



Project Title Fostering FAIR Data Practices in Europe
Project Acronym FAIRsFAIR
Grant Agreement No 831558
Instrument H2020-INFRAEOSC-2018-4
Topic INFRAEOSC-05-2018-2019 Support to the EOSC Governance
Start Date of Project 1st March 2019
Duration of Project 36 months
Project Website www.fairsfair.eu

D3.1 FAIR POLICY LANDSCAPE ANALYSIS

Work Package	WP3, FAIR Data Policy and Practice
Lead Author (Org)	DCC
Contributing Author(s) (Org)	UGOE, SPARC Europe, EUA
Due Date	30.11.2019, M9
Date	28.11.2019
Version	1.0, DRAFT VERSION NOT YET APPROVED BY THE EUROPEAN COMMISSION
DOI	https://doi.org/10.5281/zenodo.3558173

<input checked="" type="checkbox"/>	PU: Public
<input type="checkbox"/>	PP: Restricted to other programme participants (including the Commission)
<input type="checkbox"/>	RE: Restricted to a group specified by the consortium (including the Commission)
<input type="checkbox"/>	CO: Confidential, only for members of the consortium (including the Commission)

Abstract

Policies are a crucial component in the FAIR ecosystem. To this end, FAIRsFAIR Work Package 3 (WP3): FAIR Data Policy and Practice carried out an analysis of the current data policy landscape at various levels (national, funder, publisher, institutional) to provide a snapshot of the situation in 2019 and to identify policy elements that support or hinder FAIR data practice. To provide a comparative baseline for reviewing the data policies of various stakeholders, the priority and supporting actions presented in the *Turning FAIR into Reality* (TFiR) action plan were employed. To assess how well the policies of different stakeholders currently reflect TFiR's action plan, we carried out desk research to characterise policies, undertook an analysis of responses to an open consultation, and conducted a small number of interviews. This report presents the findings of these landscape assessment activities and provides an evidence base for FAIRsFAIR to build upon as work begins to define a set of practical recommendations to support policy enhancement (D3.3).



Versioning and contribution history

Version	Date	Authors	Notes
0.1	04/11/2019	Joy Davidson (DCC) Claudia Engelhardt (UGOE), Vanessa Proudman (SPARC Europe), Lennart Stoy (EUA), Angus Whyte (DCC)	Submitted for internal review
0.2	15/11/2019	Hylke Koers (SURFsara) Robin Dasler (DataCite)	Comments from internal reviewers provided on version 0.1
0.3	22/11/2019	Joy Davidson (DCC)	Revisions carried out in light of comments received during internal review
1.0	29/11/2019	Joy Davidson (DCC), Gerard Coen (DANS)	Final edit and publishing

Disclaimer

FAIRSF AIR has received funding from the European Commission's Horizon 2020 research and innovation programme under the Grant Agreement no. 831558. 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

EOSC	European Open Science Cloud (project)
FAIR	Findable, Accessible, Interoperable and Re-usable, compliance with the FAIR principles
RDA	Research Data Alliance (organization)
UGOE	University of Göttingen (organization)
EUA	European Universities Association (organization)
HEI	Higher Education Institution
TFiR	Turning FAIR into Reality (report)
TOP	Transparency and Openness Promotion
DOI	Digital Object Identifier
FsF	FAIRsFAIR (project)
DDC	Digital Curation Centre (organization)
STFC	Science and Technology Facilities Council (organization)
WP	Work Package
GEDE	Group of European Data Experts in RDA
IMI	Innovative Medicines Initiative
RDM	Research Data Management
UC3M	University Carlos III de Madrid (organization)



Executive summary

FAIRSFAR Work Package 3 (WP3): FAIR Data Policy and Practice carried out an analysis of the data policy landscape at various levels (national, funder, publisher, institutional) to provide a snapshot of the situation in 2019 and to identify policy elements that support or hinder FAIR data practice. The assessment shows that while the stakeholders' policies do reflect the actions stated in *Turning FAIR into Reality* to some extent, there is still work to be done to foster a FAIR ecosystem.

Key findings include:

- Efforts are needed to raise general awareness about the FAIR principles and the potential benefits of implementing them
- The policies of funding bodies are the key driver for many of the stakeholders developing policies – both at the national and institutional level
- The policies of all stakeholders should be described consistently using a structured data markup schema to support both human and machine readability
- Funding bodies and publishers/journal could strengthen their expectations around the sharing of both data and metadata
- Data management planning requirements should be harmonised across stakeholders and supported over the entire research lifecycle
- There is a need for clarity about which data repositories should seek certified status
- Cooperation to develop and curate a shared set of discipline-specific guidance and training resources is necessary
- Funding bodies and publishers could strengthen their requirements in relation to data citation and provide clearer guidance on how to do this in a standardised way
- Funding bodies and publishers should enforce the data sharing policies that they have put in place
- To encourage a FAIR ecosystem, the emphasis should be on providing incentives for good practice but penalties for non-compliance should be introduced where appropriate



Table of Contents

Executive summary	5
Table of Contents	6
Introduction	10
Approach	11
Methodology	13
Policy Characterisation	13
Open Consultation	16
FAIR Enough? workshop	17
FAIRSF AIR Focus Group on Research data policies and the FAIR data principles	17
Examples of good practice	18
Updating the policy landscape assessment	18
National Open Science Policies	19
Policy Characterisation mapped to Turning FAIR into Reality recommendations	27
About the policies	27
Date of policy introduction	28
Policy is machine readable	28
Policy has a persistent identifier (PID)	30
Content of the policies	30
Define - concepts for FAIR Digital objects and the ecosystem	30
Summary of landscape in relation to Define - concepts for FAIR Digital objects and the ecosystem	37
Implement - culture, technology and skills for FAIR practice	38
Summary of landscape in relation to Implement - culture, technology and skills for FAIR practice	48
Embed and Sustain - incentives, metrics and investment	49
Summary of landscape in relation to Embed and Sustain - incentives, metrics and investment	54
Results of the Open Consultation	55
Overview of respondents	55
Breakdown by organisation type	55



Breakdown by country	56
Provision of in-house support for FAIR	56
Roles of respondents	59
‘Other’ roles	60
Policy related questions	60
Assessment of drivers for policy development	60
Assessment of the influence of the research data related policies of the different stakeholder groups on researchers’ practice.	63
Assessment of policy factors	65
Conclusions	67
Key findings in relation to Define - concepts for FAIR Digital objects and the ecosystem	67
Key findings in relation to Implement - culture, technology and skills for FAIR practice	68
Key findings in relation to Embed and Sustain - incentives, metrics and investment	68
Next Steps	69
Annex 1 - Policy Characterisation Table	70
Annex 2 - FAIR Policy Analysis and indicators	74
Annex 3: Policies reviewed as part of policy characterisation	80
Annex 4 - Open Consultation - Policy related questions	83



1. Introduction

To ensure that the European Open Science Cloud (EOSC) is able to provide researchers with access to high quality and reusable data, FAIRSF AIR - Fostering Fair Data Practices in Europe - aims to support the emergence of a FAIR data culture (i.e., one where research data are findable, accessible, interoperable and reusable). To this end, FAIRSF AIR will collaboratively develop and provide practical support to a wide range of stakeholders including research communities, funding bodies, universities and research performing organisations, publishers, research infrastructure providers, data stewards and policymakers for implementing the FAIR data principles¹ over the research data life cycle.

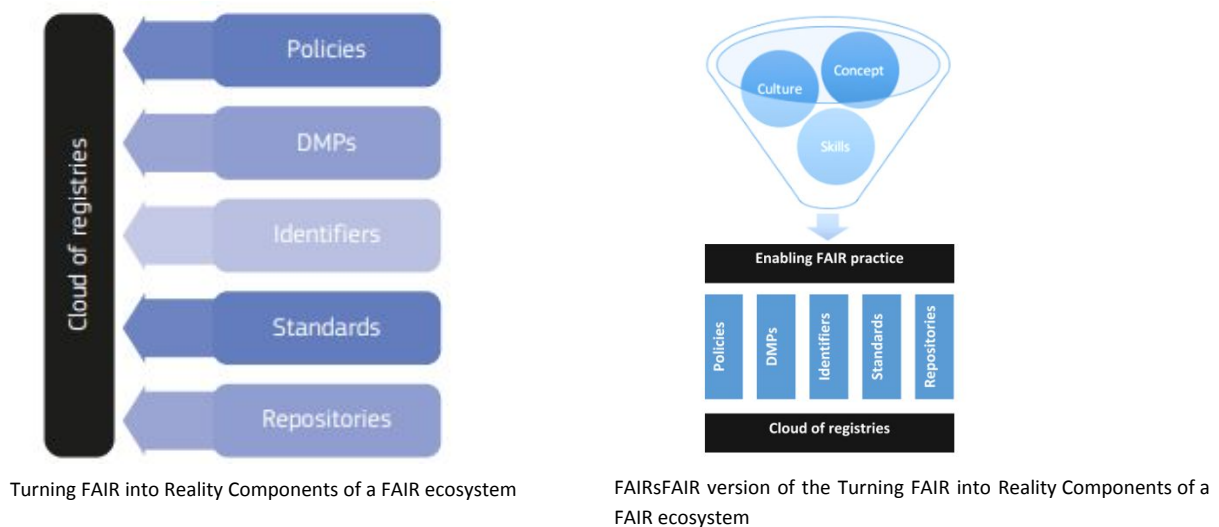
A number of publications outlining recommendations and action plans for implementing the FAIR principles have been produced in recent years. One of the most influential has been the *Turning FAIR into Reality* report² which was published in late 2018. FAIRSF AIR (FsF) has drawn upon the ecosystem model presented in the *Turning FAIR into Reality* (TFiR) report (Figure 1) to help align the activities within each of the FsF work packages with each focusing on one or more of the TFiR ecosystem components. During the initial phases of the FsF project, the partners defined an amended version of the TFiR ecosystem model to better reflect the importance of research culture on researchers' practices and to emphasise the need for training, skills and competency building as can be seen our amended version of the FAIR ecosystem below (Figure 1).

¹ Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016).

² Directorate General for Research and Innovation (European Commission). *Turning FAIR into reality*. <https://doi.org/10.2777/1524> (2018).



Figure 1. Components of a FAIR ecosystem



FAIRSFAR Work Package 3 (WP3): FAIR Data Policy and Practice has a focus on two of the FAIR ecosystem components - one on FAIR Data Policies and the other on FAIR Data Practices. According to TFiR, ‘policies define and regulate the components of the FAIR data ecosystem and their relationships.’ As a crucial component in the ecosystem, FsF carried out an analysis of current data policies at various levels (national, funder, publisher, institutional) to identify policy elements that support or hinder FAIR data practice. This report is the result of landscape assessment activities that were carried out between March and October 2019 as part of Work Package 3. Partners involved in this task included the Digital Curation Centre (DCC), SPARC Europe, Science and Technology Facilities Council (STFC), University of Göttingen (UGOE), and the European Universities Association (EUA). The analysis provides an evidence base for FAIRSFAR to define a set of practical recommendations to support policy enhancement (D3.3) and the results of the analysis will also be used to inform other areas of FAIRSFAR's work and may be of value for related initiatives’ activities.

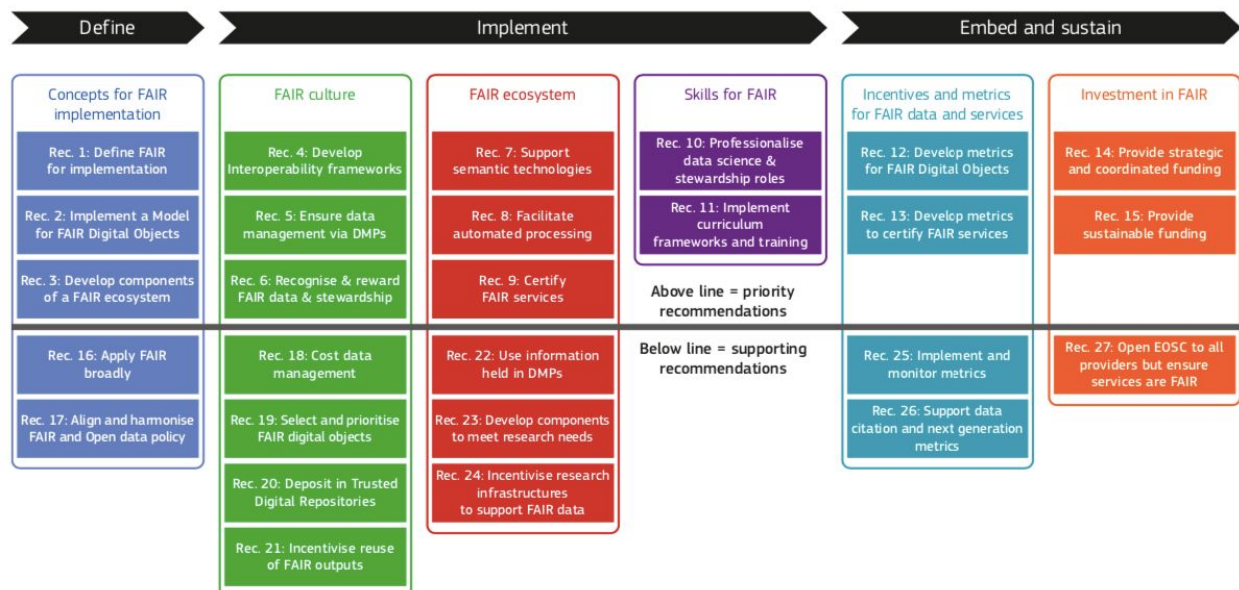
Approach

To provide a comparative baseline for reviewing FAIR data policies at the national level, and those of funders, publishers and Higher Education Institutions (HEIs), the priority and supporting actions presented in the *Turning FAIR into Reality* (TFiR) action plan have been employed. To avoid duplication of effort, the findings of the most recent Open Science policy

report prepared by SPARC Europe and the DCC³ in August 2019 was consulted to provide an overview of national level policies in Europe.

TFiR’s priority and supporting actions are grouped under three broad headings - *defining, implementing, and embedding and sustaining* (Figure 2). For consistency, this report has followed the same overall structure. Under each of the three broad headings, we present a summary stating to what extent the current European policy landscape reflects the actions outlined in TFiR. The analysis is based on desk research, an analysis of selected stakeholders’ data policies, the results of an open consultation, and a small number of interviews with stakeholders. Following the analysis of the current landscape, we present our conclusions providing an overview of the next steps to be taken in FAIRSFAR in light of these findings.

Figure 2. Turning FAIR into Reality priority and supporting recommendations



³ SPARC Europe, & Digital Curation Centre. (2019, August 28). An Analysis of Open Science Policies in Europe v4. Zenodo. <http://doi.org/10.5281/zenodo.3379705>



2. Methodology

As noted above, a picture of national level Open Science policies was derived from the recent SPARC Europe and DCC report. To get a sense of the current data policy landscape of three key stakeholders in the FAIR ecosystem - funders, publishers, and higher education institutions - and to assess how well they currently reflect TFIR's action plan, we carried out desk research to characterise policies, undertook an analysis of responses to an open consultation, and conducted a small number of interviews. Below, we provide more detail on each of the activities carried out as part of the landscape assessment.

Policy Characterisation

To support comparison across the data policies of different stakeholders, a standard set of policy features were distilled to enable policy characterisation. The policy characterisation features that were defined for FAIRsFAIR reflect and build upon the work already carried out by related initiatives already active in the field. The research data policy framework developed by the Data Policy Standardisation and Implementation Interest Group of the Research Data Alliance (RDA)⁴ was reviewed to help us align the policy features to be considered. This framework includes 14 policy features and is intended to help journal editors and publishers to create, enhance and harmonise their research data policies. The framework identifies six types of research data policy which require the inclusion of an increasing number of policy features and recommended levels of stringency as can be seen in Figure 3.

⁴ Hrynaszkiewicz I, Simons N, Hussain A, Goudie S. Developing a research data policy framework for all journals and publishers. Figshare. 2019. <https://doi.org/10.6084/m9.figshare.8223365.v1>



Figure 3. Fourteen journal research data policy features arranged as six policy types (tiers)

	Policy 01	Policy 02	Policy 03	Policy 04	Policy 05	Policy 06
Definition of the research data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exceptions to policy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Embargoes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Supplementary materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data repositories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data citation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data licensing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Researcher/ author support	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data availability statements		<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data formats and standards				<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Mandatory data sharing (specific data types)				<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Mandatory data sharing (all papers)				<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Peer review of data				<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Data Management Plans (DMPs)				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<input type="radio"/> Provide information The text for the policy feature will be included in the policy template but it is clear that the feature will not be enforced and checked as part of the publishing or peer review process	<input checked="" type="radio"/> Provide information and action The text of the policy feature is included and makes clear where applicable that the feature will be checked and enforced in the publishing or peer-review process
--	--

For the purpose of the FAIRsFAIR policy landscape assessment, we aimed to include policy features that would be common across the three stakeholder groups. We included almost all of the features identified by the RDA Interest Group in our set of policy features except for peer review of data as this is specific to publishers and is not something that we would expect to be covered by funding bodies or institutions. We also reviewed the Transparency and Openness



Promotion (TOP) guidelines⁵ which include eight modular standards, each with three levels of increasing stringency. We did not assess the policies of the three stakeholder communities in direct relation to the levels of stringency as identified by the RDA Interest Group and TOP Guidelines, but our assessment did consider variations in policy stringency.

We also looked at the policy features captured by FAIRsharing⁶. FAIRsharing aims to provide an 'Informative and educational resource that describes and interlinks community-driven standards, databases, repositories and data policies' that align with the FAIR Principles. FAIRsharing draws upon community cooperation to capture and curate information. Each record in FAIRsharing is assigned a digital object identifier (DOI) to support better discoverability and citation of the standards, databases, repositories and data policies and the content is licensed under Creative Commons Attribution ShareAlike 4.0 license (CC BY-SA 4.0). Currently, FAIRsharing provides records relating to the policies of 85 journals, 23 funding bodies, 9 Societies and 6 projects⁷.

Based on the review of the activities described above, FAIRsFAIR defined a total of 42 policy features to be reviewed during our landscape assessment. Fourteen of these features were intended to capture information about the policy itself (e.g., title, year of introduction, machine-readability) and 28 policy features were reviewed in relation to the content of the policy. The full set of policy characterisation features is available in Annex 1.

For each of the three stakeholder groups assessed in this report (funders, publishers and HEIs) we reviewed 10-15 data policies to determine how far they currently reflect the stakeholder-specific actions as recommended in TFIR. We limited the number to be reviewed to reflect the time and effort available for this task. To enable us to assess whether a specific policy element had been addressed within a given stakeholder's policy, a set of related indicators were developed and mapped to the policy features and related TFIR policy actions. These are available in Annex 2.

The policies selected for analysis as part of the landscape assessment come from stakeholders currently demonstrating a relatively high level of engagement with the FAIR principles. These have been identified through desk research and through the FAIRsFAIR consortium partners' subject knowledge and related activities. Based on their particular areas of expertise, UGOE identified publishers to be included in the review while SPARC Europe identified funding bodies

⁵ <https://cos.io/top/>

⁶ Sansone S-A, McQuilton P, Rocca-Serra P, Gonzalez-Beltran A, Izzo M, Lister AL, et al. FAIRsharing as a community approach to standards, repositories and policies. *Nat Biotechnol.* 2019;37: 358–367. doi:10.1038/s41587-019-0080-8

⁷ Number of policies registered with FAIRsharing as of 23/10/2019.



and EUA identified HEIs. The 42 policies reviewed for this task should not be considered an exhaustive inventory of stakeholders currently engaged with Open Science and the FAIR principles. The list of policies reviewed is presented in Annex 3.

