WorldFAIR

WorldFAIR: scientific and societal impact

Simon Hodson, CODATA

Alex Delipalta, RDA Europe

COMMITTEE ON DATA
CODATA
INTERNATIONAL SCIENCE COUNCIL

'Global cooperation on FAIR data policy and practice' (WorldFAIR) has received funding from the European Union's Horizon Europe project call HORIZON-WIDERA-2021-ERA-01-01, grant agreement 101058393. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union.

WorldFAIR: Global cooperation on FAIR data policy and practice

- Advances in FAIR implementation in **cross-domain** scenarios, in 11 specific **disciplines** and **globally**.
- **Global** in approach, because research domains, data and metadata standards and specifications need to be global. Leveraged **CODATA and RDA networks** to achieve this.
- Includes **authoritative international entities** (e.g. IUPAC, OneGeochemistry, GBIF, ODIS); connections with important projects or standards organisations (e.g. NanoCommons, DDI Alliance, OHDSI, TDWG, SalUrbAL).
- Considerable emphasis on **case studies** and the recommendations from these organisations.
- Leveraged links to international standards and scientific organisations, as well as reliable articulations of good (web) practice to make cross-domain recommendations. **Were able to have funded partners outside the European Union.**
- Helps reinforce **bidirectional links between EOSC and global developments**.

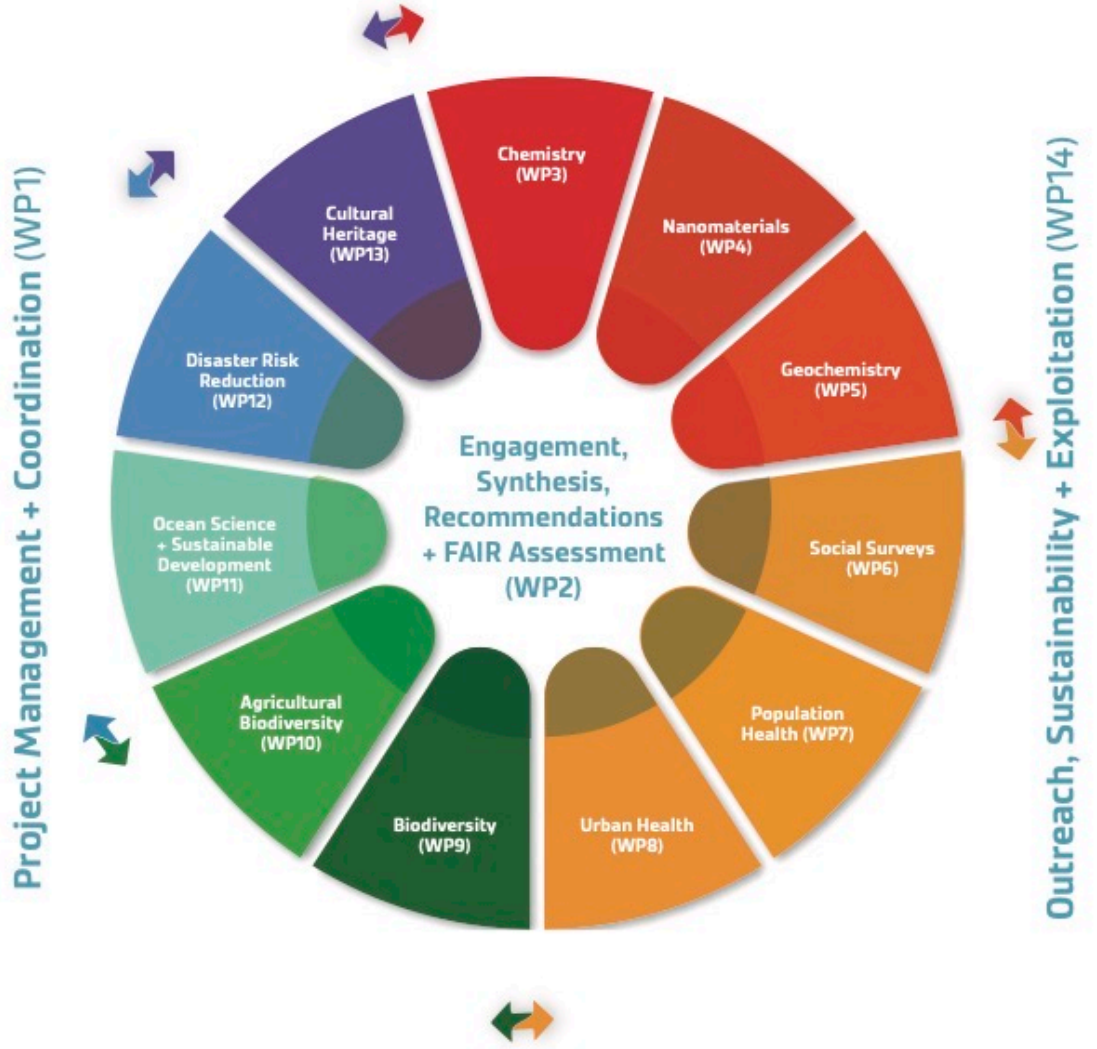
- **Funded by the European Union, HORIZON-WIDERA-2021-ERA-0 — Project: 101058393**

