WorldFAIR

Project Title	Global cooperation on FAIR data policy and practice
Project Acronym	WorldFAIR
Grant Agreement No	101058393
Instrument	HORIZON-WIDERA-2021-ERA-01
Topic, type of action	HORIZON-WIDERA-2021-ERA-01-41 HORIZON Coordination and Support Actions
Start Date of Project	2022-06-01
Duration of Project	24 months
Project Website	http://worldfair-project.eu

D11.1 An assessment of the Ocean Data priority areas for development and implementation roadmap

Work Package	WP11 - Oceanography
Lead Author (Org)	Pier Luigi Buttigieg (AWI)
Due Date	28.02.2023
Date	01.03.2023
Version	1.0 DRAFT NOT YET APPROVED BY THE EUROPEAN COMMISSION
DOI	https://doi.org/10.5281/zenodo.7682398

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Authors	Notes
0.1	20.02.2023	Pier Luigi Buttigieg	Draft for internal review
1.0	01.03.2023	Pier Luigi Buttigieg	First release following internal review

'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.

Disclaimer

WorldFAIR has received funding from the European Commission's WIDERA coordination and support programme under the Grant Agreement no. 101058393. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

Abbreviations and Acronyms

API	Application Programming Interface
CDIF	Cross-Domain Interoperability Framework
CORDIS	Community Research and Development Information Service
DDI	Data Documentation Initiative
EBV	Essential Biodiversity Variable
EML	Ecological Markup Language
EOV	Essential Ocean Variable
EOSC	European Open Science Cloud
EXIF	Exchangeable Image File
FAIR	Findable, Accessible, Interoperable, Reusable
FER	FAIR Enabling Resource
FIP	FAIR Implementation Profile
GBIF	Global Biodiversity Information Facility
GEO	Group on Earth Observations
GEO BON	Group on Earth Observations Biodiversity Observation Network
GOOS	Global Ocean Observing System
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data and Information Exchange

'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.

ISO	International Organization for Standardization
JSON(-LD)	JavaScript Object Notation (for Linked Data)
NODC	National Oceanographic Data Centre
NUTS	Nomenclature of Territorial Units for Statistics
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OBIS	Ocean Biodiversity Information System
OBON	Ocean Biomolecular Observation Network
ODIN	Ocean Data and Information Network
ODIS	Ocean Data and Information System
ODIS-Arch	ODIS Architecture
ODIS-Cat	ODIS Catalogue of Sources
ODP	Ocean Data Portal
OIH	Ocean InfoHub
Omic BON	Omic Biodiversity Observation Network
ORCID	Open Researcher and Contributor ID
OWL	Web Ontology Language
REST	REpresentational State Transfer
RDA	Research Data Alliance
RDF(S)	Resource Description Framework (Schema)
ROR	Research Organization Registry
SKOS	Simple Knowledge Organization System
UNDRR	United Nations Office for Disaster Risk Reduction
UN Ocean Decade	United Nations Decade of Ocean Science for Sustainable Development
UNESCO	United Nations Educational, Scientific, and Cultural Organization

'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.

WOD	World Ocean Database
XML(S)	eXtensible Markup Language (Schema)

Executive Summary

This report provides a synoptic overview of how cross-domain interoperability may be built from within WorldFAIR's Oceanography case study (Work Package 11). After an introduction to the Intergovernmental Oceanographic Commission of UNESCO's Ocean Data and Information System (ODIS; Section 1), the report summarises an evaluation of FAIR Implementation Profiles and FAIR Enabling Resources compiled in WorldFAIR's WP2 (Section 2). It then summarises and synthesises supplementary insights obtained through a survey distributed across WorldFAIR partners (Section 3), and identifies a pathway to implement sustainable cross-domain (meta)data flows to inform and support the development of the Cross-domain Interoperability Framework (CDIF; Section 4).

The WorldFAIR case studies on biodiversity, disaster risk reduction, chemistry, and cultural heritage were identified as focal points to bridge with ODIS, being complementary to the strategic priorities of marine science and sustainable ocean management and offering clear socio-technical interfaces compatible with ODIS's own interoperability approaches. The high-level roadmap in Section 4 of this report outlines the general approach that will be pursued in the remaining tasks in WP11, namely the expansion of ODIS interoperability conventions to interface with those prevailing in the target use cases alongside the implementation and testing of (meta)data exchanges with independent stakeholders from the target domains.

'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.

Table of contents

'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.

'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.

1. Introduction and overview

This report identifies the strategic priorities and a high-level implementation path towards greater cross-domain interoperability of the Ocean Data and Information System (ODIS)[1], the case study in WorldFAIR's work package (WP) 11. It does so by examining the FAIR Implementation Profiles compiled in WP2 and responses to an exploratory survey distributed to the WorldFAIR consortium and their communities. Its focus is restricted to identifying the most viable path towards bridging ODIS to one or more WorldFAIR use cases from other domains, and suggesting the concrete steps required to accomplish a sustainable cross-domain interoperability solution, for implementation and testing in subsequent WP11 deliverables.

1.1. A general approach to the challenge of cross-domain interoperability

WorldFAIR's efforts to cultivate sustainable cross-domain interoperability hinge on an understanding of how practitioners, projects, infrastructures, and other digital stakeholders have implemented the FAIR Principles (Wilkinson et al., 2016) within their domains. Recognising how and why *intra*domain implementations overlap and differ is the key to finding the most expedient pathway to create stable, extensible inter- and cross-domain digital exchanges. In these efforts, it is key to recognise this is both a technical and social challenge: distinct digital cultures — alongside the technologies which support them — have emerged in each domain and rarely change quickly or independently. As a result, top-down, prescriptive, or highly centralised approaches to achieving interoperability are unlikely to be meaningfully adopted or sustained in the mid- to long term. This is especially true if such approaches have not been informed by broad consultation across the practitioners they impact. Instead, successive rounds of deepening strategic and socio-technical alignment — led by well-established and trusted actors in each domain's digital landscape — will allow more organic and lasting change.

WorldFAIR Deliverable 2.1 ("FAIR Implementation Profiles (FIPs) in WorldFAIR: What Have We Learnt?"[2]) has summarised general findings on this theme through the evaluation of domain-specific FAIR Implementation Profiles (FIPs) and their bearing on the project's emerging Cross-Domain Interoperability Framework (CDIF). Viewed from the perspective of a single domain, this information can help identify priority areas for capacity development and/or alignment (e.g. overlapping FAIR Enabling Resources [FERs]), as well as which partners to engage in order to begin merging digital ecosystems across domains. The greatest opportunities lie where concrete, functional implementations of the FAIR Principles are technically similar, are capable of delivering (meta)data in domain-neutral forms, and where there is compatible usage of FERs by a diverse base of practitioners. Consequently, given an overview of a cross-domain digital exchange landscape, any

[1] https://oceaninfohub.org/odis/
[2] https://doi.org/10.5281/zenodo.7378108

'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.

one domain can begin to build and test both multi-domain and bi-domain interoperability approaches.

In this spirit, this report will provide an intradomain perspective on the challenges and opportunities identified by D2.1 alongside an initial, high-level roadmap to guide WorldFAIR's Oceanography Work Package (WP11, see Task 11.1) towards greater interoperability with the domains represented by the project's other use cases.

1.2. Initial conditions in WorldFAIR's Oceanography Work Package

Oceanography — the description and study of the oceans — is, in itself, a highly multi-disciplinary and multi-domain field. Understanding the planet's oceans requires expertise from geography in all its forms, biology, physics, chemistry, the social sciences, humanities, and many more domains. Thus, frameworks for FAIR digital exchanges must be built with approaches similar to those described above: high-level, domain-neutral standards and technologies must create a digital 'glue' between a plurality of stakeholders. Extending from this foundation, approaches that are progressively more specialised will permit the exchange of digital assets which require thematically constrained (meta)data content and structure.

The WorldFAIR Oceanography use case (WP11) centres on the Ocean Data and Information System (ODIS)[3], a system built and maintained by a diverse and global federation of digital stakeholders sharing interoperability conventions. The design principles and development philosophy which underpin ODIS are aligned to the logic outlined above: a domain-neutral approach is used to express (meta)data using web architectural standards, capable of (where needed) 'wrapping' specialist conventions/standards in those which are generic. ODIS was initiated and is coordinated by the International Oceanographic Data and Information Exchange (IODE) of UNESCO's Intergovernmental Oceanographic Commission (IOC) and is quickly gaining ground as an interoperability solution across regional infrastructures, projects, and initiatives in the UN Decade of Ocean Science for Sustainable Development (UN Ocean Decade)[4]. ODIS — in its current form — was initially scoped in 2018, and strove to move away from a centralised model of data integration, towards using linked open data[5] and web-centric approaches to connect existing resources in an interoperable federation of continuous, multilateral (meta)data exchange.

[3] https://oceaninfohub.org/odis/
[4] https://oceandecade.org/
[5] https://5stardata.info/en/

'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.

The rapidly growing ODIS federation[6] relies on a co-developed, lightweight, decentralised, and extensible (meta)data exchange architecture (ODIS-Arch)[7]. ODIS-Arch leverages the globally adopted syntactic and semantic conventions put forth by the world's major search and discovery systems (e.g., Google, Bing, Yahoo, and Yandex) and co-maintained through working groups in the World Wide Web Consortium (W3C). Within this global and domain-neutral framework, ODIS-Arch provides guidance and templates (i.e. 'patterns' or 'profiles', e.g. the pattern for Essential Ocean Variables[8]) for the exchange of (meta)data using the JavaScript Object Notation for Linked Data (JSON-LD)[9] syntax specification and the schema.org vocabulary[10] for lightweight semantics. ODIS-Arch patterns are created based on the focal areas of the IOC (e.g. experts and institutions, vessels, training opportunities) and requests from its community of users (e.g. for sensors and instrumentation, software applications, event series).

Once a partner expresses their (meta)data or other digital assets[11] onto the web (**Figure 1**), they become an independent 'node' in the ODIS network, and their information can be harvested and integrated into, e.g., open knowledge graphs which any member (or non-member) of ODIS can assemble for any purpose (**Figure 2**). The IODE is demonstrating this functionality through their Ocean InfoHub (OIH) project[12] (**Figure 3**), with partners creating or augmenting existing regional and thematic portals (e.g. The Latin American and Caribbean Clearing House Mechanism[13], Pacific Data Hub[14], and ODIN Africa[15]).

[6] Currently engaging more than 50 partners, including regional aggregators such as the European Marine Observation and Data Network (EMODnet; https://emodnet.ec.europa.eu/en), the Pacific Data Hub (https://pacificdata.org/), INVEMAR (https://www.invemar.org.co/), and ODIN Africa (http://www.odinafrica.org/). A live list of implementation partners is available here: https://github.com/iodepo/odis-arch/blob/schema-dev/config/sources.yaml

[7] https://book.oceaninfohub.org/

[8] http://book.oceaninfohub.org/thematics/variables/index.html

[9] https://json-ld.org/

[10] https://schema.org/

[11] E.g. datasets, data feeds, metadata packages, software, web services, etc.

[12] https://oceaninfohub.org/

[13] http://portete.invemar.org.co/chm#/

[14] https://pacificdata.org/

[15] http://www.odinafrica.org/en/

'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.

