WorldFAIR and FAIR Implementation Profiles:
what have we learnt?

Simon Hodson
Executive Director
CODATA
www.codata.org

WorldFAIR Workshop
EOSC Symposium, Prague
14 November 2022

WorldFAIR: Global cooperation on FAIR data policy and practice

- Two year project to advance implementation of the FAIR principles in a range of disciplines, or cross-disciplinary research areas.

- Funded by the European Union, HORIZON-WIDERA-2021-ERA-0 — Project: 101058393

- Two years from 1 June 2022.

- Nineteen partners.

- Institutions in France, Belgium, Cyprus, Denmark, Germany, Ireland, Norway; Australia, Brazil, Kenya, New Zealand, USA; UK.

- Important partnership between CODATA and Research Data Alliance.
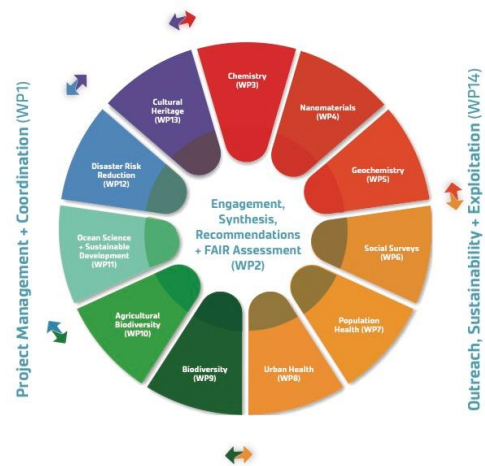
Making Data Work for Cross-Domain Grand Challenges

- **ISC Action Plan entrusts CODATA with an initiative 'Making Data Work for Cross-Domain Grand Challenges': establish a global (decadal) programme to address these issues.**

- The major, pressing global scientific and human issues of the 21st century can **ONLY** be addressed through **research that works across disciplines to understand complex systems**, and which uses **interdisciplinary and transdisciplinary** approaches to turn data into knowledge and then into action.

- **Preparatory Phase:** exploring technical issues and case studies through Dagstuhl workshops, TGs and WGs, funded projects etc, to understand the challenges and prepare the programme.

- **Core Interoperability Framework:** units, vocabularies, data structure, data description…

- **Case Studies:** in a range of domain and cross-domain research areas.

- **WorldFAIR Project:** takes this approach forward in the context of an EU-funded project.

- **Approach extended to Global Open Science Cloud (GOSC) Initiative.**

Preparatory Activities

1. **Series of exploratory workshops**

   - Two workshops in 2017, in Paris (ISC-CODATA) and London (Royal Society) launched the activity.

   - Series of Dagstuhl Workshops: 2018, 2019, 2021 (cross-domain and vocabularies) and 2022!

2. **Related outputs and activities**

   - DDI Alliance work on DDI-CDI

   - Cox et al. group (formed at Dagstuhl) produced '10 Simple Rule for Making a Vocabulary FAIR' https://arxiv.org/abs/2012.02325

   - IUSSP-CODATA WG on FAIR Vocabularies in Population Science http://bit.ly/IUSSP-CODATA-FAIR-Vocabs

   - CODATA TG on DRUM (Digital Representation of Units of Measure)

3. **Continuing engagement with various case studies**

   - Including infectious disease/epidemiology work on Ebola, COVID; urban health (including ISC UHWB Programme and Salurbal); DRR including ISC activities and CODATA TG; case studies from DDI-CDI report; and more.

The Role of DDI-CDI in EOSC: Possible Uses and Applications

The Role of DDI-CDI in EOSC: Possible Uses and Applications

Authors

Arofan Gregory, Simon Hodson, and Joachim Wackerow

Please cite this document as: Gregory, A., Hodson, S., Wackerow, J., 2021, 'The Role of DDI-CDI in EOSC: Possible Uses and Applications', https://doi.org/10.5281/zenodo.4707263

Acknowledgment

This work was supported by the EOSC Secretariat.

EOSCsecretariat.eu has received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-2018-4, Grant Agreement number 831644.

- DDI-CDI (Cross Domain Integration) is designed to interface with other standards and to help interoperability and integration of data between different data types, standards, formats.