Open Consultation

Landscaping activities have been a core activity during the initial stages of the FAIRsFAIR project and efforts to coordinate activity across the relevant FAIRsFAIR work packages were prioritised. To this end, there has been close cooperation with colleagues in Work Package 3 carrying out the FAIR data practice analysis (D3.2); Work Package 2 on assessing FAIR requirements for interoperability and persistence (D2.1); Work Package 6 on providing an overview of research communities' needs for competence centres; and Work Package 7 on mapping RDM policies and support as well as FAIR education offerings in European HEIs. In particular, efforts were made to avoid duplication of effort across the three open consultation and survey instruments developed to assess the current landscape and to define a consistent approach to presenting our findings.

The Policy and Practice open consultation included both open-ended and closed questions which sought to identify the different levels of maturity with regards to FAIR practices among disciplines, the range of policies that influence the way that researchers work, and the sources of support currently available to researchers. In developing the questions, FAIRsFAIR worked collaboratively with several related initiatives including the EOSC 5B projects⁸, the Group of European Data Experts in RDA (GEDE)⁹, and the EOSC FAIR Working Group¹⁰ and Landscape Working Group¹¹ to avoid duplication of effort in our information collection. FAIRsFAIR will continue to cooperate with these and other initiatives over the life of the project. The open consultation targeted members of the research support community to gain insights on their views and experiences in relation to implementing the FAIR principles. The open consultation questions were grouped under five broad themes:

- 1) Practice
- 2) Policy
- 3) Repositories
- 4) Skills
- 5) Competence centres

⁸ <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/infraeos-05-2018-2019>

⁹ <https://www.rd-alliance.org/groups/gede-group-european-data-experts-rda>

¹⁰ <https://www.eoscsecretariat.eu/working-groups/fair-working-group>

¹¹ <https://www.eoscsecretariat.eu/working-groups/landscape-working-group>



This report will provide an analysis of the policy-related questions. The policy-related questions included in the open consultation aimed to better understand which stakeholders' policies had the most influence on researchers' practices and what policy factors were most likely to positively or negatively influence researchers' behaviour. The questions asked in the open consultation are presented in Annex 4.

FAIR Enough? workshop

FAIRsFAIR participated in the FAIR Enough? Policies and Recommendations for the Processing of Sensitive Data workshop organised by the EOSC Life project. The workshop took place on 4-5 September, 2019 in Brussels and aimed to bring EOSC-hub¹², EOSC-Life¹³, ENVRI-FAIR¹⁴, FAIR4Health¹⁵, FAIR Plus¹⁶, FAIRsFAIR, and GO FAIR¹⁷ together to exchange information on data protection, services to handle sensitive research data, and identify gaps and synergies to develop guidance. The workshop highlighted that, in the first instance, efforts should be undertaken to arrive at a shared working definition of FAIR; that the emphasis on practical activity should be on making data findable and accessible in the short term as without these, data cannot be interoperable or reusable; and that there is a need to develop a shared set of national profiles to make clear the policy and legal framework in each country.

During the workshop, the project representatives worked to identify synergies and avoid overlaps, and, where feasible, to consider closer cooperation to align project deliverables. A key area of overlap that was identified during the workshop is the work areas of FAIRsFAIR and the IMI funded FAIR Plus project. However, while FAIRsFAIR focuses more on generic issues, FAIR Plus focuses on the life sciences and the overlaps represent an opportunity for both projects to cooperate to build a better picture of the landscape from both perspectives. As a result of the workshop, a working group has been formed, consisting of project representatives who will collaborate to develop tangible recommendations for *Turning FAIR into Reality*. The first conference call for the group took place on 18 November 2019.

FAIRsFAIR Focus Group on Research data policies and the FAIR data principles

The 2017-2018 European Universities Association (EUA) Open Access Survey Results¹⁸ revealed that while some universities are well advanced in terms of defining and implementing research data management policies and professionalising their RDM services, the majority of European

¹² <https://www.eosc-hub.eu/>

¹³ <http://www.eosc-life.eu/>

¹⁴ <http://envri.eu/envri-fair/>

¹⁵ <https://www.fair4health.eu/>

¹⁶ <https://fairplus-project.eu/>

¹⁷ <https://www.go-fair.org/>

¹⁸ <https://eua.eu/resources/publications/826:2017-2018-eua-open-access-survey-results.html>



universities are still in the early stages of setting up institutional policies and support services. As part of its role in leading FAIRSF AIR work package 7, EUA organised a focus group in cooperation with UC3M which was held on October 30, 2019 in Madrid, Spain¹⁹. The focus group brought together 25 participants from a number of European universities to consider:

- institutional challenges, drivers and good practices in establishing and implementing research data policies in universities, and
- challenges, drivers and good practices in educating and training students, academics and professional staff in research data management

This focus group was mainly aimed at university and academic leadership responsible for, or interested in, designing and implementing research data policies. Institutions at various stages in this process were encouraged to attend. Other participants, e.g. representing Research Infrastructures and RDM-related projects or initiatives were also invited to attend. The results of the focus group discussions have been reported in relevant sections of this report.

Examples of good practice

Several examples of good practice have been identified during the policy characterisation work and desk research. These have been included as short case studies under the relevant sections of this report. We anticipate that these will be a valuable starting point for developing detailed and practical guidance for enhancing data policies to better support FAIR practice (D3.3).

Updating the policy landscape assessment

FAIRSF AIR will seek to work with a broad range of stakeholders over the course of the project to help them assess their own policies in relation to how well they support FAIR practice and will provide concrete and practical recommendations on how they might enhance their existing data policies or develop new data policies. FAIRSF AIR will undertake a review of the policy landscape towards the end of the project to gain insights into whether and how the policy landscape has changed as a result of implementing our policy enhancement recommendations (this will be released as D3.8 in February 2022).

¹⁹ <https://www.fairsfair.eu/events/focus-group-research-data-policies-and-fair-data-principles-universidad-carlos-iii-de-madrid>



3. National Open Science Policies

To gain a sense of the current landscape with regard to national policies, FsF has drawn upon the findings presented in the series of Open Science Policy reports produced by SPARC Europe and the Digital Curation Centre (DCC)²⁰. These reports are released approximately every six months and aim to identify changes in the national policies of European Member States and other countries in the European Research Area. In particular, we drew from the recently released SPARC European analysis of Open Science Policies in Europe version 4 which was released in August 2019²¹. The majority of the policies reviewed for the report were those introduced by national level funding bodies but other types of policy initiatives were reviewed including national plans, strategy documents and roadmaps, codes of ethics, and in a few cases national laws.

The report states that 14 of the 28 Member States currently have national, research-data related policies in place. This number had increased by 10 since the previous update carried out in January 2018. The report notes that there has been significant activity within several countries to refine their existing policies and strategies to provide a stronger focus on FAIR (Austria, Denmark, Finland, Ireland and Sweden). In line with the European Commission aim of making the research outputs it funds as open as possible but as closed as necessary, the report states that many of the policies also make reference to the fact that in some cases data cannot be shared for ethical, commercial or security issues. A detailed set of tables and in-depth reviews of activity within each country are provided within the SPARC Europe/DCC report (see section 4). A condensed version of these tables is presented in Figure 4 below.

²⁰ <https://sparceurope.org/what-we-do/open-data/sparc-europe-open-data-resources/>

²¹ SPARC Europe, & Digital Curation Centre. (2019, August 28). An Analysis of Open Science Policies in Europe v4. Zenodo. <http://doi.org/10.5281/zenodo.3379705>



Figure 4. Summary of the National Level Open Science Policies in European Member States as presented in An Analysis of Open Science Policies in Europe, v4.

MEMBER STATE / COUNTRY	TYPE OF POLICY	YEAR POLICY CAME INTO EFFECT	SCOPE	SOFT/HARD ²²	COVERAGE OF SKILLS OR TRAINING	MONITORING AND/OR COMPLIANCE
EU						
Austria (AT)	Funder Policy	2019	Only Data	Hard	No	Yes
Belgium (BE)	Code of Ethics	2009	Protocols	Hard	No	No
Cyprus (CY)	Joint policy of Government and Funder	2016	Publications	Soft	No	No
Czech Republic (CZ)	National Strategy	2017	Publications	Soft	Yes	No
Denmark (DK)	National Plan	2015	Software, protocols	Hard	Yes	No
Spain (ES)	State Plan	2018	Covers data alongside many other RDI related issues, including OA	Soft	Yes	No
Finland (FI)	National Plan	2014	Publications, tools, methodologies	Hard	Yes	Yes

²² SPARC Europe/DCC define a ‘hard’ policy as one that employs language such as “must” or “should”, as opposed to soft policies which more gently advise or encourage.

France (FR)	Law/National Plan	2016/2018	Covers data alongside many other ICT related issues, including OA	Hard	No/Yes	No
Germany (DE)	Funder Policy	2010	Software, methods	Hard	No	No
Ireland (IE)	National Framework	2019	Publications, Infrastructure	Hard	Yes	Yes
Lithuania (LT)	Law / Funder Policy	2016	Publications	Hard	No	Yes
Netherlands (NL)	National Plan / Concordat	2017	Publications	Soft	Yes	Yes
Portugal (PT)	Funder Policy	2014	Samples, software, models	Soft	No	No
Slovenia (SI)	National Policy	2015	Publications	Hard	Yes	Yes
United Kingdom (UK)	Funder Policy / Concordat	2015/2016	Software (in the FAQs and Concordat)	Hard	Yes	No
<i>NON-EU</i>						
Norway (NO)	National Strategy	2017	Only data	Hard	No	No
Switzerland (CH)	White Paper	2014	Covers data alongside many other ICT related issues, including OA	Hard	Yes	Yes
Serbia (RS)	National policy	2018	Open Science	Soft	Yes	Yes



The SPARC Europe/DCC report states that key drivers for many of the countries developing national approaches for Open Data were the European Commission’s Open Research Data Pilot²³ which was carried out as part of Horizon 2020 and the more recent EU Directive on Open Data and the Re-use of Public Sector Information (PSI Directive)²⁴. Specifically, Article 10 which states that,

‘Member States shall support the availability of research data by adopting national policies and relevant actions aiming at making publicly funded research data openly available (‘open access policies’), following the principle of ‘open by default’ and compatible with the FAIR principles. In that context, concerns relating to intellectual property rights, personal data protection and confidentiality, security and legitimate commercial interests, shall be taken into account in accordance with the principle of ‘as open as possible, as closed as necessary’. Those open access policies shall be addressed to research performing organisations and research funding organisations’²⁵.

During the FAIR Enough? Workshop held in Brussels on 4-5 September, participants agreed that there is a need to work together to develop a set of shared national profiles that make clear the legal and policy frameworks of each European country for governing access to sensitive data to ensure that data can be as open as possible but as closed as necessary. While this was suggested to better support FAIR data production in the life sciences community, the profiles would no doubt be valuable for researchers and data stewards all disciplines.

The SPARC Europe/DCC report makes clear that Open Science remains a key priority in Horizon Europe - the European Commissions’ next Framework Programme for research and innovation²⁶. The report states that one of the operational objectives for Horizon Europe is,

‘Accelerating the transition towards open science, by monitoring, analysing and supporting the development and uptake of open science policies and practices, including the FAIR principles, at the level of Member States, regions, institutions and researchers, in a way that maximises synergies and coherence at EU level’.

²³https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

²⁴ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024>


²⁵ Article 10. Research data. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024>

²⁶ https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-research-innovation_en.pdf



Turning FAIR into Reality (TFiR) Action 16.1 states that policies must assert that the FAIR principles should be applied to research data, to metadata, to code, to DMPs and to other relevant digital objects, as well as to policies themselves. With regard to the scope of the national policies, the SPARC Europe/DCC report found that while all explicitly refer to research data, only 8 of the 18 national policies include statements indicating that this includes software, code, tools and/or models. Five of the national policies also made reference to methods, workflows or protocols, and one policy addresses physical (non-digital) samples. This implies that there is a need to work cooperatively to harmonise national policies to better reflect the broad range of research outputs that should be made FAIR and should be covered in the policies.

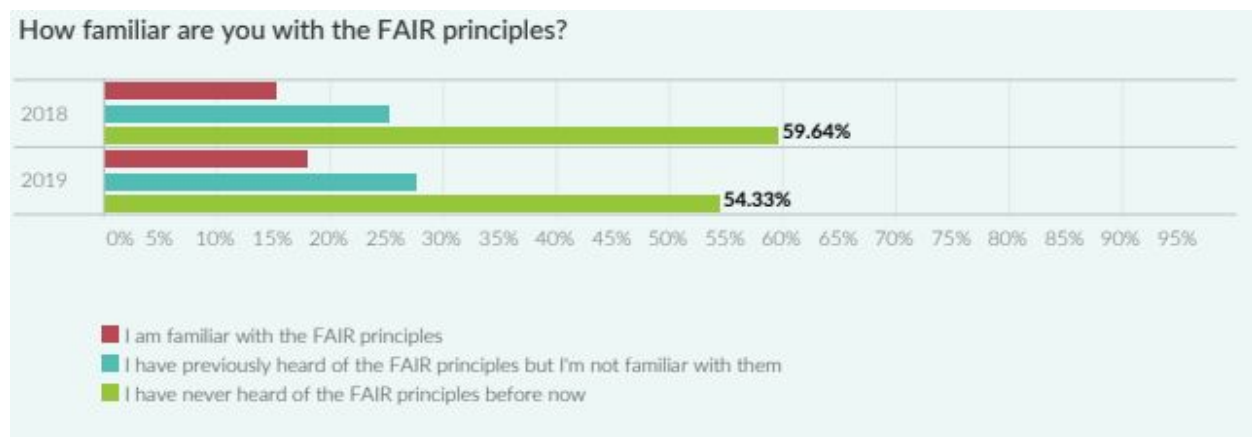
TFiR Action 26.4 calls for a broader range of metrics to be developed to recognise contributions beyond publications and citations and that these should recognise and reward Open and FAIR data practices. When it comes to monitoring compliance with data policies, only seven of the eighteen countries with national policies refer to compliance checking or mention possible sanctions for non-compliance. These include Austria, Finland, Ireland, Latvia, the Netherlands, Slovenia and Switzerland. The report found that there are a small number of countries that are beginning to include references to reward and recognition for those demonstrating good practice (Germany, France and the Netherlands).

<p>Example of good practice at the national level: Ireland</p>	
<p>In July 2019 a new National Framework on the Transition to an Open Research Environment was introduced at the ministerial level in Ireland. The framework is aligned with European Commission policy recommendations and aims to enable FAIR data production and use through the implementation of clear policies which cover the management, preservation and reuse of research data. The framework also includes details on the need for infrastructure to support open research such as improving skills and competencies and providing incentives and rewards. The framework also makes clear that research data should be as open as possible but as closed as necessary and states that funders and institutions should support requirements for data management and sharing and monitor compliance.</p> <p style="text-align: right;">An Analysis of Open Science Policies in Europe v4, 2019</p>	

The SPARC Europe/DCC report shows that there has been a significant increase in the number of national level policies emerging and that an increasing number specifically refer to the FAIR principles. National policy developments seem in line with the views of the research community which, as reported in the recently published State of Open Data 2019 report²⁷, are largely in favour of national mandates for making primary research openly available (79% of respondents to the State of Open Data 2019 survey were in favour of such mandates). The annual State of Open Data survey carried out by Figshare and Digital Science had a record number of responses from the research community this year with more than 8,400 researchers taking part. Responses were fairly equally split among different career stages (professor, tenured/tenure track, and early career).

It is worth noting that the State of Open Data 2019 report found that, of those researchers who frequently share data, 52% had never heard of the FAIR principles. The number of 2019 respondents who have never heard of the FAIR principles is only slightly lower than it was for the 2018 State of Open Data report (see Figure 5) which highlights that despite the emergence of a number of FAIR-related support initiatives in the last couple of years, there is still much work to be done in raising general awareness about the FAIR principles and the potential benefits of implementing them.

Figure 5. Familiarity with the FAIR principles 2018 versus 2019. State of Open Data 2019



With regards to the FAIR-related support initiatives themselves, the State of Open Data 2019 report found that there is a relatively low level of awareness about these and what they offer to support FAIR practice (Figure 6). The report findings suggest that there is a need to improve the visibility of these initiatives and for better cooperation between them to amplify key messages

²⁷ Science, Digital; Fane, Briony; Ayris, Paul; Hahnel, Mark; Hrynaszkiewicz, Iain; Baynes, Grace; et al. (2019): The State of Open Data Report 2019. figshare. Report. <https://doi.org/10.6084/m9.figshare.9980783.v2>



about what FAIR means in a practical sense, the potential benefits of producing and using FAIR data, and how to go about putting the FAIR principles into practice using the support provided.

Figure 6. Levels of awareness of FAIR initiatives. State of Open Data 2019²⁸.

2019	I am familiar with...	I have heard of, but am not familiar with...	I have never heard of...
GO FAIR	5%	18%	77%
FAIRdat	5%	17%	78%
MakeDataCount	11%	24%	65%
DataCite	4%	14%	82%
FORCE	11%	20%	69%

While the existence of national research data related policies is a good start, they aren't in themselves sufficient for supporting change. The SPARC Europe/DCC report states many of the national research data related policies are still in their infancy with many being introduced within the last three to five years. As the national policies become better embedded into related research administration workflows and management systems, we can expect to see a shift from simply encouraging good practice to mandating and monitoring good practice.

<p>Example of good practice: strengthening expectations for research integrity</p>	
<p>Following an inquiry led by the Science and Technology Select Committee which identified challenges with the earlier version of Concordat to Support Research Integrity (released in 2012) the concordat was revised. According to the Universities UK (UUK) website, 'significant changes have been introduced, and these include:</p>	
<ul style="list-style-type: none"> ● individual research funders have now put sanctions in place which will be enforced if the conditions of the concordat are not met ● the concordat's language has been tightened, with expectations being replaced by commitments or requirements. This will make it easier for research organisations to 	

²⁸ Research, Nature; Penny, Dan; Fane, Briony; Goodey, Greg; Baynes, Grace (2019): State of Open Data 2019. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.10011788.v2>. Question 4.9



understand what is required of them.

- in addition to publishing their annual statement, research organisations are now additionally required to send a link to the secretariat. This will mean that the sector is better able to demonstrate its commitment to the concordat.

An annual forum will convene to assess progress and work on good practice for the sector, from which a statement will be made publicly available. In addition, the signatories have committed to regularly reviewing the concordat's principles every five years to ensure its commitments are met'.

This example demonstrates a stronger commitment to policy enforcement for research integrity that could be replicated for policies more specific to data sharing. By removing ambiguity, it is anticipated that the policy will be easier to understand and, as a result, compliance rates should increase. The emphasis on transparency around demonstrating compliance is also welcome.

Universities UK statement²⁹

Indeed, in relation to the findings of State of Open Data 2019 survey, Mark Hahnel remarked that he was encouraged by the adoption of open data practices moving in the right direction but that the onus now appeared to lie with funders and government to enforce the data sharing mandates they have implemented in recent years³⁰.

²⁹ <https://www.universitiesuk.ac.uk/news/Pages/Updated-concordat-must-be-followed-if-research-is-to-be-trusted.aspx>

³⁰ <https://www.digital-science.com/blog/news/the-state-of-open-data-2019/>



4. Policy Characterisation mapped to *Turning FAIR into Reality* recommendations

The *Turning FAIR into Reality* (TFiR) report defines 27 priority and supporting actions that must be addressed to foster a FAIR data culture. The 27 actions are spread over three distinct phases that include:

Step 1: Define - concepts for FAIR Digital Objects and the ecosystem

Step 2: Implement - culture, technology and skills for FAIR practice

Step 3: Embed and sustain - incentives, metrics and investment

Under each of the three steps, there are specific actions related to the policy environment that are targeted to specific stakeholders. In this section of the report, we share the findings of our policy characterisation work in relation to a subset of the TFiR actions targeted towards policymakers, funders, publishers and institutions. It is important to note here that the *Turning FAIR into Reality* report states that the policymaker stakeholder group may also be comprised of funders, institutions and publishers.