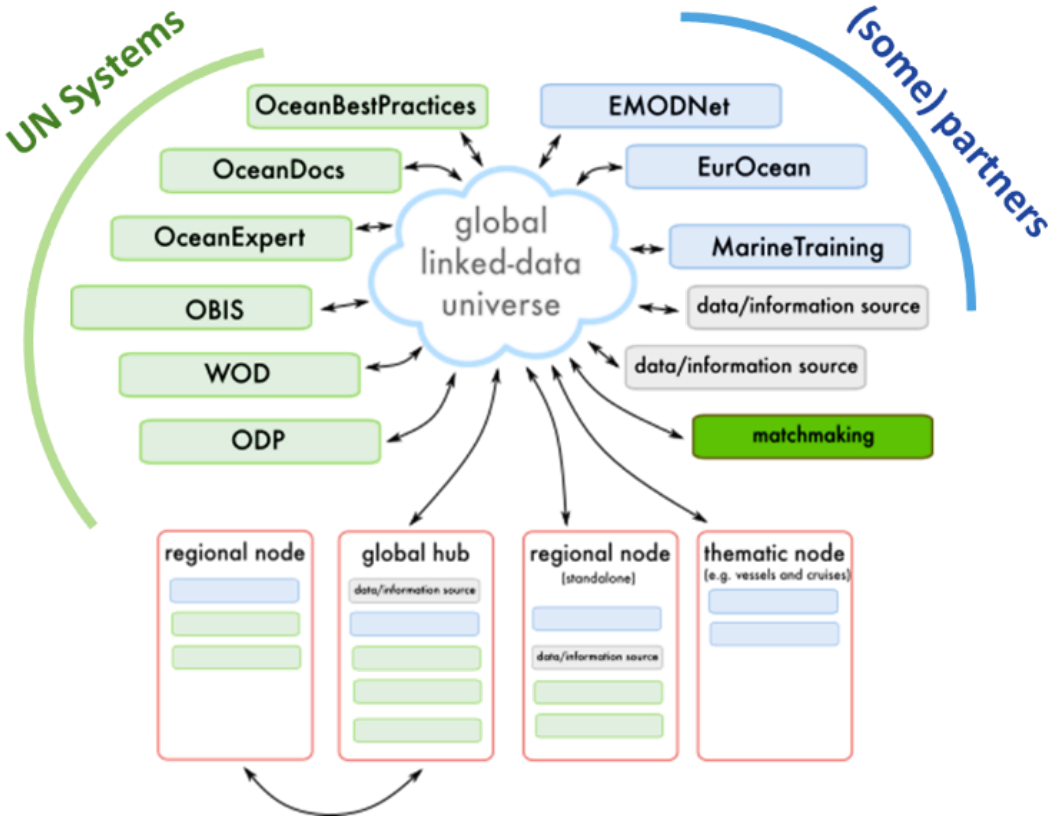

Figure 1: A schematic of the Ocean Data and information System (ODIS). ODIS is a federation of UN and non-UN data and information providers, interoperating through structured metadata exchange using web architectural patterns (https://book.oceaninfohub.org/). All content shared through ODIS is automatically discoverable across the entire ODIS federation via portals such as Ocean InfoHub (https://search.oceaninfohub.org/). Source: Image released under CC-0 courtesy of https://book.oceaninfohub.org/). Abbreviations expanded in Abbreviations and Acronyms, above.

'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.

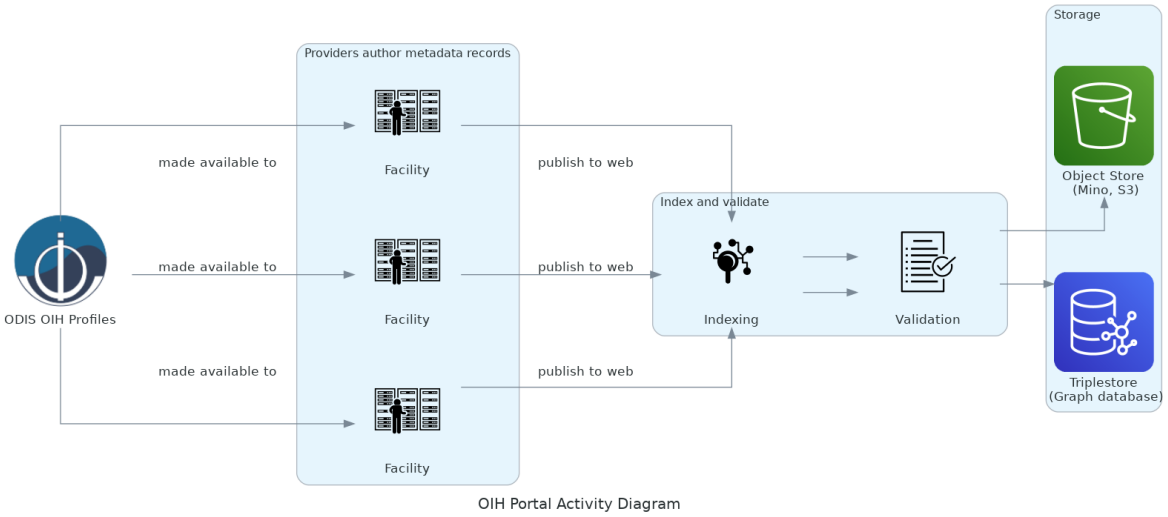

Figure 2: An illustration of (meta)data flow from ODIS nodes to harvesters or aggregation systems. Referencing the ODIS-Arch patterns and profiles — released and moderated by IODE — any provider can release (meta)data records describing their holdings to the ODIS Network and the web at large. This allows other agents to index and harvest (meta)data records of interest to them, validate their content, and rapidly integrate them into their aggregation system of choice (e.g. object stores or graph databases). The (meta)data contained in ODIS-Arch-compliant records typically links to further digital assets, including full distributions of data records or content-specific web portals to allow users to further their discovery. Source: Image released under CC-0 courtesy of https://book.oceaninfohub.org/

'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.

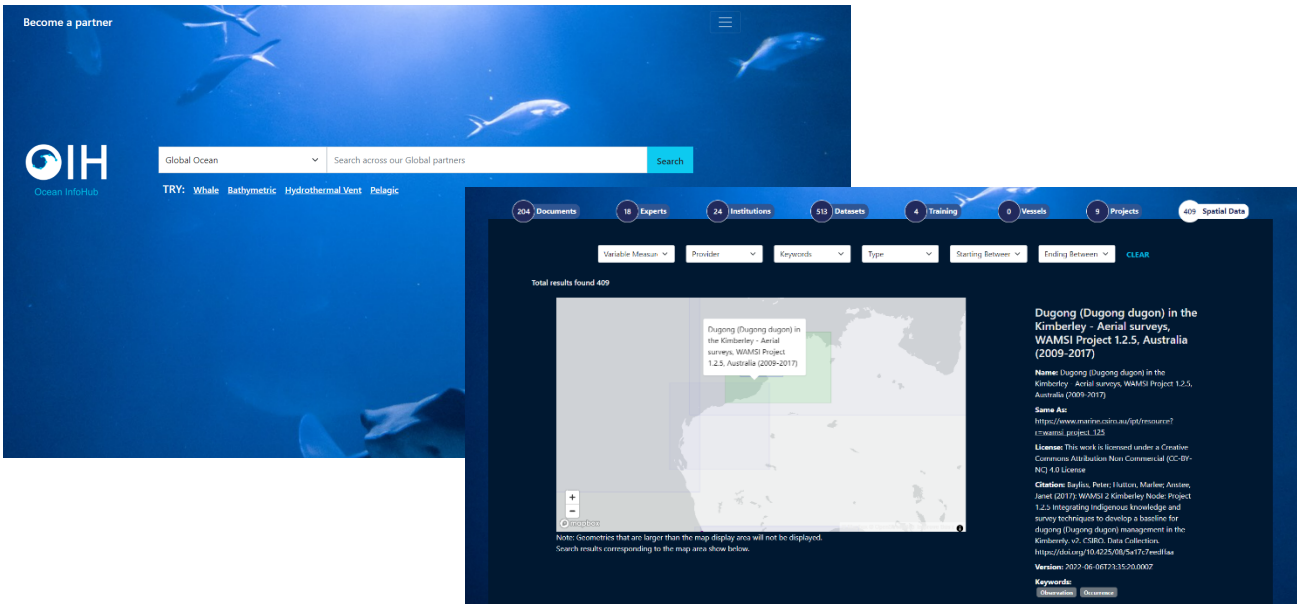

Figure 3: The Ocean InfoHub (OIH) search and discovery portal. The OIH portal is a demonstration of how (meta)data records from all ODIS nodes can be harvested and integrated to interlink and begin to harmonise digital assets across multiple ocean-related domains and stakeholder groups. By sourcing the decentralised and independent contributions of each ODIS partner, OIH provides an unprecedented scale of FAIR (meta)data exchange. This lays a solid foundation for progressive deepening of interoperability, as challenges beyond the scope of the ODIS interoperability architecture (e.g. the compatibility of data distributions in domain-specific formats) are made more transparent. Source: Images from https://search.oceaninfohub.org/ (2023-02-14).

1.3 Expanding ODIS interoperability through WorldFAIR

Through WorldFAIR's Task 11.1 ("Roadmap to expanding ODIS-Arch interoperability beyond the ocean realm") WP11 seeks to build on the success of the ODIS, and create a roadmap to align the ODIS approach with those of the project's other case studies, as well as core recommendations from CODATA and the Research Data Alliance (RDA).

Concretely, this will entail 1) the expansion of ODIS-Arch's schema.org/JSON-LD thematic profiles and patterns[16] to accommodate digital standards across more disciplines and priority topics of exchange; 2) alignment of ODIS implementation to the emerging recommendations in CDIF; and 3) testing the actionability and performance of activities 1 and 2 by running diagnostics and validation to verify successful inter-domain (meta)data transfer, centred on data sets which have value to all participating domains.

[16] https://book.oceaninfohub.org/thematics/README.html

'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.

2. Insights from FAIR Implementation Profiles and Enabling Resources

In this section, we report on the examination of the WorldFAIR FIPs (summarised in D2.1) and highlight both generic and domain-specific (meta)data exchange conventions and FERs with high potential to bridge other WorldFAIR case studies to and with ODIS. Each subsection focuses on one of the case studies and notes only those elements in the FIPs which are directly translatable into the high-level Roadmap (Section 4).

2.1. A brief position on the FAIR principles

This section outlines the high-level understanding of the FAIR Principles upon which the Oceanography use case operates. Elaborations of these positions are available in Buttigieg et al. (2022).

2.1.1. Findability

The essence of the Findability principle — as outlined in its sub-principles — is that (meta)data has been assigned unique and persistent identifiers (PIDs, such as Digital Object Identifiers [DOIs]) and that those PIDs are linked to sufficiently 'rich' metadata to make them discoverable through expected search behaviours, and both the PIDs and their metadata are indexed in a reliable and searchable resource (e.g. a web-accessible repository).

When considering findability in a cross-domain context, the creation of technology to mint and maintain PIDs and search portals is straightforward and largely a question of technical capacity and sustainability. Friction emerges, however, when the nature of 'rich' metadata and expected user behaviour is considered. Each domain — and each subdomain and/or regional group therein — invariably has a different conception of what 'rich' metadata (comprising both technical and descriptive metadata) includes, and efforts to establish minimal standards are often fraught, with noteworthy exceptions produced by bodies such as W3C and ISO. Nonetheless, any stable specification of richness maintained by an authority (e.g. a standards organisation or consortium, a regulatory body, a recognised professional society) which credibly represents a domain's requirements supports interfaces with systems such as ODIS.

2.1.2. Accessibility

Digital resources, once found, must be accessible (at least to some) to be useful. The Accessibility principle is perhaps the least variable in terms of implementation. The prevalence of standard web protocols such as HTTPS, FTP, and other TCP/IP protocols, alongside a fairly restricted range of protocols used by contemporary Application Programming Interfaces (APIs) e.g. the Open Archives

'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.

Initiative Protocol for Metadata Harvesting [OAI-PMH][17]) or the architecture of the representational state transfer [REST] style) facilitate harmonised access across most domains. The vast majority of cases which ODIS has encountered (and is likely to encounter) will be capable of providing access to content via an open protocol, and — unless an extremely compelling reason emerges — resources have not typically been allocated to creating bespoke access solutions with potential interoperability partners.