- EOSC co-creation project to explore uses and applications and make recommendations for EOSC and the specification.

- **The Role of DDI-CDI in EOSC: Possible Uses and Applications, final report** **https://doi.org/10.5281/zenodo.4707263**

- Report examines the challenge for EOSC: issues of scale and the challenge of cross-domain data reuse.

- Presents examples of how DDI-CDI can be used in data integration and cross domain use cases; describes how DDI-CDI can interact with other standards, notably DCAT; describes how DDI-CDI fits into a FAIR Ecosystem of FAIR Digital Objects; describes how DDI-CDI can be implemented in the Dataverse platform.

- Series of webinars to assist review of specification: http://bit.ly/DDI-CDI-webinar-series

- **Production version available in early 2023.**

WorldFAIR

- Among the most important, but most challenging, recommendations of the **Turning FAIR into Reality** report, is R.4:

- 'Develop **interoperability frameworks** for FAIR sharing within disciplines and for interdisciplinary research: Research communities need to be supported to develop interoperability frameworks that define their practices for data sharing, data formats, metadata standards, tools and infrastructure. To support interdisciplinary research, these interoperability frameworks should be articulated in common ways and adopt global standards where relevant.'

- UNESCO Recommendation on Open Science and ISC Action Plan make similar recommendations.

- Exploring precisely this objective through the CODATA-ISC Decadal Programme and WorldFAIR project.

European Commission

Final Report and Action Plan from the European Commission Expert Group on FAIR Data

TURNING FAIR INTO REALITY

Research and Innovation

2018

COMMITTEE ON DATA

CODATA

INTERNATIONAL SCIENCE COUNCIL

International Science Council

WorldFAIR Case Studies

- **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 in Salurbal.

- **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)
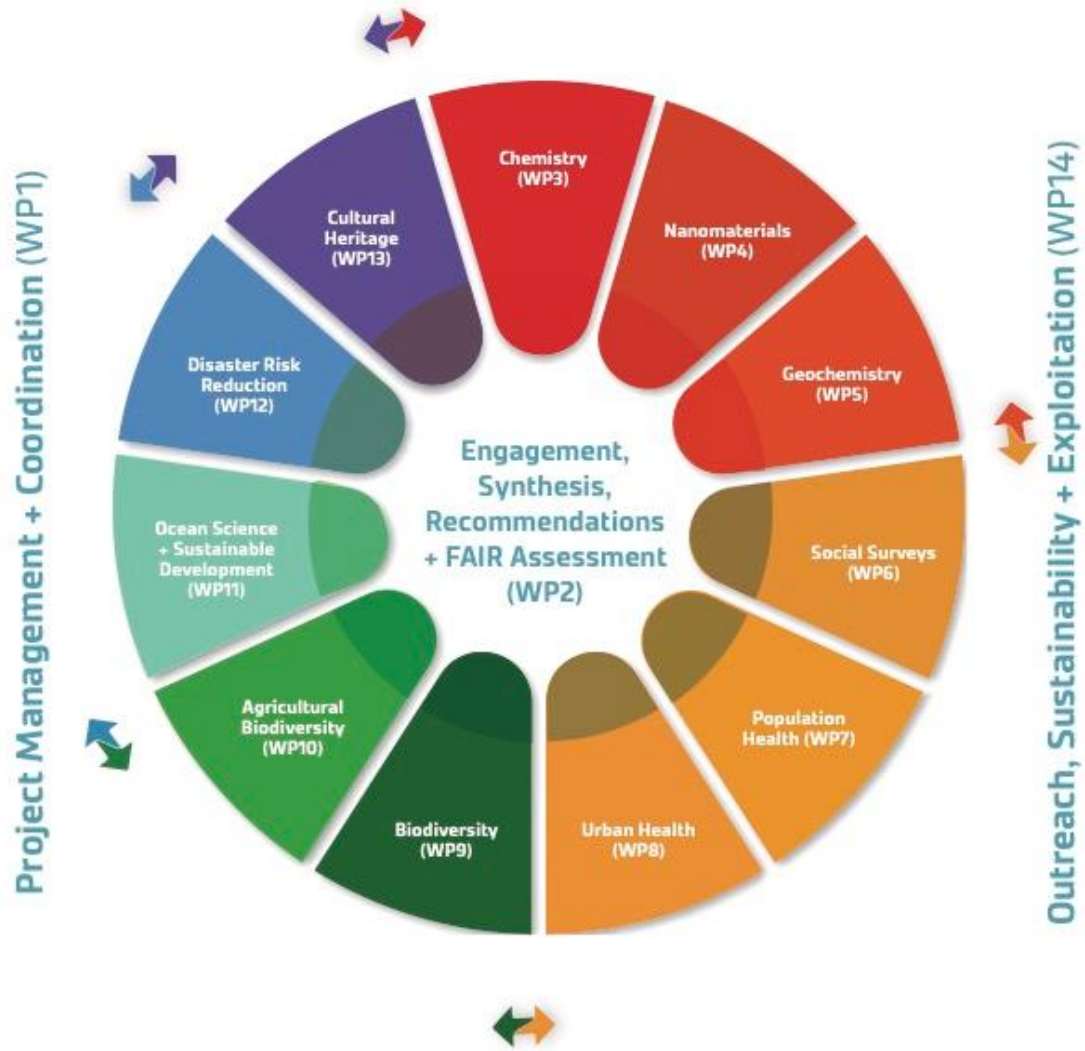
WorldFAIR and FAIR Implementation Profiles (FIPs)

