OPEN SCIENCE DALLA A ALLA Z
6 – DATI FAIR, DATI OPEN E DMP

https://doi.org/10.5281/zenodo.4483434

INRIM,
febbraio 2021

Elena Giglia
Università di Torino
elena.giglia@unito.it
@egiglia

Fatevi sentire...

https://www.menti.com/

codice mentimeter: **81 31 947**

Un invito

Bureau International des Poids et Mesures

– the intergovernmental organization through which Member States act together on matters related to measurement science and measurement standards.

Search

ABOUT US | WORLDWIDE METROLOGY | INTERNATIONAL EQUIVALENCE | SI UNITS | SERVICES

> You are here: BIPM Conference Centre > BIPM and co-organized workshops > The International System of Units (SI) in FAIR digital data

The International System of Units (SI) in FAIR digital data

22-26 February 2021 — **online**

Chair:
Prof. Joachim Ullrich (PTB) — Chair of CIPM-TG-DSI; CCU President

BIPM contact:
Andy Henson — BIPM

Organizing Committee:
UNF
INTI
CENAM
NPL
MSL
NIST
BIPM
NMIJ
PTB
KRISS

Virtual Workshop

The Task Group is now organizing a virtual workshop to bring together leading experts and groups in digitalization related to metrology and data science in order to exchange ideas about first steps to proceed in this direction aiming at agreeing on basic standards for a "Digital SI" framework.

The virtual workshop will be composed of two parts. Key-presentation sessions accompanied by plenary session are the first part that allow participants to get an insight on actions that are undertaken within various leading organizations regarding data formats, interoperability of data, FAIR principles and Artificial-Intelligence. Discussion are aiming to capture the high-level requirements for using data based on the SI across all organizations involved in, and depending on, the Quality Infrastructure. In the second part, workshop participants are kindly invited to join in panel discussions for truly international exploration of the technical approaches and promote collaboration on an international scale for a "Digital-SI" framework. The exchange on viable strategies and actions in the Qu
to define use-cases for digital measurement data based on the SI.

Registration

Register to the Workshop by using the following link:

Registration form

See also:

- CIPM Task Group on the Digital SI

Programme including registration

You're invited to provide the Expert Group of the CIPM "Digital SI" with further use-cases for interoperable data based on the SI:

Request for Use Cases

Draft - Grand Vision:
Transforming the International System of Units for a Digital World

Appendix - Glossary of Terms

https://www.bipm.org/en/conference-centre/bipm-workshops/digital-si/

Feb. 18, 2021

AISA
Associazione italiana per la promozione della scienza aperta

18 febbraio 2021: lettera aperta dell'AISA su vaccini anti-COVD-19 e brevetti

Lettera aperta al Presidente del Consiglio, Prof. Mario Draghi, alla Ministra dell'Università e della Ricerca, Prof. Cristina Messa, al Ministro della Salute, On. Dott. Roberto Speranza

Proposta per una vaccino anti-COVID-19 pubblico e aperto

1. Cooperazione invece di competizione

Notizie di stampa riferiscono di centri di ricerca italiani che stanno studiando l'elaborazione di un nuovo vaccino anti-COVID-19. Alcuni di questi centri appartengono a università o enti di ricerca pubblici che fanno capo allo Stato e sono finanziati dal contribuente. In una situazione pandemica, continuare a costringerli a reinventare la ruota in nome del culto del brevetto e della competizione intestina non è soltanto inefficiente: è letale. Occorre trovare il modo di farli cooperare, nell'interesse dell'Italia e dell'umanità.

2. Un vaccino libero

Qualche giorno fa il Ministero dell'università ha pubblicato il Programma Nazionale per la Ricerca 2021-2027, che contiene anche alcune dichiarazioni a favore della scienza aperta (§ 3.5.1). In un paese in cui il dire è sempre stato più facile del fare, università ed enti di ricerca potrebbero finalmente essere incoraggiati a collaborare per mettere a disposizione di tutti – di tutti gli italiani e di tutti gli esseri umani – le conoscenze e le tecnologie necessarie a produrre un nuovo vaccino, in una concorrenza leale, sui principi ancor prima che sui prodotti, con le multinazionali del farmaco che, grazie ai brevetti, hanno potuto costruire un sistema fondato su un'artificiale – e mortale – scarsità.

3. Finanziamento pubblico e donazioni civiche

Il finanziamento di questa ricerca potrebbe basarsi su fondi statali – per esempio del MUR e del Ministero della Salute – e sul contributo volontario dei cittadini, italiani e no.

Una simile iniziativa può sperare di avere successo solo se le sue regole di base sono chiare e pubbliche, perché pensate allo scopo di garantire, con una sorta di patto fra cittadinanza e ricerca, che da ciò che è pubblico e liberalmente donato si ottenga qualcosa che sia destinato a rimaner pubblico e liberalmente donato. Questa donazione da parte della rete della ricerca pubblica e dei cittadini che la finanziano direttamente e indirettamente non salverebbe soltanto le vite di pazienti ricchi e poveri,

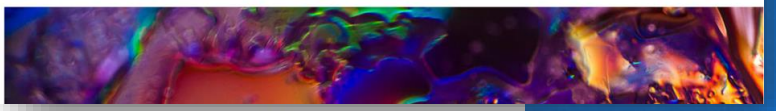
Scienza?

iai news — Feb. 1, 2021

Home | Philosophy | Science | Politics | Arts | Issue Archive

Science needs a radical overhaul

The lure of the illusion of discovery



Indeed, after 10 years as a journal editor, seeing how things work behind the scenes, I'm convinced that journals and the people who run them (editors, publishers, societies) are a bigger culprit for the spread of bad science than are individual researchers. Journals compete to be the most prestigious, but the race for prestige is not determined by who provides the best quality control. Instead, journals compete to publish the most attention-grabbing papers – the papers that are going to get the most clicks, media attention, and citations. In other words, journals are rewarding scientists for being flashy, for producing big, bold findings, and they are looking the other way when it comes to questions about whether those findings are reliable and whether the methods were rigorous. This reality is in stark contrast to the common myth about peer review – that journal-based peer review is a quality filter, and that the most prestigious journals have the most stringent filter. But the myth persists.

This misplaced faith in prestigious journals' peer review system is doing serious damage to science. Scientists continue to chase the reward of getting published in prestigious journals (because their livelihoods often depend on it,

2020

Room for everyone's talent

towards a new balance in the recognition and rewards of academics

This calls for a system of recognition and rewards of academics and research that:

1. Enables the diversification and vitalisation of career paths, thereby promoting excellence in each of the key areas;
2. Acknowledges the independence and individual qualities and ambitions of academics as well as recognising team performances;
3. Emphasises quality of work over quantitative results (such as number of publications);
4. Encourages all aspects of open science; and
5. Encourages high-quality academic leadership.

> Diversifying and vitalising career paths

We enable more diversity in career paths and profiles for academics.

> Achieving balance between individuals and the collective

We assess academics based on both their individual and their team performance.

> Focusing on quality

In our assessments of academic performance, we increasingly focus on quality, content and creativity.

> Stimulating open science

We encourage academics to share their research outcomes with society.

> Stimulating academic leadership

We stimulate good academic leadership at all levels.

TRIPLE: Team Spirit as the default approach to working in academia

2021

IMPACT

PROFESSIONAL PERFORMANCE

RESEARCH

EDUCATION

LEADERSHIP

TEAM

The TRIPLE model consists of six components: team spirit; research; impact; professional performance; leadership; and education. They describe the three domains where we generate output (research, professional performance and education), the impact they have on science and society, and leadership in academia that actively nurtures an environment in which they can flourish. The 'T' is deliberately put first. It emphasizes that, at Utrecht University, team spirit, characterized by contributing, cooperating and exchanging, is the default approach to working in academia. The order of the other abbreviations is random and mainly based on the construction of a word – TRIPLE – that can easily be remembered and that refers to something threefold (e.g., the three domains where we generate outcomes).

…r

Our remit is to give advice on indicators to foster the engagement of researchers with open science. Currently, researchers are usually not encouraged to engage in open knowledge practices. In career and research assessments open knowledge is usually not part of the performance requirements. The extra work involved may also be off-putting, especially in very competitive fields. And often it is simply unclear what "open science" should mean in practical terms. Therefore, simply taking away the current career and assessment criteria and replacing them with novel performance criteria that are oriented towards open science will not work. There are too many factors that hinder or promote open knowledge practices and they interact with each other. This creates a puzzle for the application of indicators in science and scholarship. On the one hand, there is the huge variety of scientific and scholarly practices. Universal indicators cannot address this dynamic variety. On the other hand, it is not practical to expect all scientific communities to have the technical expertise to develop and apply their own indicators in a responsible way. This explains why the alternative to universal indicators, creating large baskets of potential indicators that users can choose from as they see fit, is not advisable either.

2019

Indicator Frameworks for Fostering Open Knowledge Practices in Science and Scholarship

1. Infrastructure indicators oriented to the scientific system at national, international and disciplinary levels

The first suite of qualitative and quantitative indicators of the development of open knowledge infrastructures includes their creation, the growth of their numbers, the nature of their contribution, and their use and uptake by the research communities. This toolbox should build on the results of the Open Science Monitor and be linked to the European Open Science Cloud.

2. Indicators of open knowledge capabilities in research communities

The second toolbox of quantitative and qualitative indicators monitors the levels of open knowledge capabilities in the scientific and scholarly communities (including their support personnel). This toolbox will enable the identification of resource availability in specific communities, thus highlighting success cases as well as measures needed to redress the scarcity of capabilities in order to increase the inclusiveness, diversity and equity of the research system.

3. Indicators of pioneering open knowledge practices

The third toolbox consists of a suite of mainly qualitative, case-study based indicators, maintained and regularly updated on a public platform, that give a state-of-the-art overview of pioneering open knowledge practices. The database of case studies organized in the context of the UK Research Excellence Framework, maintained and openly accessible, might be an excellent starting point for such an international platform, provided that mechanisms are also built in for review and update on ongoing developments and initiatives. This platform may be maintained by a collective investment in the form of an annual fee by funders, publishers, and research performance organizations. Alternatively, it may be maintained in the context of an Annual Open Science Observatory (see below).

4. Individual level indicators for careers

The fourth toolbox consists of a suite of career-oriented qualitative and quantitative indicators, based on the principles of responsible metrics as formulated by the Metric Tide, the Leiden Manifesto for Research Metrics, and the DORA declaration. Again, it is not necessary to start from scratch, as several prototypes and basic design matrices for this toolboxes have already been proposed (eg. the ACUMEN portfolio, and the Open Science Career Evaluation Matrix). In relation to the use of

Indicator	OS Dimension indicated	Infrastructure	Capabilities	Champions	Career assessment	Data source	Strengths	Weaknesses	Potential
Types of data usage	A typology of different kinds of data usage	Y	N	EXEMPLARY CASES	N	Surveys among data users	Identifies developing demand for data	Must be done with a certain periodicity and with the same groups for comparability	Insight into actual data use
Accessibility of open data or code as % of all data or code produced by publicly funded projects.	Accessibility	Y	N	EXEMPLARY CASES	N	Researchers, Universities, funders	Encourages openness.	Privileges groups with money and competence to engage with research	Tracks open data infrastructure
Nr Funders requiring TOP Guidelines in publications	Adoption of TOP Guidelines	Y	N	EXEMPLARY CASES	N	Cos.io	Monitors OA among funders	Survey required	
Attitudes of researchers to data sharing	Attitudes of researchers to data sharing	N	Y	EXEMPLARY CASES	Y	Surveys	Qualifies types of data sharing behaviior; may identify best practices	Not clear categories yet exist	Inspiring examples may lead to new practices
Nr publications that can be tracked by the different altmetric sources (e.g.with a	Availability of altmetric data					Scopus, Web of	Monitors Open Data		
Nr Data Sharing Journals	Data sharing adoption	Y	N	EXEMPLARY CASES	Y	Vasilesky et al. 2017	Monitors Data Sharing	data sharing policies for practice	
Nr Open Data Repositories	Data sharing adoption	Y	Y	EXEMPLARY CASES	Y	Re3Data	Monitors Open Data		
Nr of repositories with open meta-data	Data sharing adoption	Y	Y	EXEMPLARY CASES	Y	OpenDoar	Monitors Open Data		
Nr institutes with data management infrastructure	Data sharing adoption	Y	Y	EXEMPLARY CASES	N	Surveys	Monitors Open Data		
Nr institutes with FAIR data policies	Data sharing adoption	Y	Y	EXEMPLARY CASES	N	Surveys	Monitors Open Data		
% of researchers that share data	Data sharing adoption	N	N	EXEMPLARY CASES	Y	Surveys	Tracks adoption of data sharing practices	Data sources for this indicator not available in all fields	
% Publications with data	Data sharing adoption	Y	N	EXEMPLARY CASES	Y	DataCite	Monitors data sharing practices	Does not check the quality of the data shared	Encourages data sharing

Parliamone

Backup su Google Drive (Google suite INRIM)Alcuni dati >10 anni irrecuperabili per obsolescenza softwareTutti i dati >20 anni irrecuperabili per obsolescenza hardware

Backup su circa 5 memorie diverse. Non mi è mai capitato di perderli. Non ho neanche avuto problemi di compatibilità (dati ASCII). A volte è stato un problema ricordarsi il significato delle colonne a causa di una insufficiente descrizione.

Li conservo nel mio pc o hard disk esterno. O/E NAS del laboratorio.Sì, è capitato, che dati non "backupati" si perdessero perché la macchina si era rotta...