Complexities, however, arise when access controls are required to protect sensitive, legally protected, or otherwise restricted data. Open Authorization (OAuth) protocol and standards[18] for access delegation via secure token exchanges are widely available but still not implemented by many repositories serving research data. Domains such as health and social science are well versed with this issue, while others are only beginning to face restrictions due to new regulatory frameworks and the rise of digital sovereignty and localisation. Automated user and access control — if improperly implemented or maintained — has the potential to introduce vulnerabilities to secure systems and requires considerable oversight and maintenance capacity. As a result, ODIS defaults to interoperating with publicly released (meta)data which — at most — describes the existence of sensitive data and the pathway to accessing sensitive data without releasing them.

A key element of this principle is that metadata availability and accessibility is persistent, even when the data they describe are removed, deleted, or otherwise made inaccessible. This generally requires the decoupling of data and metadata (as is native to the ODIS approach), both technically and in terms of management strategy. This requires that data and metadata records are archived independently, with their own PIDs and separate life cycles, while ensuring that every record includes the PIDs of the other (meta)data records in its constellation. Persistent metadata records which survive past the deletion of the data they describe are also a cornerstone of provenance tracing, critical to trusted (re)use (see 2.1.4).

2.1.3. Interoperability

The semantic interoperability described by the FAIR Principles is often a considerable impediment to bridging domains in depth and with machine-actionable solutions. In order for practitioners to understand what (meta)data are about (i.e. semantics), any terminology used to describe them must be digitised and available on the web, using well-accepted and stable standard specifications such as the W3C's Resource Description Framework Schema (RDFS)[19], Simple Knowledge

[17] https://www.openarchives.org/pmh/
[18] https://oauth.net/
[19] https://www.w3.org/TR/rdf-schema/

'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.

Organization System (SKOS)[20], or Web Ontology Language (OWL)[21]. Ideally, each term would be assigned a unique, dereferencable PID, and associated with a clear definition, and annotated with references and other explanatory notes where needed. In order for machine agents to compute and reason over semantics, such resources must correctly leverage the expressivity of SKOS, RDFS, OWL, or similar specifications to provide relationships between terms that machines can process. The capacity to develop and use mature knowledge representation technologies (e.g. fully-fledged ontologies which support machine-driven reasoning) are patchily distributed, with lightweight vocabularies being the more common. As domains build up the expertise needed to advance semantic interoperability, it will be key to identify and align well-built and -maintained semantic resources at the earliest possible stage to prevent semantic silos and domain-specific approaches that will be prohibitively costly to generalise or map more broadly.

2.1.4. Reusability

Even if (meta)data are shared in such a way that all the previous principles have been met, (meta)data (re)use depends on trust in and clarity on their origin and any restrictions or conditions on their use. Hence, the Reusability principle requires that (meta)data are accompanied with information specifying the licence(s) under which they are released, as well as information on their provenance. The 'richness' of this information is likely to vary - even within domains - based on the usage scenario (informal and recreational to legally consequential). However, documentation of any relevant licences and the processes which generated, modified, or curated the data should be present in order for (meta)data to be trusted and appropriately used. Advanced approaches to sharing such information, such as structured provenance data using conventions from well-known systems (e.g. the W3C's PROV[22]) are not frequently encountered, but offer a highly desirable level of precision and — if well-serialised and encoded — machine actionability across domains. Further, in the ocean domain — as well as in many others — it remains a challenge to identify and precisely communicate authoritative, dereferenceable PIDs for a given licence (e.g. Creative Commons[23]) or set of usage conditions (e.g. Traditional Knowledge and Biocultural Labels[24], supporting the CARE Principles[25]) to digital assets, their metadata, or the physical objects they describe.

2.2. Domain-focused insights

What follows is insight from our examination of the WorldFAIR project's completed FAIR Implementation Profiles (FIPs), which highlight both generic and domain-specific (meta)data

[20] https://www.w3.org/2004/02/skos/
[21] https://www.w3.org/OWL/
[22] https://www.w3.org/TR/prov-overview/
[23] https://creativecommons.org/
[24] https://localcontexts.org/labels/traditional-knowledge-labels/
[25] https://www.gida-global.org/care

'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.

exchange conventions and FERs with high potential to bridge other WorldFAIR case studies to and with ODIS. Each domain or cross-domain research area participating in the project is listed with its corresponding work package number.

2.2.1. Chemistry (WP3)

In terms of **findability**, the chemistry use case describes the use of well-known and domain-neutral PID solutions such as DOIs managed by DataCite[26] and Crossref[27], compact identifiers via identifiers.org[28], as well as ORCIDs. Domain-specific identifiers and/or metadata schema such as those supported by the Crystallographic Information Framework (CIF)[29] and the IUPAC Compendium of Chemical Terminology (the "Gold Book"[30]) are also noted, and provide fine-grained thematic schema to richly describe content.

Accessibility is provided through generic web protocols (HTTPS, FTP, etc.) with the possibility of content negotiation[31]. Access control is noted for proprietary resources and metadata persistence beyond data lifetimes specified in policies of the PID registries used as well as in regulatory frameworks and good practice in the domain itself.

Knowledge representation and **interoperability** rely on the conventions of the CIF and the formatting conventions which underpin it. Other data structures bear their own internal semantics, such as those present in the elements of chemical-data file formats (e.g. chemical structure-data files [SDFs]). The value of more generic serialisations such as RDF are noted, but implementation of these is uneven in the domain. While not necessarily released using W3C recommendations for knowledge representation, very well-adopted terminologies and nomenclatures from the IUPAC Gold Book boost interoperability between systems in this domain. More expressive and computable semantic resources, such as the Chemical Entities of Biological Interest (ChEBI) ontology, were noted.

In terms of **reusability**, the FIP notes that permissive licences such as those in the Creative Commons framework are common for metadata, but handling licence information on subject data itself is a challenge, with licencing information often not present. Provenance information frameworks are provided in both domain-specific conventions (e.g. the CIF) as well as regional/national conventions such as the United States Library of Congress' Metadata Object Description Schema (MODS)[32].

[26] https://datacite.org/
[27] https://www.crossref.org/community/
[28] https://docs.identifiers.org/
[29] https://www.iucr.org/resources/cif
[30] https://goldbook.iupac.org/
[31] https://www.w3.org/blog/2006/02/content-negotiation/
[32] https://www.loc.gov/standards/mods/userguide/recordinfo.html

'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.

2.2.2. Nanomaterials (WP4)

The FIP completed by WorldFAIR's nanomaterials use case reported a balanced range of domain-neutral and domain-specific PID and/or more transient identifier maintenance systems, metadata specifications, and discovery services supporting **findability**. Generic solutions such as those provided by DataCite, schema.org, Dublin Core, Zenodo, and Frictionless Data[33] are juxtaposed with disciplinary solutions hosted and maintained by bodies such as the International Chemical Identifier (InChI) Trust[34] (with sub-specifications for nanomaterials) as well as localised, database-specific identifiers and metadata specifications. Dedicated databases and thematic solutions such as the FAIR-aligned NanoCommons Knowledgebase[35] support (meta)data exchange which more completely describes particles, their characterisation, and key metadata such as those for safety codes. This FIP also noted the rise of peer-reviewed data journals for nanomaterials, offering their own solutions for PID maintenance and metadata.

Accessibility, again, is provided using standard web protocols (HTTPS, FTP, etc.) as well as SOAP (Simple Object Access Protocol) and REST APIs. Access policies were noted to be generally open, and specified by the (meta)data hosting service/infrastructure (e.g. EOSC, Zenodo) or governed by institutional or funder policies.

Multiple approaches aimed at providing semantic **interoperability** across systems in this domain were described in this FIP with locally (i.e. restricted to one project or institutional resource) focused technologies also present. Knowledge representation technologies serving lightweight vocabularies (e.g. the EU's NANoREG harmonised terminology[36]; Gottardo et al., 2016) as well as fully-fledged ontologies expressed in OWL (e.g. the eNanoMapper ontology and the Elementary Multiperspective Material Ontology [EMMO]) were noted, alongside semantically enabling data exchange conventions that are adopted in this and closely related domains (e.g. ISA-TAB[37] used across in research data management in several biomolecular sciences).

Licensing conventions in support of **reusability** noted the use of the Creative Commons framework and data policies/licences used by aggregators (e.g. Zenodo, data journals). The FIP responses also cautioned that complex data ownership arrangements may be in effect within a given consortium. Regarding provenance tracking, this FIP noted the use of the CODATA Uniform Description System for Materials on the Nanoscale (UDS)[38], which includes specifications for nanomaterial production, as well as support for provenance information in exchange formats such as ISA-TAB. Fine-grained workflow recording was also noted, compatible with the domain-centric archives noted above (e.g.

[33] https://specs.frictionlessdata.io/#overview
[34] https://www.inchi-trust.org/
[35] https://nanocommons.github.io/user-handbook/
[36] https://nanoreg.eu/
[37] https://isa-specs.readthedocs.io/en/latest/isatab.html
[38] https://codata.org/initiatives/previous-codata-working-groups/nanomaterials/

'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.

workflows such as that used by Martinez et al., 2020 and stored in the NanoCommons Knowledge Base).

2.2.3. Geochemistry (WP5)

To address **findability**, the geochemistry FIP reiterated the use of generic DOIs, ORCID, ROR, and DataCite schemas to provide persistent links to, and metadata for, digital assets. Additionally, systems such as Dataverse[39] that are designed to simultaneously support local findability and long-term archiving were noted, alongside some sub-domain specifications such as those used by data archiving and access systems e.g. EarthChem[40].

In terms of **accessibility**, the FIP noted REST services as a primary approach to push and pull geochemistry data. FERs supporting **interoperability** including JSON(-LD) and XML schemas for serialisation, with semantics provided by Dublin Core, and the Observations Data Model 2 (ODM2)[41], as well as internal (i.e. institutional, group-level) vocabularies and conventions. Licencing information in support of **reusability** noted the Creative Commons framework, defaulting to more open licences (CC-0, CC-BY). No frameworks for the recording or transfer of provenance information were noted.

2.2.4. Social Surveys (WP6)

In terms of **findability**, this case study noted the use of internal Unique Reference Numbers (URNs) by several participating systems, as well as dereferenceable DOIs such as PIDs. Metadata specifications to enhance findability follow the specifications in the DDI-Lifecycle 3.3[42], the DDI-Codebook, Dublin Core, and through metadata schema provided by services such as DataCite. (Meta)data are exposed for search and discovery through GraphQL APIs, published through Colectica[43] web services, Dataverse implementations[44] or through portals such as EOSC.

Accessibility is provided through generic protocols (e.g. HTTPS), as well as GraphQL endpoints and APIs associated with Microsoft Azure implementations.

[39] https://dataverse.org/
[40] https://www.earthchem.org/
[41] https://www.odm2.org/ - a US-led approach to specify an information model and supporting software ecosystem for feature-based earth observations, designed for interoperability among disciplines.
[42] https://ddialliance.org/Specification/DDI-Lifecycle/3.3/
[43] https://colectica.com/
[44] https://dataverse.ada.edu.au/

'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.

Semantic **interoperability** is supported through the use of semantics native to JSON and GraphQL conventions, as well as Parquet data types[45]. Once again, the DDI-Lifecycle is referenced as a source of structured codelists, with similar classifications and codelists from international standards bodies used to stabilise terminology and semantics (e.g. ISO3166-1 for country names, ISO639-2 for languages, NACE Rev 2 for Industry, ISCO08 for occupation, and the Nomenclature of Territorial Units for Statistics (NUTS) for regions). Further vocabularies are DDI Controlled Vocabularies, and the CESSDA vocabulary service and Social Science Thesaurus[46].