Interoperability Frameworks

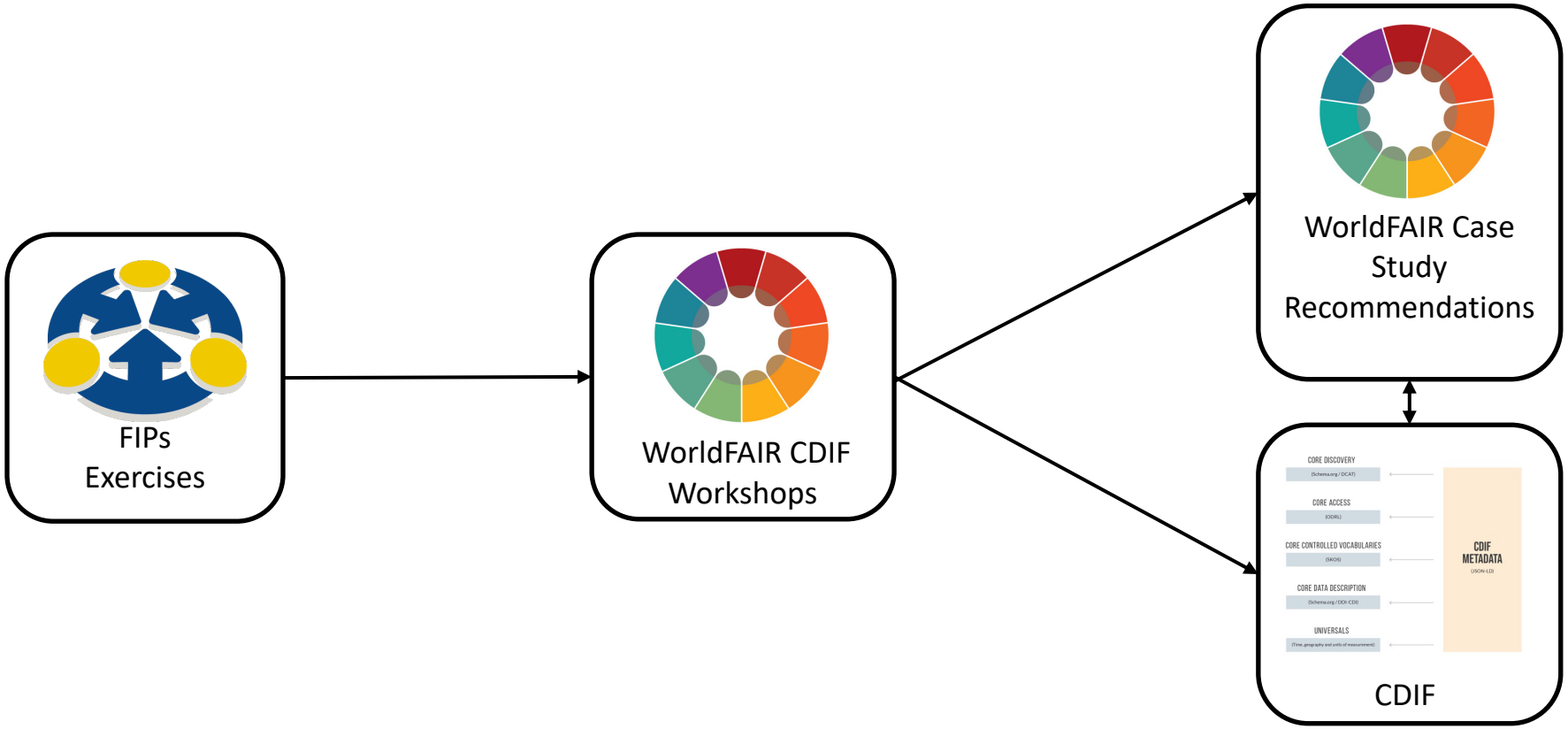
- Among the most important, but most challenging, recommendations of the **Turning FAIR into Reality** report, is R.4:
- ‘Develop **interoperability frameworks** for FAIR sharing within disciplines and for interdisciplinary research: Research communities need to be supported to develop interoperability 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.**’
- Very strong focus on the I and the R of FAIR.
- Core driver of WorldFAIR project