About the policies

We aimed to examine between 10-15 policies from each of the three stakeholder groups (funders, publishers/journals, and HEIs) and selected those which we feel currently demonstrate a relatively high level of engagement in promoting good research data management and adoption of the FAIR principles. We reviewed a total of 42 policies comprised of 17 funding body policies, 14 publisher/journal policies³¹, and 11 institutional policies. These should not be considered an exhaustive inventory of stakeholders currently engaged with Open Science and the FAIR principles. For a list of the policies consulted, please see Annex 3.

In addition to analysing the content of these policies, we also carried out some characterisation of the policies themselves primarily in light of TFiR action 3.2 which states that policies should be machine-readable and actionable and action 17.3 which states that policies should be versioned, indexed and semantically annotated in a policy registry. Findings are presented below.

³¹ To assess the publishing landscape, a mix of eight publisher policies, five journal policies and one community initiative (Author Guidelines of the COPDESS Enabling FAIR Data Project) were analysed.



- Action 3.2: By default, the FAIR ecosystem as a whole and each of its individual components should work for humans and for machines. Policies and DMPs should be machine-readable and actionable.
- Action 17.3: Policies should be versioned, indexed and semantically annotated in a policy registry to enable broad reuse within the FAIR data ecosystem. Resources mandated by policies (e.g. consent forms) should be treated the same way.

Date of policy introduction


To support both human interpretation and machine actionability as outlined in TFIR action 3.2, it is crucial that policies make clear the period of time to which they apply. Accordingly, policies should state when the policy was introduced and make clear if, and when, any reviews/updates are planned.

Of the 17 funding body policies reviewed, 10 include a date of introduction. Among those funding body policies that do include a date of introduction, the majority were introduced over the last three years. Three of the funders' policies include information about planned updates along with indicative timeframes. Only 3 publisher/journal policies explicitly state the date they were introduced or last updated. For another 4, the date of introduction was obtained through other sources (interviews or associated information). None of the publisher/journal policies reviewed include information on versions or references to scheduled reviews. In contrast, all of the 11 institutional policies we examined include a date stating when the policy was introduced or updated. Three of the institutional policies also make clear that the policy will be updated although only one institution provides a timeframe for this.

Policy is machine readable

While all of the policies we reviewed are machine readable in as much as they are available in HTML and/or PDF formats from the web, we did not find any evidence that any of the policies were described using a structured data markup schema. To enable machine-actionable workflows to be carried out in a FAIR ecosystem, it will be crucial that the status of policies governing data sharing can be interpreted unambiguously by both humans and machines. The FAIRsharing registry asks contributors to make clear the status of the resource being described. Options include *'Ready'*, *'In Development'*, *'Uncertain'*, and *'Deprecated'*. FAIRSF AIR's policy assessment also included these options as we feel that a clearly defined status is crucial for supporting effective machine actionability. The majority of policies we reviewed were deemed to be *'Ready'* although it was not always easy to make a clear decision. Indeed, the FAIRsharing options seem more appropriate for describing the status of standards and databases than policies.



<p>Example of good practice: exposing digital content</p>	
<p>‘Schema.org is a joint effort, in the spirit of sitemaps.org, to improve the web by creating a structured data markup schema supported by major search engines. On-page markup helps search engines understand the information on web pages and provide richer search results. A shared markup vocabulary makes easier for webmasters to decide on a markup schema and get the maximum benefit for their efforts. Search engines want to make it easier for people to find relevant information on the web. Markup can also enable new tools and applications that make use of the structure’³².</p> <p>The approach has been adopted in the life sciences community who have come together since 2015 to develop BioSchemas which aims to ‘improve the findability of data in the life sciences. It does this by encouraging people in the life sciences to use Schema.org markup in their websites so that they are indexable by search engines and other services. BioSchemas encourages the consistent use of markup and the structured information then makes it easier to discover, collate, and analyse distributed data’³³.</p> <p>Schemas.org is community driven and communities of practice can work together to define a common structure for describing digital content as shown in the BioSchemas example. This approach could be extended not just to research data policies, but to other components in the FAIR ecosystem as well including DMPs, repositories, support and training (as illustrated in Figure 1).</p>	

The use of a structured markup schema offers great potential to make machine readable policies more visible. By adopting this approach, the policies would not need to be manually entered as records into registries - an approach which depends upon the record owner and/or content curator keeping the information up to date and accurate. Instead, the policies could be maintained locally, updated as necessary and simply harvested by registries and any API that requires the information. A small-scale pilot carried out by the DCC in 2015 tested the viability of extending the Organisational Profile Document, which was introduced by Equipment.data to expose institutional information about research equipment bought with public funds in a machine readable way, to consistently describe hard and soft RDM infrastructure components within HEIs to make them more visible, machine readable, and discoverable³⁴. The pilot produced a Research Data Management (RDM) profile³⁵ which could be extended to produce RDM and FAIR data profiles.

³² <https://schema.org/docs/faq.html>

³³ <https://bioschemas.org/>

³⁴ <http://www.dcc.ac.uk/projects/opd-for-rdm>

³⁵ http://www.dcc.ac.uk/webfm_send/2077

Policy has a persistent identifier (PID)

As automated processing and machine actionable workflows are defined and developed to support the FAIR ecosystem, it will become increasingly important that the right version of any given stakeholders' policy can be found and fed into these pipelines. To this end, we examined whether various stakeholders currently assign a PID such as a DOI to their policies. In all cases, we were unable to find a PID for the policies reviewed.

With regard to TFiR action 17.3 (i.e., being indexed and semantically described in a policy registry), we checked to see whether the policies we reviewed had records in FAIRsharing³⁶. Several of the publishers and a few of the funding bodies we looked at did have records registered with FAIRsharing. However, in many cases, the records in FAIRsharing referred to different policies than the ones we were reviewing. For example, while FAIRsharing did include several records for the publisher Elsevier, these were records for the information for authors for their specific journals rather than the Elsevier Research Data Guidelines that was reviewed for this study. In the few cases where publisher/journal policies we reviewed were also registered with FAIRsharing and had been allocated a DOI for their record, the DOI was not made visible from the publishers' own web pages. For example, F1000 Research Data Guidelines does not have an associated PID for its policy as presented via the F1000 website but has been allocated a DOI for its FAIRsharing record.³⁷

An aspect of good practice that we noticed among several of the publishers that we looked at was the introduction of tiered policies. Springer Nature and Elsevier for instance have multiple variations of their data policies to reflect the fact that data sharing requirements vary across the journals they publish which reflects the recommended approach suggested by the RDA Data Policy Standardisation and Implementation Interest Group. In these cases, each of the policies is clearly titled to reduce ambiguity.

Content of the policies

Define - concepts for FAIR Digital objects and the ecosystem

Defining what FAIR means in a practical sense is a work in progress with several European Open Science Cloud (EOSC) related projects currently trying to determine what FAIR means within specific disciplines³⁸. Policymakers have a key role to play in working closely with these

³⁶ <https://fairsharing.org/policies/>

³⁷ F1000Research Data Policy <https://doi.org/10.25504/FAIRsharing.nb9zzm>

³⁸ <https://www.eosc-portal.eu/about/eosc-projects>



initiatives to better define what FAIR means in a practical sense and to ensure that this is reflected back in their policies and related guidance. There is also a need for policymakers to work together to ensure that there is alignment across these definitions and that their policies cover the full range of research outputs including data, code, workflows, models, and other digital research objects as well as their curation and maintenance. In this section, we explore the current policy landscape in relation to relevant TFIR actions under the broad heading of *Define*.

This broad action of defining FAIR for implementation includes the largest number of actions that directly relate to the policies of funding bodies, publishers and HEIs. These actions relate to the scope of the policy, expectations in relation to sharing, and the need to justify not sharing data. The specific actions considered in this assessment include:

- Action 1.3: The relationship between FAIR and Open should be clarified and well-articulated as the concepts are often wrongly conflated. FAIR does not mean Open. However, in the context of the EOSC and global drive towards Open Science, making FAIR data a reality should be supported by policies requiring appropriate Openness and protection, which can be expressed as ‘as Open as possible, as closed as necessary’.
- Action 16.1: Policies must assert that the FAIR principles should be applied to research data, to metadata, to code, to DMPs and to other relevant digital objects, as well as to policies themselves.
- Action 17.1: The greatest potential reuse comes when data are both FAIR and Open. Steps should be taken to ensure coherence across data policy, emphasising both concepts and issuing collective statements of intent wherever possible.
- Action 17.4: Data and other FAIR Digital Objects (e.g. code, models) that directly underpin, and provide evidence for, the findings articulated in published research must also be published unless there are legitimate reasons for protecting and restricting access
- Action 17.6: Policies should require an explicit and justified statement when (publicly funded) data cannot be Open and a proportionate and discriminating course of action should be followed to ensure maximum appropriate data accessibility, rather than allowing a wholesale opt-out from the mandate for Open data.

To assess how well funding bodies, publishers/journals and HEI's data policies currently reflect the TFIR actions listed above, the following policy characterisation elements were assessed:

- Definition for *data* is provided



- Policy specifically references FAIR
- Policy requires data sharing
- Exceptions to data sharing are allowed
- If exceptions are allowed, justifications are required
- Policy requires metadata sharing

Below, we summarise the results of the policy characterisation in relation to these elements. Please note that in cases where the policy itself did not contain the specified element clearly, we have labelled these as ‘not specified’. Figures have been provided for some of the elements listed above while others are addressed only in the accompanying text.

Figure 7. Policy defines data

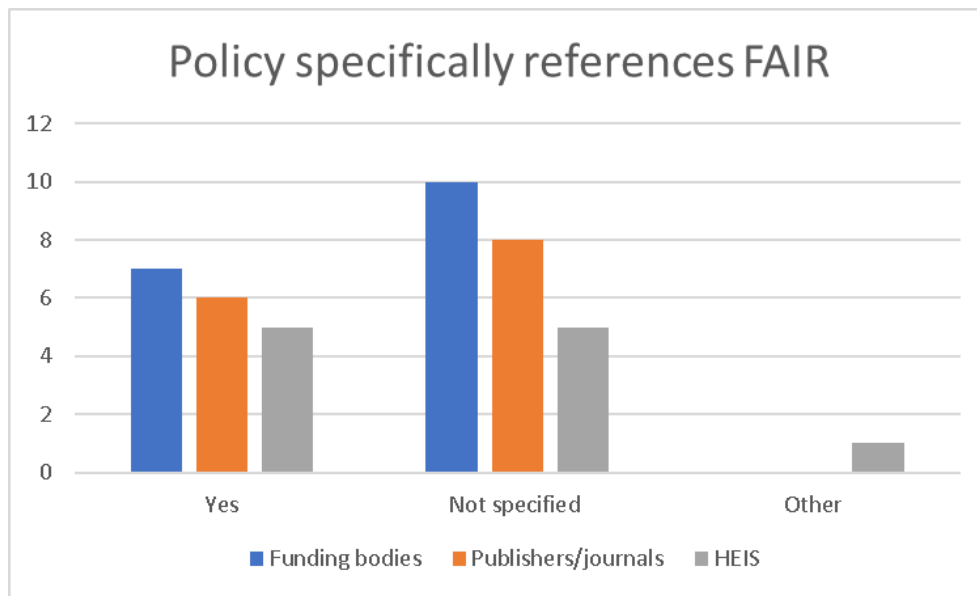


To avoid confusion, it is essential that policymakers make clear which research outputs are covered by their data policies. In this respect, providing a clear definition of what is meant by the term *data* is good practice. This is particularly true for policies governing research in the Arts and Humanities where the term *data* is not commonly used to refer to the digital outputs produced.

The majority of funder policies we reviewed do include a definition of data as do the HEI policies. However, while less than half of the policies from the publishing field specify what they mean by the term *data* - either in the policy itself or in accompanying guidance material - the majority (11 of 14) do specifically reference other types of research outputs covered by the

policy in addition to research data. In most cases, references include software, algorithms, and code. The majority of policies across the different stakeholders make clear that they apply to those data that underpin published findings or support validation. This limits the scope of research data to be made FAIR and is considered good practice as it helps to dispel the misconception of some researchers that data sharing requirements apply to all data generated within a research activity.


Figure 8. Policy specifically references FAIR



While many of the funding body and publisher/journal data policies reviewed contain FAIR-related elements, less than half (7 out of 17, and 6 out of 14 respectively) specifically refer to the FAIR principles. Concerning the relationship between FAIR and open, more publisher/journal policies reference the concept of open access to research data than FAIRness of data. In interviews conducted with a few publishers, it was mentioned that some had implemented their data policies before the FAIR principles were developed which explains why there is no specific reference. While a couple of the interviewees stated that they may align their policy to FAIR in future updates, one stated that they may not specifically reference the FAIR principles - even if their policy is consistent with them - because the term is confusing for researchers. Indeed, the recent State of Open Data 2019 report states that the concept of FAIR is still largely unknown among researchers. As new legal instruments such as the EU Directive on Open Data and the Re-use of Public Sector Information (PSI Directive) make reference to FAIR, it is advisable that even where the policies of stakeholders do not specifically reference FAIR, that their associated guidance does make clear how their policy aligns with these

principles to avoid even greater confusion among researchers and to show where there is commonality across various stakeholder requirements despite the variations in terminologies used.

Among the HEI policies reviewed, 10 out of 11 policies reference ‘open access’ to research data whilst only 5 reference the FAIR principles specifically. There are several European HEIs that have been keen advocates of the Open Science movement for many years and are demonstrating leadership in the area (for example, the University of Leiden, University College London). In these cases, it is sensible for their data policies to specifically reference FAIR. For other HEIs though, it may not be necessary for their data policies to specifically reference FAIR but rather to state that they expect researchers to meet funding body and publishers’ expectations. Indeed, our open consultation findings indicate that the policies of funding bodies and publishers/journals are considered to be most influential on researchers’ behaviour while those of HEIs are seen as significantly less so. Several participants of the FAIRsFAIR focus group on Policies and FAIR Data also felt that institutional data policies may be viewed more positively by researchers if the HEI was not the one making demands but instead focusing on supporting researchers in adhering to external mandates.

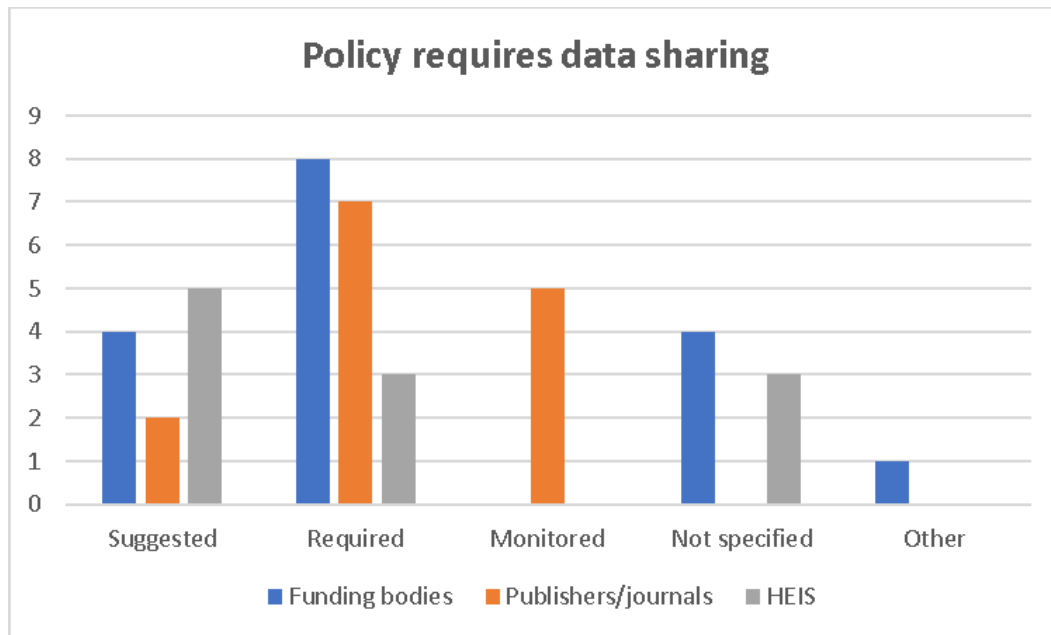
Example of good practice: explaining the FAIR principles in practical terms	
<p>In its DMP Guidelines for researchers³⁹, The Swiss National Science Foundation (SNSF) starts by stating “Managing and sharing research data as openly as possible is one of the principles of good scientific practice.” SNSF also provides helpful advice on how to interpret the FAIR principles⁴⁰ in its guidance to support the implementation of the policy and makes clear its expectations on the part of researchers in receipt of funding as well as the data repository they select to deposit their data.</p> <p style="text-align: right;">Swiss National Science Foundation Open Research Data</p>	

³⁹http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/data-management-plan-dmp-guidelines-for-researchers.aspx

⁴⁰http://www.snf.ch/SiteCollectionDocuments/FAIR_principles_translation_SNSF_logo.pdf



Figure 9. Policy requires data sharing



Eight out of the 17 funder policies we reviewed currently require data sharing and 4 only suggest it. Four of the funder policies reviewed were not clear on their stance in relation to data sharing. As our open consultation showed that funders’ policies are the most influential on researchers’ practices, it seems that there could be stronger and clearer requirements for data sharing coming from funding bodies.

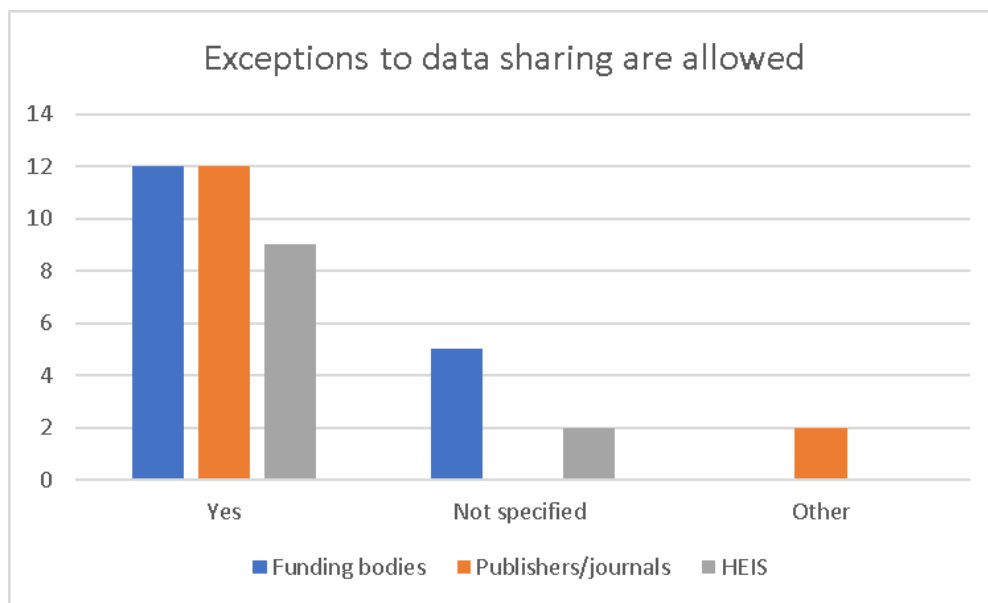
Of the policies of publishers/journals that were reviewed, data sharing is suggested in 2 policies, required in 7 and monitoring is referenced in 5 of the policies. When assessing these findings, it should be kept in mind that the policies we chose for analysis are from those relatively advanced in terms of engaging with data sharing and the FAIR principles. As such, these findings do not necessarily represent the current state of the publishing landscape in general. However, recent findings indicating that publishers who only encourage data sharing see a far lower rate of compliance⁴¹ emphasise the need for publishers/journals who have not already done so, to strengthen their position. In two of the four interviews conducted with publishers as part of the desk research for this analysis, interviewees stated that they would like to increase the rigor of their policies and that this was something that would be implemented in the future. However, these interviewees also stated that they are reluctant to progress too

⁴¹ Hrynaszkiewicz, I. Building Trust to Break Down Barriers in State of Open Data 2019. <https://doi.org/10.6084/m9.figshare.8223365.v1>

quickly with requiring data sharing and monitoring compliance as there are still large variations in the data sharing cultures across authors’ domains.