To facilitate **reusability**, this FIP notes the use of the Creative Commons licencing framework, provenance conventions documented in the DDI-Lifecycle, and the W3C PROV standard.

2.2.5. Population and Urban Health (WP7, WP8)

The FIPs for these closely related case studies were reviewed together due to their high overlap in FERs. As in other domains, **findability** relies on the use of DOIs, with metadata specifications aligned to the Data Documentation Initiative's Codebook (DDI-Codebook)[47] for survey data and adhering to guidelines used by Inter University Consortium for Political and Social Research (ICPSR)[48]. The population health FIP noted that schema.org conventions are being used to interfaces with Google Dataset Search, with domain-specific extensions provided by groups such as the Observational Health Data Sciences and Informatics (OHDSI)[49] organisation. Further, domain-centric (meta)data schema (which also contain numerous generically applicable fields) were reported, including the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)[50]. Dedicated platforms for global health concerns were noted, such as the Platform for Evaluation and Analysis of COVID-19 Harmonised data (PEACH)[51], boosting interoperability in response to the global pandemic.

These FIPs did not outline any non-standard protocols for **accessibility**, and noted the use of the domain-neutral assurances/certifications of long-term access via the CoreTrustSeal Certification, originating from the RDA Repository Audit and Certification DSA–WDS Partnership Working Group. Further, the DDI-Codebook provides dedicated (meta)data properties for access rights sourced from the globally adopted Dublin Core vocabulary.

[45] https://github.com/apache/parquet-format/blob/master/LogicalTypes.md
[46] https://vocabularies.cessda.eu/ ; https://elsst.cessda.eu/
[47] https://ddialliance.org/Specification/DDI-Codebook/
[48] https://www.icpsr.umich.edu/web/pages/
[49] https://www.ohdsi.org/
[50] https://www.ohdsi.org/data-standardization/
[51] https://inspiredata.network/

'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.

These FIPs noted that semantic **interoperability** relies on the use of authoritative and international term and code lists for diseases (e.g. the International Classification of Diseases Version 10 [ICD10], maintained by the World Health Organization[52]) and the use of ISO Codes for globally standardised lists of country names and sub-national entities (e.g. ISO 3166-1:2013[53]). Additionally, vocabularies administered by OMOP, schema.org, and the Athena vocabulary management system[54] were noted as domain standards. Encoding and serialisation for interoperable exchange noted XMLS, RDF, and the commercial STATA[55] conventions.

As with most others, these FIPs noted the Creative Commons licencing framework to support **reusability**, encouraging open sharing with attribution via the stipulations in the licence CC-BY 4.0. Provenance (meta)data is recorded following specifications in the OMOP CDM, which in turn references the PROV convention and the DDI-Codebook, the latter of which provides properties supporting the expression of derivation, version control, and referencing[56], often imported from Dublin Core and other cross-domain resources.

2.2.6. Biodiversity and Agricultural Biodiversity (WP9, WP10)

The FIPs for these closely-related case studies were also reviewed together due to their high overlap in FERs. In these cases, **findability** is anchored by PIDs issued and maintained by 1) domain-neutral systems including DataCite, CrossRef, ROR, Zenodo, GitHub, and ORCID, 2) super-domain systems such as those which include biodiversity as a dimension (e.g. Earth systems, served by PANGAEA[57] and Dryad[58]) and 3) domain-specific infrastructures and services such as GBIF, the Global Biotic Interactions (GLoBI)[59], and the transnational Biological Collection Access Service (BioCASe)[60] network of primary biodiversity repositories, and 3) infrastructure-neutral or local approaches using universally/globally unique identifiers (U/GUIDs).

The metadata specifications and standards accompanying these PIDs are closely tied to their issuing authorities, which also provide the search and discovery interfaces for the digital assets they dereference to. For example, GBIF's alignment to the Darwin Core (DwC)[61] and the metadata/semantics entailed by the use of Ecological Markup Language (EML)[62], alongside the

[52] https://icd.who.int/browse10/2010/en
[53] https://www.iso.org/standard/63545.html
[54] https://athena.ohdsi.org/vocabulary/list
[55] https://www.stata.com/
[56] e.g. properties such as isReplacedBy, hasVersion, and derivation
[57] https://pangaea.de/
[58] https://datadryad.org/stash
[59] https://www.globalbioticinteractions.org/
[60] https://www.biocase.org/
[61] https://dwc.tdwg.org/
[62] https://eml.ecoinformatics.org/

'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.

Access to Biological Collection Data (ABCD) Schema[63] for specimens and observations used by the BioCASe partnership. Findability through Google, Microsoft, and other major providers leverages schema.org.

Measures to provide and secure **accessibility** noted universal protocols (e.g. HTTPS, REST, OAI-PMH) and policies (e.g. CoreTrustSeal), but also those of major aggregators such as GBIF's publisher agreements.

Semantic **interoperability** in this case study is closely tied to the metadata conventions used to promote findability, discussed above. Conventions such as EML, ABCD, Darwin Core, and others offer relevant and often overlapping semantics to describe samples, specimens, and other dimensions of biological and ecological entities. Further semantic specification is offered by drawing from vocabularies such as the Plant-Pollinator Interactions Vocabulary[64] or the resources These are complemented or integrate generic conventions and standards, such as ISO 3166-1 (Codes for the representation of names of countries and their subdivisions – Part 1: Country codes) and GADM[65] for maps and mapping data.

These FIPs noted that semantic conventions are serialised in common web exchange formats such as XML(S), RDF(S), JSON, as well as embedded in flat formats such as the CSV format.

In terms of **reusability**, this case study also noted the heavy use of the Creative Commons framework (particularly the CC0 and CC BY licences) and provenance tracking by use of the conventions in EML, custom fields supported by GBIF's archival process, and the more generic PROV Ontology (PROV-O)[66], which expresses the W3C PROV convention[67].

2.2.7. Disaster Risk Reduction (WP12)

The disaster risk reduction case — being embedded within Earth science and observation — noted (meta)data conventions with high alignment to many of those in oceanography. In terms of findability, this case study's FIP noted the use of generic PID systems such as DOIs and PURLs, with indications that Archival Resource Keys (ARKs)[68] are gaining ground due to their low cost and greater flexibility. Metadata associated with such PIDs followed conventions found in schemata and structures such as the Generic Earth Observation Metadata Standard (GEOMS; required for satellite validation via ground-based measurements)[69], Dublin Core, the Network Common Data Form

[63] https://www.tdwg.org/standards/abcd/
[64] https://ppi.rebipp.org.br/
[65] https://gadm.org/
[66] https://www.w3.org/TR/prov-o/
[67] https://www.w3.org/TR/prov-overview/
[68] https://arks.org/
[69] https://avdc.gsfc.nasa.gov/index.php?site=1925698559

'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.

(NetCDF)[70], the Geographic information specifications in ISO 19115[71] and ISO 19139[72]. Examples of aggregators and their portals to support findability included the Global Earth Observation's System of Systems (GEOSS) Portal[73], regional data aggregators such at the Pacific Data Hub[74], Global Atmosphere Watch Station Information System (GAWSIS)[75], and the World Meteorological Organization (WMO) search and discovery systems, increasingly linked to the WMO Information System (WIS)[76].

In addition to the common routes to (meta)data **accessibility** (HTTPS etc.), this FIP noted the Open Geospatial Consortium Catalogue Services[77], which offer standards, bindings, and profiles to publish and access structured geospatial metadata over the web. This FIP highlighted the use of the XML-based Security Assertion Markup Language (SAML) framework for managing access control, entitlement, and authentication.

Overlapping with the conventions used to structure metadata for findability, this FIP noted semantic **interoperability** is granted through inherent semantics in specifications such as ISO 19115/19139, netCDF, and JSON. This FIP also noted vocabularies for knowledge representation common across much of the Earth observation community, such as the Climate and Forecast Standard Names (CF SN) vocabulary[78].

This FIP was somewhat terse on **reusability** measures, but did cite open licencing frameworks such as the GNU General Public License (GPL)[79]. This licencing framework - originally developed for software and source code - provides licencing conditions that are more appropriate to source code and software, protecting copyright while disambiguating conditions on freedom to access, modify, and (re)distribute source code over networked servers.

2.2.8. Cultural Heritage (WP13)

In terms of **findability**, the cultural heritage FIP noted the adoption and use of generic and globally oriented PID registry and discovery systems such as DataCite, ORCID, Dublin Core, and Research Organization Registry (ROR)[80]. Regional systems such as the Europeana Data Model for Cultural

[70] https://www.unidata.ucar.edu/software/netcdf/
[71] https://www.iso.org/standard/53798.html
[72] https://www.iso.org/standard/67253.html
[73] https://www.earthobservations.org/geoss.php
[74] https://pacificdata.org/
[75] https://gawsis.meteoswiss.ch/GAWSIS/#/
[76] https://public.wmo.int/en/wmo-information-system-wis
[77] https://www.ogc.org/standard/cat/
[78] https://cfconventions.org/Data/cf-standard-names/current/build/cf-standard-name-table.html
[79] https://www.gnu.org/licenses/agpl-3.0.en.html
[80] https://ror.org/

'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.

Heritage (EDM)[81] and national systems such as that provided by the Digital Repository of Ireland (DRI)[82] were also noted, in line with the case study's localisation.

Accessibility is provided via standard web protocols and web architectural patterns supporting API-based exchanges, with OAuth access delegation noted as well as user control via manual login interfaces. CoreTrustSeal-certified archives were noted, ensuring longevity of access.

Semantic **interoperability** once again leveraged standard encodings/serialisations such as JSON-LD, RDF, and XML, with notable absence of OWL and SKOS. Given the focus on image-based digital assets (meta)data specifications bundled into image formats such as JPEG, TIFF, and PNG were noted as knowledge representation languages. An impressive range of structured vocabularies, thesauri, and artefacts of greater semantic expressivity were noted including generic resources such as Dublin Core, the Nomenclature of Territorial Units for Statistics (NUTS)[83], and the EXchangeable Image File (EXIF) format for storing interchange information in digital images, alongside domain-specific vocabularies (e.g. PeriodO[84] for identifying historical periods, and Getty's Architecture and Art Thesaurus[85]). Further, thesauri such as UNESCO's subject thesaurus[86] represent a form of semantic 'middle ground' in a multidomain setting, allowing topic identification within the broad scope of UNESCO's mandate.

Licencing information for **reusability** once again referenced the Creative Commons framework, and also the Open Data Commons framework, with reference to more narrowly scoped (although not domain-specific) licences such as the Open COVID Pledge[87], intended to facilitate mitigation responses to the pandemic by time-limited loosening of reuse restrictions. Provenance approaches in this FIP referenced the United States Library of Congress' Data Dictionary for Preservation Metadata (PREMIS)[88] which consists of a data dictionary, an XML schema, and supporting documentation and is distributed in OWL to express its semantic structure.

2.3. Synthesis and recommendations for ODIS

Each FIP conducted during WorldFAIR has revealed potential routes towards interoperability with ODIS. Below, a number of persistent patterns have been synthesised into recommendations for the development of ODIS during WorldFAIR and its eventual alignment to CDIF. These inform the strategic and implementation-level considerations in Section 4. As in the subsections above, these

[81] https://pro.europeana.eu/page/edm-documentation
[82] https://www.dri.ie/
[83] https://ec.europa.eu/eurostat/web/nuts/background
[84] https://perio.do/en/
[85] https://www.getty.edu/research/tools/vocabularies/aat/
[86] https://vocabularies.unesco.org/browser/thesaurus/en/page/concept1740
[87] https://opencovidpledge.org/v1-1-ocl-pc/
[88] http://www.loc.gov/standards/premis/

'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.

are grouped with the FAIR Principle they most pertain to, prefaced with a set of general considerations.

