website updates.
- **Laura Molloy**, Senior Research Lead: WorldFAIR and RDA Tiger projects, vocabularies and terminologies, project and research development.
- **Hana Pergl**, Operations Manager: strategic and management activities, membership relations, GOSC.
- [info@codata.org](mailto:info@codata.org)



*Simon Hodson*



*Laura Molloy*



*Asha Khandelwal*



*Arofan Gregory*



*Matti Heikkurinen*



*Hana Pergl*

## CODATA Team on INSPIRE / DSWB / CDIF Projects

- **Jay Greenfield**, Data, Metadata and Systems Expert, DSWB and INSPIRE MH Projects.
- **Dorothy Mailosi**, Research Assistant, DSWB and INSPIRE MH Projects.
- **Doug Fils**, Data and Software Craftsman with a focus in semantics, linked data and web architecture, DSWB project.
- **Steve Richard**, preparing CDIF documentation, managing CDIF WG.

## Global Open Science Cloud, International Programme Office

- **Lili Zhang**, Executive Director.
- **Xueting Li**, GOSC Programme Officer.



**Thank you for your attention!**

Matti Heikkurinen, CODATA  
matti@codata.org  
www.codata.org



@codata-isc.bsky.social

LinkedIn: [codata-isc-committee-on-data](https://www.linkedin.com/company/codata-isc-committee-on-data)



# CODATA Officers

## CODATA President

- **Mercè Crosas**, head of Computational Social Sciences, Barcelona Super Computing Centre.
- See <https://codata.org/about-codata/message-from-president-merce-crosas/>



## CODATA Officers

- **Richard Hartshort, Vice President**, Professor of Chemistry, University of Canterbury, New Zealand.
- **Daisy Selematsela, Vice President**, Director of Wits Libraries, South Africa.
- **Christine Kirkpatrick, Secretary General, Greenfield**, Data, Metadata and Systems Expert, INSPIRE projects and DSWG.
- **Steve McEachern, Treasurer**, Director of the Australian Data Archive.
- **Barend Mons, Past President**, Professor of Biosemantics, Leiden University Medical Centre.
- See <https://codata.org/about-codata/executive-committee/>

# CODATA Executive Committee

## Elected Members

- **Toshihiro Ashino**, Professor, Toyo University, Japan.
- **Tyng-Ruey Chuang**, associate research fellow at the Institute of Information Science, Academia Sinica, Taiwan.
- **Leo Lahti**, professor in Data Science at the Department of Computing, University of Turku, Finland.
- **Pam Maras**, Past President IUPsyS, Emerita Professor of Psychology at the University of Greenwich, UK
- **Audrey Masizana**, Senior Lecturer in Computer Science University of Botswana.
- **Virginia Murray**, Head of Global Disaster Risk Reduction, UK Health Security Agency; ISC Fellow.
- **Marc Nyssen**, IUPESM, professor emeritus, medical informatics and biomedical engineering.
- **Elena Rovenskaya**, Advancing Systems Analysis (ASA) Program Director at the International Institute for Applied Systems Analysis (IIASA), Austria.
- **Andrew Young**, Chief Research Scientist at the Commonwealth Scientific and Industrial Research Organisation, Australia; director of National Research Collections.
- **Cyrus Walther**, IUPAP, President of the International Association of Physics Students; ISC Fellow.

## Ex Officio Members

- **Francis Crawley**, Good Clinical Practice Alliance, Belgium; Chair of the CODATA International Data Policy Committee.
- **Shaily Gandhi**, co-lead of the CODATA Connect Early Career and Alumni initiative; ISC Fellow.
- **Mark Parsons**, Research Scientist for NASA at the University of Alabama in Huntsville; co-editor-in-chief of the CODATA Science Journal.
- **Lili Zhang**, research scientist in the scientific data center at the Computer Network Information Center of Chinese Academy of Sciences; Executive Director of the CODATA International Programme Office for the Global Open Science Cloud.
- See <https://codata.org/about-codata/executive-committee/>



# Gates Foundation

## Enabling access to data in agricultural development grants

### Lessons from building the FAIR Process Framework for the Gates Foundation

Martin Parr, Director, Data Policy & Practice, CABI  
([m.parr@cabi.org](mailto:m.parr@cabi.org))

# What is wrong with how we work with data now?

*What problems are caused for actors in agricultural transformation by inadequate current data practices?*



Billions of USD\$ are invested in research on the assumption it is adding to the knowledge base and can be used to support pro-poor projects.



Research isn't used or built on beyond the lifetime of funded projects – research is done over-and-over again – wasting time and resources.



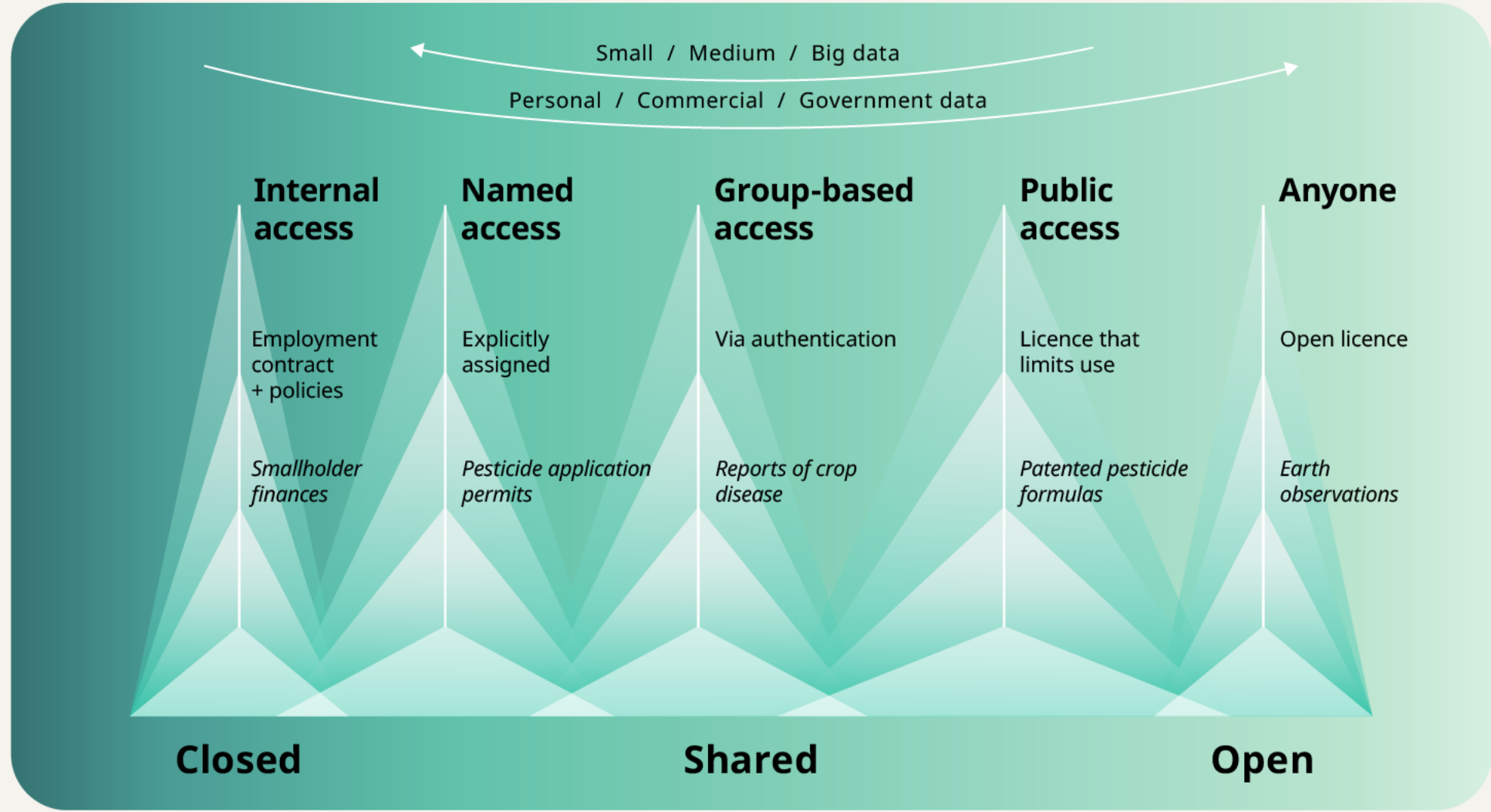
Funders and governments have made investments without necessarily requiring good practice in data management and sharing and too little attention to the people dimension.