The data policies of the HEIs assessed mostly suggest data sharing (5) rather than require it (3). Again, as the findings of the open consultation indicate that the policies of institutions tend to have less influence over researchers’ behaviour, it may be more effective for HEIs to focus on helping researchers to meet external funding body and publisher/journal expectations rather than to try to enforce data sharing from an institutional perspective.


Figure 10. Exceptions to data sharing are allowed



A key misconception is the notion that FAIR data must also be open data. This is not the case and there can be many valid reasons for not sharing research data such as ensuring data protection of research participants or respecting the intellectual property rights of commercial partners. Most of the funder policies reviewed do allow for exceptions to data sharing (12 of 17) whereas 5 do not currently make their stance on exceptions to data sharing clear. Where exceptions are permitted, many of the funders do require justifications for not sharing data (9 of 17 policies). The majority of these expect that the reasons for not sharing will be made clear in the associated data management plan (DMP). While this is a sensible approach as constraints for data sharing should be considered at the outset of any new project, funders should also make clear that these reasons are included in the metadata for the eventual dataset. Whilst the majority of funders require or suggest sharing data (12), only 8 require or suggest metadata sharing.

The vast majority of publisher/journal policies reviewed (12 out of 14) allow exceptions to data sharing. In one case, the policy made clear that while data does not necessarily have to be shared openly (e.g. because of privacy issues), they must be made available to peer reviewers and a data availability statement has to provide information about how an interested researcher can request access. Ten of the publisher/journal policies reviewed require a justification for not sharing the data that underpins the article. The majority of publisher/journal policies either monitor (5) or at least require (7) metadata sharing, with only one suggesting it.

The large majority of HEIs policies (9) allow exceptions to data sharing. Justifications for not sharing data are required by 3 institutions with the remainder not specifying it further. Encouragingly, a high number of the HEI policies reviewed (7 out of 11) require the sharing of metadata.

Example of good practice: providing clear guidance on legitimate exceptions to data sharing	
<p>Funders such as the European Commission do reference FAIR specifically and provide detailed guidance to help with implementation. The guidance for the European Commission’s H2020 Programme states that "If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions." The guidance makes clear that researchers can have valid reasons for not sharing the resulting project data but that these should be clearly stated. The guidance helpfully separates out legal or contractual barriers to data sharing, as opposed to other reasons, which helps avoid easy opt-outs and unfounded data closure.</p> <p style="text-align: right;">Guidelines for FAIR Data Management in Horizon 2020⁴²</p>	

Summary of landscape in relation to Define - concepts for FAIR Digital objects and the ecosystem

Overall, the policies of the stakeholders do reflect the priority and supporting actions outlined under TFiR’s broad heading of *Define - concepts for FAIR Digital objects and the ecosystem* to varying degrees. However, there is still a need to provide clearer definitions of what data means and to harmonise these definitions across stakeholders. In addition, there is a need for

⁴² https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf



funding bodies and publishers/journal to strengthen their expectations around the sharing of metadata. As the policies of funding bodies and publishers tend to have the greatest influence over researcher's practices, these stakeholders should take a lead in refining their policies to provide stronger expectations for data and metadata sharing and provide clearer advice about when and how to justify not sharing data.

Implement - culture, technology and skills for FAIR practice

In this section, we explore the current policy landscape in relation to relevant *Turning FAIR into Reality* (TFiR) actions under the broad heading of *Implement*. TFiR states that data management plans (DMPs) are an essential mechanism for ensuring that research data management becomes a core part of research culture and can lead to the increased production of FAIR data. Ensuring that researchers understand how and when to include eligible costs related to RDM in their grant applications is also considered crucial for fostering a FAIR data culture. To ensure that FAIR data emerging through research is findable and accessible, researchers should have access to trusted digital repositories where they can deposit their data. The provision of generic and, where needed, discipline specific guidance at the right time in the research lifecycle is also essential for realising culture change.

The specific TFiR actions that are addressed in this section include:

- Action 5.1: Research communities must be required, supported and incentivised to consider data management and appropriate data sharing as a core part of all research activities. They should establish a Data Management Plan at project outset to consider the approach for creating, managing and sharing all research outputs (data, code, models, samples etc.)
- Action 5.2: Data Management Plans should be living documents that are implemented throughout the project. A lightweight data management and curation statement should be assessed at project proposal stage, including information on costs and the track record in FAIR. A sufficiently detailed DMP should be developed at project inception. Project end reports should include reporting against the DMP.
- Action 18.1: Questions about the costs of data management, curation and publication should be included in all DMP templates. Information from existing and completed projects should be used to retrospectively identify costs and develop examples and guidelines based on these. Funders, institutions and data services should collaborate on retrospective analysis, including the cost of long-term curation.
- Action 20.1: Policy should require data deposit in certified repositories and specify support mechanisms (e.g. incentives, structural funding and/or funding for deposit fees, and training) to enable compliance.

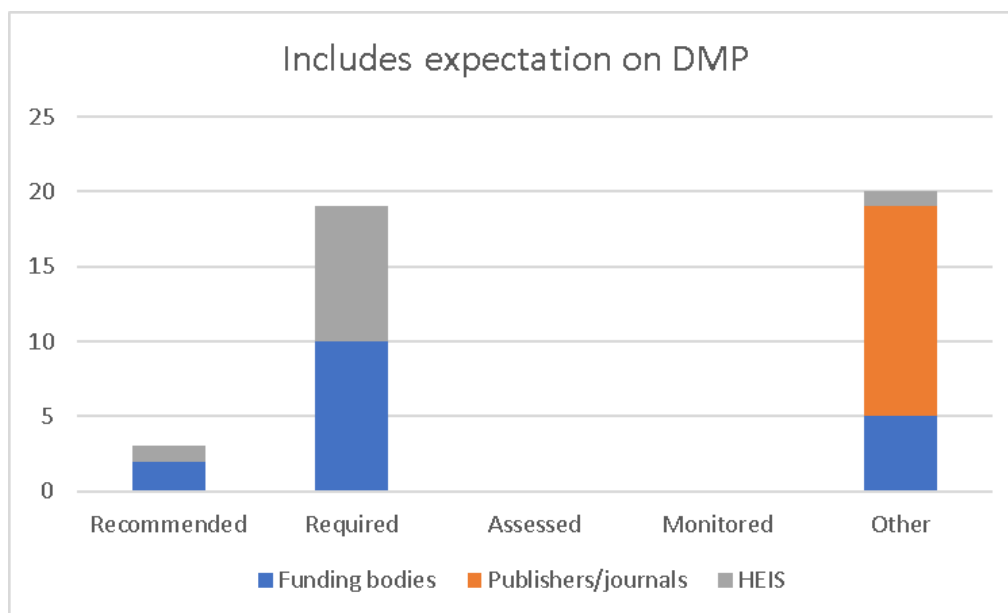


To assess how well funding bodies, publishers/journals and HEI's data policies currently reflect the TFiR actions listed above, the following policy characterisation elements were assessed:

- Policy includes an expectation on data management plan (DMP)
- Policy states at which stage the DMP should be produced
- Policy states that justified costs relating to RDM and making data FAIR are supported
- Policy references specific data repositories or scientific databases for deposit

Below, we summarise the results of the policy characterisation in relation to these elements.

Figure 11. Policy includes an expectation on DMP




Over half of the funder policies reviewed require the development of a data management plan (DMP). Most funders include guidance about what type of information needs to be included in such plans and many provide DMP templates or recommend the use of tools such as DMPonline⁴³ and DMPTuuli⁴⁴. Of those funders who require or recommend a DMP, about one third ask for this at the pre-award stage (5) and about one third at the post award stage (6). Some funders assess DMPs submitted with grant applications whereas others are not clear about whether or how the DMP will be assessed. The Wellcome Trust, for example, reviews its output management plan as part of the funding decision. Other funders state that they will

⁴³ <https://dmponline.dcc.ac.uk/>

⁴⁴ <https://www.dmptuuli.fi/>

evaluate DMP compliance but not the scientific merit of the plan itself which could result in demotivating researchers to put effort into submitting high quality plans. Some funding bodies have DMP requirements that span the project lifecycle and expect an initial DMP to be prepared during the pre-award phase and a full plan submitted post award (e.g., the Natural Environment Research Council in the UK). Others only require the plan to be submitted once the project is underway (e.g., European Commission). Encouragingly, some funding bodies are beginning to require end stage DMPs be prepared (such as the Swiss National Science Foundation). This is good practice as the end stage DMP will provide information on what actually happened with the data rather than what was planned. The more accurate information contained in end stage DMPs will make these valuable outputs for providing provenance and context to support data reuse and will provide useful data that can be mined to inform future policy development.

Example of good practice: providing clarity on DMP assessment	
<p>The Wellcome Trust recently introduced outputs management plans as part of its application process. The outputs management plan and associated guidance⁴⁵ makes clear that in addition to research data, the effective management of other digital and non-digital outputs should be considered. Their guidance makes clear what is included, that the plan should be reviewed and updated over the life of the project, and that the plan will be reviewed as part of the proposal assessment.</p> <p>‘Anyone applying for Wellcome funding must consider their approach to managing and sharing anticipated outputs at the research proposal stage. In cases where these outputs are significant – generating data, software or materials that will hold clear value as a resource for others in academia or industry – applicants will need to include an outputs management plan explaining their planned approach. We will review this plan when making our funding decision. We will fund any justified costs for delivering the plan as part of funding the research... Researchers’ approach to outputs management should be dynamic. Plans should reflect established best practice in the respective research field.’</p> <p style="text-align: right;">How to complete an outputs management plan, Wellcome</p>	

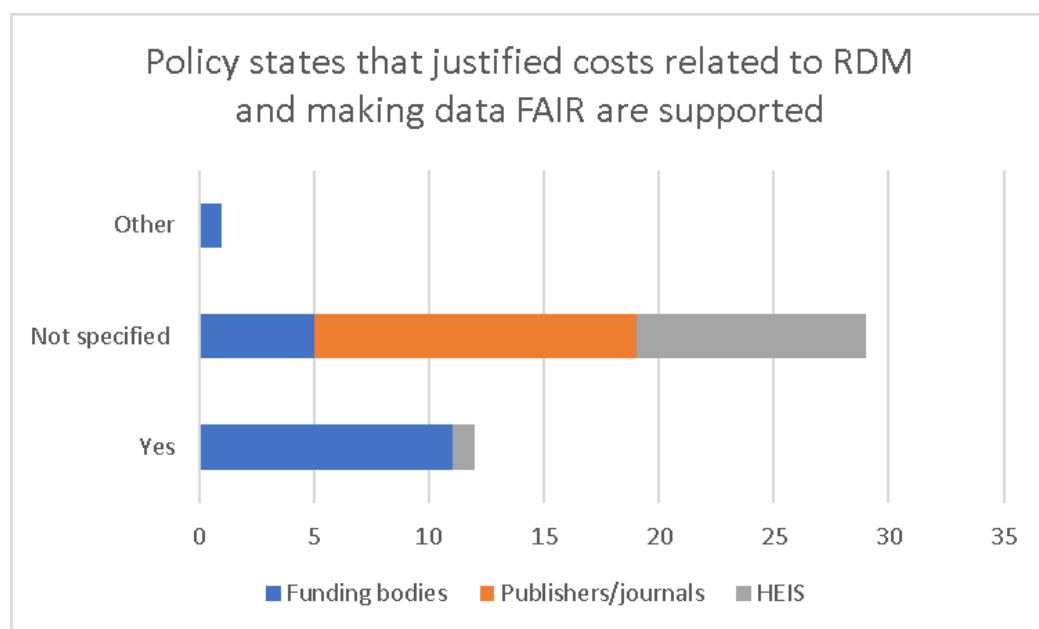
⁴⁵ <https://wellcome.ac.uk/funding/guidance/how-complete-outputs-management-plan>



None of the journal or publisher policies in the cohort reviewed mention DMPs. However, the Enabling FAIR Data Author Guidelines⁴⁶ for researchers in the earth, space and environmental sciences identifies detailed roles and responsibilities for all actors across the research lifecycle including researchers, institutions, funding bodies, repositories and publishers, as a basis for joining up services and workflows. This type of information would be relevant for inclusion in DMPs.

Nine of the 11 HEI policies do require a DMP. Of these, 3 expect the DMP to be prepared during the pre-award phase and 1 during the post-award phase. Some institutions have developed their own DMP templates with local guidance (e.g., University of Bath postgraduate template in DMPonline) and others point to funders’ templates for developing DMPs that will help researchers comply with the institutional policy.


Figure 12. Policy makes clear that justified costs related to RDM and making data FAIR are supported



Eleven of the funder policies reviewed indicate that the funding body is prepared to meet costs associated with making research data available whereas 5 do not make their stance clear. Some funders provide extensive guidelines detailing exactly what kinds of costs can be requested. It is worth noting that some of the funding body policies whose policies were reviewed will meet costs incurred in relation to RDM and making data FAIR although this is not clearly stated in the policy or associated guidance.

⁴⁶ <https://copdess.org/enabling-fair-data-project/author-guidelines/>



Example of good practice: providing clarity on eligible costs	 UK Research and Innovation
<p>UK Research and Innovation (UKRI) provides extensive information about which RDM and data sharing costs can be recovered in grant applications. In its guidance it states that:</p> <p>‘Research Organisations are encouraged to seek to recover costs that will be incurred in relation to research data arising from Research Council grants. Such costs could be associated with the production, curation and sharing of the research data according to the guidelines (which may be data-specific) published by the Research Councils, and may for example include (but is not limited to) staff time, software, hardware and third-party storage services.’</p> <p style="text-align: right;">UKRI Guidance on best practice in the management of research data ⁴⁷</p>	

None of the 14 publisher/journal policies that were assessed support any costs related to RDM and making data FAIR. As most publishers are for-profit organisations, this is not surprising. However, as the reputation of journals - and ultimately their financial viability - depends upon their perceived quality, publishers have a vested interest in supporting FAIR data for peer review and validation purposes and should bear some of the associated costs.

Only one out of the 11 HEI policies clearly states that costs associated with RDM and FAIR data are supported by the institution when justified. Given that establishing RDM processes, infrastructure and capacity within an HEI comes at an additional cost, the role of funders in supporting costs for RDM within project budgets is crucial to enable HEI-based researchers to follow RDM and FAIR requirements.

However, many researchers are reluctant to allocate budget to support RDM and data sharing in new grant applications - partly because it is seen as detracting from the funding available to do the research but also in many cases because researchers fear that their bid may be viewed as less competitive. Accordingly, the less ambiguity there is surrounding what costs can and should be requested in grant proposals the better.

While many have called for funding bodies to recommend a general amount that should be included to support RDM in grant proposals (e.g., 5% of the total budget), it is advisable that costs be considered on a case by case basis. This viewpoint is echoed in the policy of the Dutch funder NWO which states that ‘the importance and the value of reuse on the one hand and the

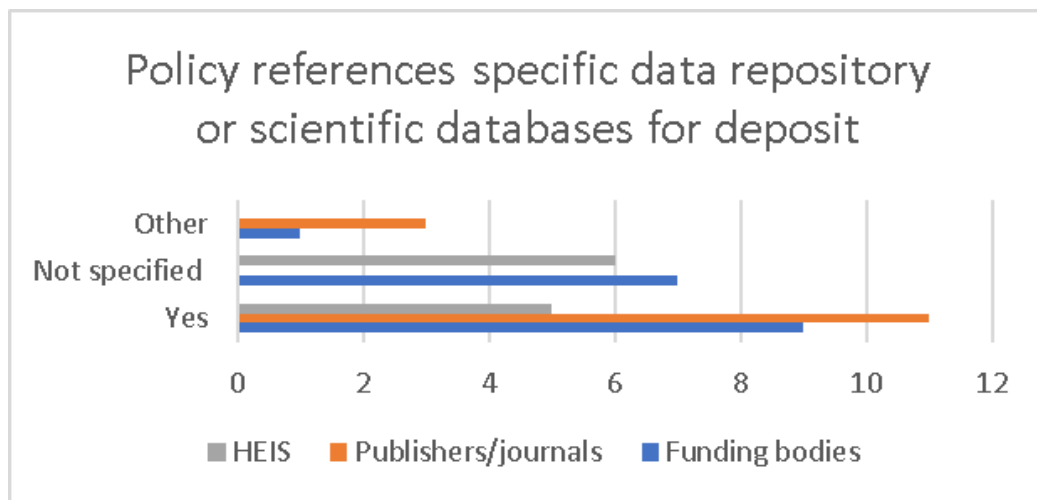
⁴⁷ <https://www.ukri.org/files/legacy/documents/rcukcommonprinciplesondatapolicy-pdf/>



costs and feasibility of data storage on the other should be in reasonable balance with each other and have a bearing on the volume of the data to be stored’.⁴⁸

For HEIs, another challenge is that the majority of institutional support services available to help researchers to manage and share their data are classified as indirect costs (Library, IT) and researchers often worry that adding in directly incurred costs to buy out support staff time for RDM may be viewed as double dipping. While some publishers have clearly defined chargeable services related to RDM (e.g., Nature⁴⁹) the majority of HEIs have not yet identified or priced specific value-added support services making it difficult to easily include relevant costs in grant proposals.

Figure 13. Policy references specific data repository or scientific databases for deposit




Nine funders provide guidance about where they expect researchers to deposit research data and refer to specific data repositories or scientific databases. Some of the funders reviewed placed a higher emphasis on data reuse than others. For example, in its data management section⁵⁰, the NWO focuses on the importance of data re-use, stating that ‘NWO only requests storage of reusable relevant data’. The Swiss National Science Foundation policy⁵¹ links the eligibility of RDM and sharing costs in grant proposals to the deposit of resulting data in recognised scientific, digital data archives that meet the FAIR principles. More alignment across funders around making payment of eligible RDM costs dependent on the deposit of data with FAIR aligned repositories could really help to ensure that research practices begin to change.

⁴⁸ <https://www.nwo.nl/en/policies/open+science/data+management+chapter>

⁴⁹ <https://www.springernature.com/gp/authors/research-data/research-data-support/pricing-for-research-data-support>

⁵⁰ <https://www.nwo.nl/en/policies/open+science/data+management+chapter>

⁵¹ http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/data-management-plan-dmp-guidelines-for-researchers.aspx

Example of good practice: taking the guesswork out of finding a suitable data repository	 Netherlands Organisation for Scientific Research
<p>The NWO data management section⁵² of its policy provides a link to repositories with a data Seal of Approval or CoreTrustSeal repositories⁵³. In its DMP Guidelines for researchers⁵⁴, researchers can find examples of repositories that comply with the FAIR Data Principles and are non-commercial. NWO provides a checklist for repositories that are FAIR as well as a separate list for non-commercial ones, which provides practical guidance for researchers in selecting a suitable place of deposit. By removing the guesswork, NWO helps to ensure that the researchers it funds can meet their requirements and ensures that FAIR data production and reuse is supported.</p> <p style="text-align: right;">NWO data management section</p>	

All publishers and journals in the cohort reviewed provide guidance to help researchers choose a suitable repository. Eleven of the publisher/journal policies included requirements for deposit with specific repositories. Three did not name a specific repository, but rather a type of repository that should be used (e.g., demonstrating compliance with international standards for data repositories). The policies and related guidance include a number of references to help researchers select a suitable data repository. Nine of the publisher/journal policies specify that the repository used for depositing data should be compliant with international standards for data repositories or standards that are community-credited. Subject or discipline-specific repositories are mentioned eight times in publisher/journal policies. An extensive list of subject-specific repositories for a wide range of disciplines is provided by Springer Nature⁵⁵. Eight of the publisher/journal policies point researchers to resources like re3data or FAIRsharing as tools to find an appropriate repository for their data. In nine cases, general repositories such as Dryad, Zenodo or Figshare are recommended, usually for instances where no appropriate subject-specific repository exists. In two instances, the repositories are run by the journal and in another two cases, publishers specifically refer to repositories that are run in collaboration with them.