General considerations:

- Multiple resources noted in the FIPs above have already implemented data exchange interfaces leveraging JSON(-LD) and/or schema.org, particularly the Dataset type[89]. If the (meta)data records shared through these interfaces are valid, opportunistic bridges may be readily built to the ODIS federation. Naturally, the content shared through these links would require filtering for relevance to ocean issues before integration into ODIS.

- Several FIPs (e.g. Chemistry, Health) noted that much of their digital exchange is stabilised by authoritative, international organisations to standardise terminologies, metadata conventions, and other cornerstones of interoperability. These provide rallying points for content harmonisation, even if local implementations use different/incompatible technologies at present. Where possible, systems such as ODIS should leverage such scenarios, as forging community consensus on content may take decades, while implementing technologies to relay content in a given form is more a matter of years.

- Domain-specific conventions for (meta)data exchange (e.g. DwC, GEOMS, OMOP CDM) are necessary to cover the precise needs of specific communities of practice. However, these are difficult for human and machine agents in other domains to understand and act on without considerable investment in establishing and maintaining a plurality of mappings (semantic, syntactic, conceptual, etc.) and adaptor software. This is further hindered by uneven capacity, resourcing, know-how, and prioritisation of external interoperability in each of these domains. Systems such as ODIS are not poised to resolve such discrepancies; however, they can generate specifications to allow domains to encapsulate and transmit their conventions in generic JSON-LD/schema.org patterns[90]. This grants visibility and transparency to the number of digital resources that are at play in digital ecosystems, which provides a continuous landscape analysis and scoping of the interoperability challenge for more targeted action.

Findability:

- ***Establish partnership with domain-centric hubs****:* While DataCite, CrossRef, Zenodo, Dataverse, and other domain-neutral systems to issue PIDs, host data, and provide discoverability services are prevalent across many case studies, domain-specific

[89] https://schema.org/Dataset

[90] e.g. the embedding of term sets https://github.com/iodepo/odis-arch/blob/b2e3b85e928b2acaee622dddbd1110b38a71e5b7/book/thematics/terms/graphs/term.json#L1-L33

'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.

infrastructures which have sustainable support and implement well-adopted and technically sound standards and conventions are more likely to provide deep interoperability with a given domain. Establishing partnerships — and, where possible, (meta)data exchange — with these entities will provide ODIS with the socio-technical orientation and consultation needed to create lasting interoperability. This insight is needed to approach more generic data aggregators and understand how domain standards have been implemented in their systems.

- ***Identify and leverage overlap between ODIS-Arch/schema.org and domain-specific conventions****:* As observed in the general considerations above, many FIPs noted the international organisations and de facto standards consortia which stabilise the type of information exchanged within a domain. Much of the content in such specifications (e.g. the OMOP CDM) overlaps with generic schema.org types used in the ODIS Architecture, and could thus offer ready crosswalks for interoperable digital exchange.

Accessibility:

- ***Prepare for authentication and access delegation:*** Overall, access protocols reported across the FIPs are homogenous and conform to near-universal web standards. A number of WorldFAIR FIPs note the use of access control and authentication systems for sensitive data and/or restricted repositories. ODIS has not implemented such capacities; however, it is likely these will become increasingly prevalent across domains and preparing accordingly is advisable.

- ***Increase transparency of preservation policies:*** Multiple FIPs noted the use of repositories certified by CoreTrustSeal and other frameworks designed to regularise access over the long term. Machine-actionable indicators that state that (meta)data records will persist — and which themselves are associated with those (meta)data — are important. In the absence of these, where long-term access is not evidently assured, stakeholders may wish to allocate resources to mirroring content. Cataloguing and encoding access policies is a large task; however, the ODIS federation is in a position to approach it by providing guidance to its partners on how to encode such policies in ODIS-Arch.

Interoperability:

- ***Deconvolute implicit and explicit semantics.*** Many FIPs noted data structuring / data formatting conventions as knowledge representation approaches and/or resources for semantic interoperability. While such conventions can be said to bear meaning, that meaning is often only accessible to those agents with pre-existing familiarity with the convention(s) used. In multi-domain systems such as ODIS, such familiarity cannot be assumed, and the use of dedicated and generic knowledge representation languages (e.g. SKOS, OWL) is to be encouraged to accelerate cross-domain semantic exchange. Where the

'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.

latter form of semantic technology is identified, more effort in bridging domains can be justified.

- ***Identify and prioritise usage of the most sustained, well-curated, and expressive semantic resources used in the domain***. Closely following the preceding item, the semantic FERs identified in the FIPs above show a considerable range of maturity and technical potential. The more semantically expressive these are (i.e. terms interlinked with well-defined relations following SKOS, RDF, or OWL conventions), the more "graph-like" and machine-actionable they become. As the JSON-LD/schema.org records shared across the ODIS federation can be assembled into a semantically qualified knowledge graph[91], it is highly desirable to interlink and bridge mature semantic resources to the ODIS graph to allow access to domain-specific knowledge representations, reasoning, and querying capability to discover new content and associations[92].

- ***Prioritise semantic resources aligned to globally standardised terminologies***. Following from the similar point in the general considerations, more sustained cross-domain bridges are to be expected if semantic resources express well-established and well-governed terminologies or classifications. Multiple FIPs noted that international bodies such as IUPAC, the WMO, the WHO, and ISO provide standardised and well-adopted conventions/term lists for semantic alignment. Where these are expressed through mature knowledge representation systems and standards, ODIS is likely to find more lasting value in creating broadly impactful cross-domain bridges.

Reusability:

- ***Establish conventions for communicating licence information at high granularity:*** While the FIPs share a relatively small set of licencing norms, no FERs provide authoritative PIDs for the various licences noted. Free text or URL-based identification of licences is an option, but one that is difficult to regularise across many stakeholders, potentially leading to errors. Further, many (meta)data records do not have licence information included, at times assuming the policies of the hosting infrastructure will be sufficient. This poses risks, as (meta)data records can easily be dissociated from the repository in which they were initially published. ODIS is not in a position to address this issue across all domains (deferring to the CDIF), but can and should begin to propose more rigorous approaches as a trial for more globally coordinated action.

[91] http://graph.oceaninfohub.org/ ; https://book.oceaninfohub.org/users/query.html

[92] An example of an ontology being linked to a JSON-LD/schema.org graph in an ODIS record is available in the code snippet, available here: https://book.oceaninfohub.org/thematics/variables/index.html#variablemeasured. schema.org propertyIDs are populated with PIDs to ontology terms, which bridge ODIS graph to, e.g., OWL artifacts such as the Environment Ontology (http://purl.obolibrary.org/obo/ENVO_01001374)

'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.

- ***Extend support for W3C PROV to support inter-domain provenance exchange.*** Several FIPs noted domain-specific or regional/national provenance models (e.g. those in PREMIS and the DDI-Codebook). The ODIS interoperability architecture could, in principle, support the embedding of such models; however, their utility across the system's stakeholders will be limited. Simultaneously, the schema.org properties that support provenance tracking[93] are limited relative to W3C's PROV conventions. To provide a neutral bridge, ODIS should explore the embedding of PROV into its interoperability architecture. This may offer a relatively domain-neutral solution, which systems using narrower conventions can project their provenance metadata into with relative ease and clear added value.

3. Survey results and summary of insights

To augment the information provided in the FIPs, a brief, exploratory survey was circulated to the WorldFAIR consortium as well as to their wider communities through professional mailing lists and social media channels. The goal of this survey was to detect further opportunities to extend ODIS in aid of cross-domain interoperability. As with the examination of the FIPs in Section 2 of this report, promising opportunities are synthesised and summarised in this section. Note that a large proportion of respondents self-assigned to the oceanography domain. While their input is valued, here emphasis is placed on the responses from other domains.

3.1. Respondent profile

A total of 40 responses were gathered, with the domains of oceans, biodiversity, nanomaterials, and cultural heritage providing the highest number of responses (11, 5, 4, 3 responses, respectively: see Figure A1 in Appendix 1). Respondents' areas of operation showed concentration in Northern America (11), Northern Europe (9), Western Europe (7), and Australia and New Zealand (5), with none in Southern Asia and Central Asia (Figure A2). The number and distribution of the responses preclude this survey being taken as representative of any domain, however, the insights have value — particularly in combination with the FIP assessments above — in guiding ODIS to domain-specific stakeholders for further engagement.

The majority of respondents (c.74%, Figure A3) self-identified as very aware of the overall state of data and digital capacity in their domain, as well as high familiarity with the FAIR Principles (c.90%, Figure A4). The majority of respondents (c. 64%, Figure A5) responded that typical practitioners in their domain would have low to middling familiarity with the FAIR Principles.

[93] https://book.oceaninfohub.org/thematics/identifier/id.html?highlight=prov#about

'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.

3.2. Summary of responses

Across all responses, respondents indicated that the **most likely route to finding digital assets** (Figure A6) was through conducting searches across academic journals, followed by conducting generic web searches, and then by consulting colleagues on an *ad hoc* basis. Other likely routes included consulting *a few* well-known domain-specific portals, consulting colleagues, or consulting *several* domain-specific portals and data catalogues. The use of specialised data searches using generic engines like Google Dataset Search were primarily ranked as neither likely nor unlikely, with symmetrical distributions of responses indicating greater or less likelihoods.

Overall, the largest percentage of respondents (38.5%, Figure A7) indicated that fewer than five **prevalent (meta)data standards or conventions** were dominant in their domain, with the next largest percentages (12.8% each) indicating that 5-10 standards/conventions prevailed, that there is no appreciable standardisation in their domain, or that each research / working group uses its own, local standards.

In terms of the inclusion of **licence information supporting reus**e (Figure A8), respondents generally indicated that it was neither likely nor unlikely that licencing information was included in any modality (e.g. in metadata records, in documentation linked to data, in general policies used by data sources), with a response of "likely" similarly distributed across modalities.

When asked to estimate the average **digital fluency** of practitioners in their domains, respondents predominately ranked this as low to middling (44.7% as "2" and 23.7% as "3" on a five-point scale, with "5" as high fluency).

When asked to identify the top three **repositories** which diverse, independent practitioners in their domain used for trusted, long-term (>10 years) archival content, respondents provided a wide range of responses. These included global, national and institutional resources, reflecting a considerable variation in scale and localisation and suggesting local silos within domains. Three respondents were unable to identify any trusted long-term repositories. Reflecting the response counts per domain (see subsection 3.1.) biodiversity repositories such as the National Center for Biological Information (NCBI) with its membership in globally synchronised database consortia such as the International Nucleotide Database Collaboration (INSDC) and biodiversity data aggregators such as the Ocean Biodiversity Information System (OBIS) and the Global Biodiversity Information Facility (GBIF) were noted most frequently, alongside Zenodo as a domain-neutral archive (n=4). Other repositories noted included PANGAEA (n=3), FigShare[94], Dryad, eNanoMapper[95], and the National Centers for Environmental Information[96] (n=2). Despite only one response registered, key hubs of Disaster Risk Reduction information were noted in the UN Space-based Information for Disaster Management

[94] https://figshare.com/
[95] http://enanomapper.net/
[96] https://www.ncei.noaa.gov/

'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.

and Emergency Response (SPIDER) knowledge portal[97], the UNDRR's DesInventar system[98], and the UN Office for the Coordination of Humanitarian Affairs (OCHA)'s ReliefWeb system[99]. The three responses self-assigned to the Cultural Heritage domain noted the Digital Archaeological Record (tDAR)[100], the archives of the Scotland-based Canmore National Record of the Historic Environment[101], the archaeological content of the Netherlands-based Data Archiving and Networked Services (DANS) system[102], the UK-based Archaeology Data Service[103], and the cultural features recorded in the British Geological Survey's offshore Geoindex[104].