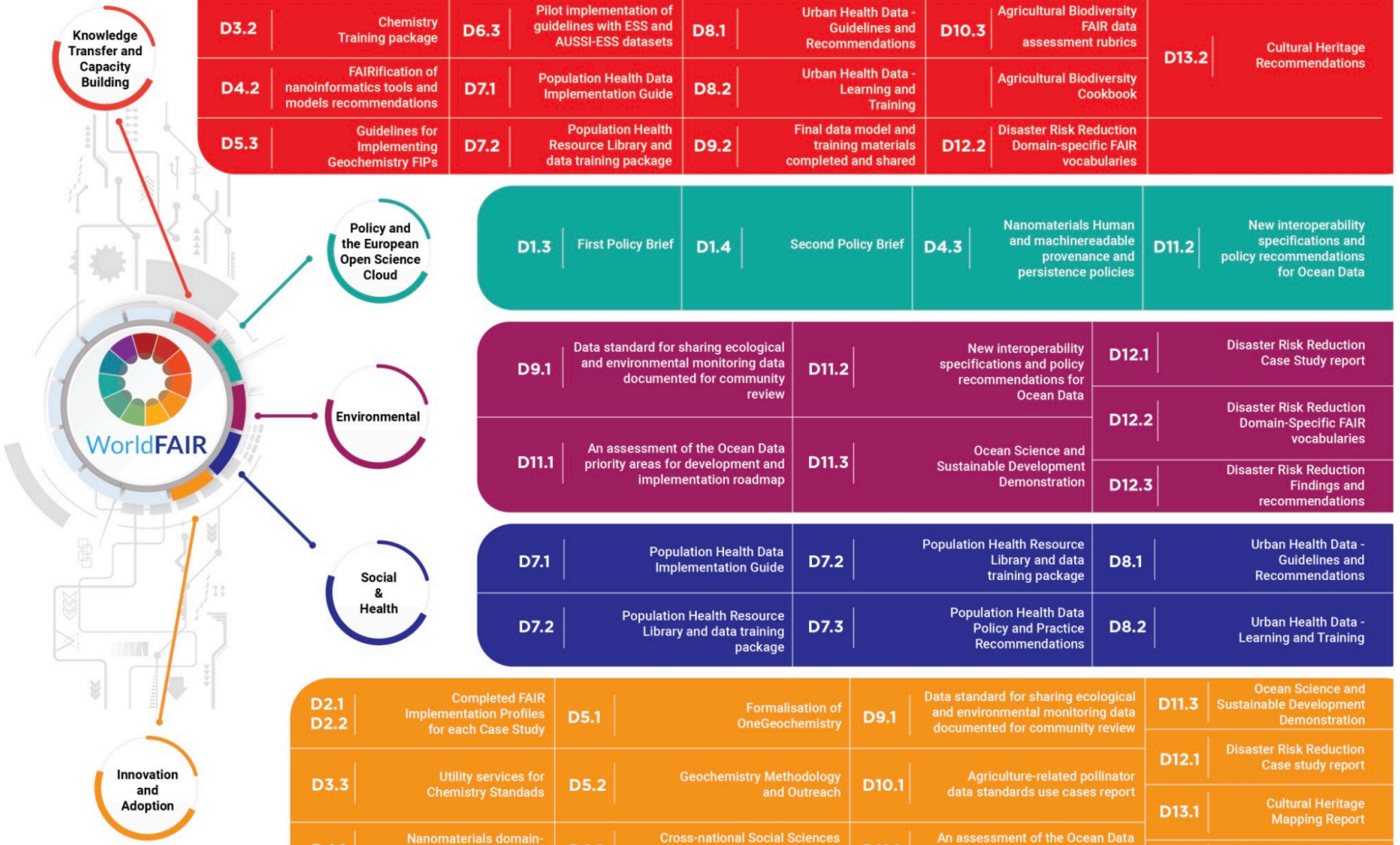
WorldFAIR Case Studies

- **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
- **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 Methodology



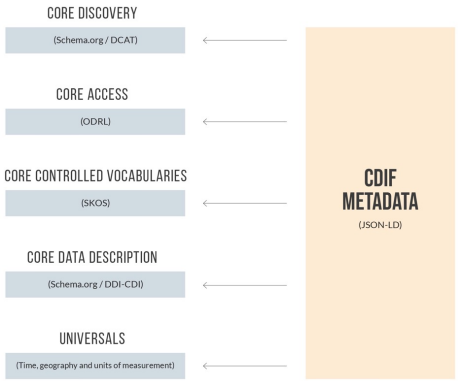

Knowledge Transfer and Capacity Building
Policy and the European Open Science Cloud
Environmental
Social & Health
Innovation and Adoption
WorldFAIR
D3.1 Digital recommendations for Chemistry FAIR data policy and practice
D6.2 Cross-national Social Sciences survey best practice guidelines
D7.3 Population Health Data Policy and practice recommendations
D10.2 Agricultural Biodiversity Standards, Best Practice and Guidelines Recommendations
D12.3 Disaster Risk Reduction Findings and recommendations
D3.2 Chemistry Training package
D6.3 Pilot implementation of guidelines with ESS and AUSSI-ESS datasets
D8.1 Urban Health Data - Guidelines and Recommendations
D10.3 Agricultural Biodiversity FAIR data assessment rubrics
D13.2 Cultural Heritage Recommendations
D4.2 FAIRification of nanoinformatics tools and models recommendations
D7.1 Population Health Data Implementation Guide
D8.2 Urban Health Data - Learning and Training
Agricultural Biodiversity Cookbook
D5.3 Guidelines for Implementing Geochemistry FIPs
D7.2 Population Health Resource Library and data training package
D9.2 Final data model and training materials completed and shared
D12.2 Disaster Risk Reduction Domain-specific FAIR vocabularies
D1.3 First Policy Brief
D1.4 Second Policy Brief
D4.3 Nanomaterials Human and machinereadable provenance and persistence policies
D11.2 New interoperability specifications and policy recommendations for Ocean Data