Five out of 11 HEI policies reference specific data repositories or scientific databases for deposit. The term ‘Trusted Repository’ is not widely used in the 11 HEI policies analysed. As many European HEIs operate their own data repositories, there is a need for greater clarity

⁵² <https://www.nwo.nl/en/policies/open+science/data+management+chapter>

⁵³ <https://www.coretrustseal.org/why-certification/certified-repositories/>

⁵⁴ http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/data-management-plan-dmp-guidelines-for-researchers.aspx

⁵⁵ <https://www.springernature.com/gp/authors/research-data-policy/repositories/12327124>



about whether these should be certified as Trusted Repositories and, if so, where additional financial and technical support to achieve this level might be sourced. On the other hand, collaboration among universities to share trusted repository infrastructures or using existing certified repositories could reduce inefficiencies.

A crucial aspect under TFIR's *Implement* header is the provision of practical support to help researchers put FAIR into practice. Accordingly, a number of recommended actions relate to the provision of concrete guidance, templates, examples, and dedicated support. These include:

- Action 5.3: Data Management Plans should be tailored to disciplinary needs to ensure that they become a useful tool for projects. Research communities should be inspired and *empowered* to provide input to the disciplinary aspects of DMPs and thereby to agree model approaches, *exemplars* and rubrics that help to embed FAIR data practices in different settings.
- Action 16.3: *Guidelines* for the implementation of FAIR in relation to research data, to metadata, to code, to DMPs and to other relevant digital objects should be developed and followed.
- Action 17.8: *Concrete and accessible guidance* should be provided to researchers to find the optimal balance between sharing whilst also safeguarding privacy. There are many *exemplars of good practice* in providing managed access to sensitive data on which researchers can draw.
- Action 18.3: *Guidelines* should be provided for researchers and reviewers to raise awareness of eligible costs and reinforce the view that data management, long term curation and data publication should be included in project proposals. Funders should collaborate to *enhance guidance*.

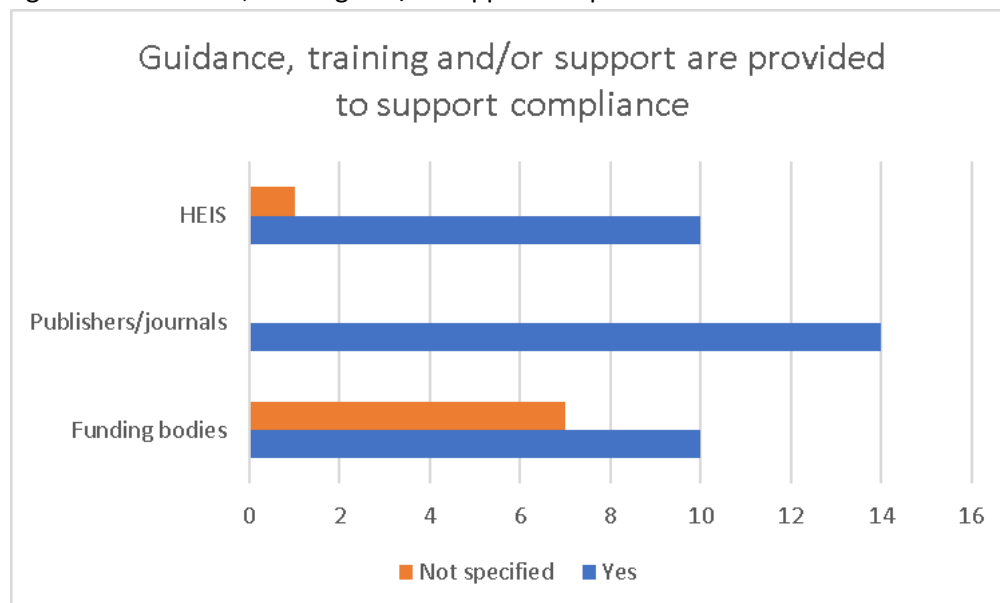
To assess how well funding bodies, publishers/journals and HEI's data policies currently reflect the TFIR actions listed above, the following policy characterisation element was considered:

- Guidance, training and/or support are provided to support compliance



Below, we summarise the results of the policy characterisation in relation to this element.

Figure 14. Guidance, training and/or support are provided



Encouragingly, the vast majority of all stakeholders reviewed do provide guidance to support researchers. All 14 of the publisher/journal policies reviewed are accompanied by guidance. Usually, these are guidelines or FAQs that aim to support authors to adhere to their data policies. Some include examples or blanket statements, usually in relation to data citation. In many cases, publishers/journals provide access to a dedicated helpdesk. Clear guidance was available for 10 of the 17 funder policies that were reviewed. However, while the numbers are not as high as one might expect, we have seen some excellent examples of guidance relating to FAIR emerging in recent years. Good examples include the detailed guidance from the Swiss National Science Foundation (SNSF)⁵⁶ and the revised Guidelines for the Deutsche Forschungsgemeinschaft (DFG) for Safeguarding Good Scientific Practice⁵⁷. Very few of the funders provide specific reference or guidance in relation to standards or protocols that should be used (3 of 17). One of those that do is the European Commission in its Guidelines on FAIR Data Management in Horizon 2020⁵⁸ which points applicants to FAIRsharing⁵⁹ to identify appropriate data standards in the life sciences. Specific standards or protocols are notably absent from the HEI policies with only 2 out of 11 referencing these.


⁵⁶ http://www.snf.ch/SiteCollectionDocuments/FAIR_principles_translation_SNSF_logo.pdf

⁵⁷ https://www.dfg.de/en/research_funding/principles_dfg_funding/good_scientific_practice/

⁵⁸ https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

⁵⁹ <https://fairsharing.org/>



<p>Example of good practice: provision of access to shared tools and guidance to support data management planning</p>	 KONE FOUNDATION
<p>The Kone Foundation’s policy⁶⁰ links to the Finnish Social Science Data Archive⁶¹ which provides excellent resources to data management planning⁶² as well as important information on rights, confidentiality, and file formats written in an easily digestible form. The guidance points researchers to the national DMPTuuli⁶³ data management planning tool. By providing access to one-stop shop with comprehensive and easy to follow guidance, the Kone Foundation is making RDM and the production of FAIR data easier for researchers. This example also shows that guidance does not have to be developed in-house. If there are good shared resources available, it makes sense to direct researchers to these rather than re-inventing any wheels.</p> <p style="text-align: right;">Kone Foundation instructions for grant recipients</p>	

The publishers/journals we reviewed were more likely to offer discipline specific guidance in relation to subject specific journals. While some funding bodies such as the European Research Council (ERC) provide discipline specific guidance⁶⁴ to grant applicants, most funding bodies provide generic guidance. At the institutional level, guidance tends to be generic however there have been recent efforts to introduce support at the faculty level. For example, ‘every TU Delft faculty has a dedicated Data Steward to answer questions, provide advice, and help develop appropriate solutions for research data management and sharing’⁶⁵. However, in many universities there may not be sufficient staff or financial resources to replicate this approach. In these instances, HEIs will need access to a set of shared domain specific resources to support better RDM practice among researchers. The OpenAIRE and FAIR Data Expert Group survey report on Horizon 2020 template for DMPs⁶⁶ identified potential to work more collaboratively with the European Strategy Forum on Research Infrastructures (ESFRIs)⁶⁷ and other disciplinary groups to develop domain-level guidance that builds upon the existing Science Europe domain protocols initiative. As many of the ESFRIs are currently involved in INFREOSC Cluster projects to support FAIR data production, this approach does offer great potential.

⁶⁰ <https://koneensaatio.fi/en/grants/forgrantrecipients/>

⁶¹ <https://www.fsd.uta.fi/en/>

⁶² <https://www.fsd.uta.fi/aineistonhallinta/en/data-management-planning.html>

⁶³ <https://www.dmptuuli.fi/>


⁶⁴ https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf

⁶⁵ <https://www.tudelft.nl/en/library/current-topics/research-data-management/r/data-stewardship/>

⁶⁶ <http://doi.org/10.5281/zenodo.1120245>

⁶⁷ <https://www.esfri.eu/>



<p>Example of good practice: providing domain specific protocols for RDM</p>	
<p>Research organisations and funders increasingly ask researchers to create data management plans (DMPs) for their work and proposals. A lack of standardisation means that these can be time-consuming to create and difficult to compare and evaluate. Science Europe presents a framework for the creation of domain-specific protocols that can be used as standardised templates, reducing the administrative burden on both researchers, research organisations, and funders.</p> <p>The core idea of the approach is that research communities will use this framework to formulate ‘protocols’ for the collection and management of data within their disciplinary domain or community. Instead of having to evaluate and monitor many individual DMPs, funders and research organisations would simply require project proposers to comply with the relevant protocol. This would result in much shorter DMPs on average, reducing the time needed to review and evaluate them, as well as the time needed for researchers to create them.</p> <p style="text-align: right;">Science Europe Guidance Document Presenting a Framework for Discipline-specific Research Data Management⁶⁸</p>	

Summary of landscape in relation to Implement - culture, technology and skills for FAIR practice

In general, the stakeholders’ policies are addressing the actions outlined under TFiRs heading of *Implement - culture, technology and skills for FAIR practice* but there is still work to be done to harmonise data management planning requirements and support across stakeholders and the research lifecycle. There is potential to explore the concept of the assessed outputs management plan being piloted by Wellcome across other funders to ensure a broader range of research outputs are managed and made FAIR. End stage DMPs could provide real value for supporting FAIR data reuse and more funding bodies should consider the approach that has been introduced by the Swiss National Science Foundation. There is still some uncertainty with regards to eligible RDM and data sharing costs. Funding bodies should provide clearer guidance spelling out precisely what costs can be requested. On the HEI side, more should be done to develop clear pricing structures for value-added services offered by central support units that can be included in grant applications as directly incurred costs. Clarity about whether HEI data repositories should become certified is needed. If these repositories should be certified, there is a need to develop shared guidance around how to progress towards certified status along with

⁶⁸ http://www.scienceeurope.org/media/nsxdyvqn/se_guidance_document_rdmeps.pdf



a clear costs/benefit analysis to support decision making. While some HEIs are introducing domain specific RDM support in faculties, the majority of universities do not have the staff or financial resources to provide this level of support. As such, there is a need to develop and curate a shared set of discipline specific guidance and training resources that can be easily found and reused. This should be done in close cooperation with the ESFRIs and related Cluster projects.

Embed and Sustain - incentives, metrics and investment

In this section, we explore the current policy landscape in relation to relevant *Turning FAIR into Reality* (TFiR) actions under the broad heading of *Embed and Sustain*. Many of the actions in this section relate to sustainable funding for FAIR data production and reuse which are not reflected in the policies of funding bodies, publishers/journals or HEIs. However, under the *Embed and Sustain* heading TFiR states that for evidence of change to be identified, metrics on FAIR data need to be collected and reported. TFiR cites the tracking and reporting of open access publication statistics over time and the move to increased automation of the process as an example that might be emulated with respect to FAIR data. The specific TFiR actions that are addressed in this section are:

- Action 26.2: Citation of data and other research outputs needs to be encouraged and supported - for example, by including sections in publishing templates that prompt researchers to reference materials, and providing citation guidelines when data, code or other outputs are accessed.
- Action 26.4: A broader range of metrics must be developed to recognise contributions beyond publications and citation. These should recognise and reward Open and FAIR data practices.

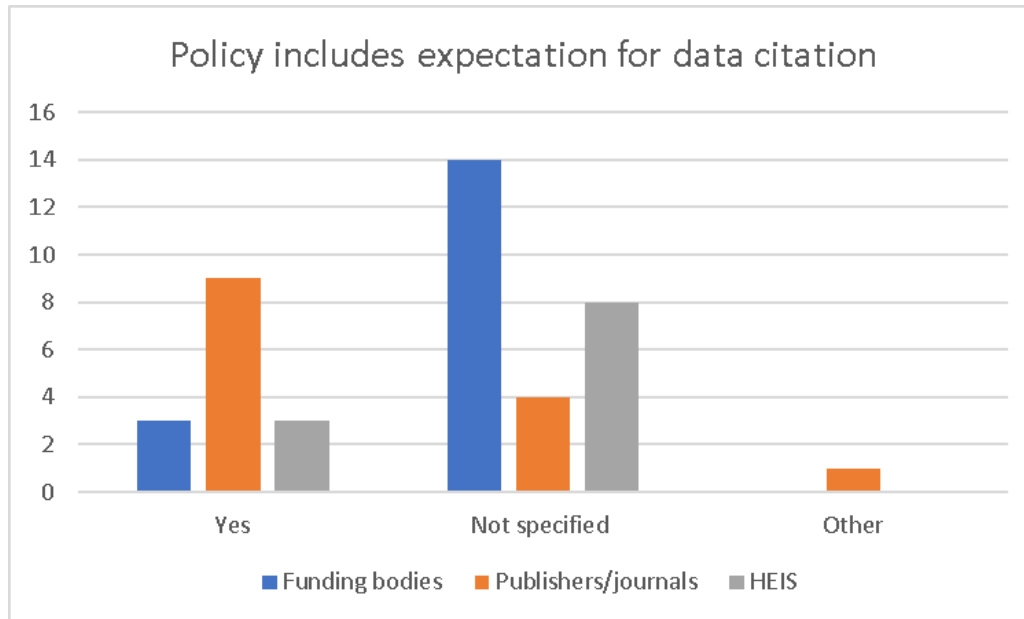
To assess how well funding bodies, publishers/journals and HEI's data policies currently reflect the TFiR actions listed above, the following policy characterisation elements were assessed:

- Policy includes an expectation on data citation
- Policy requires a data availability statement
- Policy includes specific reference to preservation (mid to longer-term)
- Policy compliance is monitored



Below, we summarise the results of the policy characterisation in relation to these elements.

Figure 15. Policy includes expectation on data citation



As noted in the recent State of Open Data 2019 report, citations are seen by researchers as the ‘holy grail in terms of reward’⁶⁹. Almost all funder policies reviewed (14/17) do not currently include expectations on data citation. Given that citations are a key motivator for researchers when considering whether to share their data, a stronger requirement from funding bodies to formally cite data along with concrete guidance on how to do so would help to provide a structure to reward authors and support more automated tracking of data reuse. Guidance on data citation should be in line with the Joint Data Citation Principles which would support TFIR Action 26.3. Data availability statements also help ensure improved access to research and can lead to increased citations, yet only 6 funders currently require these.

Data citation is required as part of 9 publisher/journal policies. One publisher/journal policy suggests data citation but does not make it mandatory. Four of the publisher/journal policies reviewed make no mention of data citation. Publishers and journals have a key role to play in embedding data citation as part of good research practice and as such there is scope to strengthen requirements. Of those publishers that do require data citation, the standards most frequently recommended include the Joint Declaration of Data Citation Principles⁷⁰, DataCite


⁶⁹ <https://www.digital-science.com/resources/portfolio-reports/the-state-of-open-data-2019/>

⁷⁰ Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 <https://doi.org/10.25490/a97f-egyk>

Recommendations⁷¹, FORCE11 Software Citation Principles⁷² and the ESIP Data Citation Guidelines for Earth Science Data⁷³. Data availability statements are required by 11 of the publisher/journal policies examined.

Three of the 11 HEI policies reviewed currently promote data citation. The relatively small number of HEI policies currently promoting data citation suggests that HEI policies could be strengthened in this respect. However, it is worth bearing in mind that as funding body and publisher requirements tend to have greater influence on researchers' behaviour, it is perhaps not essential that HEI policies address data citation explicitly but rather that they provide institutional support to help researchers to meet external expectations. In this respect, it is even more vital for funders and publishers to strengthen their requirements for data citation in the short term. Three of the 11 HEI policies reviewed currently require a data availability statement.

While a few HEIs are adapting their assessment procedures, few currently recognise FAIR data outputs and data stewardship. A recent survey conducted by EUA⁷⁴ shows that only 48% of universities consider 'other types of research output' including research data as important or very important for the assessment of researchers compared to research publications, which were considered to be important or very important by 90% of HEIs. If research assessment methodologies are to evolve to better reflect a broader range of outputs, it will be essential that reuse of data and other non-traditional outputs can be more easily tracked and used to support decision making.

<p>Example of good practice: measuring and rewarding FAIR practices</p>	
<p>University College London (UCL) has been a strong advocate of the Open Science movement for many years and is actively developing hard and soft infrastructure to support its researchers in making their data FAIR and has adapted its assessment procedures to reward good open research practices. In his foreword for the State of Open Data 2019 report, Dr Paul Ayriss explains:</p> <p>'One of the key requirements of the change of culture needed to deliver open and FAIR data is a change in the university reward and incentive system. Current practice is focused on publications and, in many cases, the impact factor of the journals in which articles are published. There is little room for</p>	

⁷¹ <https://datacite.org/cite-your-data.html>

⁷² <https://www.force11.org/software-citation-principles>

⁷³ <https://doi.org/10.6084/m9.figshare.8441816>

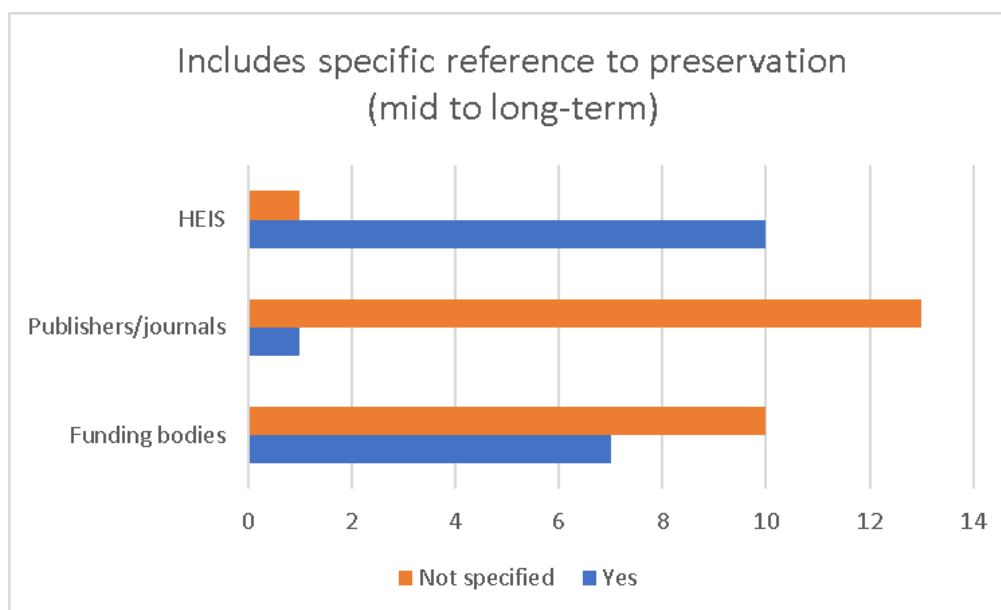
⁷⁴ 2019 Open Science survey results on research assessment. Zenodo. <http://doi.org/10.5281/zenodo.3435325>



research data in this model. Professor Bernard Rentier and a Working Group of the European Commission have recently presented a report entitled Evaluation of Research Careers fully acknowledging Open Science Practices. This report identifies 23 rounded criteria for reward, of which datasets is one. No university in Europe has yet introduced this complete matrix, but UCL has already modified its academic promotions framework to acknowledge openness as a criterion for reward’.