When asked to identify the three most used, domain-specific **search and discovery systems** in their domain, several respondents reported generic search/multi-domain portals (n=4) or were not able to provide an answer (n=9). However, several domain-specific search and discovery systems were identified. Notable examples in Biodiversity were the search and discovery interfaces used by the data archiving systems noted above, with the addition of the World Register of Marine Species (WoRMS)[105]. (Meta)data discovery systems related to chemical entities were reported and associated with repositories of chemical (meta)data such as Pubchem[106], ChemSpider[107], and UniChem[108]. Similarly, the responses for Disaster Risk Reduction directed attention to discovery systems linked to the World Bank Data Catalogue[109], the Open Data for Resilience Initiative[110], the Pacific Risk Information System[111], and the EU's Copernicus portal[112]. Respondents self-classifying as representatives of the Cultural Heritage domain directed attention to the discovery functions offered by Europeana system (noted in the corresponding FIP), the EU's Ariadne Infrastructure[113], and the UK's Archaeology Data Service[114].

When asked about **dominant serialisation and format conventions**, the most prevalent responses were comma- and tab-separated value formats (CSV, TSV; n = 31), NetCDF (13), JSON (11), RDF (8), and Microsoft Excel (7), with less prevalence of both generic and domain/application-centric

[97] https://un-spider.org/
[98] https://www.desinventar.net/
[99] https://reliefweb.int/
[100] https://www.tdar.org
[101] https://canmore.org.uk/
[102] https://dans.knaw.nl/nl/
[103] https://archaeologydataservice.ac.uk/
[104] https://www.bgs.ac.uk/map-viewers/geoindex-offshore/
[105] https://www.marinespecies.org/
[106] https://pubchem.ncbi.nlm.nih.gov
[107] http://www.chemspider.com
[108] https://www.ebi.ac.uk/unichem
[109] https://datacatalog.worldbank.org/home
[110] https://www.opendri.org/
[111] https://risk.spc.int/
[112] https://www.copernicus.eu/en
[113] https://portal.ariadne-infrastructure.eu
[114] https://archaeologydataservice.ac.uk

'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.

formats such as extensible markup language (XML; n= 3) and Shapefile (3). The mix of generic (e.g. CSV, JSON), domain-centric (e.g. NetCDF across the Earth Observation community, FASTA in the biosciences) and application-centric (e.g. Shapefiles for geospatial data) is indicative of a multi-layered challenge: multi-domain data provisioning requires stable, quality checked and persistently maintained format conversion technology in a potentially combinatorial space.

3.3. Perspectives for ODIS

This informal survey — while limited in its coverage and therefore not to be treated as representative of the domains approached — has provided valuable guidance on where and how ODIS can engage other domains in the WorldFAIR consortium to demonstrate and secure cross-domain interoperability. This section focuses on potentials to bridge ODIS to digital systems in biodiversity, chemistry, cultural heritage, and disaster risk reduction (see section 4, for further rationale).

In general, domains and/or subdomains where a small number of prevalent (meta)data standards, which have well-known portals and/or aggregators, and which have higher than average digital fluency among their practitioners are to be prioritised. These features would greatly simplify and expedite linking ODIS to trusted aggregators / infrastructures. Given the limited number of survey responses available, it is not possible to make broad conclusions on these dimensions, however, respondents did indicate multi-stakeholder infrastructures that aggregate and integrate domain (meta)data, which ODIS can readily interface with, validating and extending those suggested in the FIPs.

Supporting the findings of the **biodiversity** FIPs, infrastructures including GBIF, OBIS, INSDC/NCBI, and DataONE are natural focal points to engage, with metadata conforming to specifications such as EML, Darwin Core, and related resources from the Biodiversity Information Standards (TDWG)[115] organisation, as well as the Minimal Information about any (x) Sequence (MIxS) checklist[116] from the Genomic Standards Consortium (GSC)[117]. Raw or subject data access may prove more difficult, as a broad range of unstandardised serialisations and formats were reported.

Guidance from respondents self-classifying as **chemistry** data experts also highlighted a small number of data hubs to begin interoperability activities, with emphasis on PubChem as a prominent domain aggregator with native RDF encoding. Self-identified WorldFAIR participants responding to the survey also noted that more options would be coming via WorldFAIR Chemistry and other chemistry data initiatives.

[115] https://www.tdwg.org/
[116] http://www.gensc.org/pages/standards-intro.html
[117] http://www.gensc.org/

'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.

Cultural heritage respondents also highlighted metadata aggregators as nuclei of sustained interoperability, suggesting establishing contact with platforms such as Europeana and Ariadne, described above. The other data aggregators operating at national and subnational scale are also of interest, particularly those with direct marine relevance such as cultural features recorded in the British Geological Survey's offshore Geoindex, noted above.

The respondents self-classifying as **disaster risk reduction** experts emphasised that fruitful engagement is likely when this involves several of the digital aggregators and coordination organisations, including large agencies such as UN bodies (e.g. UN Environment, UNDRR, UN OCHA, Tsunami monitoring programmes) and regional infrastructures dedicated to hazard and disaster monitoring. As elaborated upon in Section 4, this is of dual strategic value for ODIS, addressing a pressing global need for human well-being and safety, while potentially strengthening ties to the UN data systems.

Respondents mentioned several other institutional and national repositories, discovery services, serialisation and semantic standards, and recommendations for building interoperability. While there is insufficient capacity to address them all during WorldFAIR, ODIS will attempt to identify any which a) serve JSON-LD/schema.org and b) host ocean-relevant content. Should willingness to join the ODIS federation exist, efforts will be made to facilitate connection, either directly or through existing regional partners.

4. Roadmap towards increased cross-domain interoperability

Informed by the assessment of the FIPs (Section 2) and insights from the informal survey (Section 3), this section presents a condensed and focused roadmap to cultivate greater cross-domain interoperability in ODIS. As it focuses on what can be accomplished in the timeframe of the WorldFAIR Project (June 2022-May 2024), the roadmap restricts itself to the most pragmatic pathways to secure cross-domain (meta)data flow with selected domains. A two-phased approach which relies on overlapping co-design and co-implementation is briefly described below. As stated in subsection 1.3, this roadmap focuses on 1) the expansion of ODIS-Arch's schema.org/JSON-LD thematic profiles and patterns to accommodate digital standards across more disciplines and priority topics of exchange, 2) alignment of ODIS implementation to the emerging recommendations in CDIF, and 3) testing the actionability and performance of activities 1 and 2 by both diagnostic evaluation and use on selected high-value data set examples to demonstrate added value to ocean data.

'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.

4.1. Priority areas

Through the synthesis of the FIPs, and with some insights gleaned from the informal survey summarised in section 3, this roadmap centres on the following priorities:

1. ***Developing cross-domain interoperability along existing strategic and thematic priorities***. Leveraging the insights gleaned from the FIPs and supplementary survey, and with the support of the WorldFAIR consortium, ODIS is better placed to forge new partnerships to address IOC priorities and the challenges of the Ocean Decade. In particular, effort will be focused on interoperating with domains which bear upon pressing issues in sustainable ocean management. Namely:

    a. The discovery, understanding, and protection of marine life and biodiversity, and deeper collaboration with UN Ocean Decade Programmes including the Ocean Biomolecular Observing Network (OBON)[118] and Marine Life 2030[119].

    b. Interoperability with chemical data hubs to provide information on ocean chemistry, marine pollutants, and bio(geo)chemical fluxes.

    c. Improved data flows around marine hazards and disasters, developing existing connections to systems coordinated by IOC-UNESCO, such as the Global Tsunami Warning System[120], WMO's WIMS, Global Telecommunications System (GTS)[121] and the Environmental Research Division's Data Access Program (ERDDAP)[122] utilised by the Global Ocean Observing System (GOOS).

    d. Integration of cultural data and information to increase appreciation of humanity's deep relationship with the ocean, and the multifaceted value it brings to human life and heritage. A special focus will be placed on image (meta)data, and its interoperation with emerging FAIR specifications for scientific imagery[123].

2. ***Co-development of continuous and mutualistic alignment and integration***. The remainder of WorldFAIR and the tasks in WP11 offer an opportunity to continuously align ODIS's strategy and implementation to the emerging principles and guidance of CDIF (WP2), while contributing to the CDIF itself through actively bridging domains. This will ready the members of the ODIS federation to more effectively engage with external domains converging on CDIF, with minimal disruption to their internal operations.

[118] https://www.obon-ocean.org/
[119] https://marinelife2030.org/
[120] https://ioc.unesco.org/our-work/global-tsunami-early-warning-and-mitigation-programme
[121] https://community.wmo.int/en/activity-areas/global-telecommunication-system-gts
[122] https://coastwatch.pfeg.noaa.gov/erddap/index.html
[123] https://www.nature.com/articles/s41597-022-01491-3

'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.

3. ***Establishing integration-on-demand as a foundation for a cross-domain data space***. In line with Europe's digital transformation agenda[124], ODIS seeks to create a foundation for a far more encompassing and mature digital ecosystem, where components are able to push and pull digital assets according to the standards they require, when they require them (i.e. an integration on demand 'data space'). To this end, ODIS will engage the WorldFAIR partners identified in this roadmap to establish a viable, initial path to a multi-domain data space, first through harmonised metadata exchange leveraging authoritative mappings from domain-specific to domain-neutral standards, and - where possible - through negotiated transfer of the data they describe. While a comprehensive solution is beyond the scope of WorldFAIR, the priority is to demonstrate a viable route to expand in future initiatives.

4.2. Phase 1 – Establishing focus groups and securing co-implementation partners

The ODIS federation was built and is maintained by frequent interaction across the partnership, coordinated by IODE. The operational scenarios of each implementation partner are taken into account while their needs are transferred to the ODIS interoperability architecture and (meta)data exchanges are established. In Phase 1 of implementing cross-domain interoperability, the same ODIS workflow will be deployed with representatives of the WorldFAIR case studies identified in Section 4.1. This phase will begin in March 2023, and be maintained in parallel to Phase 2 to reinforce co-implementation.

Notably, a number of ODIS partners are already active in the biodiversity and disaster risk reduction domains. This offers opportunities to expand the communities co-designing interoperability specifications, ultimately boosting their reach and impact. Convening co-development sessions with representatives from OBIS, GEO BON, OBON, Omic BON and thematically related European initiatives such as the MARine COastal BiOdiversity Long-term Observations (MARCO-BOLO) project[125], EuroGOOS, and EuropaBON will offer opportunities to broaden and solidify interoperability and data delivery to decision makers via Essential Ocean[126] and Biodiversity[127] Variable Frameworks. ODIS also has existing links to NODCs in regions where tsunami early warning systems are in place, as well as to UNDRR. These stakeholders will be re-engaged through WorldFAIR-focused activity.

The domains of chemistry and cultural heritage are not yet well represented in ODIS. The efforts now beginning through WorldFAIR will thus pioneer interoperability at these interfaces, adding value to all participants.

[124] https://digital-strategy.ec.europa.eu/en
[125] https://cordis.europa.eu/project/id/101082021
[126] https://www.goosocean.org/eov
[127] https://github.com/iodepo/odis-arch/issues/170

'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.

4.3. Phase 2 – Progressive co-implementation and strategic alignment

Earlier sections of this report provided recommendations on how to approach the implementation of cross-domain (meta)data exchange. The identification of target domains above, and the engagement approach in Phase 1, must now be complemented with concrete and testable implementation of cross-domain, FAIR (meta)data exchange. During this co-implementation phase, which will begin in May 2023, WP11 will:

1. Transfer the thematic interoperability conventions currently specified in ODIS-Arch[128] into a domain-neutral space, under co-governance by a wider multidomain consortium. This will support more collaborative and community-driven evolution and synchrony with external standards.

2. Through regular co-development sessions with the focus groups established in Phase 1, create and/or extend ODIS-Arch interoperability patterns to accommodate key digital assets from the target domains identified above. Work for hazards and disasters[129] and imagery[130] (relevant to many digitised cultural artefacts and records) is already underway.