D9.1 Data standard for sharing ecological and environmental monitoring data documented for community review
D11.2 New interoperability specifications and policy recommendations for Ocean Data
D12.1 Disaster Risk Reduction Case Study report
D12.2 Disaster Risk Reduction Domain-Specific FAIR vocabularies
D11.1 An assessment of the Ocean Data priority areas for development and implementation roadmap
D11.3 Ocean Science and Sustainable Development Demonstration
D12.3 Disaster Risk Reduction Findings and recommendations
D7.1 Population Health Data Implementation Guide
D7.2 Population Health Resource Library and data training package
D8.1 Urban Health Data - Guidelines and Recommendations
D7.2 Population Health Resource Library and data training package
D7.3 Population Health Data Policy and Practice Recommendations
D8.2 Urban Health Data - Learning and Training
D2.1 D2.2 Completed FAIR Implementation Profiles for each Case Study
D5.1 Formalisation of OneGeochemistry
D9.1 Data standard for sharing ecological and environmental monitoring data documented for community review
D11.3 Ocean Science and Sustainable Development Demonstration
D3.3 Utility services for Chemistry Standads
D5.2 Geochemistry Methodology and Outreach
D10.1 Agriculture-related pollinator data standards use cases report
D12.1 Disaster Risk Reduction Case study report
D13.1 Cultural Heritage Mapping Report
D4.1 Nanomaterials domain-specific FAIRification mapping
D6.1 Cross-national Social Sciences survey FAIR implementation case studies
D11.1 An assessment of the Ocean Data priority areas for development and implementation roadmap
D13.3 Final Report on Cultural Heritage FAIR sharing