Digital and data policy and practice is often not informed by appropriate consultation and evidence.

# The Data Spectrum: Agriculture

Source: Open Data Institute



# F A I R

---

## Findable

Can it be located easily through clear descriptions and proper identifiers?

## Accessible

Are there clear instructions on how to access the data?

## Interoperable

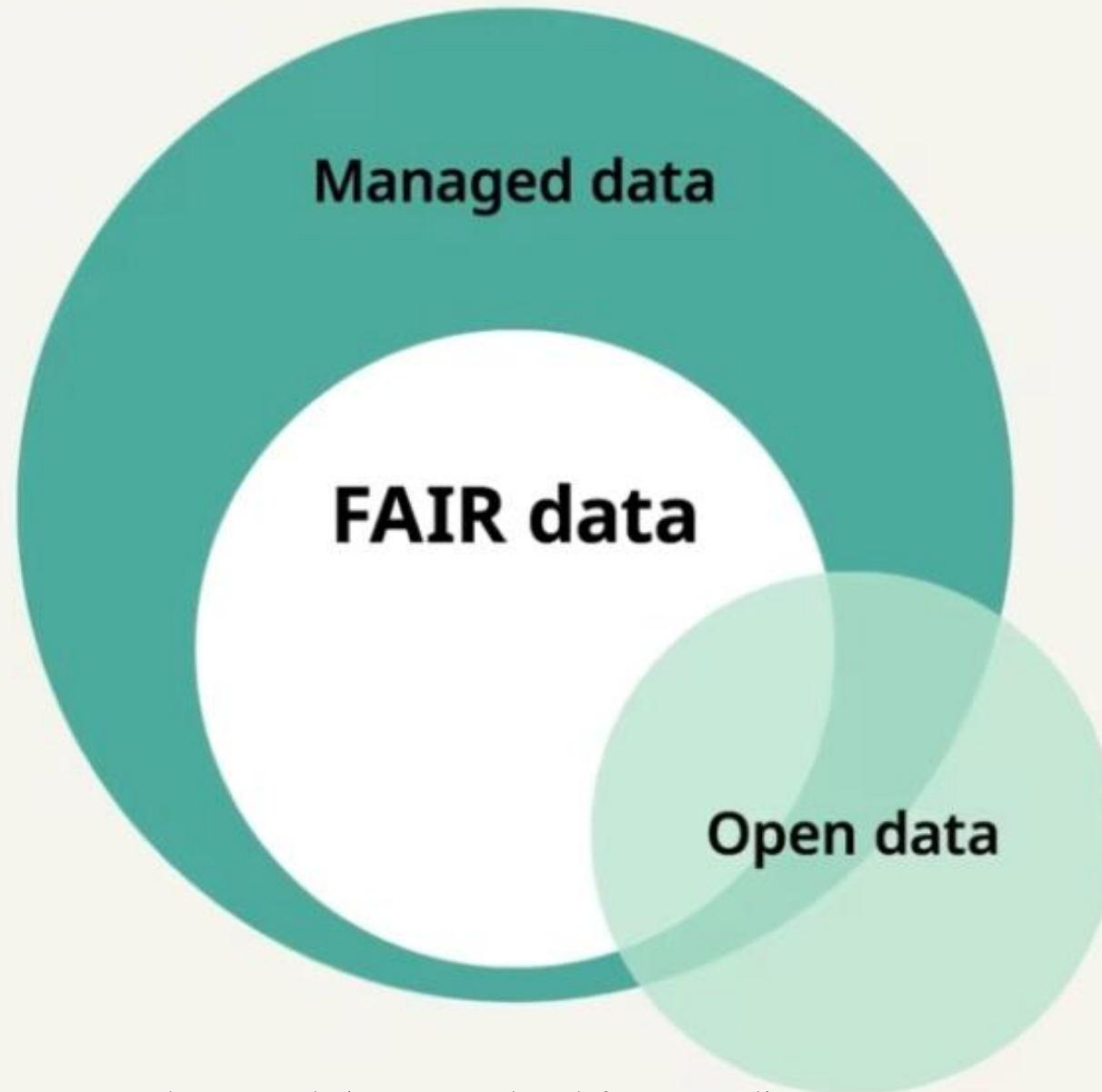
Is it compatible and able to work with other data, systems, and tools?

## Reusable

Is it well-described and structured and with a license in place?

# FAIR vs Open research data

Source: Ghent University

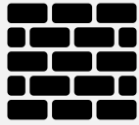


Foundation Master Template (TO EDIT THIS TEXT: select Header & Footer on Insert tab)

The value of Open and FAIR



# FOUNDATION STRUGGLES TO MAKE DATA AND EVIDENCE ACCESSIBLE FOR VALUABLE USES CONSISTENTLY AND AT SCALE



1

## Gaps in Data-Related Grant Commitments

Without specific grant commitments around what, how and when data will be shared, grant data may not be accessible

- X *Grantees are misaligned and unincentivized to make data accessible*
- X *Related policies (OA, GA) not specific enough make data move*
- X *POs lack guidance for specifying data commitments upfront in grants*

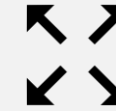


2

## Grantee Capacity Gaps Limit Access to Data

Grantees often face challenges around making data accessible that slow down or lock up data if unaddressed

- X *Grantee gaps in data stewardship expertise and capacity lock up data after the grant*
- X *Research embargos or delays in unlocking data after the grant may prevent timely data uses*



3

## Cross-Cutting Data Access Barriers Lack Ownership

Challenges to making data accessible that are too expansive, costly, or out of scope for any one grant often go unaddressed without ownership in place.

- X *Challenges to data access (like sensitive data) require ownership, and oftentimes cross-team coordination*
- X *Rolling out solutions to these challenges require leadership voice and sponsorship*

# Economic Impact of Open and FAIR Data

*Unleashing Economic Value*

**\$3-5T**

annual economic gain

## **Global Potential:**

McKinsey Global Institute estimates that open data, driven by FAIR principles, could generate \$3 trillion to \$5 trillion in annual economic gain across sectors like education, healthcare, and transportation.

# Economic Impact of Open and FAIR Data

*Failing to reduce inefficiency, replication and data-loss has a lasting impact*

£10.2B

could be saved

**European Union:**  
Implementing FAIR data principles could save a minimum of £10.2 billion annually by reducing inefficiencies.

# Economic Impact of Open and FAIR Data

*A use case for unleashing economic value through FAIR data principles*

**\$796.0B**

investment generated

## **Human Genome Project:**

A \$3.8 billion investment in making available well-structured data generated \$796 billion in economic impact and created approximately 310,000 jobs, driving the genomic revolution.

# Benefits of FAIR - Beyond the Numbers

*The ROI to funders is important, but the value of FAIR should be considered in a wider context.*

*FAIR and responsible data practices have broader impacts including enabling future innovation, enhancing decision-making, driving transparency and growth, and supporting greater equity*



## Non-Quantifiable Benefits

These include the acceleration of agricultural research, the creation of shared data infrastructures, and efficiency gains across various sectors.



## Empowering Decision-Making

FAIR data principles enhance data-driven decision-making processes in governments and industries, leading to more informed policies and strategies that can better address global challenges such as climate change, public health, and sustainable development.



## Global Impact

FAIR data fosters innovation, transparency, and collaboration, driving long-term growth and societal benefits.



## Promoting Equity and Inclusion

FAIR data supports equitable access to information, enabling researchers, policymakers, and communities in developing regions to participate more fully in global scientific and economic advancements.

Engagement with the foundation

# Why is FAIR good for data-rich projects?



**Data Silos:** Valuable data is often trapped within institutions or proprietary systems, making it hard to find and access, which hinders collaboration and limits the impact of data-driven decisions.