Impareremo in questo modulo

1. Cosa significa FAIR nella ricerca quotidiana
2. come scrivere un DMP, che diventerà obbligatorio [supporto!]

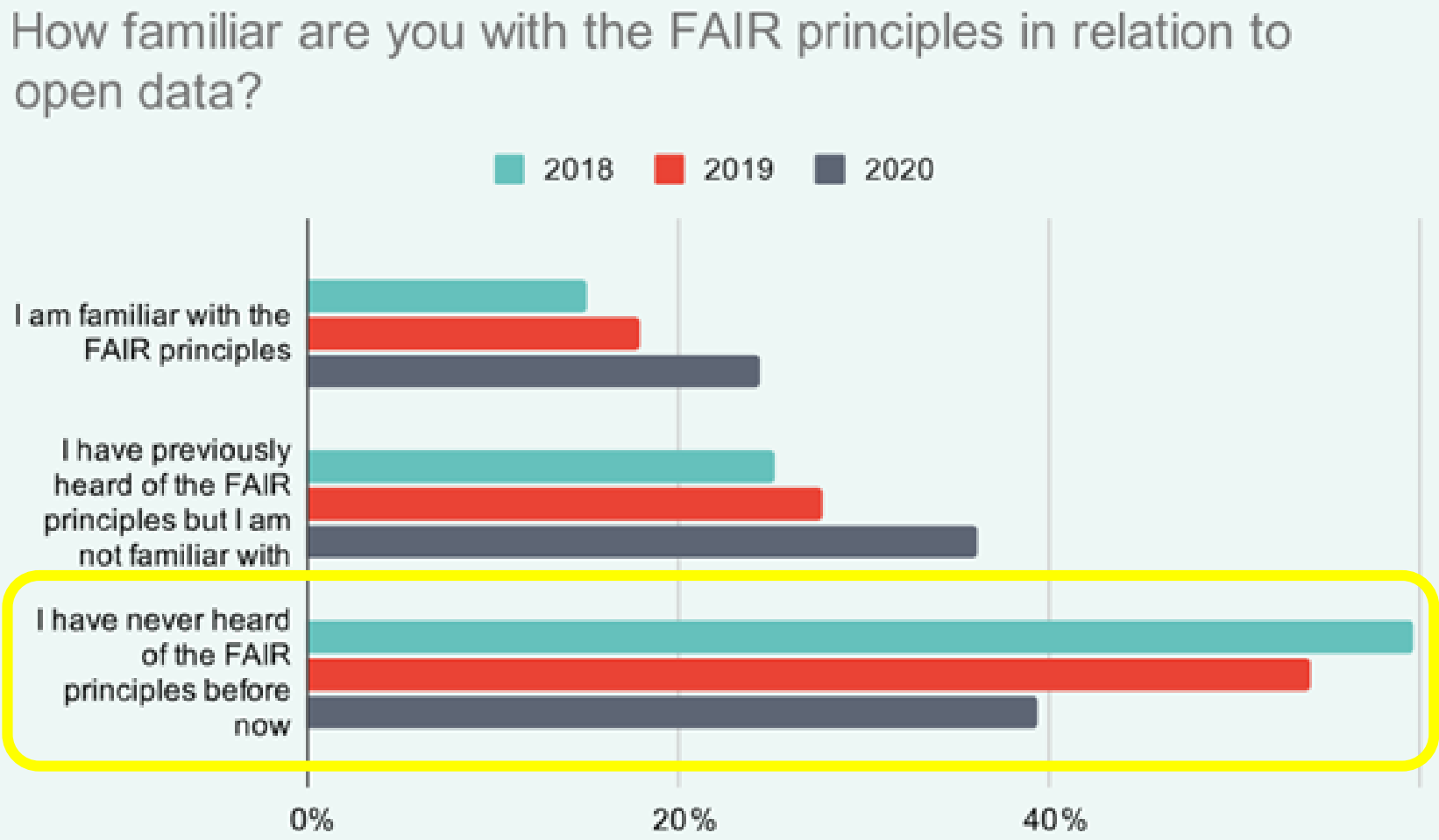
MESSAGGI CHIAVE

- FAIR è il futuro (se no si resta tagliati fuori)
- È più facile di quanto sembri, basta provarci
- DMP online/Data Wizard sono pronti all'uso, basta fare pratica

2. DATI FAIR

Dati FAIR???

Dec.2020



FAIR train – GoFAIR video

FARM DATA TRAIN

…FAIR SIGNIFICA
[anche e soprattutto per le macchine]


SCIENTIFIC DATA

We'd like to understand how you use our websites in order to imp

Open Access | Published

The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson, Michel Dumontier, [...] Barend Mons

FAIR guide, Nature, March 2016

FINDABLE

- IDENTIFICATIVI
- METADATI

ACCESSIBLE

- DOVE SONO CONSERVATI E A QUALI CONDIZIONI DI ACCESSO
- **NON SIGNIFICA «OPEN»**
- FORMATI APERTI

INTEROPERABLE

- STANDARDS
- ONTOLOGIE

REUSABLE

- LICENZE D'USO
- DOCUMENTAZIONE

IL TUTTO, LEGGIBILE DALLE MACCHINE

…FAIR SIGNIFICA ANCHE

I SINGOLI «BLOCCHI» DELLA RICERCA POSSONO ESSERE RICHIAMATI CON L'IDENTIFICATIVO, SENZA PERDERE TEMPO A RISCRIVERE (ES. UN PROTOCOLLO)

FINDABLE

- RIDUCE IL RISCHIO DI PERDITA
- FA RISPARMIARE TEMPO

ACCESSIBLE

- CONSERVA I DATI ORA E PER IL FUTURO

INTEROPERABLE

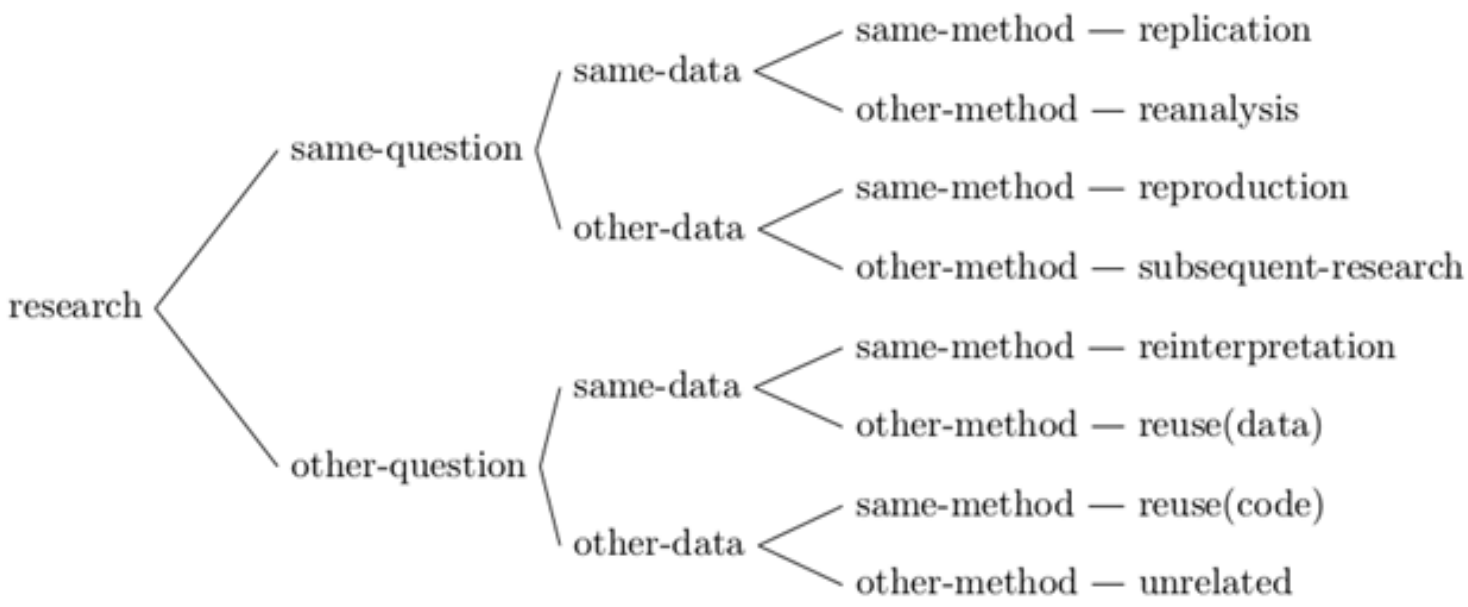
- FAVORISCE RICERCHE INTERDISCIPLINARI

REUSABLE

- EVITA SPRECO DI RISORSE
- CREA FIDUCIA

Focus

ASSICURARE IL RIUSO



Van de Sandt et al. 2019

DATA IS A RENEWABLE RESOURCE AS MUCH AS SUN AND WIND.
EVERY 18 MONTHS WE DOUBLE THE AMOUNT OF DATA WE PRODUCE.
INDUSTRIAL AND COMMERCIAL DATA, 85% OF WHICH IS NEVER USED.
THIS IS NOT SUSTAINABLE. WITHIN THOSE DATA, THERE ARE HIDDEN TREASURES AND UNTAPPED OPPORTUNITIES FOR BUSINESS AND SOCIETY [Von der Leyen 2020]

Obiettivo

DATI
FAIR BY DESIGN

FAIR principles

To be Findable:

F1. (meta)data are assigned a globally unique and eternally persistent identifier.
F2. data are described with rich metadata.
F3. (meta)data are registered or indexed in a searchable resource.
F4. metadata specify the data identifier.

TO BE ACCESSIBLE:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
A1.1 the protocol is open, free, and universally implementable.
A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
A2 metadata are accessible, even when the data are no longer available.

TO BE INTEROPERABLE:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles.
I3. (meta)data include qualified references to other (meta)data.

TO BE RE-USABLE:

R1. meta(data) have a plurality of accurate and relevant attributes.
R1.1. (meta)data are released with a clear and accessible data usage li
R1.2. (meta)data are associated with their provenance.
R1.3. (meta)data meet domain-relevant community standards.

«ACCESIBLE»
NON SIGNIFICA «OPEN».
SIGNIFICA SOLO SAPERE DOVE E A QUALI CONDIZIONI I DATI SONO ACCESSIBILI

FAIR/Open

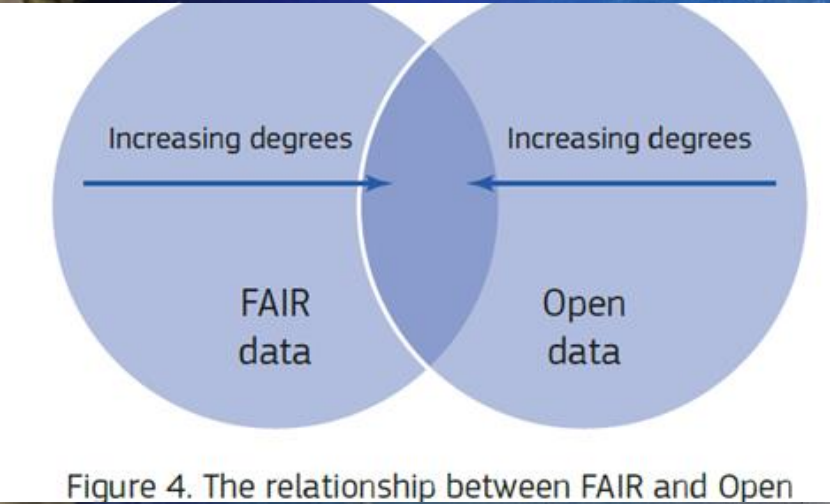
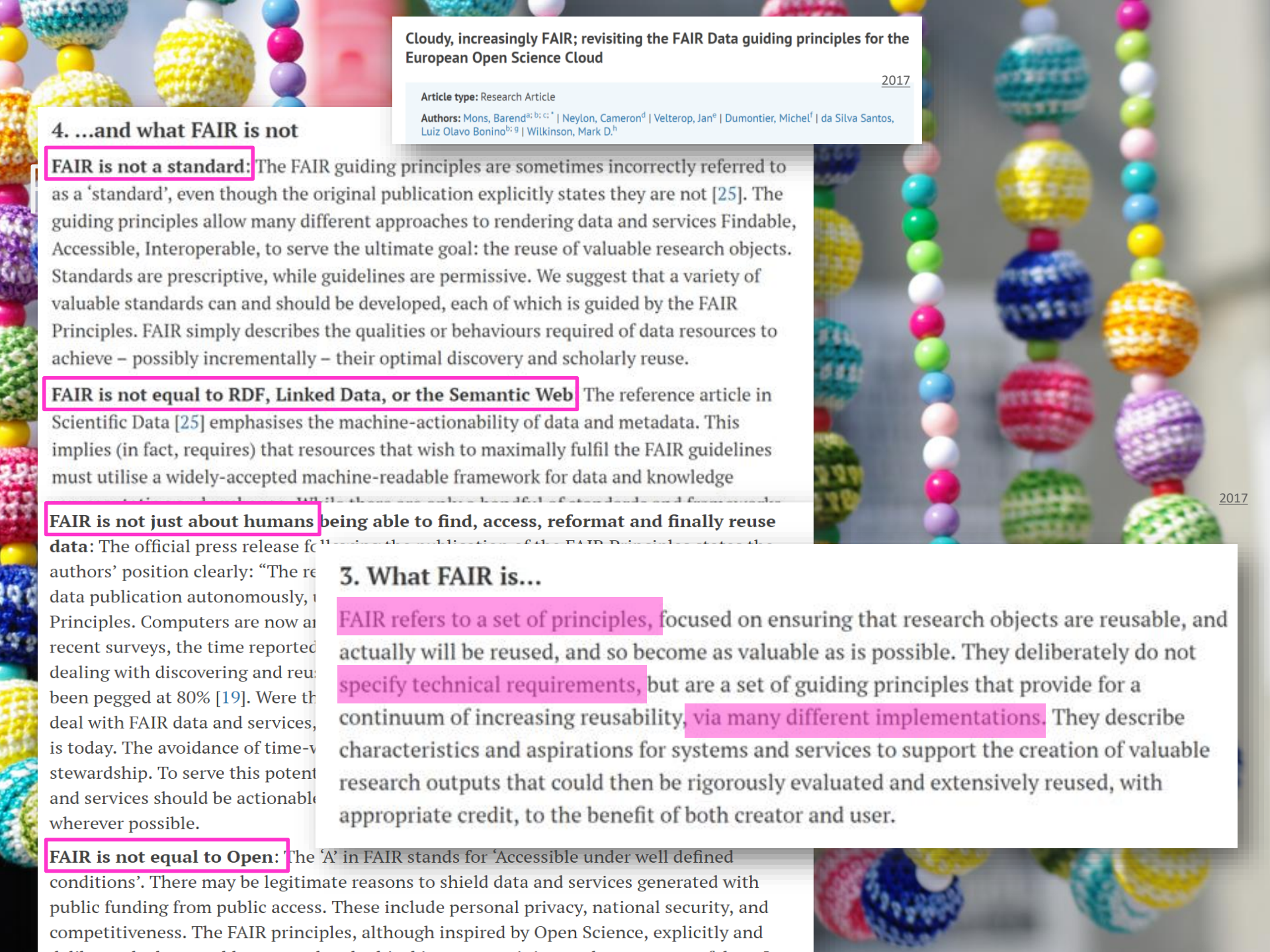


Figure 4. The relationship between FAIR and Open

Turning FAIR into reality, 2018

A TENDERE, I DUE INSIEMI SARANNO SEMPRE PIÙ SOVRAPPOSTI. MA EISTERANNO SEMPRE DATI PERFETTAMENTE FAIR CHE NON POSSONO ESSERE OPEN



Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud

2017

Article type: Research Article

Authors: Mons, Barend[a; b; c; *] | Neylon, Cameron[d] | Velterop, Jan[e] | Dumontier, Michel[f] | da Silva Santos, Luiz Olavo Bonino[b; g] | Wilkinson, Mark D.[h]

4. ...and what FAIR is not

FAIR is not a standard: The FAIR guiding principles are sometimes incorrectly referred to as a 'standard', even though the original publication explicitly states they are not [25]. The guiding principles allow many different approaches to rendering data and services Findable, Accessible, Interoperable, to serve the ultimate goal: the reuse of valuable research objects. Standards are prescriptive, while guidelines are permissive. We suggest that a variety of valuable standards can and should be developed, each of which is guided by the FAIR Principles. FAIR simply describes the qualities or behaviours required of data resources to achieve – possibly incrementally – their optimal discovery and scholarly reuse.