3. Where needed, support implementation partners in exposing interoperable (meta)data via JSON-LD/schema.org, leveraging community work to — for example — link domain-centric conventions in biodiversity (EML, biomolecular data) to the JSON-LD/schema.org approaches used in ODIS-Arch[131,132].

4. Harvest ocean-relevant content from partner systems and integrate these into the OIH graph, ideally leveraging the generic publisher[133] and aggregator[134] approaches.

5. Support co-implementation partners in discovering content in ODIS relevant to their domain and harvesting it into their systems.

5. Conclusions

In conclusion, this deliverable has drawn from the WorldFAIR FIPs and consortium insight to identify the most viable routes to establish and sustain cross-domain interoperability. In the next phase of WP11, the syntheses and phases of the high-level roadmap will be executed and shaped by the specific needs of each implementation partner. By dovetailing the efforts under WorldFAIR with

[128] https://book.oceaninfohub.org/thematics/README.html
[129] https://github.com/iodepo/odis-arch/issues/110
[130] https://github.com/iodepo/odis-arch/issues/144
[131] https://github.com/ESIPFed/science-on-schema.org/issues/238
[132] https://github.com/iodepo/odis-arch/issues/146
[133] https://book.oceaninfohub.org/publishing/publishing.html
[134] https://book.oceaninfohub.org/indexing/index.html

'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.

existing initiatives within ODIS's existing multi-domain partnership, alignment and impact will be multiplied and leveraged across several other European and international projects, consortia, and infrastructures. Subsequent WP11 deliverables will focus on the concrete outputs of these efforts, and their alignment with CDIF.

'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.

6. Bibliography

Buttigieg PL , Curdt C, Ihsan AZ, Jejkal T, Kubin M, Mannix O, et al. (2022) An interpretation of the FAIR principles to guide implementations in the HMC digital ecosystem. HMC Paper 1 . HMC-Office, GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany, 26 pp. DOI 10.3289/HMC_publ_01.

Fils D, Scott L, Buttigieg PL, Spears T, Lambert A, Provoost P, et al. Ocean InfoHub: A Global Knowledge Network for the Ocean Data and Information System (ODIS). AGU Fall Meeting Abstracts 2021, IN45H-0523

Gottardo et al (2016) NANoREG harmonised terminology for environmental health and safety assessment of nanomaterials; EUR 27808; doi:10.2788/71213

Gregory A, Hodson S. (2022). WorldFAIR Project (D2.1) 'FAIR Implementation Profiles (FIPs) in WorldFAIR: What Have We Learnt?' (1.0). Zenodo. https://doi.org/10.5281/zenodo.7378109

Martinez DST, Da Silva GH, de Medeiros AMZ, et al. (2020) "Effect of the Albumin Corona on the Toxicity of Combined Graphene Oxide and Cadmium to Daphnia magna and Integration of the Datasets into the NanoCommons Knowledge Base" Nanomaterials 10, no. 10: 1936. https://doi.org/10.3390/nano10101936

Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, et al., (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018. https://doi.org/10.1038/sdata.2016.18

'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.

Appendix 1 – Survey responses

Which domain does your work focus on? (Please choose only one. If more than one applies to you, please consider filling out this form again from that perspective)

39 responses

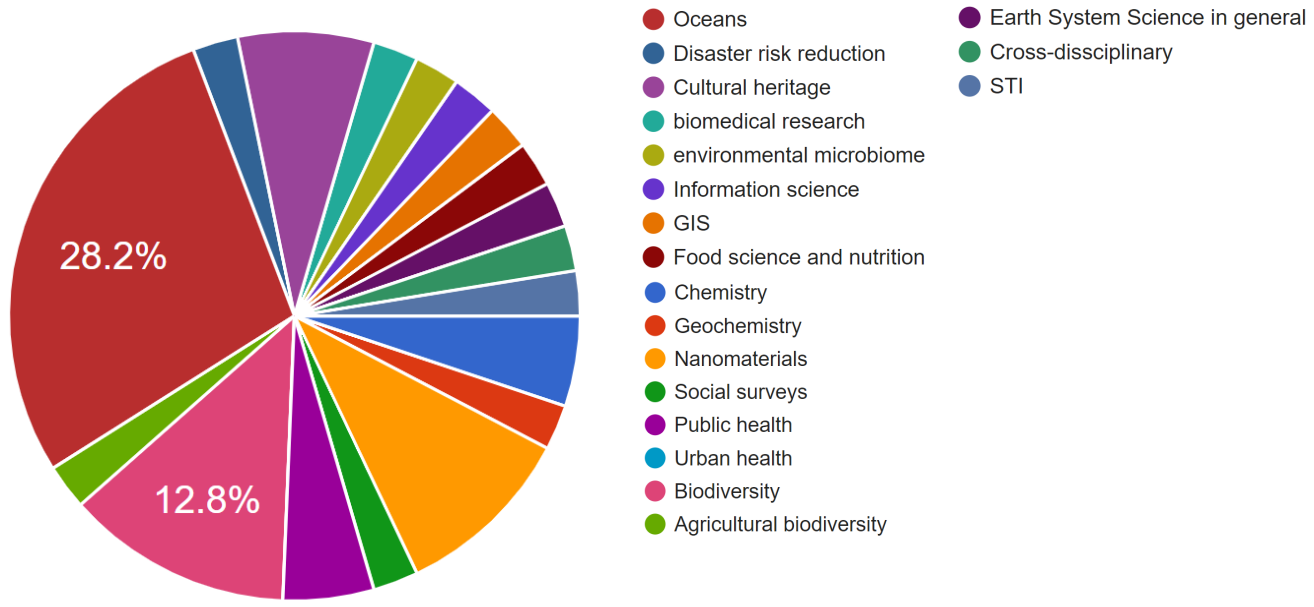

Figure A1: Self-reported domain of the respondent.

'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.

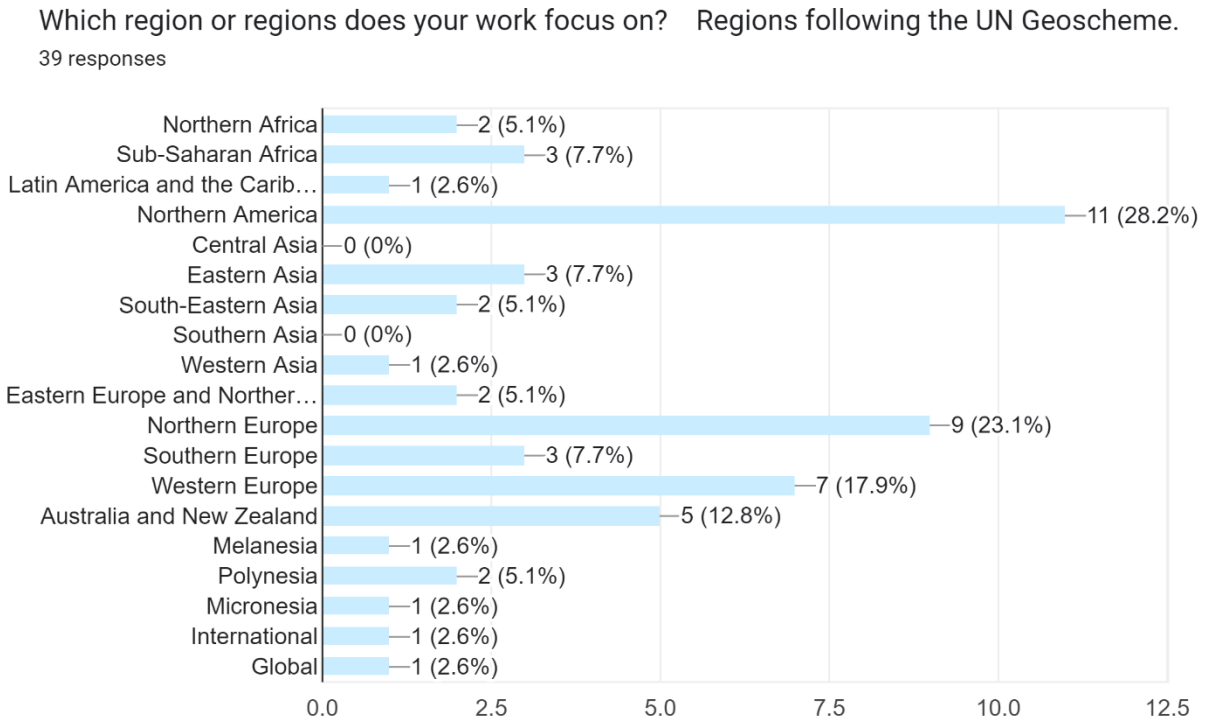

Figure A2: Overview of respondents' area(s) of operation.

'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.

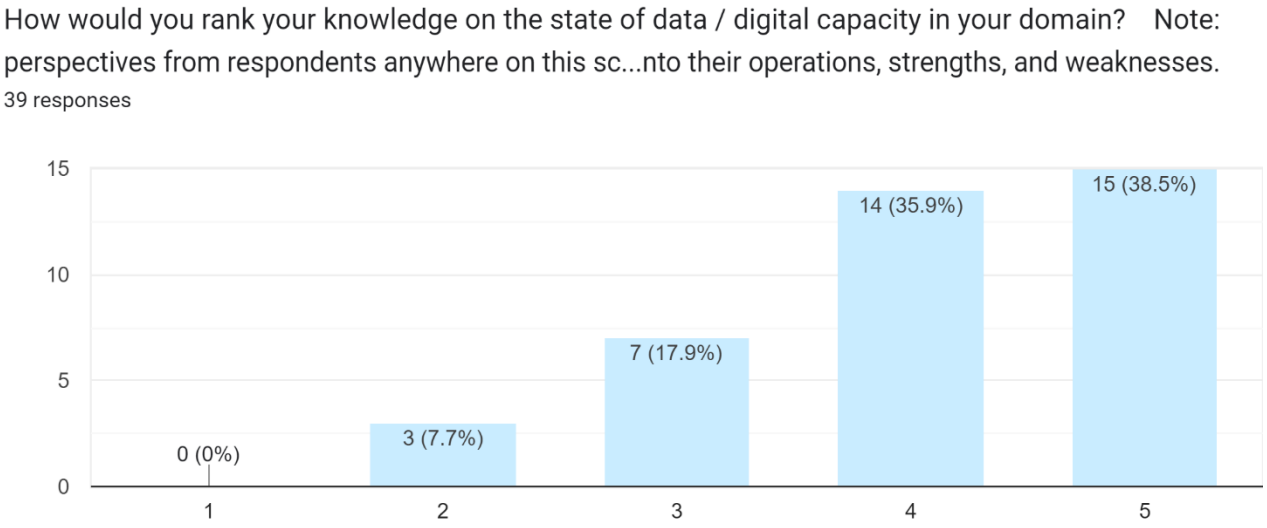

Figure A3: Respondents' familiarity with the state of data and digital capacity in their domain (1 = low, 5 = high).