- Each case study will explore aspects of FAIR implementation in their domain and cross-domain research area.

- Synthesis WP02 will explore **FAIR Implementation Profiles**, Cross-Domain Interoperability Framework and Recommendations for Domain Sensitive FAIR Assessment.

- **FAIR Implementation Profiles** (how do you implement FAIR, what are the **FAIR Enabling Resources** that you use?)

- Set of questions that allow a 'community' to state the FAIR Enabling Resources it uses to make data and metadata FAIR for each of the principles.

- Useful exploratory approach, encourages discussion in a community.

- Supported by an online tool: https://fip-wizard.ds-wizard.org/
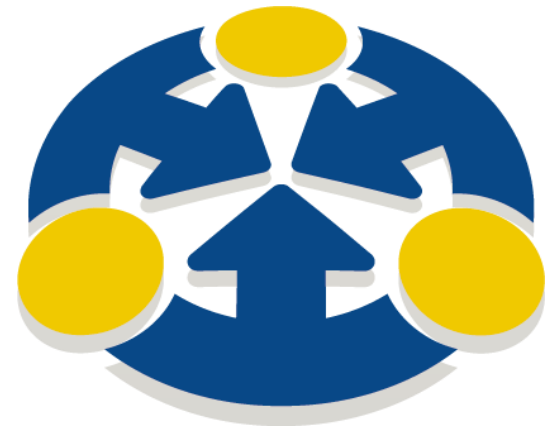
FIPs: Important Concepts

- Question set: what **FAIR Enabling Resource (FER)** (standards, tools, services) are you using for each FAIR Principle.

- For most principles an answer is requested for **data** and for **metadata**.

- Creates a de facto **typology** of FERs, **types of FERs**, based on the function they perform.

- This in turn, usefully informs a **functional breakdown** of the FAIR principles.

    - However, some existing FERs perform more than one function in relation to the FAIR principles.

- FIPs help provide a benchmark for a communities practice.

- **FIPs performed early and late in the project, will allow us to show any changes in practice or aspiration among the Case Studies.**

Cross-Domain Interoperability Framework

- FIPs activity prepares WorldFAIR for two 'WP2' outputs: **Cross-Domain Interoperability Framework** and **recommendations for domain sensitive FAIR assessment.**

- Exploring features of a **Cross-Domain Interoperability Framework** with case studies from a range of research areas.

    - Units and quantities/measurands/properties; terminologies/vocabularies/ontologies;

    - Data description/variable cascade, data structure;

    - Provenance and processing; data types; licences; protection and programmatic access…

    - …

- A number of webinars / conference sessions: GOSC DataIO Webinar; SciDataCon Session; DCMI Conference.