FAIR is not equal to RDF, Linked Data, or the Semantic Web The reference article in Scientific Data [25] emphasises the machine-actionability of data and metadata. This implies (in fact, requires) that resources that wish to maximally fulfil the FAIR guidelines must utilise a widely-accepted machine-readable framework for data and knowledge

FAIR is not just about humans being able to find, access, reformat and finally reuse data: The official press release fo
authors' position clearly: "The re
data publication autonomously,
Principles. Computers are now a
recent surveys, the time reported
dealing with discovering and reu
been pegged at 80% [19]. Were th
deal with FAIR data and services,
is today. The avoidance of time-
stewardship. To serve this potent
and services should be actionable
wherever possible.

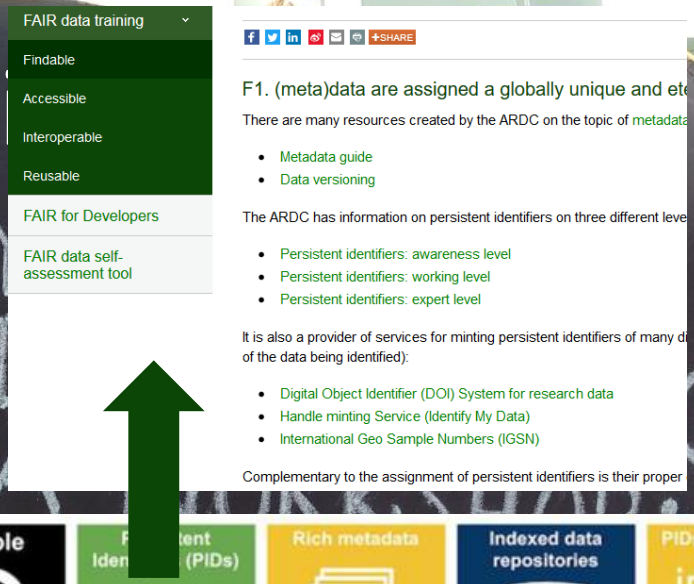
FAIR is not equal to Open: The 'A' in FAIR stands for 'Accessible under well defined conditions'. There may be legitimate reasons to shield data and services generated with public funding from public access. These include personal privacy, national security, and competitiveness. The FAIR principles, although inspired by Open Science, explicitly and

2017

3. What FAIR is...

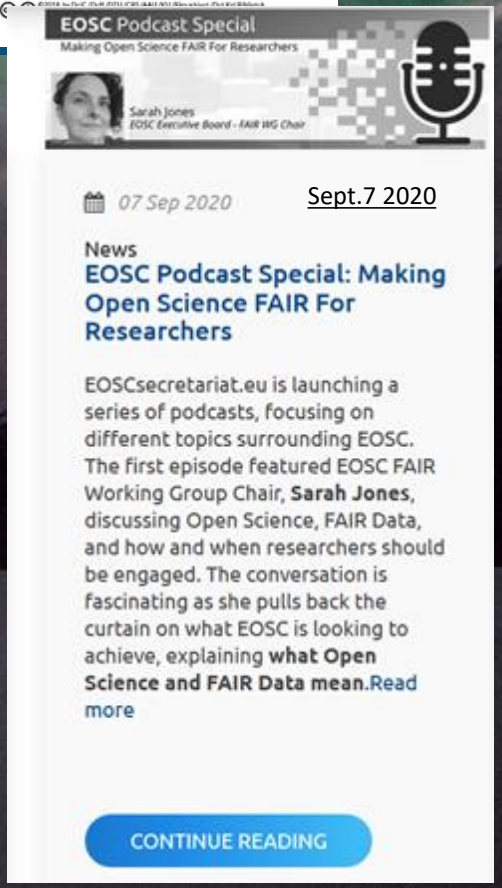
FAIR refers to a set of principles, focused on ensuring that research objects are reusable, and actually will be reused, and so become as valuable as is possible. They deliberately do not specify technical requirements, but are a set of guiding principles that provide for a continuum of increasing reusability, via many different implementations. They describe characteristics and aspirations for systems and services to support the creation of valuable research outputs that could then be rigorously evaluated and extensively reused, with appropriate credit, to the benefit of both creator and user.

FAIR in sintes



https://www.ands.org.au/working-with-data/fairdata/training

I miti su FAIR



FAIR myths

FAIR software

IOS Press Content Library

Data Science, June 12 2020

Towards FAIR principles for research software

Issue title: FAIR Data, Systems and Analysis

Guest editors: Paul Groth and Michel Dumontier

Article type: Position Paper

Authors: Lamprecht, Anna-Lena[a,*] | Garcia, Leyla[b] | Kuzak, Mateusz[c,d] | Martinez, Carlos[e] | Arcila, Ricardo[f] | Martin Del Pico, Eva[g] | Dominguez Del Angel, Victoria[h] | van de Sandt, Stephanie[i] | Ison, Jon[j] | Martinez, Paula Andrea[k] | McQuilton, Peter[l] | Valencia, Alfonso[m,n] | Harrow, Jennifer[o] | Psomopoulos, Fotis[p] | Gelpi, Josep Ll.[q,r] | Chue Hong, Neil[s,t] | Goble, Carole[u] | Capella-Gutierrez, Salvador[v,**]

2. Software is not data

Technically, software is a special kind of data. In computing, digital data (ultimately sequences of ones and zeros) are used to represent all information, including factual data as well as computer instructions. In the more abstract context of FAIR, software and data are regarded as different kinds of digital research objects next to each other. As such, they share particular characteristics that allow the
such as the possibility of having a Digit
license. However, as elaborated by Katz

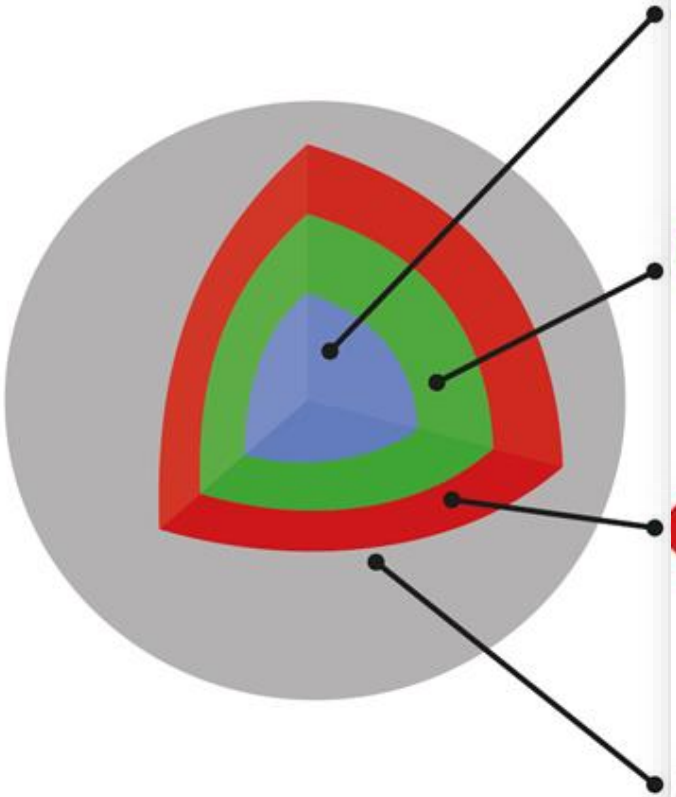
Quality aspects concerning the form of software can be considered as covered by FAIR, in particular by the interoperability and reusability principles. It is important to realise that unlike data, software is not static and can only be (re)used if it is sustainable and evolves along with the continuous development of the entire software ecosystem. The quality of its codebase is decisive for a software's ability to evolve sustainably. This characteristic is often also referred to as maintainability, and

Summary of the proposed FAIR principles for research software and how they relate to the FAIR Guiding Principles for data. It is indicated whether a given FAIR data principle has been simply rephrased to adjust it to software, extended to cover a broader scope, reinterpreted to match the different context, discarded as it does not apply, or newly proposed as it only applies for research software

	FAIR for data	FAIR for software	Operation
F1	(Meta)data are assigned a globally unique and persistent identifier.	Software and its associated metadata have a global, unique and persistent identifier for each released version.	Rephrased
F2	Data are described with rich metadata.	Software is described with rich metadata.	Rephrased
F3	Metadata clearly and explicitly include the identifier of the data it describes.	Metadata clearly and explicitly include identifiers for all the versions of the software it describes.	Rephrased and extended
F4	(Meta)data are registered or indexed in a searchable	Software and its associated metadata are included in a searchable software	Rephrased
I2	(Meta)data use vocabularies that follow FAIR principles.	–	Reinterpreted, extended and split
I2S.1	–	Software and its associated metadata are formally described using controlled vocabularies that follow the FAIR principles.	Reinterpreted, extended and split
I2S.2	–	Software use and produce data in types and formats that are formally described using controlled vocabularies that follow the FAIR principles.	Reinterpreted, extended and split
I3	(Meta)data include qualified references to other (meta)data.	–	Discarded
I4S	–	Software dependencies are documented and mechanisms to access them exist.	Newly proposed


Nov. 20, 2018
European Commission
Final Report and Action Plan from the European Commission Expert Group on FAIR Data
TURNING FAIR INTO
Define
Implement
Embed and sustain
Concepts for FAIR implementation
Rec. 1: Define FAIR for implementation
Rec. 2: Implement a Model for FAIR Digital Objects
Rec. 3: Develop components of a FAIR ecosystem
Rec. 16: Apply FAIR broadly
Rec. 17: Align and harmonise FAIR and Open data policy
FAIR culture
Rec. 4: Develop Interoperability frameworks
Rec. 5: Ensure data management via DMPs
Rec. 6: Recognise & reward FAIR data & stewardship
Rec. 18: Cost data management
Rec. 19: Select and prioritise FAIR digital objects
Rec. 20: Deposit in Trusted Digital Repositories
Rec. 21: Incentivise reuse of FAIR outputs
FAIR ecosystem
Rec. 7: Support semantic technologies
Rec. 8: Facilitate automated processing
Rec. 9: Certify FAIR services
Rec. 22: Use information held in DMPs
Rec. 23: Develop components to meet research needs
Rec. 24: Incentivise research infrastructures to support FAIR data
Skills for FAIR
Rec. 10: Professionalise data science & stewardship roles
Rec. 11: Implement curriculum frameworks and training
Above line = priority recommendations
Below line = supporting recommendations
Incentives and metrics for FAIR data and services
Rec. 12: Develop metrics for FAIR Digital Objects
Rec. 13: Develop metrics to certify FAIR services
Rec. 25: Implement and monitor metrics
Rec. 26: Support data citation and next generation metrics
Investment in FAIR
Rec. 14: Provide strategic and coordinated funding
Rec. 15: Provide sustainable funding
Rec. 27: Open EOSC to all providers but ensure services are FAIR

Oggetto FAIR ideale



DIGITAL OBJECT

Data, code and other research outputs

At its most basic level, data or code is a bitstream or binary sequence. For this to have meaning and to be FAIR, it needs to be represented in standard formats and be accompanied by Persistent Identifiers (PIDs), metadata and documentation. These layers of meaning enrich the object and enable reuse.

IDENTIFIERS

Persistent and unique (PIDs)

Digital Objects should be assigned a unique and persistent identifier such as a DOI or URN. This enables stable links to the object and support citation and reuse to be tracked. Identifiers should also be applied to other related concepts such as the data authors (ORCIDs), projects (RAIDs), funders and associated research resources (RRIDs).

STANDARDS & CODE

Open, documented formats

Digital Objects should be represented in common and ideally open file formats. This enables others to reuse them as the format is in widespread use and software is available to read the files. Open and well-documented formats are easier to preserve. Data also need to be accompanied by the code use to process and analyse the data.

METADATA

Contextual documentation

In order for Digital Objects to be assessable and reusable, they should be accompanied by sufficient metadata and documentation. Basic metadata will enable data discovery, but much richer information and provenance is required to understand how, why, when and by whom the objects were created. To enable the broadest reuse, they should be accompanied by a plurality of relevant attributes and a clear and accessible usage license.