**Lack of Standardization:** Inconsistent formats and metadata lead to inefficiencies and errors, making data from different sources difficult to combine or compare, restricting interoperability.



**Innovation Barriers:** The inability to efficiently share and reuse data across sectors slows down scientific progress and the development of new technologies.



**Poor Reusability:** Without proper documentation and licensing, data becomes unusable over time, leading to duplicated efforts and wasted resources.



**Findable:** Reduces wasteful duplication of time, effort and money, enhances efficiency and aids decision-making.

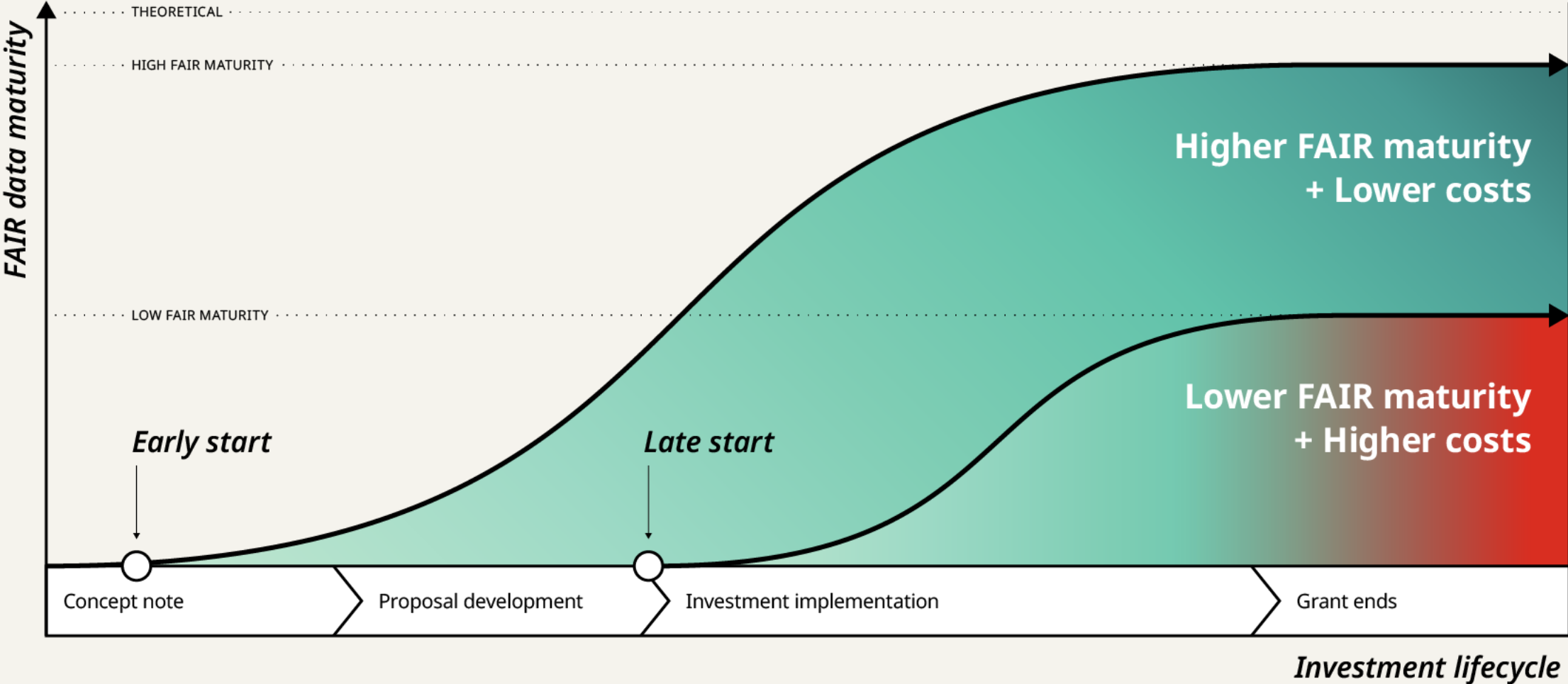
**Accessible:** Promotes inclusivity and knowledge sharing and wider benefits of sharing regardless of location or affiliation.

**Interoperable:** Enables collaboration: combining diverse datasets, increasing insights and effective solutions to food security challenges regardless of location or affiliation.

**Reusable:** By reusing data and FAIR Process tools created in previous investments, costs are reduced. Reusable data and digital assets amplify research impact, fostering innovation and allowing the foundation to utilize existing datasets to work towards sustainable agriculture goals.

# FAIR data maturity during investment lifecycles


Plan early for FAIR data implementation to simplify data management and reduce costs



# The FAIR Process Framework

Six steps towards unlocking US\$ billions of value in data



 **Enable, engage and encourage M&E activity throughout the process**

# Explore the FAIR Process Framework: Your Go-To Resource Hub

Resources to operationalise FAIR and responsible data

[Data Management and Access Plan \(DMAP\) template](#)



[Data Sharing Agreement template](#)



[Data governance policy template](#)



[Guidance on data ecosystem mapping](#)



[Guidance on data standards, licensing](#)



[Guidance on making data inventories](#)



[Lists of Repositories and other data platforms/tools](#)



[Case studies](#)



[FAIR Primer](#)



# FAIR and responsible data practices unlock value

Use this framework to generate the necessary data management and access plan for a successful project



## Is this for you?

**Why FAIR?** — Find out how adopting Findable, Accessible, Interoperable, and Reusable (FAIR) data practices maximizes your project's impact by enhancing collaboration, improving outcomes, and increasing return on investment, and the research behind this framework.

**The steps** — Adopt the six-step process designed with over 32 global partners, offering bite-sized activities to help you and your team seamlessly integrate a customized FAIR and responsible data plan from concept to implementation.

**Investments in action** — Find out about three Bill & Melinda Gates Foundation investments, and how they are implementing FAIR and responsible data strategies across diverse fields using the FAIR Process Framework.

## How does it work?

The earlier you begin planning for FAIR and responsible data value and the lower the cost of implementing the investment Access Plan (DMAP) will be. Planning early for FAIR data practices increases the impact and cost-effectiveness.

**Are you a Program Officer from the Bill & Melinda Gates Foundation?**

Please [download this Job Aid](#) to find out how to support your Open Access Policy.

**Are you seeking an investment for a project that will generate data?**

[Use this checklist to get started](#)

## Investment types and illustrative scenarios



The six steps will be applied differently based on your project type. We have identified common AgDev investments and developed three 'investments' and 'illustrative scenarios.' Each activity page includes examples and a guide to your FAIR data management and implementation.

**AgriConnect** **AgroThrive** **NGBT**

If your project is delivering a digital product or has a significant digital solution investment type throughout steps 1-3.

**AgriConnect: a digital solution investment**

Recently, a large African-led organization, AgriConnect, implemented digital solution investment processes. Its work focuses on scaling agricultural innovation, livelihoods, and ultimately increase food security across the continent.

[Read more](#)

# [www.fairprocessframework.org](http://www.fairprocessframework.org)

FAIR Process Framework

Why FAIR? The steps In action Search

## 1.0 Defining data intervention types

To lay the groundwork for a sustainable, data-driven project, define and document problems, goals, activities, challenges and solutions that enable FAIR data practices

FAIR Process Framework

Why FAIR? The steps In action Search

## 3.2 Identify output and input data assets

Identifying the output data and input data assets that produce the data

FAIR Process Framework

Why FAIR? The steps In action Search

## 5.2 Data Management and Access Plan (DMAP)

Developing a living document that outlines how data will be collected, stored, and handled throughout the project

FAIR Process Framework

Why FAIR? The steps In action Search

## 5.3 Data Sharing Agreements

Developing an agreement between two or more entities who are either providing or receiving data on how the shared data can be used

Step 1: Discovery

1.0 Defining data intervention types

1.1 Define your data problem

1.2 Develop your intention

1.3 Specify the data intervention type

1.4 Identify the data blockers

1.5 Develop data levers

1.6 Put it all together

Step 2: Understanding

Step 3: Planning

Step 4: Co-developing

Step 5: Strategy

Step 6: Implementing

Share Download

In this step you will:

Step 1: Discovery

Step 2: Understanding

Step 3: Planning

3.0 Identifying data assets

3.1 Plan the inventory

3.2 Identify output and input data assets

3.3 Establish attributes and metadata

3.4 Compiling the inventory

Step 4: Co-developing

Step 5: Strategy

Step 6: Implementing

5.0 Developing a FAIR data strategy

5.1 Data Governance Policy

5.2 Data Management and Access Plan (DMAP)

5.3 Data Sharing Agreements

Step 6: Implementing

Share Download

In this activity you will:

## Why should I do this?

To establish clear terms and conditions for the exchange of data between parties, ensuring privacy, security, and compliance with legal regulations.