Policy and EOSC

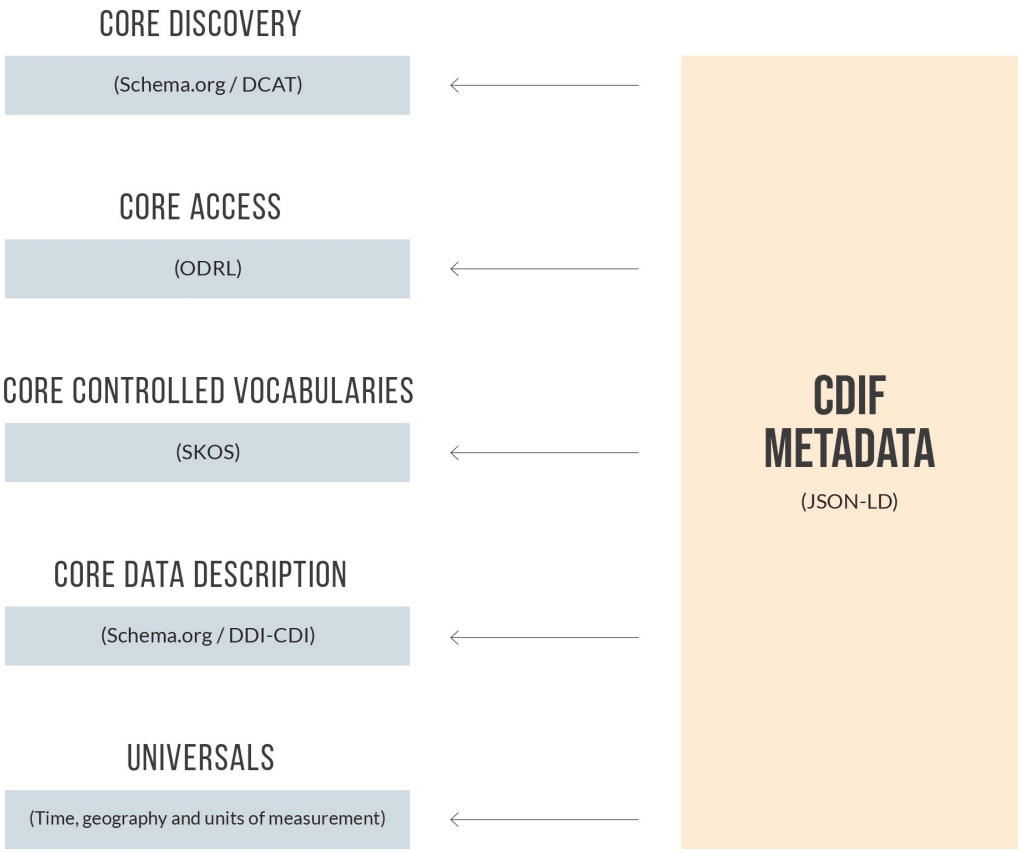
Responds to the SRIA:

- Engages with research communities.
  - ➢ See https://bit.ly/WorldFAIR-WP-Outputs for the recommendations, guidelines, demonstrations, implementations, and training materials from the WF case studies.
- Enhances and adds detail to the EOSC Interoperability Framework.
- Takes a standards-based approach.
- Interoperability for data combination.
  - Cross-Domain Interoperability Framework (CDIF): https://doi.org/10.5281/zenodo.11236871 (over 2000 downloads!)
  - A practical guide to FAIR implementation!
  - Adopt widely used web standards and use them in line with good practice.