Dr Paul Ayris, Pro-Vice-Provost, University College London Library Services

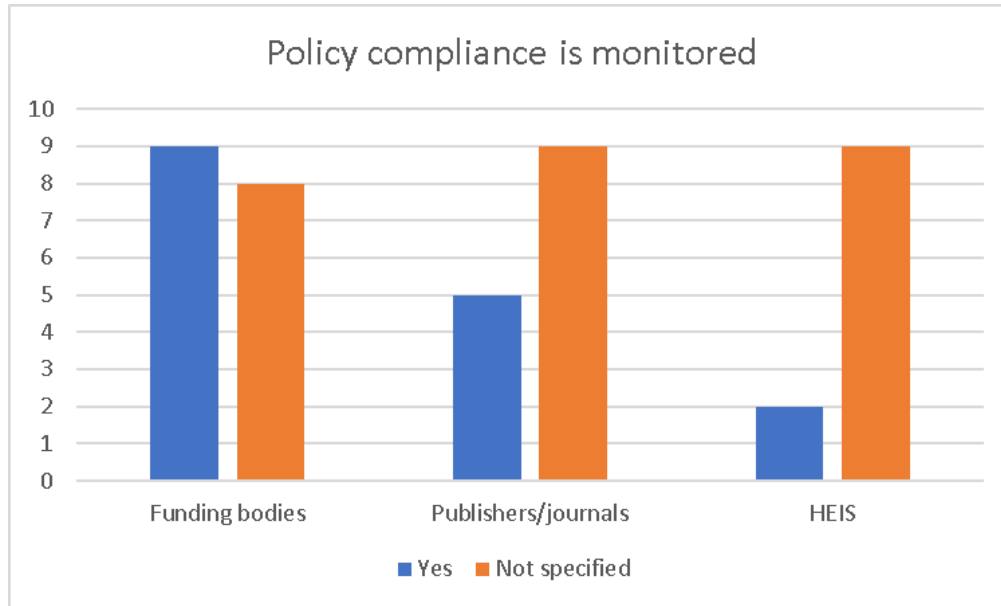
Figure 16. Policy includes specific reference to preservation (mid to long-term)



Regarding the length of data availability, the majority of funders do not specify a minimum length of data availability but 7 do make reference to longer-term preservation of the data. Although all of the analysed publisher/journal policies either suggest, require or even monitor data sharing, none make clear an expectation for the minimum length of data availability and only one makes a reference to longer-term preservation. However, in this case the period for which the data should be preserved is vague and simply states ‘a sufficiently long period of time’. Six of the HEI policies reviewed include minimum lengths for data availability and the vast majority of policies from HEIs (10 out of 11) do include specific reference to long-term preservation. While it is encouraging to see that HEIs are clearly aware of the need to preserve some data for the longer-term, it is worrying that the total cost of data preservation may fall to

HEIs to bear. Funding bodies and publisher/journals also have a key role to play in supporting the longer-term costs.

Figure 17. Policy compliance is monitored



Just over half of funding body policies reviewed (9 of 17) do indicate that compliance is monitored with three of these indicating that sanctions for non-compliance may be applied. Publishers/ journals were less active in this respect with just over a third stating that they monitor compliance. However, 6 publisher/journal state that sanctions may be applied to those who do not comply with their policies. Half of those who may impose sanctions do not clearly indicate in their policy that compliance is monitored which implies that even though compliance checking may not be referenced in the policy, it does take place. The vast majority of HEI policies reviewed (9 of 11) do not specify whether they monitor compliance and none indicate that there are any sanctions applied to those who are found to be non-compliant. As investment in staff to support RDM at the HEI level remains relatively low, the fact that compliance with the policies is not yet monitored is not surprising.

Monitoring of compliance with data sharing requirements poses several challenges. Not only is compliance monitoring extremely resource intensive, there are also concerns about introducing such requirements too quickly in research communities where data sharing has not been the norm. Questions remain about what sort of sanctions for non-compliance might be appropriate. The State of Open Data 2019 report states that 67% of survey respondents felt that funders should withhold funding or penalize researchers who do not share their data if the

funder mandated that they should⁷⁵. However, just under half (43%) of respondents to the FAIRSF AIR open consultation indicated that the introduction of such penalties for not making data FAIR would be very/quite negative. While views on the introduction of penalties vary, it does seem that without some level of compliance checking and potential repercussions for non-compliance, researchers may view RDM and data sharing as merely box-ticking exercises when applying for new grants. While realising culture change is generally better achieved with carrots rather than sticks, there is a need for policymakers to define some measures to dis-incentivise non-compliance.

Summary of landscape in relation to Embed and Sustain - incentives, metrics and investment

The actions listed under TFIR's heading of *Embed and Sustain - incentives, metrics and investment* are being addressed in the policies and guidance of the stakeholders' policies to some extent. However, if research assessment is to evolve to reflect and reward researchers for a broader range of outputs and activities, it is essential that data citation using agreed standards and norms becomes more common practice. Funding bodies and publishers have a key role to play in strengthening their requirements in relation to data citation and guiding researchers on how to do this effectively. Producing FAIR data requires an investment of time and money. To maximise the value of this investment, stakeholders need to provide clearer expectations for the long-term preservation of selected outputs. Currently, HEIs are most likely to refer to longer-term preservation in their policies and there is a real risk that they will end up bearing the brunt of the longer-term costs associated with keeping data FAIR over time. There is a need to develop more equitable models that spread the costs between all stakeholders. There is a need for funding bodies and publishers to enforce the data sharing policies that they have put in place. The emphasis should be on providing incentives to enforce compliance - such as requiring evidence of previous data sharing when applying for new grants - but must also include penalties where appropriate.

⁷⁵ Science, Digital; Fane, Briony; Ayris, Paul; Hahnel, Mark; Hrynaszkiewicz, Iain; Baynes, Grace; et al. (2019): The State of Open Data Report 2019. figshare. Report. <https://doi.org/10.6084/m9.figshare.9980783.v2>



5. Results of the Open Consultation

The open consultation ran from August 2 to September 27, 2019. A total of 106 responses were received. Anonymised and aggregated data resulting from the open consultation has been deposited in the FAIRsFAIR community in Zenodo and is openly available for use by third parties

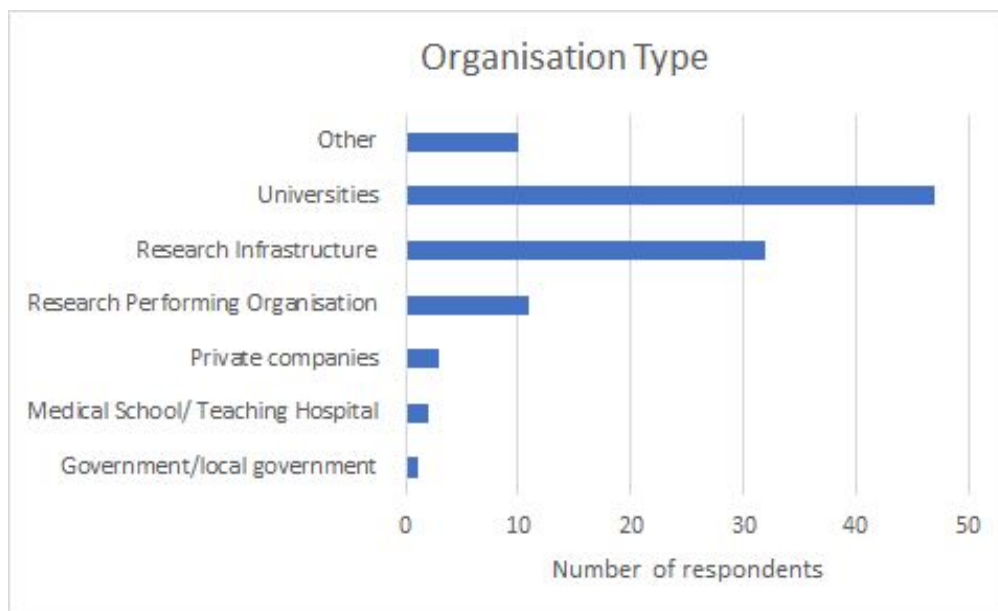
76

Overview of respondents

Breakdown by organisation type

Responses were provided by representatives of a wide range of organisations as can be seen in Figure 18. The majority of responses were received from staff working in Universities (44%) followed by responses from Research Infrastructure staff (30%), Research Performing Organisations (10%) and ‘other’ (9%). Those selecting ‘other’ included financing organisation, think tank; University Medical Center; Funding bodies, e-infrastructures, Institute of Technology; Cross-disciplinary trustworthy digital repository; cluster of research infrastructures; not-for-profit organisation.

Figure 18. Breakdown of respondents by organisation type, FAIRsFAIR open consultation on Policy and Practice



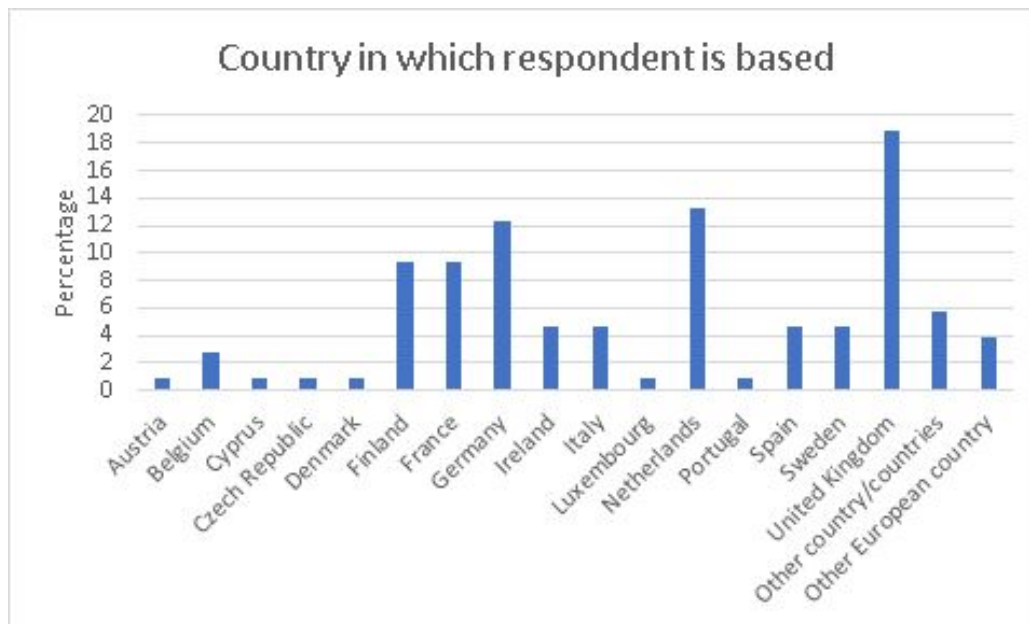
⁷⁶ FAIRsFAIR Policy and Practice Survey 2019 data for D3.1_D3.2_D6.1 <https://doi.org/10.5281/zenodo.3550529>



Breakdown by country

We received a good number of responses to the open consultation from across Europe as shown in Figure 19. The majority of responses were received from respondents based in the UK (19%), followed by the Netherlands (13%), Germany (12%), Finland and France (9% each), Ireland, Italy, Spain and Sweden (5% each). The majority of responses came from countries that currently demonstrate a relatively high engagement with Open Science and the FAIR principles.

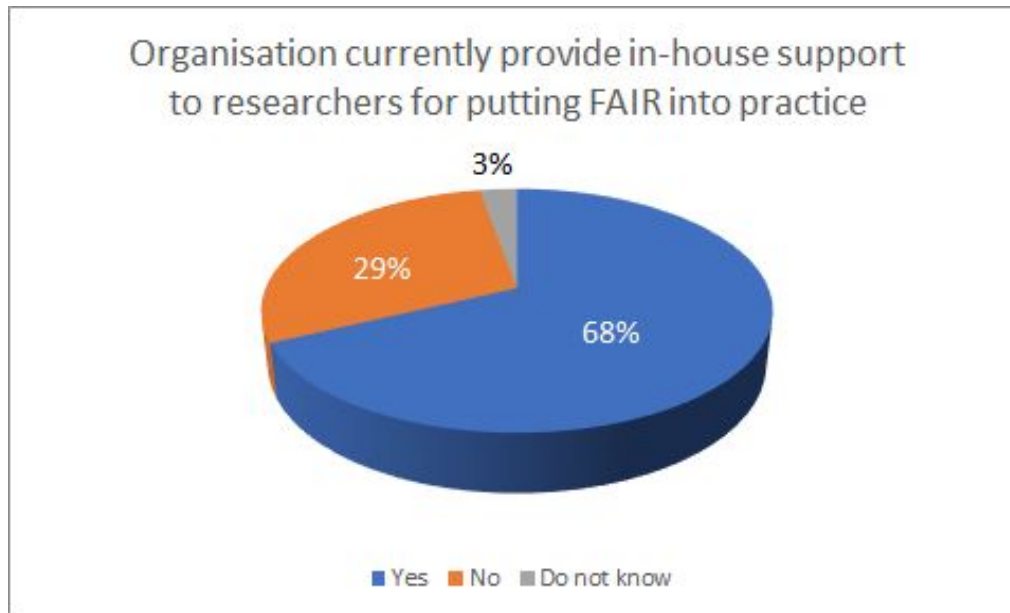
Figure 19. Breakdown of respondents by country in which they are based, FAIRsFAIR open consultation on Policy and Practice



Provision of in-house support for FAIR

In light of the need for access to guidance and support across the research ecosystem, the open consultation asked respondents to indicate whether their host organisation currently provides in-house support to researchers for putting FAIR into practice. Of the 106 respondents who provided an answer to this question, about two thirds do offer in-house support services as can be seen in Figure 20. As noted above, the majority of responses came from countries with a high level of commitment to Open Science and the FAIR principles. While the responses are too low to provide concrete evidence, the findings suggest that in-house support is most frequently provided by respondents based in Belgium (100%), Finland (90%), United Kingdom (85%), the Netherlands (71%) and Germany (69%).

Figure 20. Organisations currently providing in-house support, FAIRsFAIR open consultation on Policy and Practice

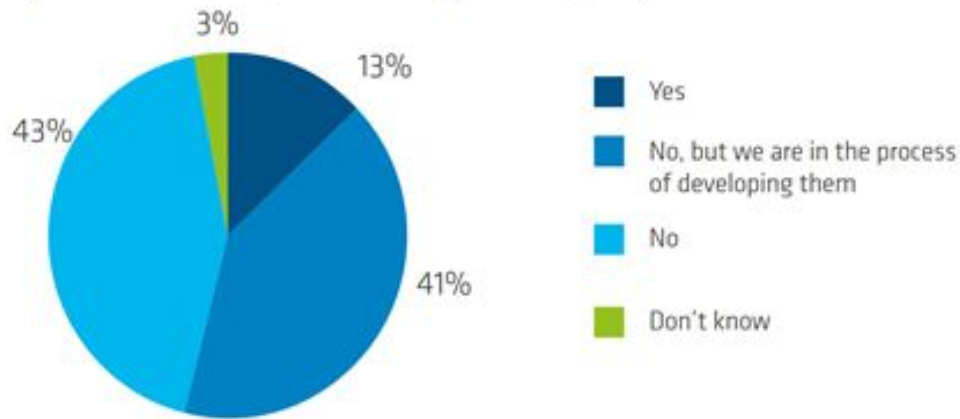


More than three quarters of respondents working in Universities provide in-house support (77%), followed by 66% of Research Infrastructures, and 64% of Research Performing Organisations. While the vast majority of HEIs responding to the open consultation do provide in-house support, these findings should be viewed in light of the findings of the 2017-2018 EUA Open Access survey which revealed that only 13% of European HEIs had developed institutional guidelines for open access to research data⁷⁷. However, as the table below indicates, 41% of EUA survey respondents were in the process of developing such guidelines (Figure 21). The EUA survey is currently being re-run as part of FAIRsFAIR WP7 and it will be interesting to see if the number of HEIs providing guidance has increased as planned.

⁷⁷ <https://eua.eu/resources/publications/826:2017-2018-eua-open-access-survey-results.html>

Figure 21. Institutional guidelines for open access to research data as presented in the 2017-2018 EUA Open Access Survey

Figure 33. Existence of institutional guidelines for Open Access to research data



Number of respondents: 318/321.

Respondents who indicated that they did provide in-house support in relation to FAIR were asked to provide a free text response outlining the nature of the support provided. Several respondents indicated that they provide training and guidance on aspects of research data management (RDM) and Open Science with a few saying that they specifically address FAIR. One respondent stated that a handbook covering all aspects of FAIR is made available to researchers. In several cases where FAIR is not explicitly referenced in training and guidance, respondents indicated that the principles are addressed implicitly.

A couple of respondents indicated that they provide access to data stewards and/or provide data stewardship training.

Some organisations stated that they provide on demand advice regarding data management planning, repositories, metadata and documentation, file formats, versioning, vocabulary, licensing and regulatory and ethics issues.

Several respondents stated that they provide generic guidance via institutional web pages. These are most frequently provided by Library or Research Support units. In some cases, these pages provide pointers to externally developed resources such as those developed by Research Infrastructures and related Cluster projects (e.g., ELIXIR, SSHOC partners).

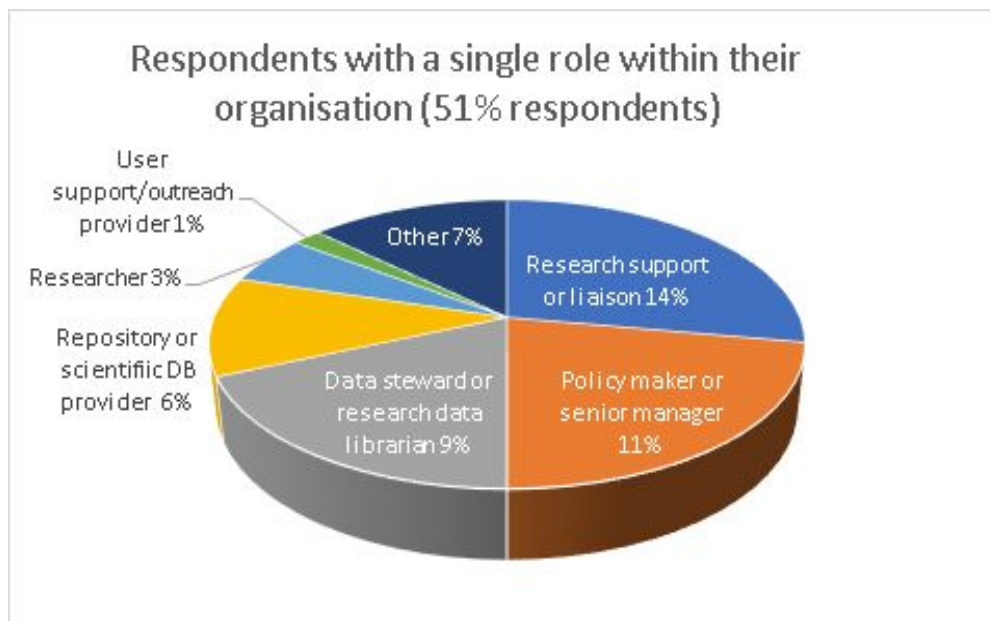
A few respondents pointed to technical infrastructure that is made available to researchers including in-project data storage, data repositories and data archiving services. One respondent stated that they provide tape storage of data, but that researchers are responsible for curating the associated catalogues which they cite as being a major gap while another specified metadata curation as a key service provided by the organisation. Some respondents indicated that they struggle to provide advice and guidance in relation to making data interoperable which reflects findings in other studies.