If the details on the data and the conditions on use are clear, it prevents misuse and miscommunication among the entities involved in the transaction.

## In this activity you will:

Develop data sharing agreements (DSA) for your investment.

What are DSAs?

Impact

# Ethiopia's Soil Data Sharing Initiative

*Transforming Agricultural Data Management in Ethiopia*

Ethiopian authorities collected extensive soil data, crucial for agricultural decision-making, but by 2018, benefits were limited due to the lack of a national data-sharing policy. Informal connections existed, but there were no formal mechanisms for data sharing, hindering the full potential of the data.

CABI was invited by the foundation to **investigate and address data management challenges**. Collaborating with the ODI and GiZ, CABI facilitated development of a soil data-sharing policy directive.

CABI **applied FAIR principles** to ensure effective data management and sharing. This guided the development of a sustainable data-sharing framework and **national policy on sharing based on FAIR**.

In 2023, the Ministry of Justice ratified the Ministry of Agriculture's directive, leading to the **publication of the National Soil Information System (NSIS) and a culture shift in agricultural data sharing in Ethiopia**.

**Rise in Open Science collaborations: ca. 145 researchers share data and publish together** (inc. Nature Publication)

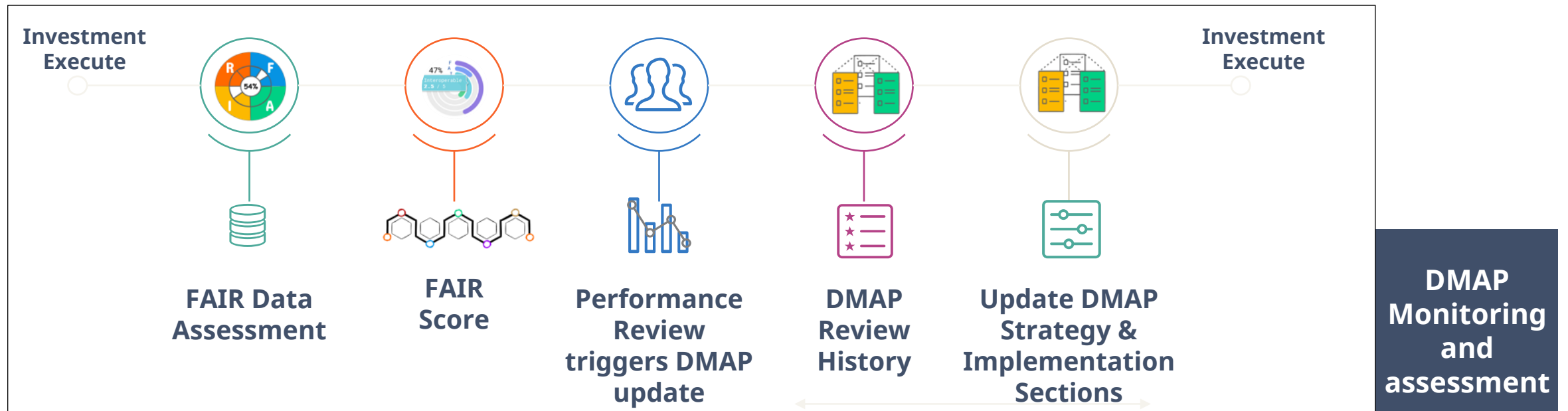
*Data policy development support leads to the ratification of a Ministerial directive on soil data sharing*



# FAIR Data Practices Across CGIAR Centers

*Tools for FAIR Data Target Setting and Assessments – with Data Management & Access Plans*

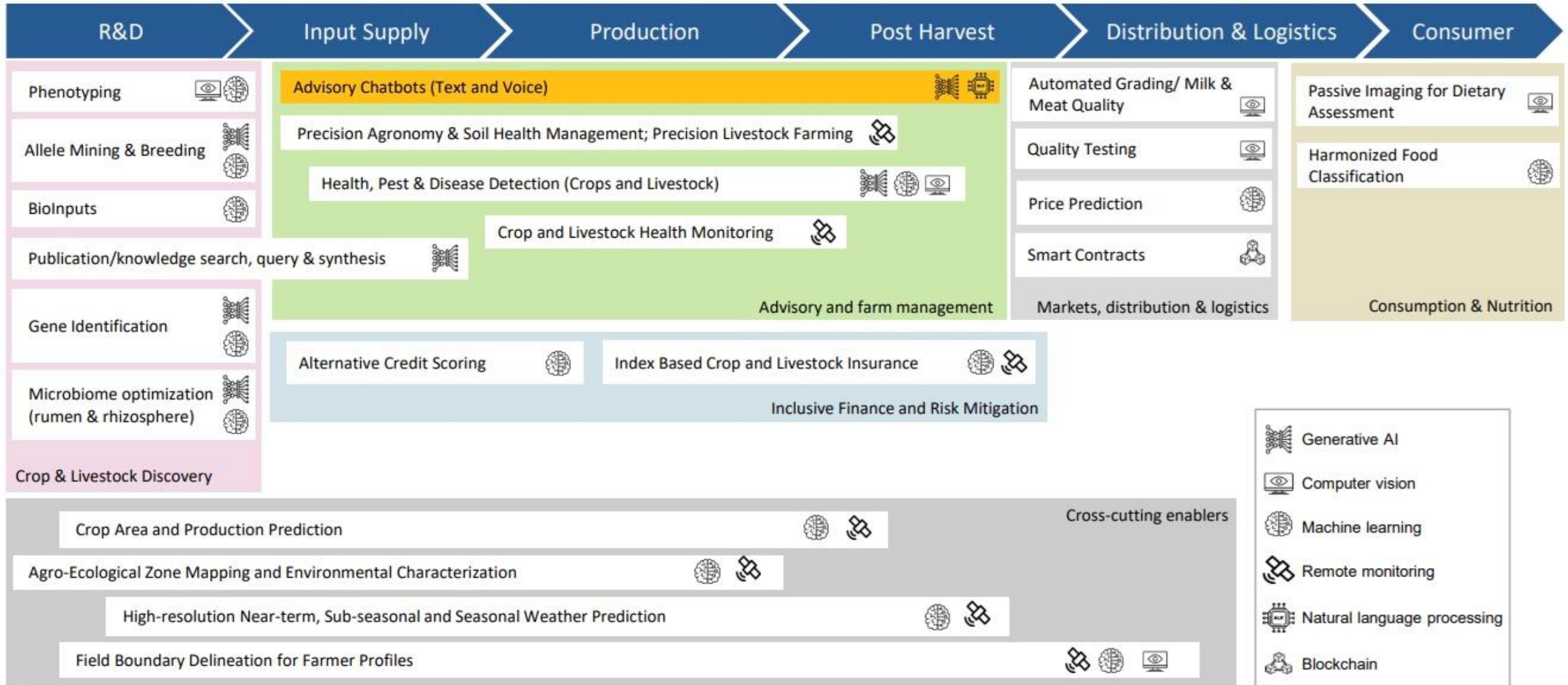
- CGIAR centres struggle to consistently implement good data practices in Gates-funded investments.
- CGIAR-specific guidelines, templates and tools, that enable data management aligned with FAIR principles.
- Resources help data practitioners, benchmark the current level of FAIR in a project, set targets for achieving FAIR, Data Management & Access Plans and then provide annual assessment reports on the FAIRness of project output.