ECD, 2015

FAIR: technology VS domain

Technical infrastructure (generic operations)
Data/metadata (domain-specific content)

Box 2 | The FAIR Guiding Principles https://www.nature.com/articles/sdata201618

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

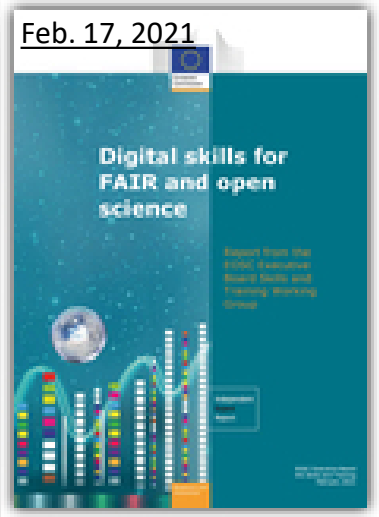
To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

E.Schultes, 2019

Digital skills for FAIR and open

Feb. 17, 2021

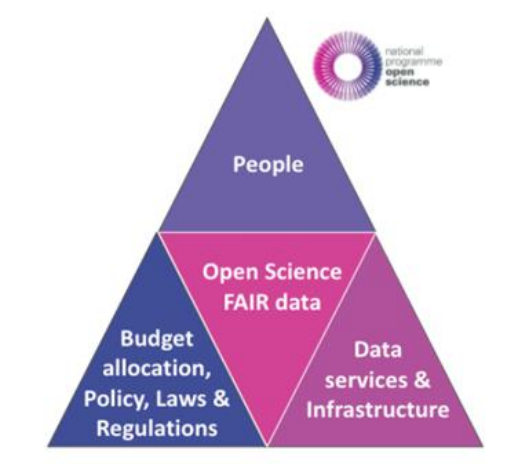


Digital s
science

Report section	Policy makers & funders	Universities & research organisations	Competence centres	EOSC Association	EOSC projects
2	Understand the importance of addressing gaps in provision of digital skills for FAIR and open science				
3	Identify actors/roles that need to be supported in the EOSC ecosystem	Consider diversity of employment and training needs of staff	Identify knowledge to be taught and skills to be learned	Understand range of roles that initiatives must address	Ensure projects consider appropriate roles for both their staff and users
4	Consider how to support competence centres and encourage collaboration	Consider how to establish competence centres	Learn about governance, business models and alignment approaches of other competence centres	Consider how to encourage and support competence centres approach	Ensure projects consider FAIR and open science training
5	Recognise the need for investment in hubs for training resources	Know how to provide and access learning and training resources	Learn how to maximise interoperability, FAIRness and usage of learning and training resources	Facilitate interoperability and discovery of learning and training resources	Identify and engage key stakeholders, and build learning and training catalogues according to their goals and actions
6	Identify actions to advance national policies on FAIR and open science	Consider how to encourage and/or implement national approaches	Understand the importance of aligning with national programmes and policies	Identify ways to increase international alignment	Not relevant

Feb. 17, 2021

Professiona ta st



Feb. 1, 2021

nationaal programma open science

Professionalising data stewardship in the Netherlands: competences, training and education

Dutch roadmap towards national implementation of FAIR data stewardship

Preamble: The urgency of a coherent approach towards professionalising data stewardship

"Invest 5% of research funds in ensuring data are reusable. It is irresponsible to support research but not data stewardship", said Barend Mons recently in a Nature article[1]. "Students in PhD programmes spend up to 80% of their time on 'data munging', fixing formatting and minor mistakes to make data suitable for analysis — wasting time and talent. With 400 such students,

Data stewardship is a new profession. To ensure a data steward's position within a research institute, the profession should become part of HR planning of the organisation and include a vision on career development. In addition, research institutes should collaborate with policy makers and research-funding organisations to realise a common approach to recognise and reward data stewards as full members in research groups, and to make sufficient budgets available to maintain data steward positions in the long run.

The interplay between these three aspects determines whether researchers in a particular research setting are sufficiently equipped and supported to be able to perform data stewardship in the context of Open Science. By having a clear view on each of these aspects, it will be finally possible to answer the question that each research-performing institute should ask itself: **how many data stewards do we need where in the organisation with which competences?**

FAIR è graduale

AS FAIR AS POSSIBLE

EOSCFAIR
Executive Board Working Group

Feb. 2020

Interim recommendations on FAIR Metrics for EOSC
February 2020
Draft for consultation

Françoise Genova, Observatoire Astronomique de Strasbourg, ORCID 0000-0002-6318-5028
Jan Magnus Aronsen, University of Oslo, Norway, 0000-0003-2593-1744
Oya Beyan,Fraunhofer FIT, Germany, 0000-0001-7611-3501
Natalie Harrower, Digital Repository of Ireland, 0000-0002-7487-4881
András Holl, Hungarian Academy of Sciences, 0000-0002-6873-3425
Rob W.W. Hooft, Dutch Techcentre for Life Sciences, 0000-0001-6825-9439
Leslie McIntosh, Research Data Alliance, 0000-0002-3507-7468
Pedro Principe, University of Minho, 0000-0002-8588-4196
Kostas Repanas, European Commission, 0000-0002-7848-2834
Ana Slavec, InnoRenew CoE, 0000-0002-0171-2144
Sarah Jones, Digital Curation Centre, 0000-0002-5094-7126

Edited by Françoise Genova & Sarah Jones, Co-Chairs of the EOSC FAIR Working Group

Inclusiveness: consider FAIR as a process

If FAIR is not seen as a continuum, we risk losing communities who are not well advanced in sharing their data in a FAIR way, as well as advanced communities for whom the effort to attain optional indicators doesn't outweigh the effort required. In addition to avoiding "mandatory" criteria, using multi-step maturity scales to measure the FAIRness level of a resource, instead of a yes/no evaluation for each criterion, would provide an inclusive system, and a way to set up goals and measure progress. A bar representation of the levels of compliance of all the criteria would provide a visualisation of the level of FAIRness of a given resource.

… sfumature di FAIR

Findable

Does the dataset have any identifiers assigned? — No identifier

Is the dataset identifier included in all metadata records/files describing the data? — No

How is the data described with metadata? — The data is not described

What type of repository or registry is the metadata record in? — The data is not described in any repository

Accessible

How accessible is the data? — No access to data or metadata

Is the data available online without requiring specialised protocols or tools once access has been approved? — No access to data

Will the metadata record be available even if the data is no longer available? — Unsure

UTILISSIMO PER PORSI LE DOMANDE GIUSTE… MA SOGGETTIVO…

https://www.ands-nectar-rds.org.au/fair-tool

FAIR aware

- DOMANDE
- VERIFICA INTENZIONI
- VERIFICA CONOSCENZA
- FORNISCE INFORMAZIONI

FAIR Aware

Data Archiving and Networked Services DANS

FAIRaware

Let's assume you have research data almost ready for uploading to a repository: do you already know how you and the repository can work together to make the data as findable, accessible, interoperable and reusable (FAIR) as possible? By guiding you through the assessment process, the FAIR-Aware tool can help you to better understand the FAIR Principles and how making data FAIR can increase the potential value and impact of your data.

FAIR-Aware is an disciplinary-agnostic online tool developed by the FAIRsFAIR project. Different scientific communities can adapt it to their own use. You should, however, have a target dataset in mind to be able to answer the questions and complete the assessment.

1. Are you aware that a dataset should be assigned a globally unique and persistent identifier when deposited with a data repository?

Selected datasets should be assigned a globally unique and persistent identifier (PID) so they can be located unambiguously by humans or machines on the web. Persistent identifiers are maintained and governed so that they remain stable and direct the users to the same relevant object consistently over time. Examples of PIDs include Digital Object Identifier (DOI), the Handle System, identifiers.org, w3id.org and Archival Resource Key (ARK).

Identifiers are normally assigned by data repositories (or other service providers) when data and/or metadata are made available through their services. Repositories ensure that the identifier continues to point to the same data or metadata, according to the specified access terms and conditions. For example, you can search for data repositories providing DOIs on registries such as Re3data or FAIRsharing (see related databases).

It is worth noting here that not all data you produce during your research will need a PID. In general, those that underpin published findings or have longer term value are worth assigning a PID. If in doubt about which data should be allocated a PID, speak to your local research data management support team.

Want to know more?

F

FINDABLE

1. Are you aware that a dataset should be a globally unique and persistent identifier whe with a data repository?

2. Are you aware that when you deposit a d repository, you will need to provide some de as discovery metadata) in order to make the understandable and reusable to others?

3. Are you aware that the repository providing access to your dataset should make the metadata describing your datasets available in a format readable by machines as well as humans?

Yes
No

To what degree do you intend to comply with this?

Very likely 5 4 3 2 1 Very unlikely

FAIR aware

FAIRaware

FAIR Aware

Data Archiving and Networked Services DANS

Let's assume you have research data almost ready for uploading to a repository: do you already know how you and the repository can work together to make the data as findable, accessible, interoperable and reusable (FAIR) as possible? By guiding you through the assessment process, the FAIR-Aware tool can help you to better understand the FAIR Principles and how making data FAIR can increase the potential value and impact of your data.

FAIR-Aware is an disciplinary-agnostic online tool developed by the FAIRsFAIR project. Different scientific communities can adapt it to their own use. You should, however, have a target dataset in mind to be able to answer the questions and complete the assessment.

ACCESSIBLE

4. Are you aware that access to your dataset may need to be controlled and that metadata should include licence information under which the data can be reused? ○ Yes ○ No

5. Are you aware that metadata should remain available over time, even if the data is no longer accessible? ○ Yes ○ No

INTEROPERABLE

6. Are you aware that the metadata describing your datasets should use controlled vocabularies? ○ Yes ○ No

REUSABLE

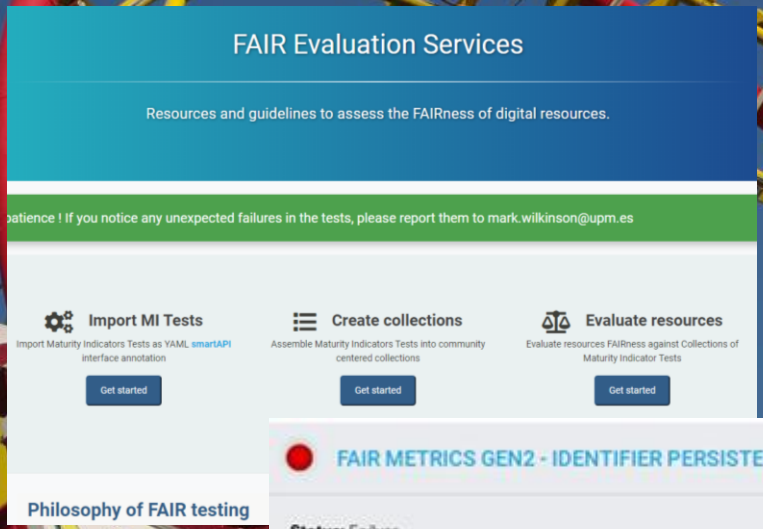
7. Are you aware that provenance information about the collection and/or generation of data should be included in the metadata? ○ Yes ○ No

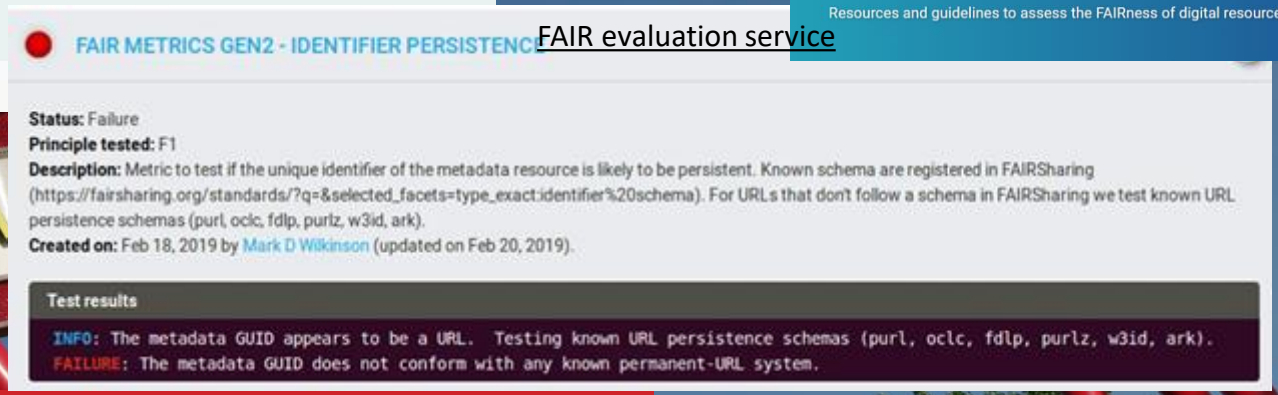
8. Are you aware that metadata describing your data should follow the specifications of a community-endorsed standard? ◉ Yes ○ No

9. Are you aware that data should be deposited preferably in a file format that is open – to support reuse – and supported by the repository for long-term preservation? ○ Yes ◉ No

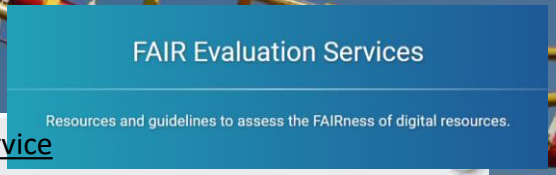
10. Are you aware that maintaining your dataset FAIR over time requires professional data curation and preservation? ○ Yes ○ No