Roles of respondents

The open consultation included a question asking about the respondent's role(s) within their organisation. Respondents could select as many of the options as they felt were relevant. Just over half of those responding (51%) classified their role using just of the options provided as is presented in Figure 22. Of these, the majority of respondents classify their role as Research Support or Liaison (14%), followed by Policymaker or Senior Manager (11%), and by Data Stewards or Research Data Librarian (9%). Almost half of respondents (49%) indicated that they held more than one role within their organisation. In many cases, respondents had between three and five roles within their organisation. This finding resonated with many of the participants of the FAIRSF AIR Policy and FAIR data focus group held in Madrid on October 30, 2019. The long tail of hybrid and potentially niche roles may reflect disparity among organisational structures and specific institutional needs. While the sample rate is too small to provide any concrete evidence, the long tail of respondents carrying out a number of different roles implies that there is a lack commonly agreed career profiles for many of those working in the broad field of data stewardship.



Figure 22. Breakdown of respondents by role, FAIRSFAR open consultation on Policy and Practice



‘Other’ roles

Data manager; Digital Humanities Officer, Project Manager, Librarian and Repository Manager (Publications repository), Special Advisor and Project Manager, All-purpose librarian, Publisher, Public Affairs Officer, IT Staff, Academic journals publisher in addition to repository provider, Data analyst for policy support, scientific Project Manager (mix of logistics and scientific work), IT expert, Data Steward on departmental level in the research domain of a University Medical Center, Public IT Infrastructure Designer and Developer.

Policy related questions

To keep the burden on respondents to a minimum, the consultation aimed only to collect information that would help us to assess how the policies of various stakeholders influence the way that researchers work. Accordingly, only three questions relating to FAIR data policies were presented to respondents all of which aimed to better understand how policies may affect researchers' practices.

Assessment of drivers for policy development

The first question asked respondents to indicate what, if any, were the main drivers for data policy development in their organisation. This was presented as a free text question with a limit of 250 characters for responses. A total of 84 respondents provided an answer to this question (79% of total respondents) and 22 respondents left this question blank. The key drivers

identified by the respondents have been grouped under broad headings and are described below.

[Funding body requirements/mandates:](#)

The most frequently cited driver for data policy development among the respondents is funding body requirements and mandates. Of these, those of the European Commission were frequently mentioned as being a key driver. A few respondents indicated that their institutional policy is being developed to align with major funders' requirements. One respondent noted that funder requirements are driving the development of their research data management and sharing policy, but that they are placing an equal emphasis on ensuring that their policy will be workable with regards to researchers' practices and are working to develop adequate support to embed good practice. One respondent ranked the drivers that are influencing their policy development stating that funder requirements were the main driver followed by publisher requirements and then by good research practice.

[General good practice \(RDM, Open Access, Open Science\):](#)

A large number of respondents indicated that the key driver for their organisation is to support good research practice and to support Open Science objectives. A few respondents cited the need to provide free and open access to data. Other respondents made reference to the fact that by having such free and open access to research data, researchers could avoid duplicating data, ensure maximal impact for publicly funded research, and avoid having to spend more time or money than necessary on pre-processing and preparing research data. A few respondents also mentioned that their organisational policy was developed to ensure longer-term access to research information including data and software.

[Drivers relating to research data protection, integrity, reproducibility, quality, ethics:](#)

Several respondents made specific reference to compliance with the General Data Protection Regulation (GDPR) as the key driver for their policy development. Many respondents also stated that they were developing policies to support best practice with regards to research integrity codes of conduct. Some respondents cited legal and ethical issues as being the key drivers for developing policies related to research data. And a few respondents made specific reference to reproducibility as being the key driver for their policy development.

In relation to the policy characterisation work described above, all of the HEI policies referenced data protection rules such as GDPR. In 9 cases, research data policies also referenced research integrity as a backdrop or component of the policy. The visibility of legal



regulations such as GDPR in the 11 analysed HEI policies reflects findings⁷⁸ of European University Association (EUA) on the main elements of universities' research data policies. Based on the mandatory and optional elements of these policies, it appears that legal considerations factor in the strongest in their development. Both provisions for handling personal and sensitive data are most frequently mentioned as mandatory elements. Data storage and guidelines for specific disciplines are mostly optional or not included. Seen in this light, policies may be primarily an instrument to ensure compliance with existing regulatory frameworks for data protection.

Drivers at the Institutional level:

At the institutional level, many indicated that policy development was undertaken to ensure that researchers meet funding body requirements. In a few instances, respondents stated that policy development was driven by researchers' needs (bottom-up). One respondent indicated that their policy development was driven by the Research Excellence Framework (REF). One respondent stated that their policy was developed to provide a framework for the support services they are developing. Another respondent indicated that their policy development was linked to Institutional incentives and rewards systems.

Drivers relating to Publishers:

Several respondents indicated that their policy development was driven by the mandates of publishers. As publications remain the primary source for researcher performance assessment, this is not surprising.

Drivers at the National / International level:

For several respondents, policy development was undertaken to ensure alignment with the policies of government in relation to research integrity and Open Science. One respondent made specific reference to their policy work being driven by an aim to align with EU policy making in relation to the European Open Science Cloud (ESOC).

Competitive advantage:

A few respondents stated that their policy development was driven by the aim of obtaining a competitive advantage. In one instance this was based around a desire to know what data is being produced in-house and how such knowledge could inform future data acquisition and support data reuse. Another respondent stated that the policy development was undertaken to support career advancement of researchers.

⁷⁸ <https://eua.eu/downloads/publications/2017-2018%20open%20access%20survey%20results.pdf>



Drivers relating to technical requirements and development activity:

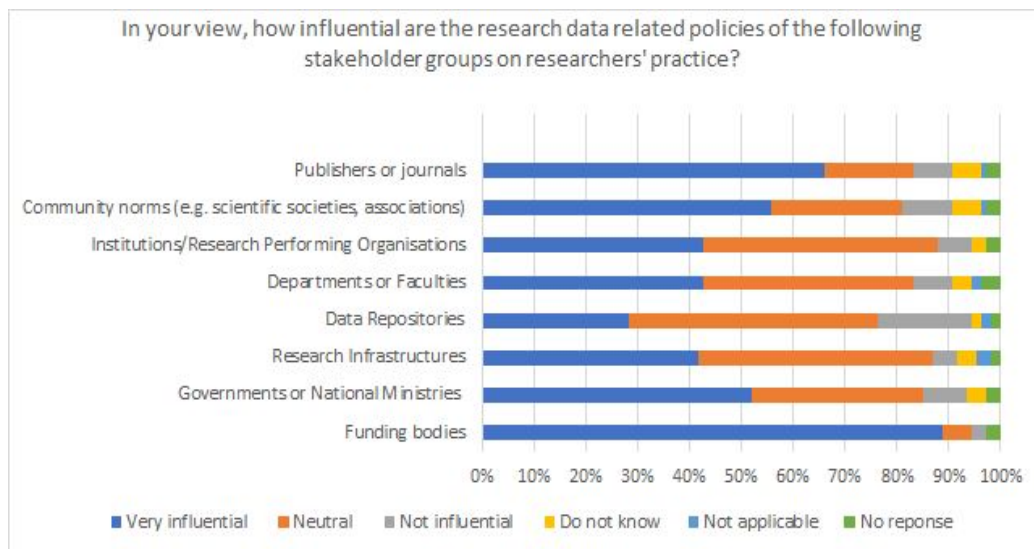
A few respondents made reference to technical requirements driving policy development at their organisation. In a couple of instances, the need to manage a very high volume of data generation was cited and, in another instance, the need to provide remote data access and analysis was mentioned.

Assessment of the influence of the research data related policies of the different stakeholder groups on researchers' practice.

Respondents were then asked to rate how influential they felt the research data related policies of the different stakeholder groups were on researchers' practice. This question was presented as a matrix and the options presented included *Very influential*, *Neutral*, *Not influential*, *Don't know*, *Not applicable*. Below we highlight the findings collectively and for each stakeholder group.

Overall, the policies of funding bodies were considered to have the greatest influence on researchers' practices (90%) as can be seen in Figure 23. This tallies with the findings of the recent State of Open Data 2019 report which found that 'funders are perceived as important in changing attitudes to engagement with research data, as 69% of respondents thought that funders should make the sharing of research data part of their requirements for awarding grants'.

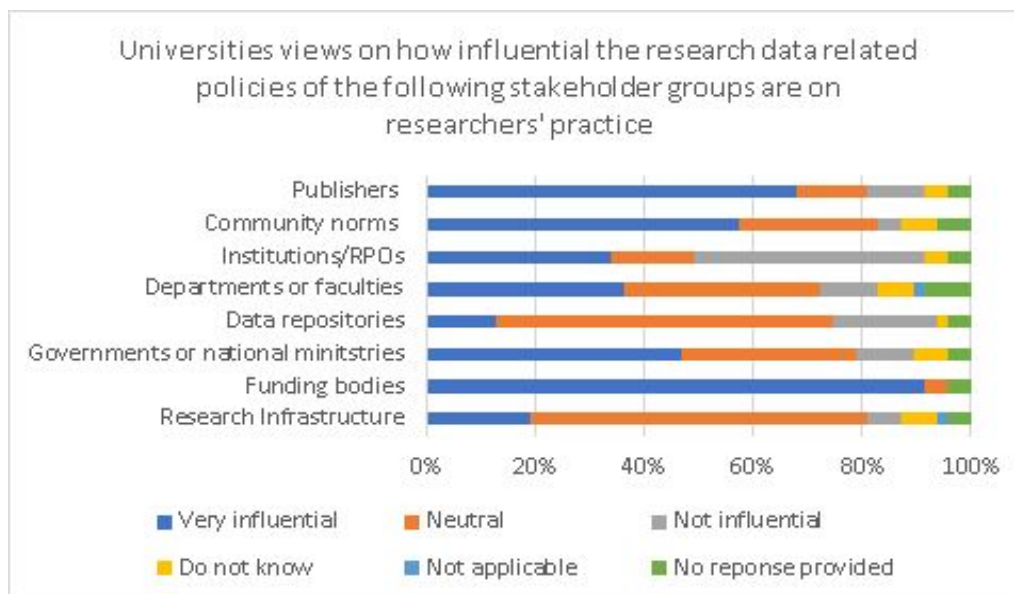
Figure 23. Influence of stakeholders' policies on researchers' practice, FAIRSF AIR open consultation on Policy and Practice



Publisher and journal policies were considered to be the next most influential on researchers' practice (66%). Community norms were also seen as having significant influence on researchers' practices (56%) as were the policies of Government or National Ministries (52%). At the organisational level (Institutions/RPOs, Departments or Faculties) policies were seen to have slightly less influence over researchers' practices (43% each). The policies of Research Infrastructures were considered to be very influential by 42% of respondents. The least influential policies appear to be those of data repositories with just 28% of respondents stating that these were very influential on researchers' practices.

When looking at how each of the stakeholder groups rated the influence of different policies, funding bodies were universally identified as having the most influence on researchers' practice. However, it seems that Universities don't feel that their own policies have very much influence over researcher's behaviour with more than 40% of HEI respondents stating that their policies were not influential (Figure 24).

Figure 24. Universities view on influence of stakeholders' policies on researchers' practice, FAIRSFAR open consultation on Policy and Practice



Respondents were given the option to select 'other' in the matrix above, and six respondents provided information. Of these, four were rated 'Very influential' and included:

- Keeping large data sets requires storage space; who pays for that space and how much drives how much can be kept

- Peers
- Specialised IT Infrastructures (e.g. in Biomedical)
- In different circumstances, all these bodies set policies (standards, recommendations, terms and conditions) for different activities (funding receipt, deposit, publication etc.). Funding bodies (money) and publishers (recognition) have most direct influence.

Two respondents provided information for ‘other’ but did not rate them. These include:

- Funding bodies and national ministries are closely interrelated in Finland, since the most important funder, Academy of Finland, is funded by the Ministry of Education.
- These are all influential sources of policy and practice guidance. Coordination and alignment are therefore vital. Guidance should be globally applicable where possible and only designed for local contexts when local conditions apply.

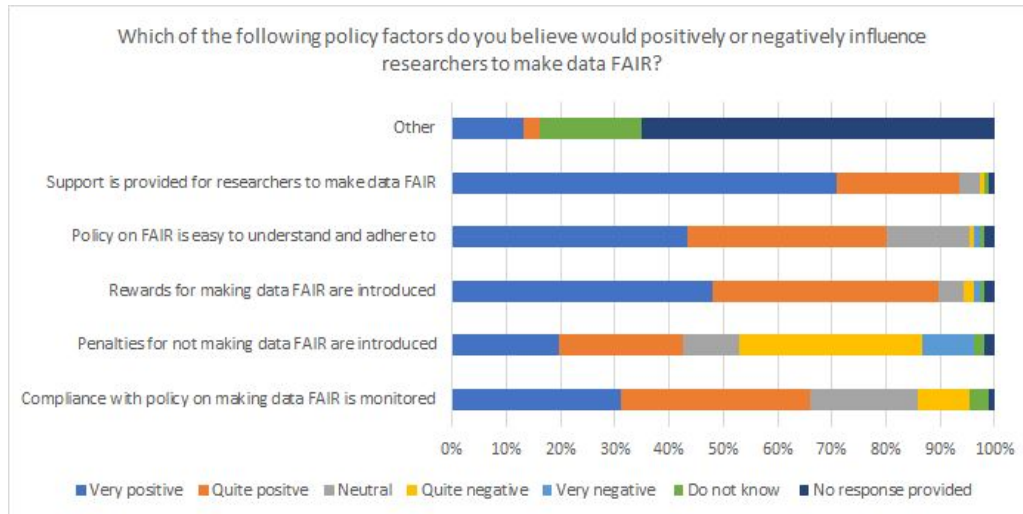
Assessment of policy factors

Respondents to the open consultation were asked to consider a range of policy factors and indicate whether they felt the factor might positively or negatively influence researchers to make their data FAIR. This question was presented as a matrix and the options presented included Very positive, Quite positive, Neutral, Quite negative, Very negative, or Do not know. Below we highlight the findings in relation to each of the policy factors.

The findings of the open consultation show that the availability of support to help researchers make their data FAIR is considered to be the most positive policy factor influencing researchers’ behaviour with more than 93% of respondents rating this factor as ‘very’ or ‘quite’ positive (Figure 25). This is encouraging, as the results also show that more than two thirds of respondents currently provide in-house support. Just under 90% of respondents also felt that the introduction of rewards for making data FAIR would have a positive influence on how researchers work. Policies that are easy to interpret and follow were also rated as a positive influence by 80% of respondents.



Figure 25. Influence of policy factors on researchers’ practice, FAIRSF AIR open consultation on Policy and Practice



With regards to monitoring compliance with policies, 66% felt that this would positively influence researchers’ behaviour. However, almost 10% of respondents thought compliance monitoring would actually be quite negative. The policy factor that was seen as the most negative was the introduction of penalties for not making data FAIR with more than 43% of respondents saying this would be very/quite negative. This does not seem to tally with the findings in the State of Open Data 2019 which reported that 67% of respondents thought that funders should introduce penalties for non-compliance with data sharing mandates (State of Open Data, 2019).

6. Conclusions

FAIRSF AIR’s analysis of the data policy landscape in 2019 has shown that the priority and supporting actions outlined in the *Turning FAIR into Reality* (TFiR) report are being reflected in the policies of funding bodies, publishers/journals and HEIs to some extent. However, there is still much to be done by policymakers to foster the emergence of the FAIR ecosystem that is envisaged by TFiR and is necessary to support the aims of the European Open Science cloud. Key findings in relation to each of the three stages outlined by *Turning FAIR into Reality* are presented below.

Key findings in relation to *Define - concepts for FAIR Digital objects and the ecosystem*

- Efforts are needed to raise general awareness about the FAIR principles and the potential benefits of implementing them
- The visibility of FAIR related initiatives needs to be improved and there should be better cooperation between them to amplify key messages about what FAIR means in a practical sense
- The policies of funding bodies are the key driver for many of the stakeholders developing policies – both at the national and institutional level
- Collaboration is needed to develop a set of shared national profiles that make clear the legal and policy frameworks of each EU country
- To support both human interpretation and machine actionability, it is crucial that policies make clear the period of time to which they apply
- The policies of all stakeholders should be described consistently using a structured data markup schema to support both human and machine readability
- Policies should be assigned PIDs to ensure that the right version can be found and fed into machine actionable pipelines
- Clearer definitions of what data means are needed and definitions should be harmonised across stakeholders
- Funding bodies and publishers/journals could strengthen their expectations around the sharing of both data and metadata



Key findings in relation to *Implement - culture, technology and skills for FAIR practice*

- Data management planning requirements should be harmonised across stakeholders and supported over the entire research lifecycle
- Clarification is needed on eligible RDM and data sharing costs
- There is a need for clarity about which data repositories should seek certified status
- Support for making data FAIR has a positive influence on researchers' behaviour. Cooperation to develop and curate a shared set of discipline-specific guidance and training resources is necessary

Key findings in relation to *Embed and Sustain - incentives, metrics and investment*

- Funding bodies and publishers could strengthen their requirements in relation to data citation and provide clearer guidance on how to do this in a standardised way
- More equitable business models are needed to ensure that the costs of making and keeping data FAIR is split more equally between stakeholders
- Funding bodies and publishers should enforce the data sharing policies that they have put in place
- To encourage a FAIR ecosystem, the emphasis should be on providing incentives for good practice but penalties should be introduced for non-compliance where appropriate



7. Next Steps

Based on this initial landscape assessment and the forthcoming work of other projects funded under the INFRAEOSC-05-2018-2019 call⁷⁹, FAIRSF AIR will work to define a series of practical recommendations for policy enhancements to ensure that the components of the FAIR ecosystem are better aligned. A key aim for FAIRSF AIR is to identify and amplify existing policy recommendations wherever possible rather than to develop yet another set that duplicates what has already been done. To this end, we will work closely with colleagues in the INFRAEOSC 5 National Policies and Governance Task Force, the EOSC Life Turning FAIR into Reality working group which includes representatives from EOSC hub, EOSC pilot, FAIR4Health, and FAIR Plus. We will also seek to work closely with other initiatives such as FAIRsharing, the RDA Data Policy Standardisation and Implementation Interest Group as well as the EOSC Landscaping, FAIR, and Rules of Participation Working Groups.

Once the recommendations have been defined, FAIRSF AIR will provide practical support to the community to implement policy enhancements or to develop new policies. The instruments developed to carry out the initial landscape assessment will be reused including the FAIR data policy characterisation and related indicators. In this way, we will be able to provide an update on this initial snapshot of activity captured in 2019 to see if, and how, the policy landscape has changed over the life of the FAIRSF AIR project.

⁷⁹ <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/infraeosc-05-2018-2019>



Annex 1 - Policy Characterisation Table

The table below shows the policy elements that were considered in relation to the policies of stakeholders for Task 3.1.