DMAP = Data Management & Access Plan

# Traditional AI combined with generative AI presents new opportunities

LLM's and Transformer Architecture is being applied to a rapidly increasing number of domains beyond language and text



Note: This not an exhaustive list of AI solutions in Ag or AI technologies but illustrative of some of the areas AI is being used and GenAI has potential

# Explore the FAIR Process Framework: Your Go-To Resource Hub

Resources to operationalise FAIR and responsible data

[Data Management and Access Plan \(DMAP\) template](#)



[Data Sharing Agreement template](#)



[Data governance policy template](#)



[Guidance on data ecosystem mapping](#)



[Guidance on data standards, licensing](#)



[Guidance on making data inventories](#)



[Lists of Repositories and other data platforms/tools](#)



[Case studies](#)



[FAIR Primer](#)



# ORD from a Swiss perspective

**Dr. Igor Sarman, USI Research Data Manager (Research Service)**

**EUPHRESKO III open data webinar, May 13, 2025**

## About myself

Research Data Manager at Università della Svizzera italiana (USI, Lugano, southern Switzerland)

Master degree in Economics at Uni Trieste (Italy) and PhD in Economics at USI

I support USI researchers in all the Research Data Management issues they may have (in collaboration with other services)



# SNSF policy on Open Research Data

## SNSF Open Research Data Policy

- The Swiss National Science Foundation (SNSF) supports scientific research in all academic disciplines under different funding schemes
- The SNSF adopts an Open Research Data policy for the projects it funds
  - Research data should be **freely accessible to everyone** (scientists and the general public)
- The SNSF values research data sharing as a fundamental contribution to the **impact**, **transparency** and **reproducibility** of scientific research
  - Ever heard of Reproducibility crisis and how it is (slowly) leading to significant change in academic research?

## SNSF Open Research Data Policy

- SNSF expects data is carefully **curated** and **stored**, and **shared as openly as possible**
- Therefore, funded researchers are requested to:
  - ✓ **store the research data** they have worked on and produced in their research work
  - ✓ **share data** with other researchers, unless they are bound by legal, ethical, copyright, confidentiality or other clauses
  - ✓ **deposit data and metadata** onto existing public repositories in formats that anyone can easily find, access and reuse
- **SNSF grants public funds for research and expects that research artefacts become public!**

## Data sharing according to SNSF

- The SNSF expects funded researchers to share at least all data **underlying a publication**
  - As soon as the publication is available
  - Data that is necessary to **reproduce the results** presented in a publication
  - Researchers decide whether to share additional data
- SNSF expects that the same principles are followed for **code or software as well**
  - The code/software should be shared with the **underlying publication**
- **Reproducibility of results** and **reuse of research artefacts** are the guiding principles for data sharing

# PROs (and CONs) of ORD

## PROs of Open Research Data

- PROs from opening and sharing data:
  - ✓ Increase transparency and reproducibility of results
  - ✓ Generate research efficiencies, data reuse, new collaborations, projects and networks
  - ✓ Simplify and speed-up the research flow
  - ✓ Increase exposure of academic efforts and citations
  - ✓ Researcher's visibility do not only come from publications but from data too
  - ✓ The reuse of your data can be valued and acknowledged by other users
  - ✓ Old scientific metrics are questioned, new research assessment metrics will likely rely on OS/ORD practices
  - ✓ Data sharing can be seen as means of academic prestige
  - ✓ Sharing data is sharing knowledge, it benefits the community
  - ✓ Many Funders and Journals now request it

---

## CONs of Open Research Data

- CONs from opening and sharing data:
  - ✓ It needs collaboration efforts (and agreement) if data belongs to multiple actors
  - ✓ Time consuming and extra work
  - ✓ Needs knowledge on how to share data and make it interoperable
  - ✓ Needs good quality metadata
  - ✓ No concrete incentives (yet)

# Final considerations

## Some FAQs from researchers

- Where shall I publish my data? I have no idea!
  - Data repositories! You can choose from generalist and discipline-specific ones
  - <https://www.re3data.org/>
- I cannot publish my data because it is sensitive/confidential
  - **As open as possible, as closed as necessary**
  - Publish metadata and documentation so that people are aware of it
- I would really like to share my data but given its sensitivity I don't want everybody to access it
  - Apply a license or access conditions to your data (i.e., presentation of a research project, signature of a Data Sharing Agreement)

---

## Some FAQs from researchers

- I have too much data, I don't know what may be useful or not
  - Share at least the data underlying a publication (so that it can be reproducible)
  - At the beginning of your project think about what data deserves to be retained and shared, especially unique and non-reproducible information
- My data will eventually get lost in the plethora of databases
  - Data repositories assign unique DOIs to data collections, you can easily use them in publications to refer to your data
- What formats shall I use to share my data?
  - Consider open/standard file formats (.txt, .csv, .jpeg, .flac, .tiff, ...) and the use of community standards (if any) to grant long-term reusability

## General considerations and recommendations

- Start to work on ORD and with an ORD mentality since the beginning of the project
- Draft a Data Management Plan and use it as an awareness tool, it makes you think and consider data taken in a broader context!
  - DMPs make researchers realize if and to what extent they are **aware about what they are doing** with project data
- Make an effort to visualize ORD since the beginning of the project
- Regard data as an **important resource** that has to be carefully managed (just like research funds)
- Regard ORD as an **investment** and an **opportunity**, not as a burden
- **Make data open for yourself, your team and the community, not for the funder or the journal!**

## Contact

Dr. Igor Sarman, Research Data Manager

USI Research Service

[igor.sarman@usi.ch](mailto:igor.sarman@usi.ch)

058 666 4790

[USI Research Data Management webpage](#)

# Q&A

# Thank you!



# Global Agricultural Research Archive

Damian Bird, May 2025

# Open science

**STANDARDIZATION  
AND BEST  
PRACTICE**

**COLLABORATION  
AND INCLUSIVITY**

**TRANSPARENCY  
AND ACCESSIBILITY**

# Open science

**STANDARDIZATION  
AND BEST  
PRACTICE**



**COLLABORATION  
AND INCLUSIVITY**



**TRANSPARENCY  
AND ACCESSIBILITY**



# GARA

- Archives agricultural knowledge from around the world
- 8,000 full text PDFs including conference papers, journal articles, annual reports and bulletins
- Developed in collaboration with partners and contributors
- Available on CABI Digital Library, with advanced features and millions of users each month

The screenshot displays the GARA website interface. At the top, the CABI Digital Library logo is visible on the left, and navigation links for Home, About, Browse, and Help are on the right. The main header features the title "Global Agricultural Research Archive" in large white text against a background of golden wheat. Below the title is a subtitle: "An agricultural knowledge archive of scientific research to support knowledge extension and practical application in developing regions." A search bar is positioned below the subtitle, with a "Search" button and a link to "Advanced Search". Two buttons, "Browse Full Text" and "Browse Open Access", are located at the bottom of the search area.

The main content area is divided into two columns. The left column, titled "Latest Content", lists three recent abstract records with their titles and authors. The right column, titled "Overview", provides a brief description of GARA's purpose and a "Learn More" button.

Below the "Latest Content" section, a "View All Content" button is present. At the bottom of the page, a grid of seven thumbnail cards is displayed, each representing a regional archive: IAALD, Malawi, Myanmar, Pakistan, Philippines, RUFORUM, and CTA.

# Discovery and accessibility



Digitizes and protects less available material, adding value and context



Accessible through a modern platform, and integrated with other content



Standardised data means consistency downstream and better integration



Expert indexing drives the search and discovery for CABI users

# EUPHRESCO III site

- Launched in 2024
- 38 records to date
- Usage is growing steadily, and benefits from millions of users to platform as a whole

The screenshot shows the EUPHRESCO III website interface. At the top, there is a navigation bar with the CABI Digital Library logo, authentication status, a Bookshop button, a Browse dropdown, a search icon, a shopping cart icon, and a Login button. Below the navigation bar, the main header features the EUPHRESCO III title and a subtitle: "A knowledge repository of reports on national and international research activities conducted in EUPHRESCO III countries". A search bar with the placeholder text "Search EUPHRESCO III" and a Search button is prominently displayed. Below the search bar, there is a link for "Advanced Search". The main content area is divided into two columns. The left column is titled "Latest Content" and lists three abstract records with their titles, authors, and dates. The right column is titled "Overview" and provides a brief description of the EUPHRESCO III network and its purpose, with a "Learn More" button. At the bottom of the page, there is a "View All Content" button and the EUPHRESCO III logo.