FAIR maturity evaluator





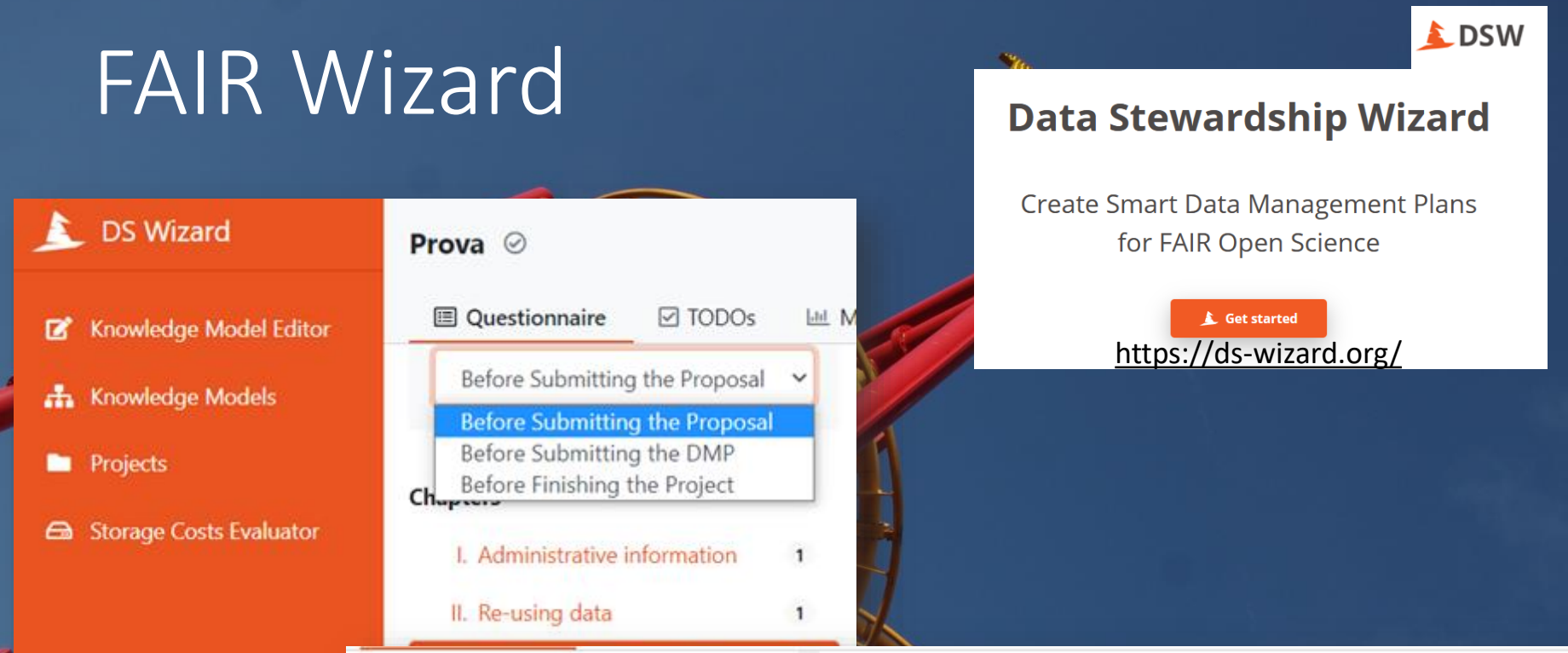
Sept. 20, 2019

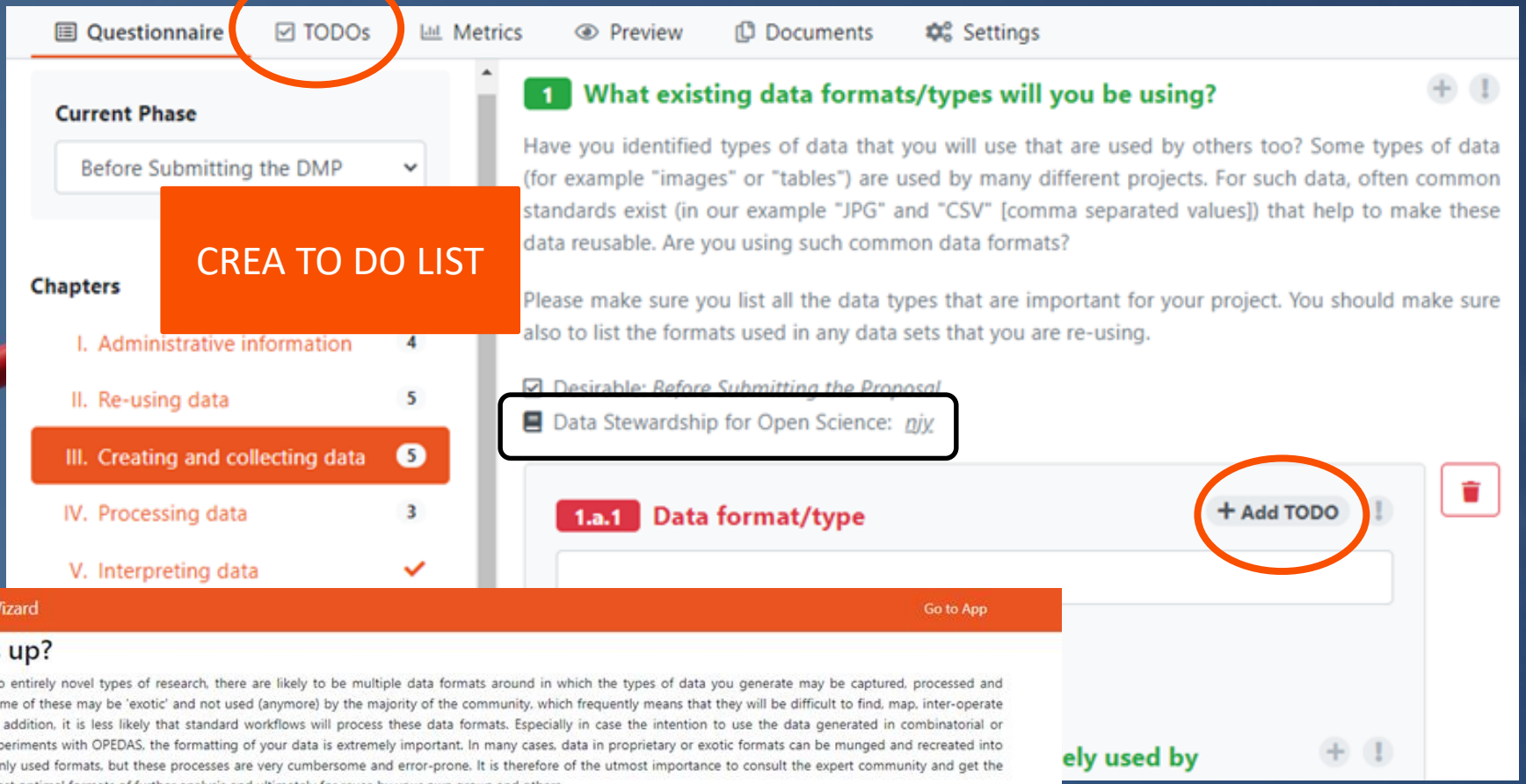


FAIR evaluation service

- OGGETTIVO
- LEGGIBILE DALLE MACCHINE – COME I DATI FAIR

FAIR Wizard


DSW
Data Stewardship Wizard
Create Smart Data Management Plans for FAIR Open Science
Get started
https://ds-wizard.org/
DS Wizard
Knowledge Model Editor
Knowledge Models
Projects
Storage Costs Evaluator
Help
Elena Giglia
Prova
Questionnaire
TODOs
Before Submitting the Proposal
Before Submitting the Proposal
Before Submitting the DMP
Before Finishing the Project
I. Administrative information 1
II. Re-using data 1
Current Phase
Before Submitting the DMP
Chapters
I. Administrative information 4
II. Re-using data 2
III. Creating and collecting data 5
IV. Processing data 3
V. Interpreting data
1 Is there any pre-existing data?
Are there any data sets available in the world that are relevant to your planned research?
Desirable: Before Submitting the Proposal
Data Stewardship for Open Science: ata
External links: Google dataset search, Datacite Search
a. No
b. Yes
Clear answer
LINK ESTERNI
SI CREA IN BASE ALLE RISPOSTE



DS Wizard — Go to App

What's up?

Unless you do entirely novel types of research, there are likely to be multiple data formats around in which the types of data you generate may be captured, processed and formatted. Some of these may be 'exotic' and not used (anymore) by the majority of the community, which frequently means that they will be difficult to find, map, inter-operate and reuse. In addition, it is less likely that standard workflows will process these data formats. Especially in case the intention to use the data generated in combinatorial or integrated experiments with OPEDAS, the formatting of your data is extremely important. In many cases, data in proprietary or exotic formats can be munged and recreated into more commonly used formats, but these processes are very cumbersome and error-prone. It is therefore of the utmost importance to consult the expert community and get the data in the most optimal formats of further analysis and ultimately for reuse by your own group and others.

Do

- Always use community-compliant, supported and sustainable data formats whenever possible.
- Turn to experts to tell you what are the best formats to use for the particular data types you will create.
- Ensure you are prepared to answer questions on the use of the data (for instance, which workflows will they be subjected to).
- Choose the formats with the richest expression possibility. It is easier to leave things blank then extending a poor data format later.

Don't

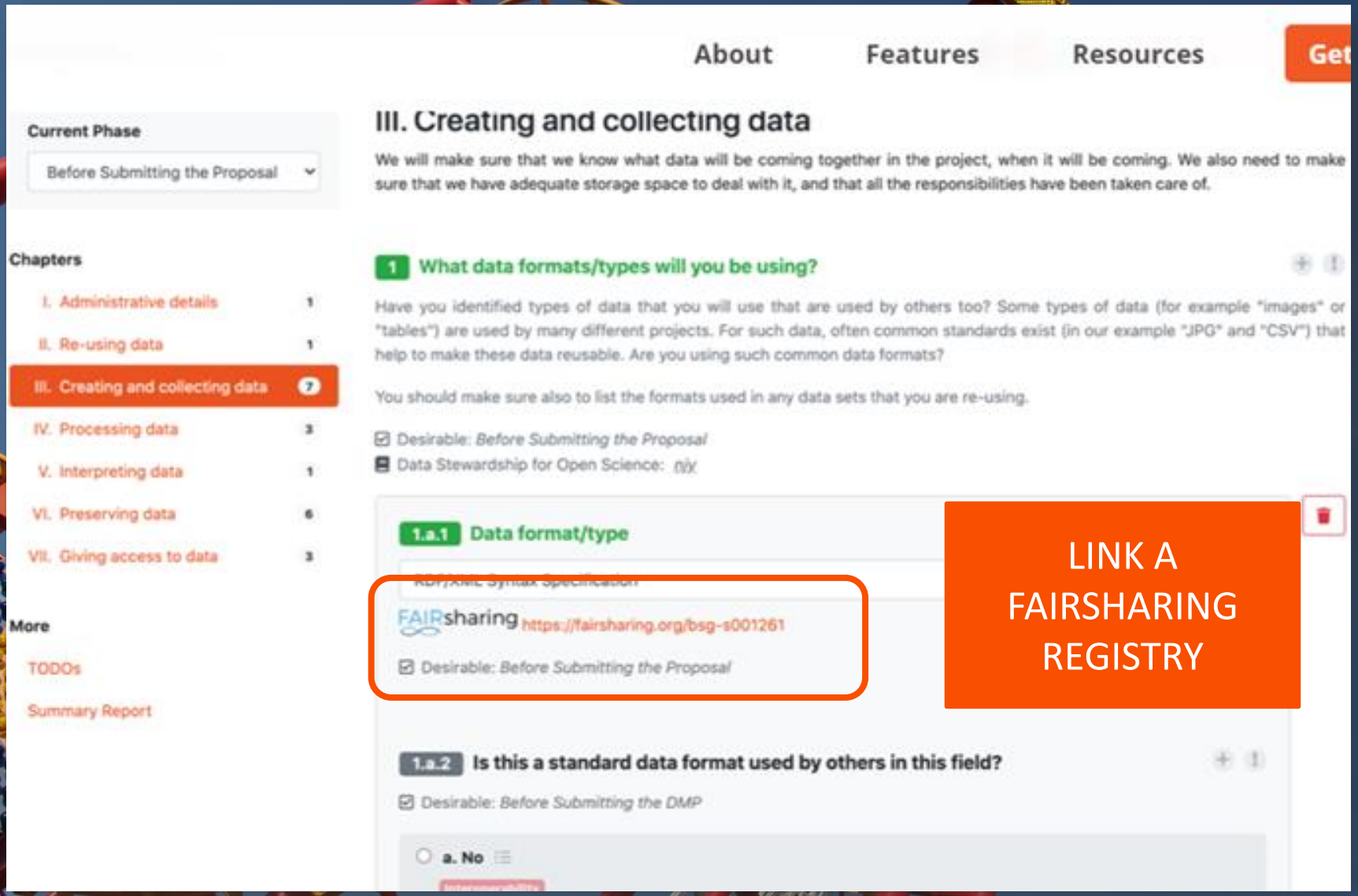
- Assume that your data is so unique that it needs an entirely new format.
- Think that a spreadsheet with free text labels or your locally developed database is the best way to store and reuse your data.
- Format and store data in any format without keeping rich and relevant metadata and provenance.
- Throw away the original data unless you are absolutely sure that storing them has no further added value, for example for review of experimental and analytical procedures. Not having certain pre-formatted data available may actually preclude the publication, reuse and citation of your (original) data by others and might also jeopardise the publication of accompanying articles.

Links

- DS Question GitHub resources repository: njy

APRE IL CAPITOLO DEL LIBRO DI BAREND MONS CORRISPONDENTE ALL'OGGETTO DELLA DOMANDA

FAIR Wizard


About
Features
Resources
Get
Current Phase
Before Submitting the Proposal
III. Creating and collecting data
We will make sure that we know what data will be coming together in the project, when it will be coming. We also need to make sure that we have adequate storage space to deal with it, and that all the responsibilities have been taken care of.
Chapters
I. Administrative details 1
II. Re-using data 1
III. Creating and collecting data 7
IV. Processing data 3
V. Interpreting data 1
VI. Preserving data 6
VII. Giving access to data 3
More
TODOs
Summary Report
1 What data formats/types will you be using?
Have you identified types of data that you will use that are used by others too? Some types of data (for example "images" or "tables") are used by many different projects. For such data, often common standards exist (in our example "JPG" and "CSV") that help to make these data reusable. Are you using such common data formats?
You should make sure also to list the formats used in any data sets that you are re-using.
Desirable: Before Submitting the Proposal
Data Stewardship for Open Science: n/a
1.a.1 Data format/type
FAIRsharing https://fairsharing.org/bsg-s001261
Desirable: Before Submitting the Proposal
1.a.2 Is this a standard data format used by others in this field?
Desirable: Before Submitting the DMP
a. No
LINK A FAIRSHARING REGISTRY

FAIR tools

The table also shows how the individual tool fits in with the FAIR dimensions. You can use the overview to find the right tools to make your research data (more) FAIR - depending on the data you have and what you want to do with it.

Availability	Discipline	FAIR dimension	Phases in research life cycle	Service Name	Description
International	Generic (Tabular data)	__I_	Process/Analyze	OpenRefine	OpenRefine is a standalone open source desktop application for data cleanup and transformation to other formats (i.e. data wrangling)
International	Generic	FA(I)(R)	Publish/Disseminate, Archive, Discover & Re-use, Release, Preserve	Zenodo	Zenodo is a general-purpose open access research data repository, hosted by CERN (Switzerland) that provides a place for researchers to deposit datasets. Researchers in any subject area to are able to upload files up to 50 GB. It has an integration with GitHub to make code hosted in GitHub citable. Support: zenodo.org/support
International	Generic - Highly recognized in Social Sciences	FA(I)(R)	Publish/Disseminate, Archive, Discover & Re-use, Release, Preserve	Harvard Dataverse	Dataverse is a data repository that is widely used within the Social Sciences. Researchers can login with their institutional credentials via WAYF. Data can be made findable by applying discipline-specific metadata schemes and digital object identifiers (DOIs). Data is
International	Generic	_A(I)(R)	Publish/Disseminate, Archive	B2SHARE	B2S and trustworthy way for researchers, scientific communities and citizen scientists to store and share small-scale research data from diverse contexts
International	Generic	F___	Discover & Re-use	B2FIND	B2FIND is a discovery service based on metadata steadily harvested from research data collections from EUDAT data centres and other repositories.
International	Generic (Tabular data)	__IR	Document	Data Package Creator	Data Package is a simple container format used to describe and package a collection of data. It can be used to package any kind of data. At the same time, for specific

DK FAIR portal

FAIR convergence matri

The FAIR principles articulate the behaviors expected from digital artifacts that are Findable, Accessible, Interoperable and Reusable by machines and by people. Although by now widely accepted, the FAIR Principles by design do not explicitly consider actual implementation choices enabling FAIR behaviors. As different communities have their own, often well-established implementation preferences and priorities for data reuse, coordinating a broadly accepted, widely used FAIR implementation approach remains a global challenge. In an effort to accelerate broad community convergence on FAIR implementation options, the GO FAIR community has launched the development of the FAIR Convergence Matrix. The Matrix is a platform that compiles for any community of practice, an inventory of their self-declared FAIR implementation choices and challenges. The Convergence Matrix is itself a FAIR resource, openly available, and encourages voluntary participation by any self-identified community of practice (not only the GO FAIR Implementation Networks). Based on patterns of use and reuse of existing resources, the Convergence Matrix supports the transparent derivation of strategies that optimally coordinate convergence on standards and technologies in the emerging Internet of FAIR Data and Services.