- Dagstuhl Workshop: discussion paper, report and links at https://worldfair-project.eu/

- WG and AG being set up to support the WorldFAIR work on this.
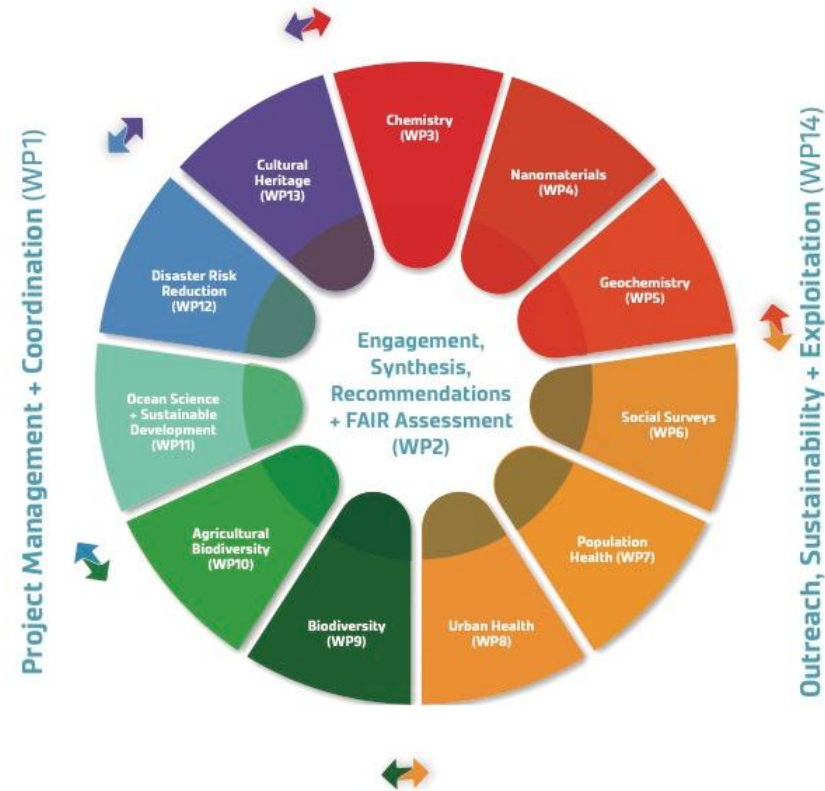
Interoperability for Cross-Domain Research

- Recent CODATA-DDI workshop at Schloss Dagstuhl, Leibnitz Foundation Centre for Informatics.

- Continued the methodology of exploring interoperability and data integration in a set of case studies, while also trying to develop the CDIF.

- Included some of the WorldFAIR case studies.

- **Cross-domain data combination (social science and environmental data)**

- **Sensitive data (fine grained data management)**

- **Combining oceans and assets-loss data in DRR**

- **Cross-Domain Interoperability Framework**

- Short summary report available at https://worldfair-project.eu/

Domain Sensitive FAIR Assessment

- Recommendations for **domain-sensitive FAIR assessment**.

- Need for more targeted, domain-sensitive recommendations / examples of good practice.

- Derive domain-sensitive FAIR assessment from the FIPs.
    - E.g. what is current practice, what would a recommendation for better practice be?
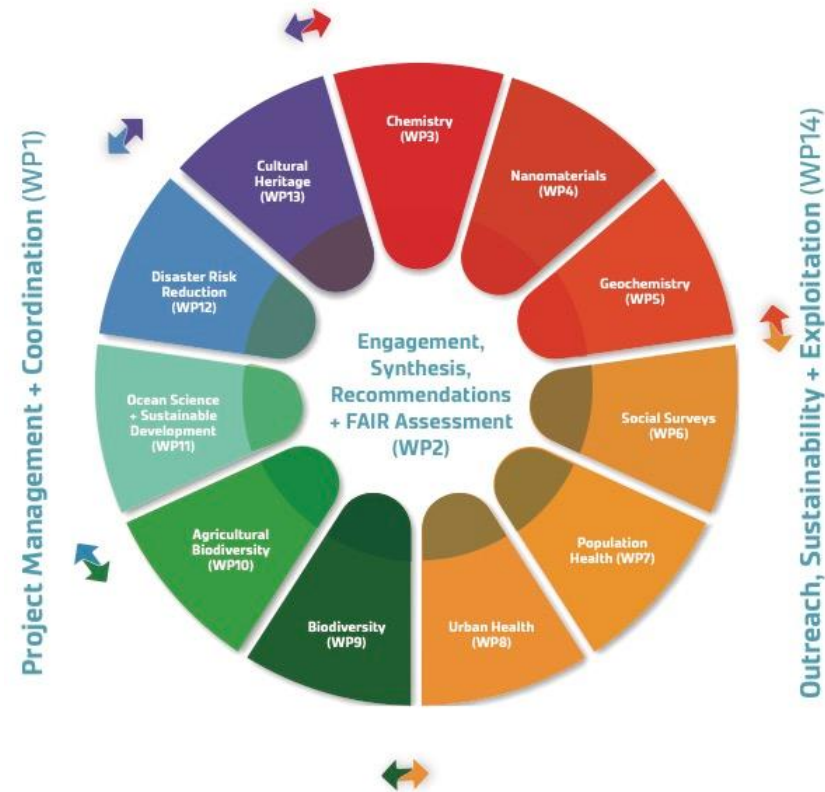
FIPs: what have we learnt?