**Latest Content**

Abstract record | 24 January 2025  
**Guide for determination of 12 *Agriilus* species native in Austria.**  
Werner Hinterstoisser, James Connell, Gernot Hoch

Abstract record | 18 July 2024  
**Modern tools and approaches to tackle pests in the Mediterranean crop production systems.**  
N. T. Papadopoulos

Abstract record | 17 July 2024  
**The International Association for the Plant Protection Sciences.**  
G. Norton, E. A. Heinrichs, M. Tamo

Abstract record | 17 July 2024  
**Plant Health research coordination: an international endeavour. XX International Plant Protection Congress Healthy Plants Support Human Welfare, 3 July 2024.**

[View All Content](#)

**Overview**

EUPHRESCO III is an international network of organizations that develop, own and manage phytosanitary research programmes, fund and undertake research activities and use research provisions for policy development and implementation. This resource provides a repository of national and international phytosanitary research activities conducted in EUPHRESCO III countries.

[Learn More](#)

EUPHRESCO III

# CTA

- CTA: Technical Centre for Agricultural and Rural Cooperation (Wageningen, the Netherlands)
- EU project ended in 2020
- CABI digitized and indexed material and made available to the world
- More than 3000 articles, chapters and reports

The screenshot shows the CABI News website. At the top, there is a navigation bar with the CABI logo and links for 'About CABI', 'News and blogs', 'Careers', 'Contact us', 'Members', and 'CABI centres'. A 'Sign up for updates' button is in the top right corner. Below the navigation bar, there are tabs for 'What we do', 'Products and services', and 'Projects'. A search bar is located in the top right corner. The main content area features a large image of people carrying baskets, with the title 'CABI News' overlaid. Below the image, there is a breadcrumb trail: 'You are here: Home / News Articles / CTA's 35-year legacy of agricultural knowledge safe in CABI's hands'. The date 'November 17, 2020' is displayed. The main headline is 'CTA's 35-year legacy of agricultural knowledge safe in CABI's hands'. Below the headline, there is a navigation bar with links for 'All news and blogs', 'CABI News', 'CABI Blog', 'PlantwisePlus Blog', 'Invasives Blog', and 'Media Centre'. The 'CABI News' link is highlighted. Below the navigation bar, there is a large image of a library with bookshelves. To the right of the image, there is a 'Join our mailing list' section with a 'Sign up' button. Below that, there is a 'Media requests' section with contact information for Wayne Coles. At the bottom, there is an 'Archives' section with a 'Select Month' dropdown menu and a 'Categories' section.

**CABI News**

You are here: Home / News Articles / CTA's 35-year legacy of agricultural knowledge safe in CABI's hands

November 17, 2020

## CTA's 35-year legacy of agricultural knowledge safe in CABI's hands

All news and blogs CABI News CABI Blog PlantwisePlus Blog Invasives Blog Media Centre

### Join our mailing list

Sign up to receive the latest news, information, updates and offers from CABI.

Sign up

### Media requests

If you are a journalist and are looking for more information or for interviews please contact Wayne Coles or visit the Media Centre

### Archives

Select Month

### Categories

# Content types

Content can be gathered into one place for all outputs, including:

- Journal articles
- Workshop and conference proceedings
- Success stories and case stories from joint projects
- Research reports and publications

The screenshot shows the EUPHRESKO III website. At the top, there is a navigation bar with the CABI Digital Library logo, authentication status, a Bookshop button, a Browse dropdown, a search icon, a shopping cart icon, and a Login button. Below the navigation bar, the main header features the text 'EUPHRESKO III' in large white letters, followed by the subtitle 'A knowledge repository of reports on national and international research activities conducted in EUPHRESKO III countries'. A search bar with the placeholder text 'Search EUPHRESKO III' and a 'Search' button is prominently displayed. Below the search bar, there is a link for 'Advanced Search'. The main content area is divided into two columns. The left column is titled 'Latest Content' and lists three items: 1. 'Abstract record | 24 January 2025' with the title 'Guide for determination of 12 *Agriilus* species native in Austria.' and authors 'Werner Hinterstoisser, James Connell, Gernot Hoch'. 2. 'Abstract record | 18 July 2024' with the title 'Modern tools and approaches to tackle pests in the Mediterranean crop production systems.' and author 'N. T. Papadopoulos'. 3. 'Abstract record | 17 July 2024' with the title 'The International Association for the Plant Protection Sciences.' and authors 'G. Norton, E. A. Heinrichs, M. Tamo'. At the bottom of this list is a 'View All Content' button. The right column is titled 'Overview' and contains a paragraph describing EUPHRESKO III as an international network of organizations that develop, own and manage phytosanitary research programmes, fund and undertake research activities and use research provisions for policy development and implementation. It also states that this resource provides a repository of national and international phytosanitary research activities conducted in EUPHRESKO III countries. Below the text is a 'Learn More' button. At the bottom right of the page, there is a 'Feedback' button and the EUPHRESKO III logo.

# GARA

**OPEN ACCESS**



**DIGITISED  
MATERIALS**



**REGION AND  
PROJECT  
COLLECTIONS**







**EUPHRESKO III**

Network for phytosanitary  
research coordination and funding

Euphresco as a  
knowledge hub for the  
exchange of information

Baldissera GIOVANI, Euphresco/EUPHRESKO III



This project has received funding from the European Union's HORIZON-WIDERA-2023-ERA-01-01 under grant agreement No 101130467



# The objectives of Euphresco since 2006

Annex 1. Description of work

04/08/2006

Proposal: 036212 EUPHRESKO

## 1. PROPOSAL SUMMARY

### Strategic Objectives Addressed

EUPHRESKO will address the strategic objective of the Framework 6 Programme *'Integrating and Strengthening the European Research Area'* by *'Supporting the Cooperation and Coordination of Research Activities Carried out at National and Regional Levels'*.

*In Europe, despite there being coordination at the EU level for the development of EU Plant Health Policy and complementary EU initiatives to increase phytosanitary capacity, there is no coordination of national phytosanitary research programmes at the EU level to underpin and support these activities.*

EUPHRESKO therefore aims to increase cooperation and coordination of national phytosanitary research programmes at the EU level through networking of research activities and mutual opening of national programmes, to achieve three main over-arching strategic goals:

Develop phytosanitary (statutory plant health) research policy at the EU-wide level.

- Optimise the research provision that underpins EU quarantine plant health policy development and policy implementation, in an era of increasing biosecurity threats from alien plant pests, diseases and invasive species.
- Increase the capacity of European phytosanitary science and research, in order to prevent the disappearance of EU expertise in this field and maintain Europe's competitiveness in the global market.

Beca

DataCite Annual

## NWO & FAIR research and workflows

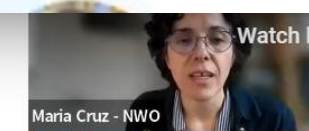
Grant requirements and guidance for researchers

Incentives and rewards for researchers

Adopting PIDs in own funder workflows

[Sans titre]

<https://datacite.org/event/the-role-of-funders-in-building-a-robust-and-trustworthy-output-tracking-mechanism-using-pids-and-open-metadata/>



makers

UPHRESO III

# Why?

Because it is the pertinence of research funders and policy makers to set research policy

- **Principles and Standards that underpin research**
  - Ethical, legal and professional norms such as data protection, consent, confidentiality, plagiarism, authorship, conflict of interest, social responsibility
- **Resources and support**
  - Human, financial, **technical and physical resources that enable** the conduction and **dissemination of research**
- **Processes and procedures**
  - **Methods to design, conduct, analyze and report research**
- **Collaboration and innovation**
  - Strategies and practices to foster creativity, diversity and excellence in research
- **Communication and impact**
  - **Approaches and channels to share, disseminate and translate research findings, outputs and outcomes**

# Key benefits since 2006

EUPHRESKO will enable better coordination and structuring of phytosanitary research in Europe (Figure 2.1), with the following key benefits realised:

- Efficient deployment of national funds for both national and trans-national (joint) research.
- Better coordination of research priorities suitable for EU Framework programme funding.
- Better availability of research data for Plant Health Policy makers.
- Better consultation with EC policy makers and other European plant health bodies, e.g. EPPO/EFSA, Non-EU bodies (e.g. RPPOs), and other stakeholders (including 'industry' bodies).