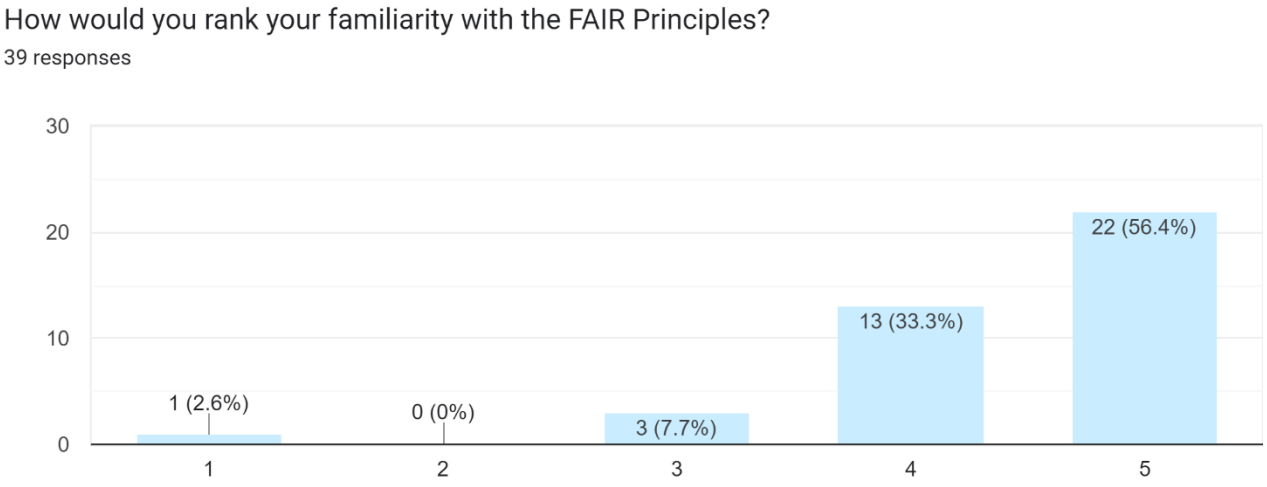

Figure A4: Respondents' familiarity with the FAIR Principles (1 = low, 5 = high).

'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.

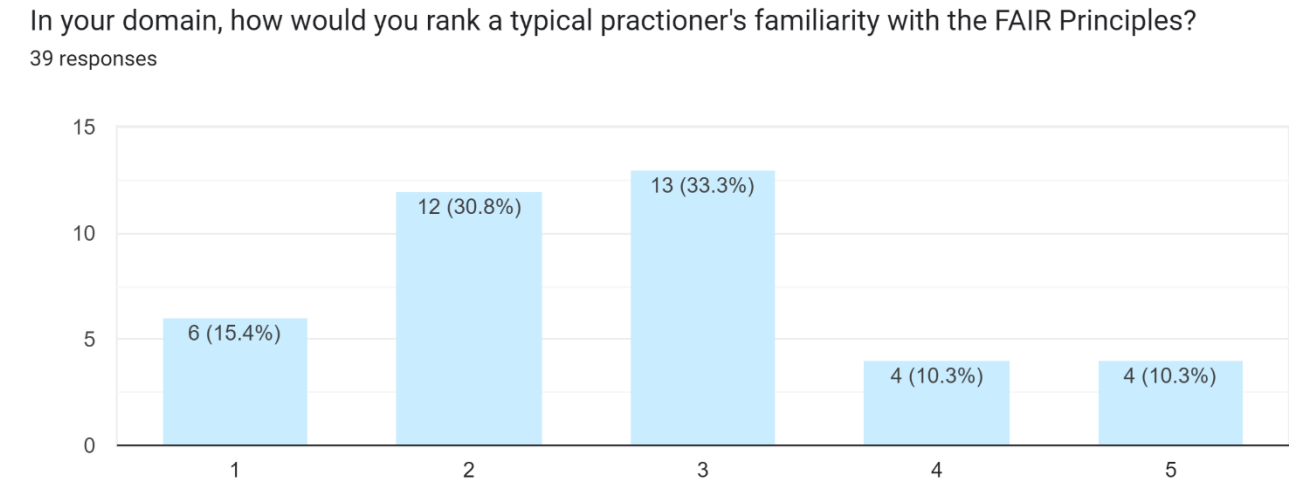

Figure A5: Respondents' impression of a typical practitioner's familiarity with the FAIR Principles, in their domain (1 = low, 5 = high).

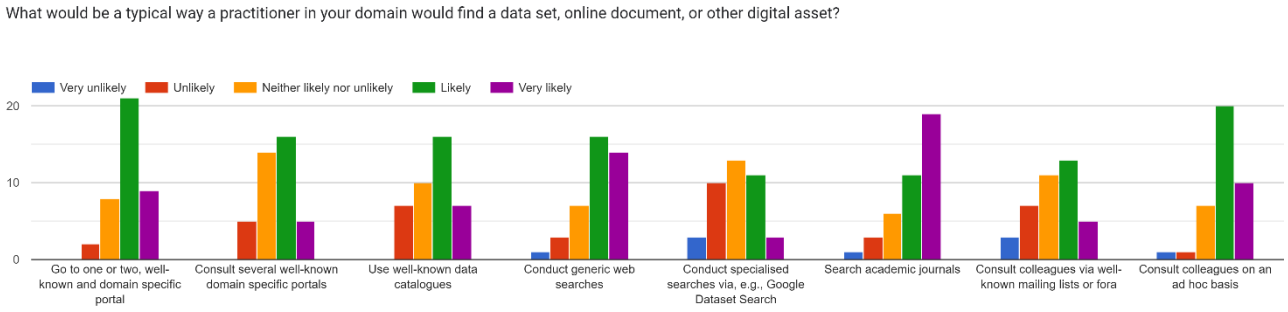

Figure A6: Routes to (meta)data discovery

'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.

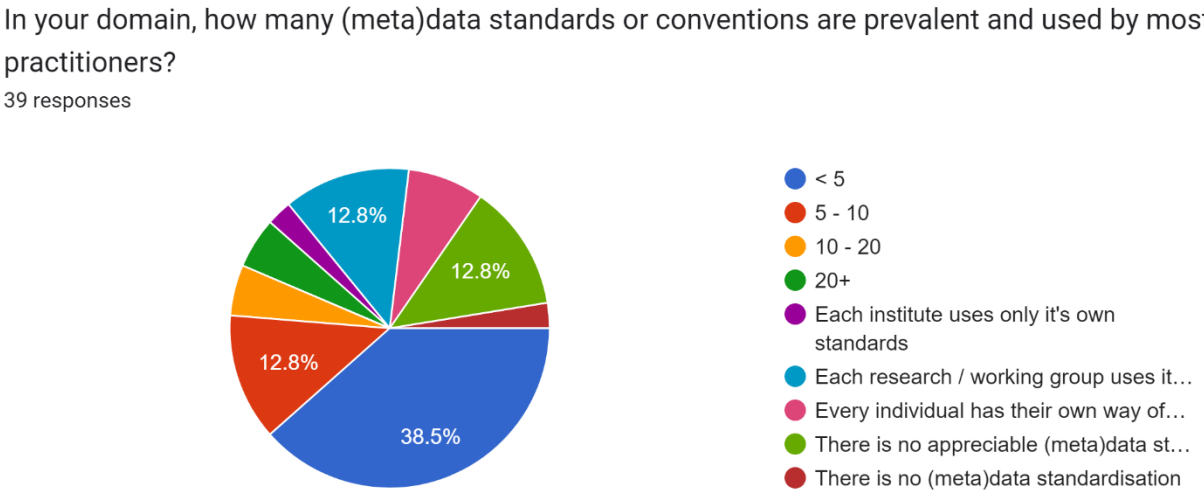

Figure A7: Plurality of prevalent metadata standards and conventions

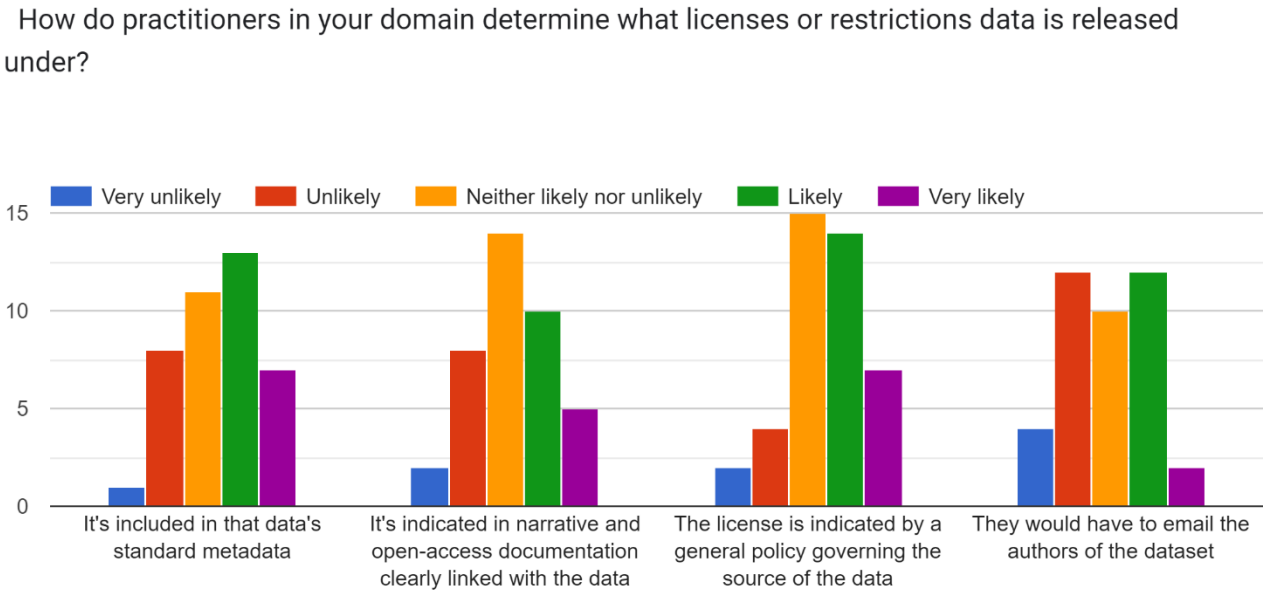

Figure A8: Inclusion of licencing metadata

'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.

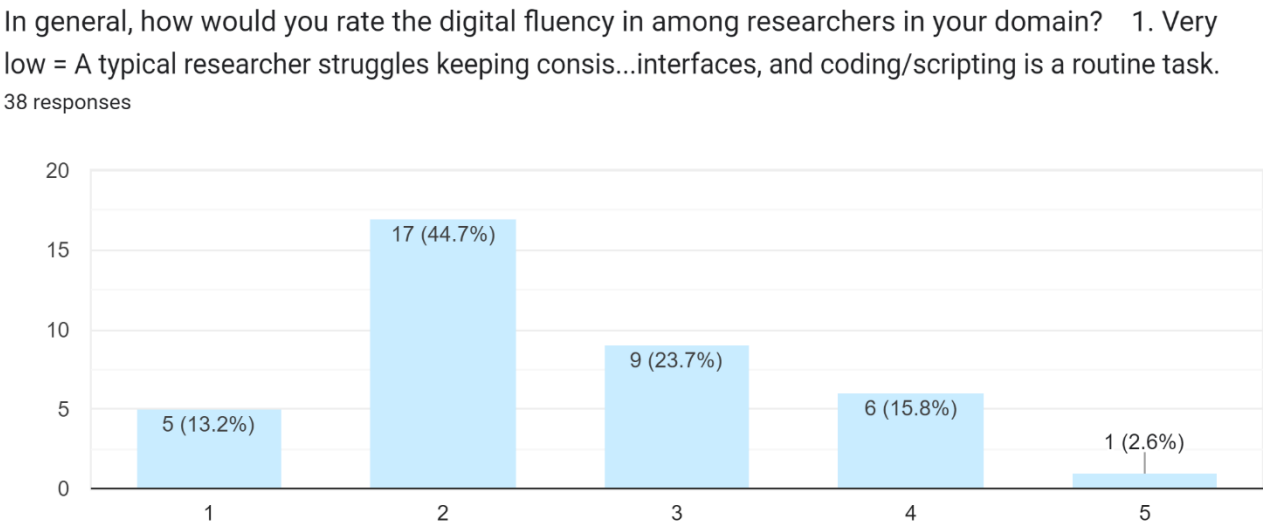

Figure A9: Estimation of digital fluency of domain practitioners

'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.