1. What have you learnt from the process?

2. Has using FIPs helped you describe practices around FAIR in your case studies?

3. Has it helped identify any gaps or areas which would benefit from further attention?

4. Has the process identified ways in which the FIPs methodology and the tools around it can be improved?

5. What have you learnt about the FAIRness of your community or domain?

6. Have you identified any next steps in response to what you have learnt?
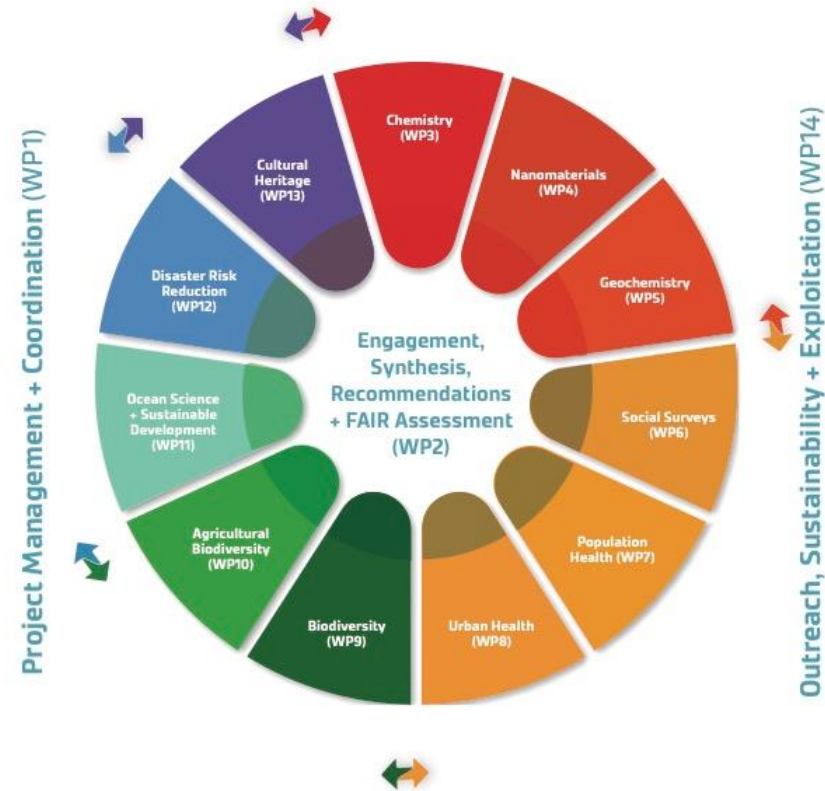
Exercise with the FIPs Questionaire

1. Look at the questions from the perspective of a research project/infrastructure that you are involved with.

2. Try and identify the **FAIR Enabling Resources** you use in response to some of the questions.

3. Use this spreadsheet template https://bit.ly/EOSC-Symposium-FIPs

4. Rename / create new sheet.

5. Discuss and ask questions!

COMMITTEE ON DATA
CODATA
INTERNATIONAL
SCIENCE COUNCIL

International
Science Council

Thank you for your attention

Simon Hodson, CODATA
www.codata.org
simon@codata.org
@simonhodson99 ; @CODATAnews