## Optimisation of Results through Exploitation and/or Dissemination Plans

A communications strategy will be developed early in the project. However, the activities outlined below are envisaged.

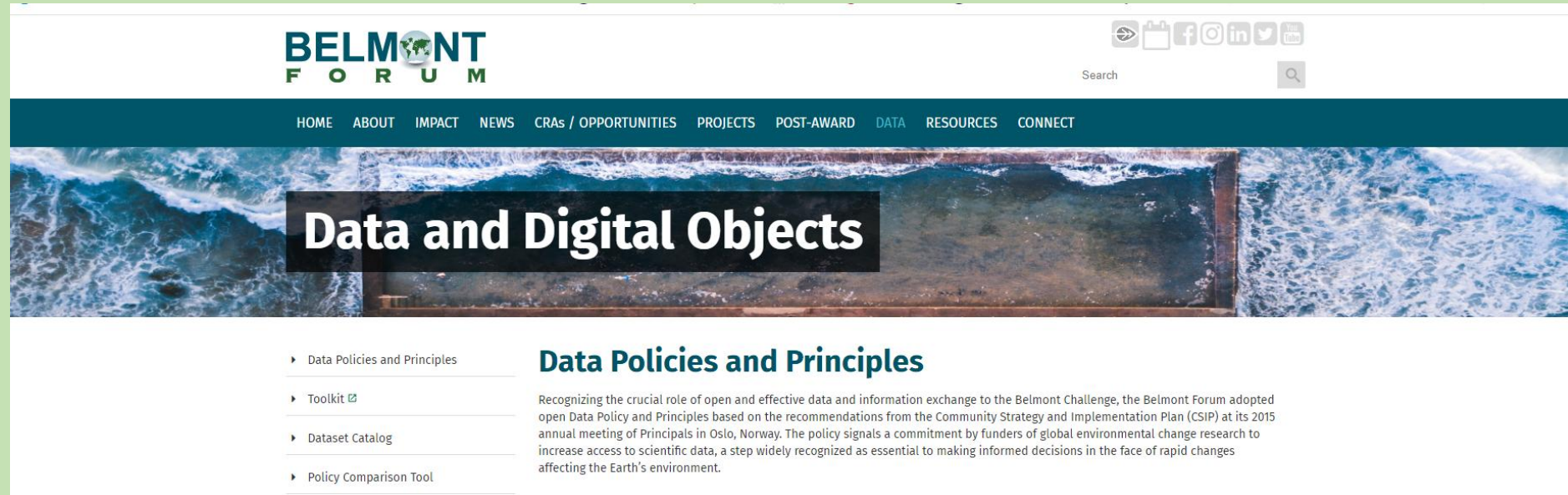
- EUPHRESKO will, for the first time, collate information on national programmes, resources and infrastructures, and make this publicly available via a database on its website. This will be updated yearly throughout the project and beyond the end of the ERA-Net. Where possible, national research reports will also be made available via the website, either directly or through links to national websites. Data and information will therefore be more readily accessible to policy makers and scientists.

# Why?

Because it is the mandate of Euphresco to produce research evidence for plant health stakeholders

Because sharing information on research activities is an effective way to coordinate research

# Why?



The screenshot shows the Belmont Forum website. The header includes the Belmont Forum logo, social media icons, and a search bar. The navigation menu lists: HOME, ABOUT, IMPACT, NEWS, CRAs / OPPORTUNITIES, PROJECTS, POST-AWARD, DATA, RESOURCES, and CONNECT. The main banner features a blue ocean background with a dark box containing the text "Data and Digital Objects". Below the banner, there is a list of links on the left and a main content area on the right.

**BELMONT FORUM**

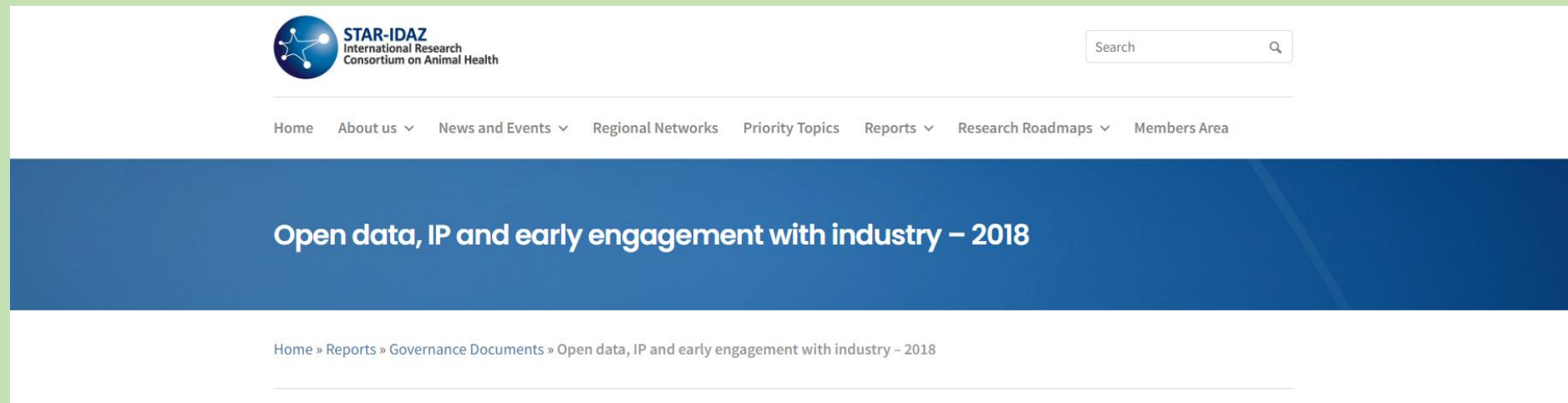
HOME ABOUT IMPACT NEWS CRAs / OPPORTUNITIES PROJECTS POST-AWARD DATA RESOURCES CONNECT

## Data and Digital Objects

- ▶ Data Policies and Principles
- ▶ Toolkit [↗](#)
- ▶ Dataset Catalog
- ▶ Policy Comparison Tool

### Data Policies and Principles

Recognizing the crucial role of open and effective data and information exchange to the Belmont Challenge, the Belmont Forum adopted open Data Policy and Principles based on the recommendations from the Community Strategy and Implementation Plan (CSIP) at its 2015 annual meeting of Principals in Oslo, Norway. The policy signals a commitment by funders of global environmental change research to increase access to scientific data, a step widely recognized as essential to making informed decisions in the face of rapid changes affecting the Earth's environment.



The screenshot shows the STAR-IDAZ website. The header includes the STAR-IDAZ logo, a search bar, and a navigation menu. The main banner features a dark blue background with white text. Below the banner, there is a breadcrumb trail and a horizontal line.