DATA Intelligence

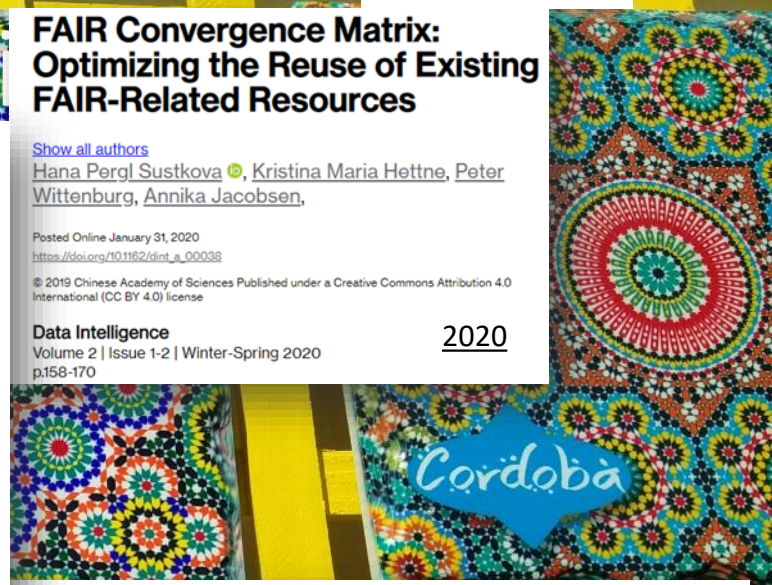
FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources

Show all authors
Hana Pergl Sustkova, Kristina Maria Hettne, Peter Wittenburg, Annika Jacobsen,

Posted Online January 31, 2020
https://doi.org/10.1162/dint_a_00038

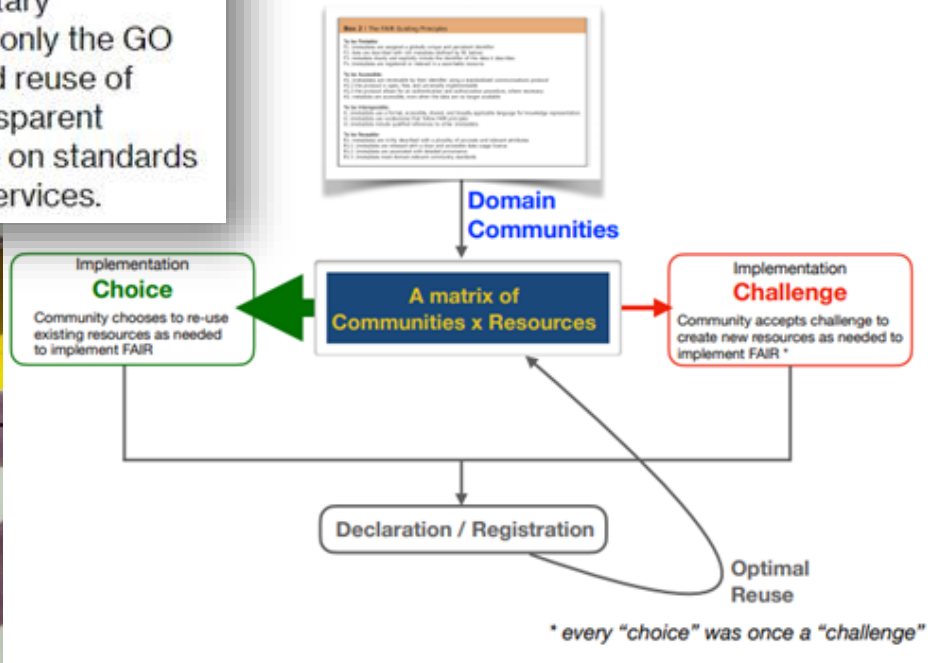
Data Intelligence
Volume 2 | Issue 1-2 | Winter-Spring 2020
p.158-170

2020



FAIR Implementation Profiles

2020



…FAIR

Oct. 2020

...FAIR per enti di ricer

Policy: Does your organization (institute / university (of applied sciences)) have a FAIR research data policy?

1. **Are the FAIR principles explicitly mentioned in the data (or research data) policy of your organization?**	The FAIR principles are not explicitly mentioned, but there	The F and A principle are explicitly mentioned.	All FAIR principles are explicitly mentioned (or there is an explicit	☐ Beginner ☐ Intermediate ☐ Advanced
2. **Is the general research data policy translated into faculty/institutes specific data policies?**	helpdesk.		new researchers are informed of the data policy upon arrival.	
3. **Is the general research data policy supported by dedicated service units (e.g. DCC) ?**				
4. **Which means and channels are used to communicate the research data policy and services to the researchers?**				

Services: Does your organization have a (virtual) DCC which provides services (including infrastructure) to allow comply with FAIR?

5. **Which services does your organization provide in order for researchers to comply with the F principles**	We provide or refer to a service to deliver a PID for a data set	We provide or refer to service for PID and adding metadata (including reference to the dataset).	On top of adding PIDs and metadata, we provide or refer to a service to make the data findable through indexed resources.	☐ Beginner ☐ Intermediate ☐ Advanced
6. **Which services does your organization provide in order for researchers to comply with the A principles?**	We provide or refer to a service with a standard and open communication protocol as to access the data.	We provide or refer to a service to enable restricted access to data and meta data (authentication protocol)	Our organization as an archiving policy, both on data and meta data.	☐ Beginner ☐ Intermediate ☐ Advanced
7. **Which services does your organization provide in order for researchers to comply with the A principles?**	We provide or refer to a service that enriches the dataset metadata, referencing to other datasets and articles (I3).	We provide or refer to a service to attribute (meta)data that use a formal, accessible, shared, and broadly applicable language for knowledge representation.	We provide or refer to a service to attribute (meta)data that use vocabularies that follow FAIR principles in order to facilitate machine readability.	☐ Beginner ☐ Intermediate ☐ Advanced
8. **Which services you're your organization provide in order for researchers to comply with the R principles?**	We provide advice on licenses and provide or refer a service to add licenses in the metadata	We monitor the reuse of our data, both within the institute as well as outside	We perform quality control on the (meta)data, e.g. adequate read me file as to assure that somebody else can actually reuse the data. We assure that (meta)data meet domain-relevant community standards	☐ Beginner ☐ Intermediate ☐ Advanced ☐ Not yet

Sssssht…forse ALCUNI editori si stanno muovendo….

An Academic Publishers' GO FAIR Implementation Network (APIN)

Jan Velterop[a,*] and Erik Schultes[b,*]

Jan.6, 2021

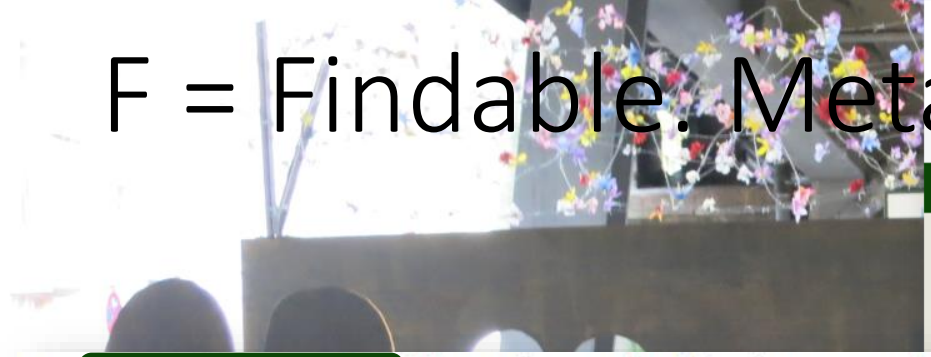
Abstract

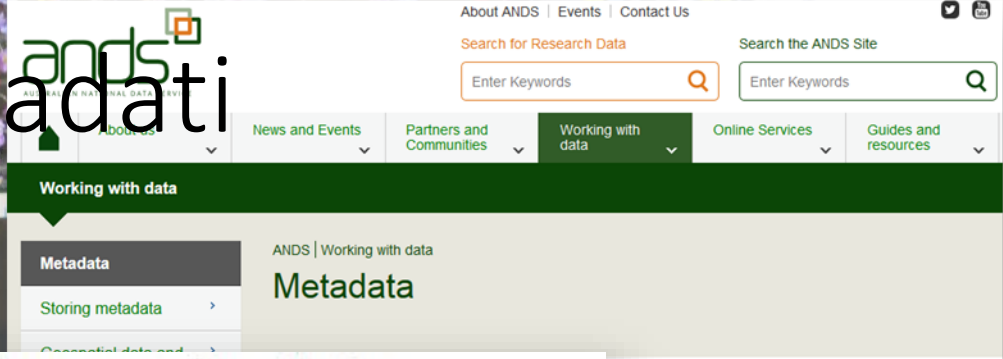
Presented here is a proposal for the academic publishing industry to get actively involved in the formulation of protocols and standards that make published scientific research material machine-readable in order to facilitate data to be findable, accessible, interoperable, and re-usable (FAIR). Given the importance of traditional journal publications in scholarly communication worldwide, active involvement of academic publishers in advancing the more routine creation and reuse of FAIR data is highly desired.

CREAZIONE DI ARTICOLI SECONDO I PRINCIPI FAIR

FINDABLE

F = Findable: Metadati

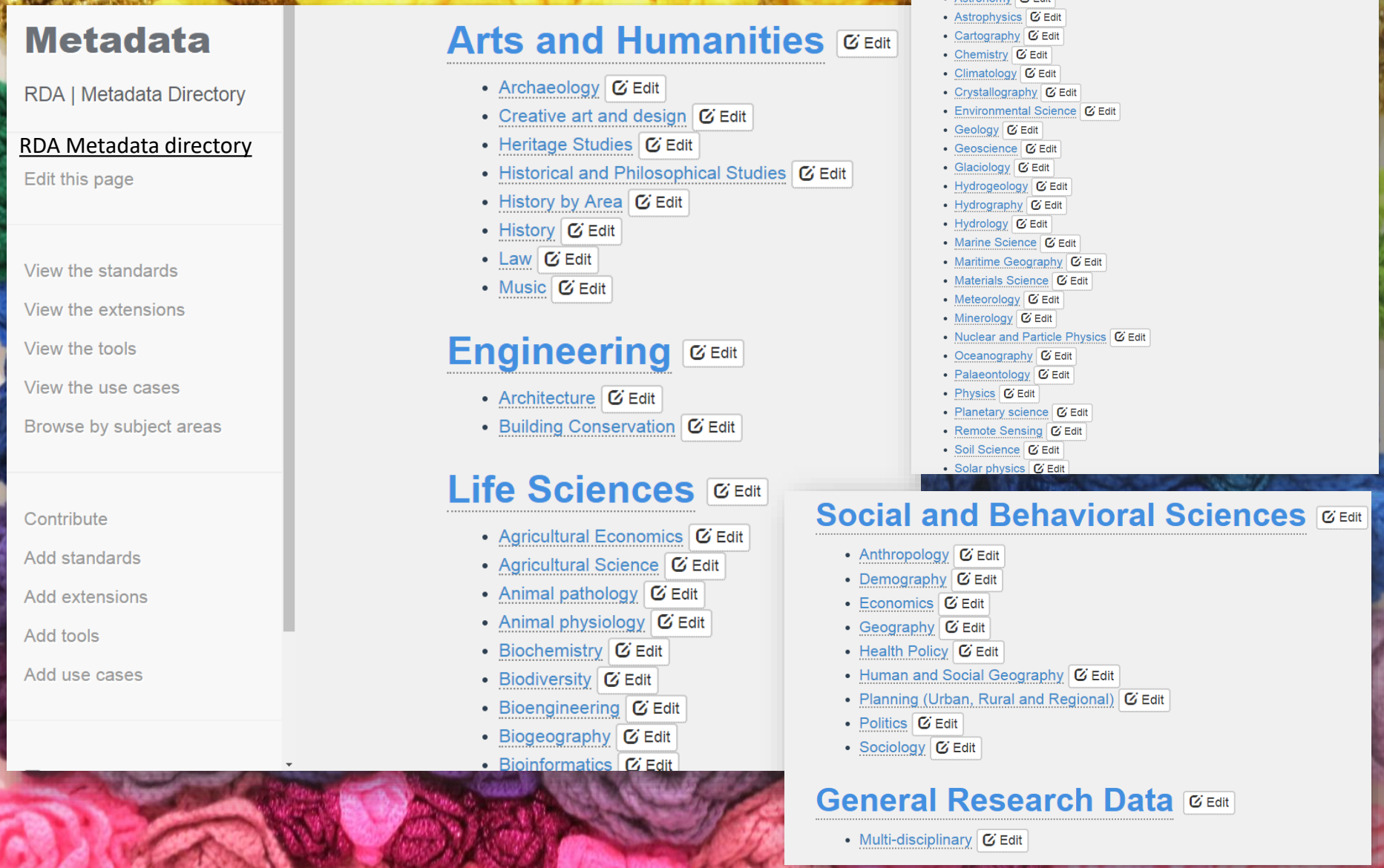




- **Descriptive metadata**: information required for discovery and assessment of the collection,
  - e.g. title, contributors, subject or keywords, study description, and location and dates of the study.
- **Provenance metadata**: this relates to the origins and processing of the data, and enables interpretation and reuse of the data. It ranges from the human to the highly technical, and usually requires some knowledge of the domain to create.
  - e.g. Where did the data come from? Why was it collected? Who collected it, when and where? What instruments/technologies were used to collect the data, and how were they set up? How has the data been processed?
- **Technical metadata**: fundamental information for a person or a computer application to read the data.
  - e.g. How is the data set up? What formats, and versions of formats, are used? How is the database configured? How does it relate to other data?
- **Rights and access metadata**: information to enable access, and licensing or usage rules.
  - e.g. How can someone access the data? Who is allowed to view or modify the data, or the metadata, and under what conditions? Who has some kind of authority over the data? Are there costs associated with access? Under what licence is the data being made available?
- **Preservation metadata**: this builds on the history from the Provenance, Rights and Technical metadata, and also includes information to allow the data to be managed for long-term accessibility.
  - e.g. Has there been any restructuring or other changes to the files, e.g. due to migration to new file formats? What software has been used to access the data?
- **Citation metadata**: information required for someone to cite the data
  - e.g. Creator(s), Publication Year, Title, Publisher, Identifier.