Structure for policy analysis	Options
Title of policy	Short text
Link for policy	Short text
Organisation/body	Short text
Country/Countries if international	
Year policy was introduced	YYYY If no year is given, please state unspecified.
If date is given, what does date refer to (e.g., introduced, updated)	Short text
Status of policy	-Ready -In development -Deprecated -Uncertain -Other
Scheduled for review	-Yes -Unspecified -Other
If scheduled for review, when?	Short text
PID for policy	-Yes -Unspecified -Other
Machine readable	-Yes -Unspecified -Other
Subject area/discipline covered	-Health and Food -Social and Cultural Innovation -Physical Science and Engineering -Energy -Environment -Multiple areas/general research
Sub-discipline	Short text
Type of policy	-Research performing organisation (non HEI) -Institutional (HEI) -Funder -Ministry -Statute -Government -Repository -Journal/Publisher -Research Infrastructure



	<ul style="list-style-type: none"> -Society -Project level -Concordat -Roadmap -Informal/community approach -Other -Informal/community approach -Roadmap -Concordat
Scope of the policy	<ul style="list-style-type: none"> -Publications -Research data -Software -All of the above -Other
Definition of data provided	<ul style="list-style-type: none"> -Yes -Unspecified -Other
Policy references Open Access to research data	<ul style="list-style-type: none"> -Yes -Unspecified -Other
Policy requires data sharing	<ul style="list-style-type: none"> -Suggested -Required -Monitored -Not specified -Other
Policy requires metadata sharing	<ul style="list-style-type: none"> -Suggested -Required -Monitored -Not specified -Other
Exceptions to data sharing are allowed	<ul style="list-style-type: none"> -Yes -Unspecified -Other
If exceptions are allowed, justifications required	<ul style="list-style-type: none"> -Yes -Unspecified -Other
Includes expectation on data management plan (DMP)	<ul style="list-style-type: none"> -Recommended -Required -Assessed -Monitored -Other
Stage at which DMP should be produced	<ul style="list-style-type: none"> -Pre-award -Post-award -Other -NA
Reference to data protection (e.g., GDPR)	<ul style="list-style-type: none"> -Yes -Unspecified -Other



Reference to research integrity	-Yes -Unspecified -Other
Reference to related legislation (e.g., PSI directive)	-Yes -Unspecified -Other
Reference to specific standards or protocols	-Yes -Unspecified -Other
Requires a data availability statement	-Yes -Unspecified -Other
Allows embargo period	-Yes -Unspecified -Other
Specifies a preferred license type for outputs	-Yes -Unspecified -Other
Addresses Intellectual Property (IP) rights of researchers (i.e., not as an exception to data sharing but more about innovation).	-Yes -Unspecified -Other
Specifies use of researcher identifier (e.g., ORCID)	-Yes -Unspecified -Other
Includes expectation on data citation	-Yes -Unspecified -Other
Includes minimum length of data availability	-Yes -Unspecified -Other
Includes specific reference to preservation (mid to longer term)	-Yes -Unspecified -Other
Specifically references FAIR?	-Yes -Unspecified -Other
Justified costs associated with RDM and making data FAIR (explicitly or implicitly) are supported	-Yes -Unspecified -Other
References specific data repositories or scientific databases for deposit	-Yes -Unspecified -Other
References specific Research Infrastructures	-Yes -Unspecified -Other
Policy compliance is monitored	-Yes -Unspecified -Other



Sanctions applied to those found to be non-compliant with the policy	-Yes -Unspecified -Other
Guidance, training and/or support are provided to support compliance	-Yes -Unspecified -Other



Annex 2 - FAIR Policy Analysis and indicators

The table below shows a subset of the *Turning FAIR into Reality* priority and supporting actions that were considered in relation to the policies of stakeholders for Task 3.1.

D3.1 FAIR Policy Analysis and indicators			
Relevant FsF D3.1 stakeholder group	Please note that several actions are not in scope for D3.1 as they focus on activity beyond the scope of policies.	Indicators	Corresponding policy characterisation element
DEFINE Rec. 1 Define FAIR for implementation			
Polymakers	Action 1.3: The relationship between FAIR and Open should be clarified and well-articulated as the concepts are often wrongly conflated. FAIR does not mean Open. However, in the context of the EOSC and global drive towards Open Science, making FAIR data a reality should be supported by policies requiring appropriate Openness and protection, which can be expressed as 'as Open as possible, as closed as necessary'	Legitimate exceptions to data sharing are required to restrict access to data. Examples and guidance provided on these. Require metadata to be made open	<ul style="list-style-type: none"> • Policy requires data sharing • Exceptions to data sharing are allowed • If exceptions are allowed, justifications required • Policy requires metadata sharing
Rec. 3 Develop components of a FAIR ecosystem			
Polymakers	Action 3.2 By default, the FAIR ecosystem as a whole and each of its individual components should work for humans and for machines. Policies and DMPs should be machine-readable and actionable.	Policies themselves should be both human and machine readable	<ul style="list-style-type: none"> • PID for policy • Machine readable policy
Rec. 16: Apply FAIR broadly			
Polymakers	Action 16.1: Policies must assert	Policy specifies a broad	<ul style="list-style-type: none"> • Scope of the



	that the FAIR principles should be applied to research data, to metadata, to code, to DMPs and to other relevant digital objects, as well as to policies themselves.	range of outputs that are covered and expected to be managed and shared beyond data	<p>policy is clearly defined</p> <ul style="list-style-type: none"> Specifically references FAIR
Funders	Action 16.3: Guidelines for the implementation of FAIR in relation to research data, to metadata, to code, to DMPs and to other relevant digital objects should be developed and followed	Availability of guidance, examples and templates.	<ul style="list-style-type: none"> Guidance, training and/or support are provided to support compliance
Rec. 17: Align and harmonise FAIR and Open data policy			
Policymakers, Funders, Publishers	Action 17.1: The greatest potential reuse comes when data are both FAIR and Open. Steps should be taken to ensure coherence across data policy, emphasising both concepts and issuing collective statements of intent wherever possible.	<p>Legitimate exceptions to data sharing are required to restrict access to data. Examples and guidance provided on these.</p> <p>Require metadata to be made open</p>	<ul style="list-style-type: none"> Policy requires data sharing Exceptions to data sharing are allowed If exceptions are allowed, justifications required Policy requires metadata sharing
Policymakers	Action 17.3: Policies should be versioned, indexed and semantically annotated in a policy registry to enable broad reuse within the FAIR data ecosystem. Resources mandated by policies (e.g. consent forms) should be treated the same way	<p>Policy includes information on version and dates of applicability.</p> <p>Policy is available in registry such as FAIRsharing</p>	<ul style="list-style-type: none"> PID for policy Policy is machine readable Policy makes clear when policy was introduced Policy makes clear its status Policy review dates are specified
Funders,	Action 17.4: Data and other FAIR	Legitimate exceptions	<ul style="list-style-type: none"> Policy requires



Policymakers, Publishers	Digital Objects (e.g. code, models) that directly underpin, and provide evidence for, the findings articulated in published research must also be published unless there are legitimate reasons for protecting and restricting access	to data sharing are required to restrict access to data. Examples and guidance provided on these. Require metadata to be made open	<ul style="list-style-type: none"> • data sharing • Exceptions to data sharing are allowed • If exceptions are allowed, justifications required • Policy requires metadata sharing
Funders, Policymakers	Action 17.6: Policies should require an explicit and justified statement when (publicly-funded) data cannot be Open and a proportionate and discriminating course of action should be followed to ensure maximum appropriate data accessibility, rather than allowing a wholesale opt-out from the mandate for Open data.	Legitimate exceptions to data sharing are allowed. Policy requires exceptions to data sharing be justified Require metadata to be made open	<ul style="list-style-type: none"> • Exceptions to data sharing are allowed • If exceptions are allowed, justifications required • Policy requires metadata sharing
Institutions, Publishers	Action 17.8: Concrete and accessible guidance should be provided to researchers to find the optimal balance between sharing whilst also safeguarding privacy. There are many exemplars of good practice in providing managed access to sensitive data on which researchers can draw.	Availability of guidance, examples and templates.	<ul style="list-style-type: none"> • Guidance, training and/or support are provided to support compliance
<p>IMPLEMENT Rec. 5 Ensure Data Management via DMPs</p>			
Funders, Institutions, Publishers	Action 5.1: Research communities must be required, supported and incentivised to consider data management and appropriate data sharing as a	Policy includes requirement for DMP	<ul style="list-style-type: none"> • Policy includes expectation on DMP • Definition of data provided



	core part of all research activities. They should establish a Data Management Plan at project outset to consider the approach for creating, managing and sharing all research outputs (data, code, models, samples etc.)		in policy
Funders, Institutions	Action 5.2: Data Management Plans should be living documents that are implemented throughout the project. A lightweight data management and curation statement should be assessed at project proposal stage, including information on costs and the track record in FAIR. A sufficiently detailed DMP should be developed at project inception. Project end reports should include reporting against the DMP	Policy includes statement on updating DMP during life of project	<ul style="list-style-type: none"> • Policy indicates at which stage(s) the DMP should be produced
Funders	Action 5.3: Data Management Plans should be tailored to disciplinary needs to ensure that they become a useful tool for projects. Research communities should be inspired and empowered to provide input to the disciplinary aspects of DMPs and thereby to agree model approaches, exemplars and rubrics that help to embed FAIR data practices in different settings.	Availability of discipline-specific guidance, examples and templates.	<ul style="list-style-type: none"> • Discipline-specific guidance, training and/or support are provided to support compliance
Rec. 18: Cost data management			
Funders, Institutions	Action 18.1: Questions about the costs of data management, curation and publication should be included in all DMP	Eligible costs relating to RDM and data sharing are allowed and encouraged.	<ul style="list-style-type: none"> • Justified costs associated with RDM and making data



	<p>templates. Information from existing and completed projects should be used to retrospectively identify costs and develop examples and guidelines based on these. Funders, institutions and data services should collaborate on retrospective analysis, including the cost of long-term curation.</p>		<p>FAIR (explicitly or implicitly) are supported</p>
<p>Funders, Institutions</p>	<p>Action 18.3: Guidelines should be provided for researchers and reviewers to raise awareness of eligible costs and reinforce the view that data management, long term curation and data publication should be included in project proposals. Funders should collaborate to enhance guidance.</p>	<p>Guidance on what may be included and what is not eligible.</p>	<ul style="list-style-type: none"> • Guidance, training and/or support are provided to support compliance
<p>Rec. 20: Deposit in Trusted Digital Repositories</p>			
<p>Policymakers; Funders; Publishers</p>	<p>Action 20.1: Policy should require data deposit in certified repositories and specify support mechanisms (e.g. incentives, structural funding and/or funding for deposit fees, and training) to enable compliance.</p>	<p>Policy requires deposit within a certified repository.</p>	<ul style="list-style-type: none"> • References specific data repositories or scientific databases for deposit
<p>EMBED AND SUSTAIN</p> <p>Rec. 26: Support data citation and next generation metrics</p>			
<p>Publishers, Institutions</p>	<p>Action 26.2: Citation of data and other research outputs needs to be encouraged and supported - for example, by including sections in publishing templates that prompt researchers to reference materials, and providing citation guidelines when data, code or other</p>	<p>Provision of guidance, examples and templates.</p>	<ul style="list-style-type: none"> • Guidance, training and/or support are provided to support compliance



	outputs are accessed.		
Funders, Publishers, Institutions	26.4: A broader range of metrics must be developed to recognise contributions beyond publications and citation. These should recognise and reward Open and FAIR data practices.	Policy indicates that researcher assessment covers more than publications	<ul style="list-style-type: none"> • Policy references DORA, Leiden Manifesto or similar



Annex 3: Policies reviewed as part of policy characterisation

Below is a list of funders, publishers/journals and HEIs whose policies were reviewed as part of this landscape assessment. The resulting policy characterisation data is available for reuse from the FAIRsFAIR Zenodo community⁸⁰.

Funders:

- Austrian Science Fund (FWF) Open Access to Research Data accessed at <https://www.fwf.ac.at/en/research-funding/open-access-policy/open-access-to-research-data/>
- Netherlands Organisation for Scientific Research (NWO) Open (FAIR) Data accessed at <https://www.nwo.nl/en/policies/open+science/data+management>
- Research Foundation Flanders Data Management Plan (DMP) accessed at <https://www.fwo.be/en/the-fwo/organisation/data-management-plan/>
- European Commission DG RTD Unit Open Science Open Access and Data Management accessed at http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm
- Parkinson's UK Data sharing & preservation: Policy & guidelines accessed at <https://www.parkinsons.org.uk/sites/default/files/2017-06/Data%20sharing%20policy%20and%20guidelines%20May%202017.pdf>
- UK Research and Innovation Common principles on data policy accessed at <https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/>
- Swiss National Science Foundation Open Research Data accessed at http://www.snf.ch/en/theSNF/research-policies/open_research_data/Pages/default.aspx
- Kone Foundation Research funding; Ethical considerations, reuse of research materials and open access accessed at <https://koneensaatio.fi/en/grants/forgrantrecipients/>
- The Research Council of Norway Open access to research data accessed at <https://www.forskingsradet.no/en/about-the-research-council/forskningpolitikk/open-science/open-access-to-research-data/>
- Research Council Formas Good to know before you apply: Open access to research results and data accessed at <https://formas.se/en/start-page/applying-for-funding/how-it-works/good-to-know-before-you-apply.html>
- The French National Research Agency Open Science accessed at <https://anr.fr/en/anrs-role-in-research/values-and-commitments/open-science/>
- Fundação para a Ciência e a Tecnologia Data Availability Policy and other Results of FCT-funded R&D Projects accessed at https://www.fct.pt/documentos/PoliticaAcessoAberto_Dados.pdf

⁸⁰ D3.1_FAIRsFAIR_Policy Characterisation Data_20191122_v1.0 <https://doi.org/10.5281/zenodo.3550544>



- The Royal Society Conditions of Award accessed at <https://royalsociety.org/-/media/grants/schemes/Conditions-of-Award.pdf?la=en-GB&hash=73BEAFC806D50AFC020B11165953A3DF>
- Arcadia Fund Open Access and Digital Preservation Policy accessed at <https://www.arcadiafund.org.uk/about-peter-baldwin-lisbet-rausing/open-access-and-digital-preservation-policy/>
- Research Council of Lithuania Resolution on the approval of open scientific publications and data guidelines accessed at <https://www.e-tar.lt/portal/en/legalAct/dceeeb10e05711e59cc8b27b54efaf6e>
- Wellcome Trust Data, software and materials management and sharing policy accessed at <https://wellcome.ac.uk/funding/guidance/policy-data-software-materials-management-and-sharing>
- European Research Council (ERC) Open Access Guidelines accessed at https://erc.europa.eu/sites/default/files/document/file/ERC_Open_Access_Guidelines-revised_feb_2016.pdf

Publishers:

- Coalition for Publishing Data in the Earth and Space Sciences (COPDESS) Enabling FAIR Data Author Guidelines accessed at <https://copdess.org/enabling-fair-data-project/author-guidelines/>
- Taylor & Francis Data Sharing Policy Open and Fully FAIR accessed at <https://authorservices.taylorandfrancis.com/understanding-our-data-sharing-policies/>
- Springer Nature Research Data Policy Type 4 accessed at <https://www.springernature.com/gp/authors/research-data-policy/data-policy-types/12327096>
- Nature (Research) Policy on reporting standards and availability of data, materials, code and protocols accessed at <https://www.nature.com/authors/policies/availability.html#data>
- Wiley Data Sharing Policy - Mandates Data Sharing and Peer Reviews Data accessed at <https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/data-sharing-policy.html>
- Elsevier Research Data Guidelines Option E accessed at <https://www.elsevier.com/authors/author-resources/research-data/data-guidelines>
- Sage Research Data Sharing Policies - Requires data sharing accessed at <https://uk.sagepub.com/en-gb/eur/research-data-sharing-policies>
- Hindawi Research Data Policy accessed at <https://www.hindawi.com/research.data/>
- American Economic Association Data and Code Availability Policy accessed at <https://www.aeaweb.org/journals/policies/data-code>
- F1000 Research accessed at <https://f1000research.com/for-authors/data-guidelines>
- The Cryosphere Data Policy accessed at https://www.the-cryosphere.net/about/data_policy.html
- American Journal of Political Science Verification Policy accessed at <https://ajps.org/ajps-verification-policy/>



- Biostatistics Information for Authors accessed at https://academic.oup.com/biostatistics/pages/General_Instructions
- Public Library of Science Data Availability accessed at <https://journals.plos.org/plosone/s/data-availability>

Institutions:

- Free University Amsterdam, Faculty of Behaviour and Movement Sciences Research Data Management Policy accessed at https://www.fgb.vu.nl/nl/Images/data_policy_plan_fgb_v2_14012019_tcm263-903861.pdf
- Forschungsdaten-Leitlinie der Universität Göttingen accessed at <https://www.uni-goettingen.de/de/01-juli-2014-forschungsdaten-leitlinie-der-universitaet-goettingen-einschl-umg/488918.html>
- TU Munich Guidelines for Handling Research Data accessed at <https://www.it.tum.de/en/projects/research-data-management/>
- Medical University of Vienna Policy für Forschungsdaten-management accessed at https://www.meduniwien.ac.at/web/fileadmin/content/serviceeinrichtungen/itsc/it4science/Policy_fuer_Forschungsdaten-Management_v1.0.pdf
- KU Leuven Policy Plan Research Data accessed at <https://www.kuleuven.be/english/research/scholcomm/rdm/policy-plan-rdm-ku-leuven-2014>
- NUI Galway Research Data Management Policy accessed at [http://www.nuigalway.ie/media/staffsub-sites/researchoffice/files/Research-Data-Management-Policy-\(QA509\).pdf](http://www.nuigalway.ie/media/staffsub-sites/researchoffice/files/Research-Data-Management-Policy-(QA509).pdf)
- Leiden University Research Data Management Regulations accessed at https://www.bibliotheek.universiteitleiden.nl/binaries/content/assets/ul2ub/research--publish/research-data-management-regulations-leiden-university_def.pdf
- University of Helsinki Research Data Policy accessed at <https://www.helsinki.fi/en/research/research-environment/research-data/research-data-policy>
- University of Bath Research Data Policy accessed at <https://www.bath.ac.uk/corporate-information/research-data-policy/> and <https://www.bath.ac.uk/guides/research-data-policy-guidance/#section-1>
- TU Delft Research Data Framework Policy accessed at <https://d1rkab7tlqy5f1.cloudfront.net/Library/Themaportalen/RDM/researchdata-framework-policy.pdf>
- University of Cambridge Research Data Management Policy Framework accessed at <https://www.data.cam.ac.uk/university-policy>



Annex 4 - Open Consultation - Policy related questions

The FAIRsFAIR open consultation on Policy and Practice was made available via EU Survey⁸¹ from August 2 - September 27, 2019. The consultation was also made accessible from the FAIRsFAIR website⁸². A total of 106 responses were submitted by research support staff from across Europe and beyond. The open consultation included two matrix questions designed to provide insights into which stakeholder groups' policies have the most influence on researchers' practices and which policy factors positively or negatively influence researchers' behaviour. The full open consultation data is available for reuse from the FAIRsFAIR Zenodo community⁸³.

⁸¹ <https://ec.europa.eu/eusurvey/home/welcome>

⁸² <https://www.fairsfair.eu/fairsfair-open-consultation-fair-data-policies-and-practices>

⁸³ FAIRsFAIR Policy and Practice Survey 2019 data for D3.1_D3.2_D6.1 <https://doi.org/10.5281/zenodo.3550529>



FAIR related policies

In this section, we'd like to find out about the range of policies that may affect researchers' practices.

What, if any, are the main drivers for data policy development in the infrastructure organisation you work for?

250 character(s) maximum

In your view, how influential are the research data related policies of the following stakeholder groups on researchers' practice?

	Very influential	Neutral	Not influential	Don't know	Not applicable
Funding bodies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Governments or national ministries	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Research Infrastructures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data repositories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Departments or faculties	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Institutions/Research Performing Organisations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Community norms (e.g. scientific societies, associations)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Publishers or journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you selected 'other' above, please specify.

250 character(s) maximum

Which of the following policy factors do you believe would positively or negatively influence researchers to make data FAIR?

	Very positive	Quite positive	Neutral	Quite negative	Very negative	Don't know
Compliance with policy on making data FAIR is monitored	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Penalties for not making data FAIR are introduced	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rewards for making data FAIR are introduced	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Policy on FAIR is easy to understand and adhere to	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Support is provided for researchers to make data FAIR	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you selected 'other' above, please specify.

75 character(s) maximum