**STAR-IDAZ**  
International Research Consortium on Animal Health

Home About us News and Events Regional Networks Priority Topics Reports Research Roadmaps Members Area

## Open data, IP and early engagement with industry – 2018

Home » Reports » Governance Documents » Open data, IP and early engagement with industry – 2018

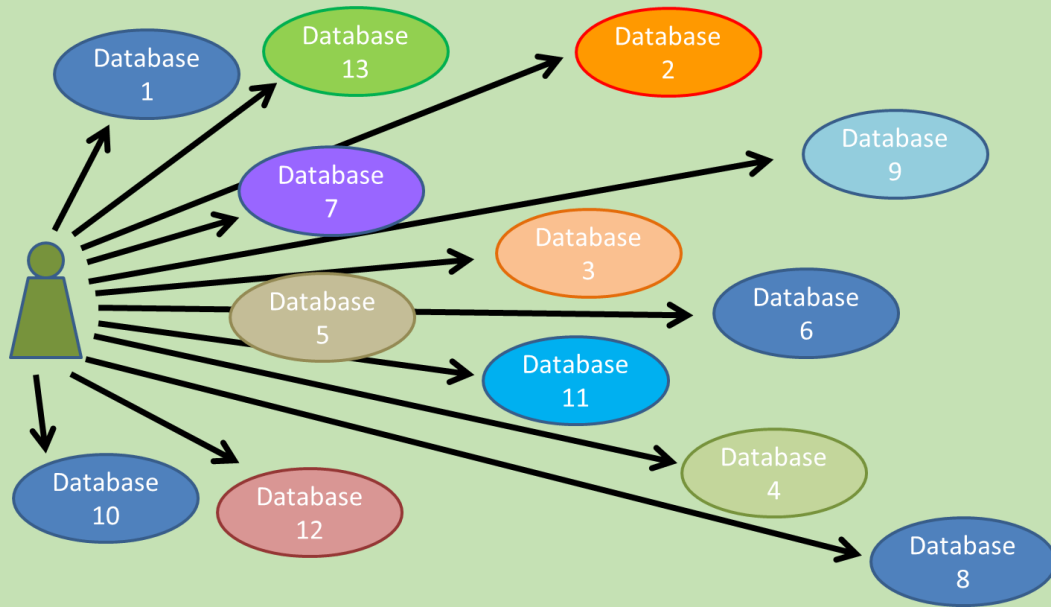
# Recommendations evaluation 2014

In the future, as EPPO takes on the role of hosting the EUPHRESKO network we recommend that they consider:

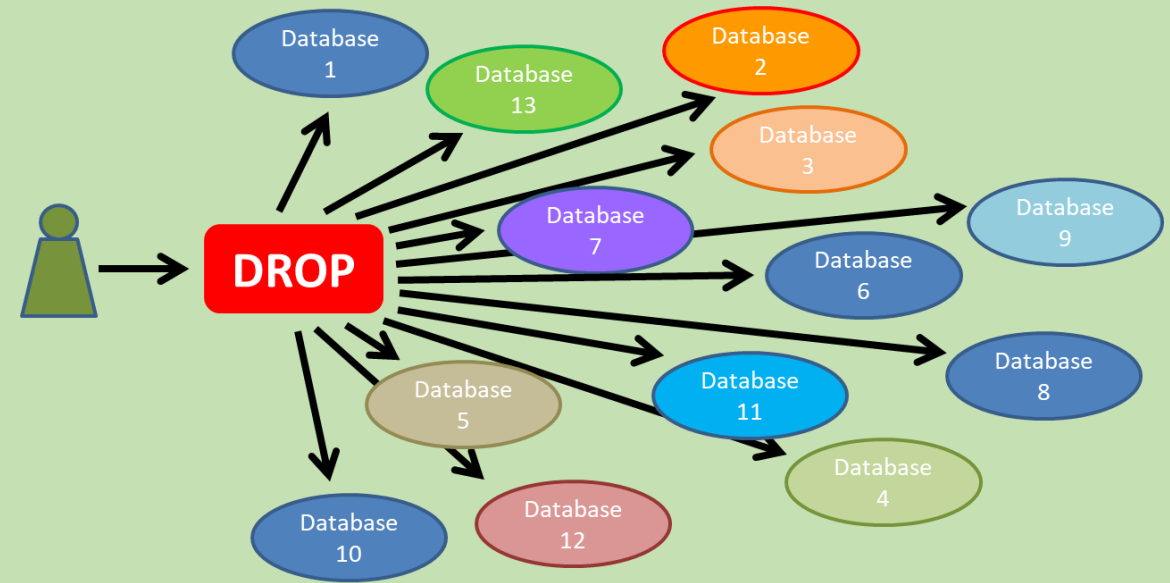
- (a) How it can become a proactive knowledge exchange hub for the collection and dissemination of research findings across Europe which is accessible to and effectively utilised not only by scientists but also by practitioners, policy makers and stakeholders.

.....

# Digital Research Object Portal (DROP)



The large number of diverse infrastructures has the counter-effect to scatter information which may be more difficult to reach, or which is 'hidden'



DROP works as a **node**, a unique entry point that **facilitates retrieval** of digital research objects (data & documents)

# Open digital research objects



A digital object is an information object, of any type of information or any format, that is expressed in digital form

- Simple: made up of a single file, such as a PDF or an image
- Complex: made up of multiple files, such as a website

**Text**  
Protocols  
Articles

**Dataset**  
Lists  
Tables  
Spreadsheets  
Databases

Software

Model

**Image**  
Photographs  
Videos  
Graphics  
Maps

**Event**  
Conferences  
Workshops  
Webinars

# Visibility of data: DROP vs Zenodo

December 13, 2016 (v1) Conference paper Open

## Transcriptome profiling of two olive cultivars in response to *Xylella fastidiosa* infection

Giampetruzzi A., MORELLI M., Saponari M., Loconsole G., Chiumenti M., Boscia D., Savino V.N., Martelli G.P., Saldarelli P.,

Olive Quick Decline Syndrome (OQDS) is a destructive disease of olive (*Olea europaea*) described in Apulia (Italy) to which the CoDiRO strain of *Xylella fastidiosa* subsp. *pauca* (Xfp) is associated. Differently from the severely affected cv. Ogliarola salentina, Xfp-infected plants of the cv. Leccino develop reduced symptoms, which are compatible...

Uploaded on September 1, 2018 | Published in: Proceedings of 1st Annual Meeting of H2020 POnTE (Pest Organisms Threatening Europe) Project. Madrid, Spain, 12-13 December 2016, 58, 2016.

36 Views 20 Downloads

**38 views on ZENODO**

December 1, 2019 (v1) Conference paper Open

## Spatial Bayesian modelling applied to the surveys of *Xylella fastidiosa* in Alicante

Cendoya, Martina; Martínez-Minaya, Joaquín; Ferrer, Amparo

Spatial Bayesian modelling applied to the surveys of *Xylella fastidiosa* in Alicante

Uploaded on December 1, 2019

36 Views 19 Downloads

**60 views on ZENODO**

## Olea europaea Transcriptome LC6

EUPHRESKO SECRETARIAT 636 VIEWS

Dataset

Authors

Giampetruzzi, Annalisa; Saponari, Maria; [Saldarelli, Pasquale](#)

**636 views on DROP**

## Data and codes for the *Xylella fastidiosa* remote sensing study

EUPHRESKO SECRETARIAT 974 VIEWS

Collection

Authors

Zarco-Tejada, P.J.; Camino, C.; Beck, P.S.A.; Calderon, R.; Hornero, A.; Hernández-Clemente, R.; Kattenborn, T.; Montes-Borrego, M.; Susca, L.; Morelli, M.; Gonzalez-Dugo, V.; North, P.R.J.; Landa, B.B.; Boscia, D.; Saponari, M.; Navas-Cortes, J.A.

**974 views on DROP**

# Benefits of using DROP

- **Reduces fragmentation**
  - Makes accessible information available in various specialized databases
  - By referencing data using agreed metadata, DROP smoothens the diversity associated to different databases
- **Increases knowledge accessibility**
  - Makes data more accessible and exploitable
  - Makes expertise and scientific production more visible
- **Enhances data operations**
  - Encourages (multidisciplinary) collaboration
  - Instructs policymaking and research funding
- **Harmonizes practices**
  - Facilitates the standardization of data sharing practices within the phytosanitary community
  - Promotes discussions on important data to open, appropriate infrastructures and metadata ontology (consensus building)

[www.euphresco.net](http://www.euphresco.net)

[www.phrescoglobal.net](http://www.phrescoglobal.net)



#Euphresco

For information:

[bgiovani@euphresco.net](mailto:bgiovani@euphresco.net)