Metadata

F = Findable. Standards di metadati



Metadata

RDA | Metadata Directory

RDA Metadata directory

Edit this page

View the standards

View the extensions

View the tools

View the use cases

Browse by subject areas

Contribute

Add standards

Add extensions

Add tools

Add use cases

Arts and Humanities Edit

- Archaeology Edit
- Creative art and design Edit
- Heritage Studies Edit
- Historical and Philosophical Studies Edit
- History by Area Edit
- History Edit
- Law Edit
- Music Edit

Engineering Edit

- Architecture Edit
- Building Conservation Edit

Life Sciences Edit

- Agricultural Economics Edit
- Agricultural Science Edit
- Animal pathology Edit
- Animal physiology Edit
- Biochemistry Edit
- Biodiversity Edit
- Bioengineering Edit
- Biogeography Edit
- Bioinformatics Edit

Physical Sciences & Mathematics Edit

- Astronomy Edit
- Astrophysics Edit
- Cartography Edit
- Chemistry Edit
- Climatology Edit
- Crystallography Edit
- Environmental Science Edit
- Geology Edit
- Geoscience Edit
- Glaciology Edit
- Hydrogeology Edit
- Hydrography Edit
- Hydrology Edit
- Marine Science Edit
- Maritime Geography Edit
- Materials Science Edit
- Meteorology Edit
- Minerology Edit
- Nuclear and Particle Physics Edit
- Oceanography Edit
- Palaeontology Edit
- Physics Edit
- Planetary science Edit
- Remote Sensing Edit
- Soil Science Edit
- Solar physics Edit

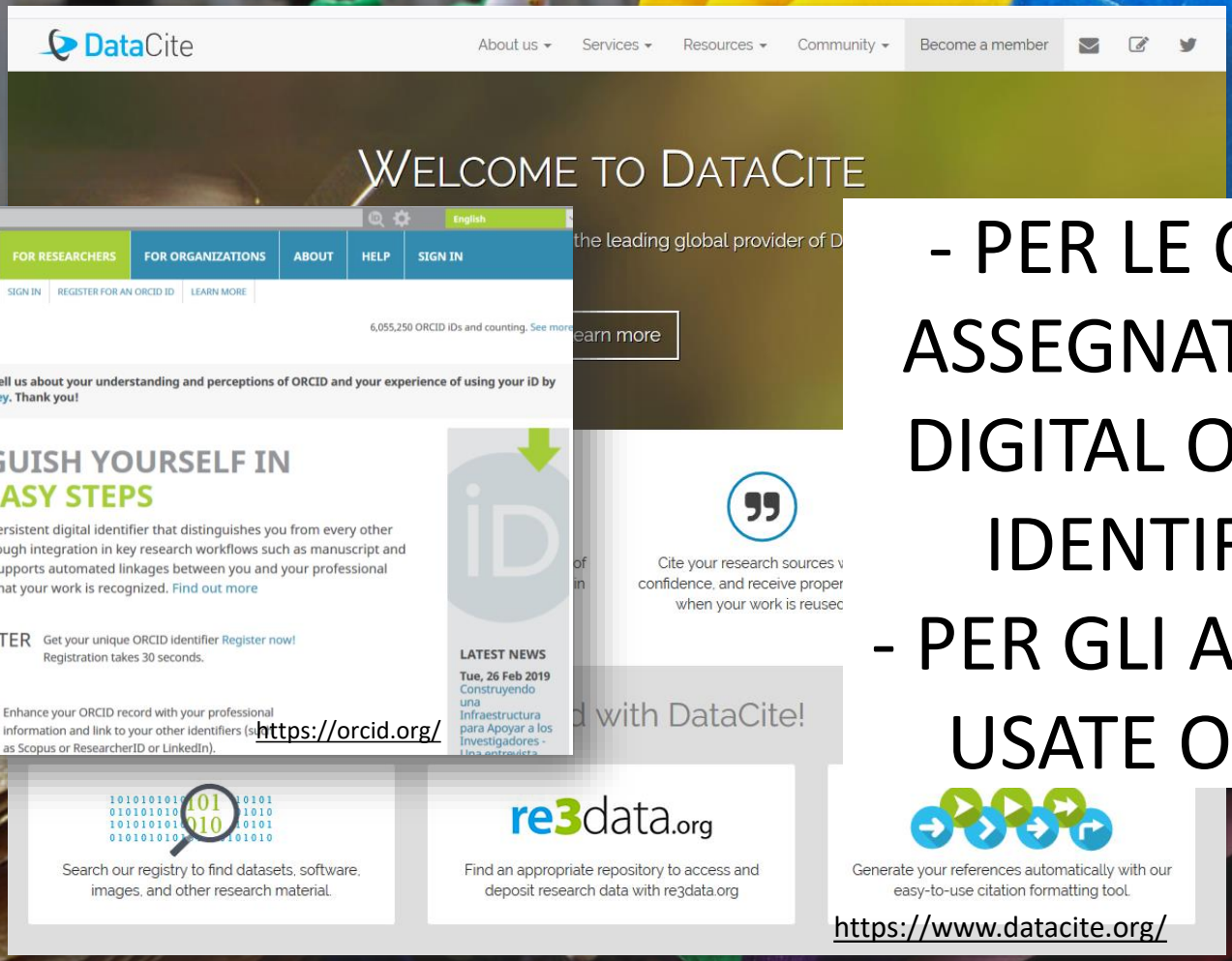
Social and Behavioral Sciences Edit

- Anthropology Edit
- Demography Edit
- Economics Edit
- Geography Edit
- Health Policy Edit
- Human and Social Geography Edit
- Planning (Urban, Rural and Regional) Edit
- Politics Edit
- Sociology Edit

General Research Data Edit

- Multi-disciplinary Edit

F = Findable. Identificativi persistenti

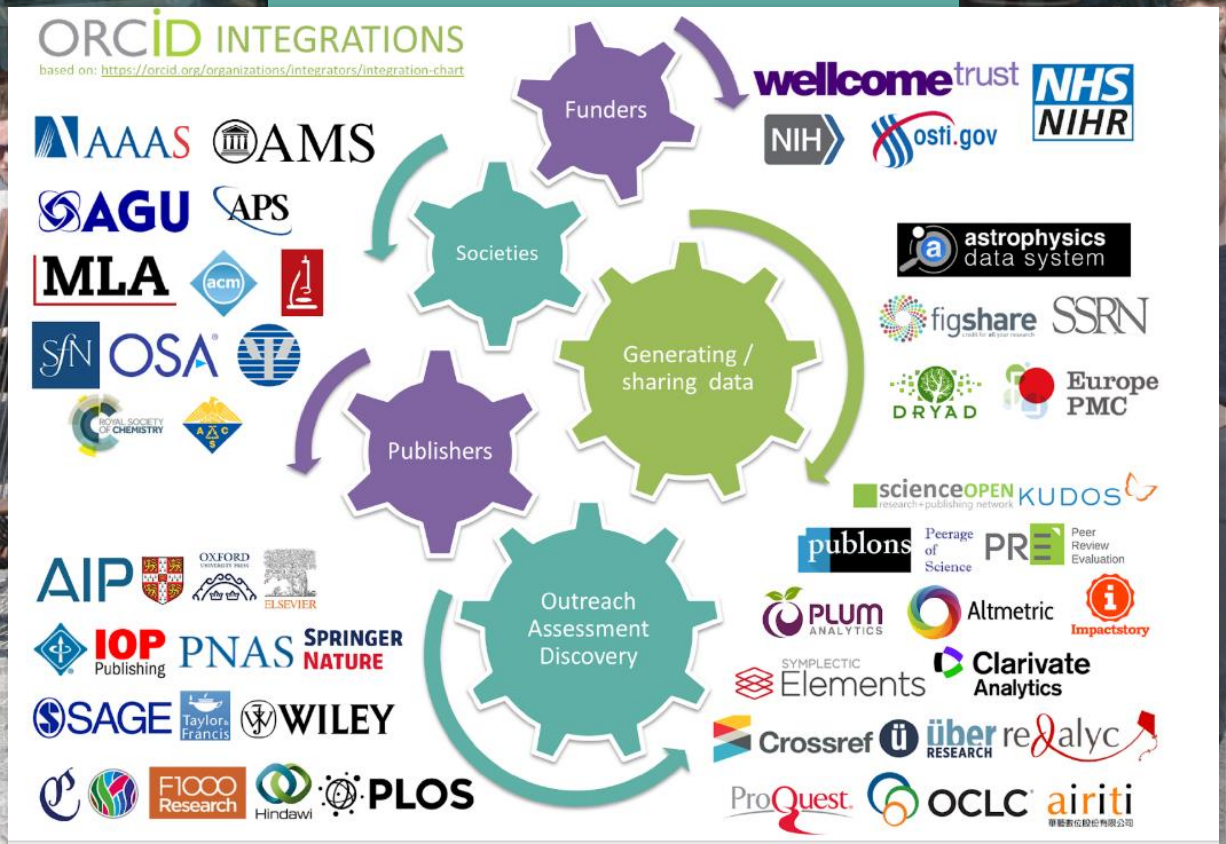


https://orcid.org/

- PER LE COSE: ASSEGNATE DOI DIGITAL OBJECT IDENTIFIER
- PER GLI AUTORI: USATE ORCID

ORCID – ID. Tutto intorno a me

CREATE E MANTENETE UN PROFILO ORCID. ATTIVA TUTTI I MECCANISMI!



ACCESSIBLE

A = Accessible

ACCESSIBLE≠OPEN
«ACCESSO» PUÒ ANCHE ESSERE RISERVATO O SOTTO EMBARGO

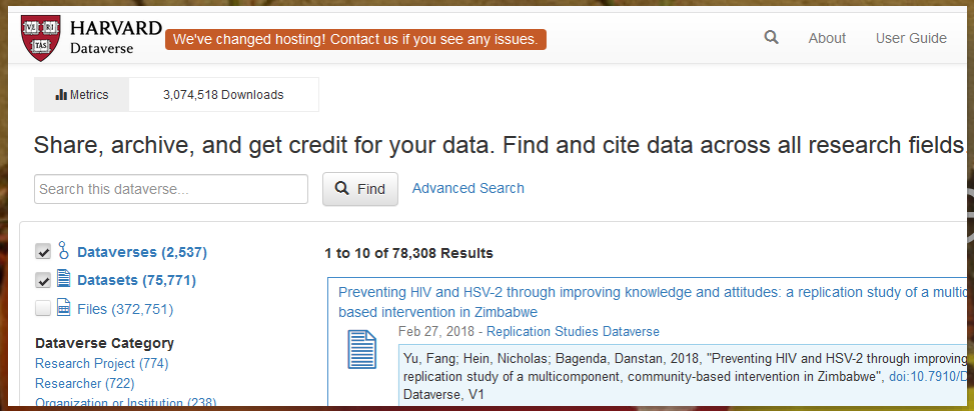
- **Open access**
  Data that can be accessed by any user whether they are re[illegible]
  Data in this category should not contain personal information unless consent is given (see 'Informed consent').

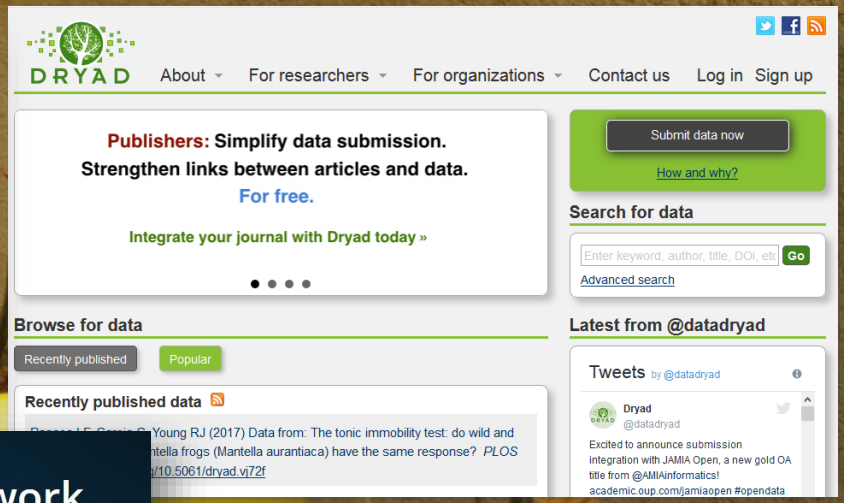
- **Access for registered users (safeguarded)**
  Data that is accessible only to users who have registered with the archive. This data contains no direct identifiers but there may be a risk of disclosure through the linking of indirect identifiers.

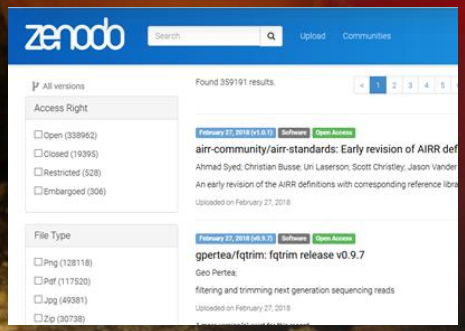
- **Restricted access**
  Access is limited and can only be granted upon request. This access category is for the most sensitive data that may contain disclosive information.
  Restricted access requires the long-term commitment of the researcher or person responsible for the data to handle the upcoming permission requests.