WorldFAIR



CDIF, the Cross-Domain Interoperability Framework



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.

"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**".

Policy and EOSC

- **Policy recommendations:**
  - Second WorldFAIR Policy: https://doi.org/10.5281/zenodo.11242702
  - To achieve the objectives of EOSC…
  - ➢ We need to **shift from a bibliographic to an engineering approach to data stewardship**.
  - ➢ We need **metadata uplift** to support interdisciplinary research, fine-grained access management, machine-actionability and increased automation and the responsible use of AI.
  - ➢ **Enable this** through CDIF, empowering research data infrastructures, supporting standards organisations and efforts for international coordination.

Innovation and Adoption

- Direct involvement of **authoritative international entities** (e.g. IUPAC, OneGeochemistry, GBIF, ODIS-OIH).
- Connections with **standards organisations** (e.g. DDI Alliance, OHDSI, various W3C groups, TDWG, GloBI); or influential projects, institutions (NanoCommons, INSPIRE, APHRC, SALURBAL).
- **Bi-directional learning:** CDIF learns from and recommends approaches from ODIS, SOSO, INSPIRE… In turn, ecommendations are already being implemented by these initiatives and other projects.

WorldFAIR

- Chemistry: project has enabled an important step in making IUPAC assts FAIR.
- Agricultural biodiversity: recommendations incorporated by GloBI.

Innovation and Adoption: WorldFAIR+

- Endorsed by as part of its portfolio of activities: https://bit.ly/ISC-WorldFAIR-PLUS
- Federation of case studies and CDIF implementation projects.
- **Five new projects with CDIF implementation:** population health; disaster risk reduction (earthquake, floods/cholera); XAS data; geology data; climate adaptation (urban heat, ocean/coasts, soils/built environment).
- Existing and new case studies.
- Open to discuss participation and collaboration: contact me!
- Still accepting applications for the CDIF AG and WG: https://bit.ly/CDIF-AG-WG-Apply

Knowledge transfer and capacity building

- Significant knowledge transfer among the WPs, with international organisations (GBIF, ODIS, IUPAC), which fed into CDIF.
- **Social surveys:** knowledge transfer and collaboration between European and Australian social surveys.
- **Chemistry:** D3.2 Training Package: FAIR Chemistry Cookbook https://doi.org/10.5281/zenodo.10711950
- **Population Health:** (D7.2) Population health resource library and training package https://doi.org/10.5281/zenodo.10010936

WorldFAIR

Environmental

- **Nanomaterials:** D4.2 ‘FAIRification of nanoinformatics tools and models recommendations’ is being used in Horizon Europe project INSIGHT to increase FAIRness of predictive models for chemical and nanomaterials environmental risk assessment. WP also feading into PARC and PINK.
- **Biodiversity:** contributed directly to aspects of the new GBIF data model.
- **Agricultural biodiversity:** plant-pollinator data model will help address an important gap identified by IPBES report.
- **Oceans:** enabling ODIS to plan and extend its cross-domain / interdisciplinary extensions
    - D11.2, ‘New interoperability specifications and policy recommendations’ https://doi.org/10.5281/zenodo.10219933; and D11.3, ‘Ocean Science and Sustainable Development Demonstration’, https://doi.org/10.5281/zenodo.11242798

Social and health

- **Social Surveys:** Great harmonization and interoperability between European Social Survey and Australian Social Survey.
- **Population Health:** Contribution to and implementation of CDIF recommendations for INSPIRE population health data platform (pan African), uses CDIF, DDI, OHDSI standards.
  - ➢ See summary recommendations https://doi.org/10.5281/zenodo.11242767
- **Urban Health:** adoption of DDI standards by SALURBAL, guidance and training on FAIR and CARE.
- **Cultural Heritage:** recommendations on steps to adopt the FAIR principles for cultural heritage / GLAM institutions; adoption and demonstration by DRI; input to the RDA Collections as Data IG.
  - ➢ See recommendations https://doi.org/10.5281/zenodo.7897244 and implementation report https://doi.org/10.5281/zenodo.10850009