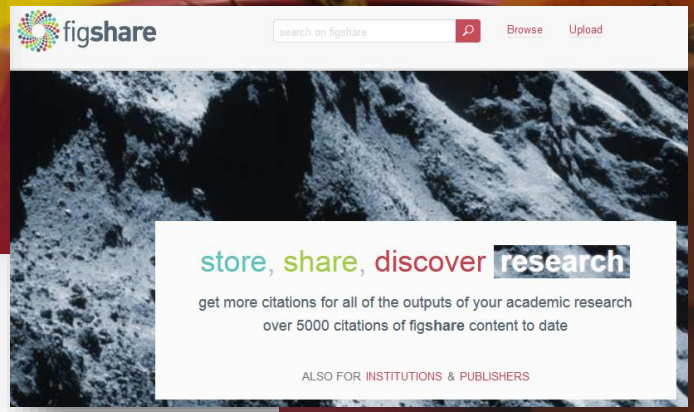
- **Embargo**
  Besides offering the opportunity for restricted access 'for eternity' most data repositories allow you to place a temporary embargo on your data. During the embargo period, only the description of the dataset is published. The data themselves will become available in open access after a certain period of time.

CESSDA Guide









General depositories for research data

The following depositories are of interest to researchers in all domains:

- Zenodo (not-for-profit, hosted by CERN): https://zenodo.org:
- Dryad (not-for-profit membership organisation): http://www.datadryad.org
- Figshare (free service provided by private company): https://figshare.com
- Open Science Framework (not-for-profit, developed and maintained by the Center for Open Science[1]): https://osf.io
- Harvard Dataverse (not-for-profit, hosted by the Institute for Quantitative Social Studies IQSS at Harvard University): https://dataverse.harvard.edu

A = Accessible. Data lab [in progress]

Data lab

Save your data during research

Many research projects involve collaboration with several researchers, and a large amount of research data is generated. To manage such quantities of data and to allow the researchers to share them with each other, 4TU.Centre for Research Data offers the possibility of establishing a 'data lab'.

A data lab is a secure online environment (with or without screening) for storing, processing and sharing dynamic research data, software, visualisations and other items with fellow researchers.

DataverseNL

Members of the three technical universities (Eindhoven, Twente and Delft) can draw upon DataverseNL through 4TU.ResearchData. The objective of this open source application is to make research data accessible to others. You can store data in the DataverseNL environment and grant multiple individuals controlled access to them.
Once you have requested the project environment or dataverse from 4TU.ResearchData, you are free to arrange this environment as you wish.

Why DataverseNL?

- Organization of data files in dataverses and datasets
- Addition of metadata and documentation
- Version management
- Management of access rights
- Easy collaboration with fellow researchers or project partners, even beyond your own university or research institute
- Centralized professional storage and backup
- To fulfil funding agency requirements for data management and sharing during your research.

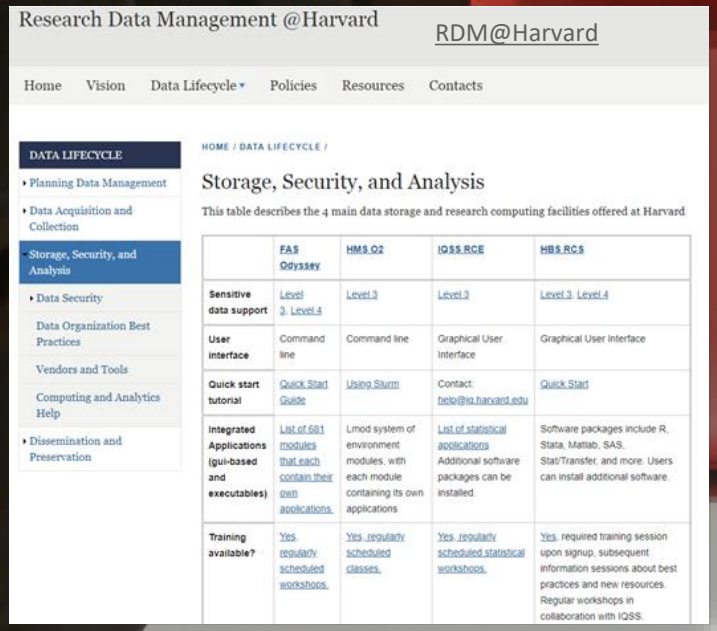
See our quick guide for getting started or contact us when you need additional support.

Costs

Members of the three technical universities (Eindhoven, Twente and Delft) can store up to 100 GB of data free of charge. Storage of more than 100 GB of data costs € 150 per additional 50 GB per year.

[gentle reminde
…RISK ASSESSM

RICORDATEVI ANCHE DELLA SICUREZZA


Research Data Management @Harvard

RDM@Harvard

Home | Vision | Data Lifecycle | Policies | Resources | Contacts

DATA LIFECYCLE
- Planning Data Management
- Data Acquisition and Collection
- Storage, Security, and Analysis
- Data Security
- Data Organization Best Practices
- Vendors and Tools
- Computing and Analytics Help
- Dissemination and Preservation

HOME / DATA LIFECYCLE /

Storage, Security, and Analysis

This table describes the 4 main data storage and research computing facilities offered at Harvard.

	FAS Odyssey	HMS O2	IQSS RCE	HBS RCS
Sensitive data support	Level 3, Level 4	Level 3	Level 3	Level 3, Level 4
User interface	Command line	Command line	Graphical User Interface	Graphical User Interface
Quick start tutorial	Quick Start Guide	Using Slurm	Contact: help@iq.harvard.edu	Quick Start
Integrated Applications (gui-based and executables)	List of 681 modules that each contain their own applications	Lmod system of environment modules, with each module containing its own applications	List of statistical applications Additional software packages can be installed.	Software packages include R, Stata, Matlab, SAS, Stat/Transfer, and more. Users can install additional software.
Training available?	Yes, regularly scheduled workshops.	Yes, regularly scheduled classes.	Yes, regularly scheduled statistical workshops.	Yes, required training session upon signup, subsequent information sessions about best practices and new resources. Regular workshops in collaboration with IQSS.

Level	Data Classification and Examples (abridged version)
5	***Information that would cause severe harm to individuals or the University if disclosed.***
	• Research information classified as Level 5 by an IRB or otherwise required to be stored or processed in a high security environment and on a computer not connected to the Harvard data networks • Certain individually identifiable medical records and genetic information, categorized as extremely sensitive
4	***Information that would likely cause serious harm to individuals or the University if disclosed.***
	• High Risk Confidential Information (HRCI) and research information classified as Level 4 by an IRB • Personally identifiable financial or medical information • Information commonly used to establish identity that is protected by state, federal, or foreign privacy laws and regulations • Individually identifiable genetic information that is not Level 5 • National security information (subject to specific government requirements) • Passwords and Harvard PINs that can be used to access confidential information
3	***Information that could cause risk of material harm to individuals or the University if disclosed.***
	• Research information classified as Level 3 by an IRB • Information protected by the Family Educational Rights and Privacy Act (FERPA) to the extent it is not covered under Level 4 including non-directory student information and directory information about students who have requested a FERPA block • HUIDs associated with names or any other information that could identify individuals • Harvard personnel records (employees may discuss terms and conditions of employment with each other and third parties) Level 4 including non-directory student information and directory information about students who have requested a FERPA block • HUIDs associated with names or any other information that could identify individuals • Harvard personnel records (employees may discuss terms and conditions of employment with each other and third parties) • Institutional financial records • Individual donor information • Other personal information protected under state, federal and foreign privacy laws not classified as Level 4 or 5
2	***Information the disclosure of which would not cause material harm, but which the University has chosen to keep confidential.***
	• Unpublished research work and intellectual property not in Level 3 or 4 • Research information classified as Level 2 by an IRB • Patent applications and work papers, drafts of research papers • Building plans and information about the University physical plant
1	***Public information.***
	• Research data that has been de-identified in accordance with applicable rules • Published research • Published information about the University • Course catalogs • Directory information about students who have not requested a FERPA block • Faculty and staff directory information

Harvard security

A = Accessible. Data repositories

3. Will the data be safe in legal terms?

For this criterion we first consider the basic legal terms and conditions to chec
Here only one capability level is given, as a repository either will or will not comply. Then we consider licensing, disclosure risk and access control, when repository may offer different levels of capability you can match to your needs
Note that Re3data provides relevant details for the repositories it lists, under 'terms'.

Legal terms and conditions
Level 1
Personal data or data which may identify individuals when linked to other data should not be stored outside the European Economic Area, unless in a legal jurisdiction that ensures personal data is adequately protected ❑
By agreeing to the terms and conditions the depositor will not be breaching other Data Protection principles, or the terms of any confidentiality agreement with data subjects or owners (e.g. consent form, consortium agreement) ❑

Findable, accessible and interoperable		
Level 1	Level 2	Level 3
Metadata publishing: Data collections are catalogued in a repository according to funder expectations so that they are discoverable by title, creator, and date of deposition ❑	Repository publishes other pertinent information as metadata fields to enhance cross-disciplinary discovery ❑	Metadata is catalogued to enhance reuse according to sector-leading standards, or to fulfil domain-specific purposes ❑
Stable identifiers: Enables a DOI or other open standard identifier to be assigned to a landing page for each ingested dataset/ collection ❑	Supports assignment of related persistent IDs per dataset/ collection ❑	Supports assignment of multiple persistent IDs at different levels of granularity within dataset/ collection ❑
Discovery metadata:Provides Datacite mandatory metadata and exposes	Provides metadata elements to enable broader discovery (e.g. geo-spatial) to	Exposes discovery metadata as Linked Open Data to optimise automatic

Checklist: is it the right repository for your data?

The checklist that follows addresses the five key questions posed in this guide:

1. is the repository reputable?
2. will it take the data you want to deposit?
3. will it be safe in legal terms?
4. will the repository sustain the data value?
5. will it support analysis and track data usage?

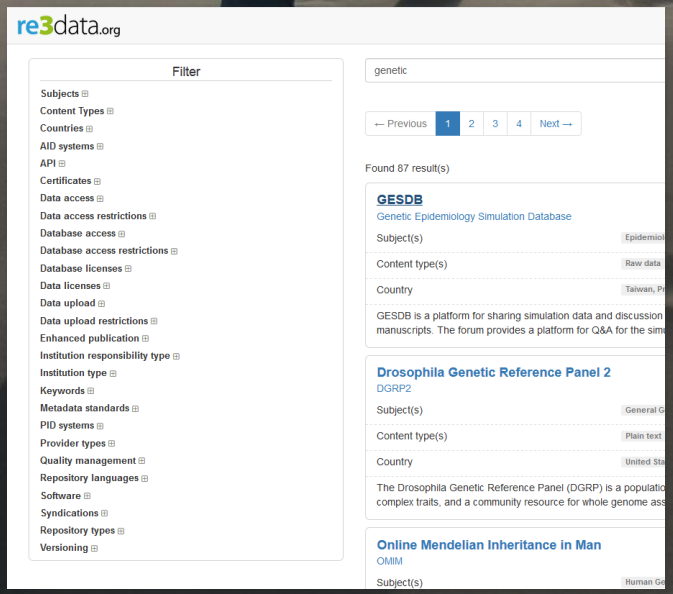
«PRIORITÀ»

1. DISCIPLINARY
2. ZENODO
3. INSTITUTIONAL

(A MENO CHE SIA DATAVERSE, CHE VIENE USATO DURANTE IL PROGETTO)

A = Accessible.
Cercate un archivio?

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES
Search... Search
https://www.re3data.org/



2,000 Data Repositories and Science Europe's Framework for Discipline-specific Research Data Management

By offering detailed information on more than 2,000 research data repositories, re3data has become the most comprehensive source of reference for research data infrastructures globally. Through the development and advocacy of a framework for discipline...

Read more

Three new DOI Fabrica features to simplify account management

Last month month we launched DOI Fabrica, the modernized version of the DataCite Metadata Store (MDS) web frontend. It is the one place for DataCite providers and their clients to create, find, connect and track every single DOI from their organization...

Read more

One step closer towards instant DOI search results

Art Art? You might be wondering, what this pink and green picture illustrates? A few months ago we couldn't show you this picture; the data that we used to created it, did not exist. And the answer to what this illustrates – this is simply a distorted...

Read more

Criteria for the selection of a trustworthy repository

TRUSTWORTHY REPOSITORIES

Trustworthy repositories should meet the following minimum criteria:

1. Provision of Persistent and Unique Identifiers (PIDs)

a. Allow data discovery and identification
b. Enable searching, citing, and retrieval of data
c. Provide support for data versioning

2. Metadata

a. Enable finding of data
b. Enable referencing to related relevant information, such as other data and publications
c. Provide information that is publicly available and maintained, even for non-published, protected, retracted, or deleted data
d. Use metadata standards that are broadly accepted (by the scientific community)
e. Ensure that metadata are machine-retrievable

3. Data access and usage licences

a. Enable access to data under well-specified conditions
b. Ensure data authenticity and integrity
c. Enable retrieval of data
d. Provide information about licensing and permissions (in ideally machine-readable form)
e. Ensure confidentiality and respect rights of data subjects and creators

4. Preservation

a. Ensure persistence of metadata and data
b. Be transparent about mission, scope, preservation policies, and plans (including governance, financial sustainability, retention period, and continuity plan)

A = Accessible. Data journals

Title	URL	Charge	Notes for authors (N.B. we suggest checking in particular for policy on submission of data already published)	Publisher	Notes on Subject Area
Journal of Open Archaeology Data	http://openarchaeologydata.metajnl.com/		http://openarchaeologydata.metajnl.com/about/submissions	Ubiquity Press	Archaeology
Open Health Data	http://openhealthdata.metajnl.com/		http://openhealthdata.metajnl.com/about/submissions#authorGuidelines	Ubiquity Press	Public Health
Journal of Open Psychology Data	http://openpsychologydata.metajnl.com/		http://openpsychologydata.metajnl.com/about/submissions#onlineSubmissions	Ubiquity Press	Psychology
Scientific Data	http://www.nature.com/scientificdata/				
Genomics Data	http://www.journals.elsevier.com/genomics-data/				
Geoscience Data Journal	http://www.geosciencedata.com/				
Earth System Science Data	http://earth-system-science-data.net/				
Ecological Archives	http://esapubs.org/archive/				
Journal of Chemical and Engineering Data	http://pubs.acs.org/journal/jceaax				

Data journals list

Dataset Description

Object Name

- *walkers* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for records made by individual walkers during stage-one fieldwalking.
- *counts* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for pot-sherds countedduring stage-one fieldwalking.
- *pottery* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for the main pottery database, assembled various artefact specialists.
- *petrography* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for those sherds sampled for thin section petrography.
- *lithics* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for the main lithics database.
- *other* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for the main database of all non-ceramic and non-lithic finds.
- *structs* – three files providing the data, metadata and field type definitions (.csv, .txt, .csvt respectively) for the main database of all standing remains, except for terraces.
- *coast* – a vector polygon dataset (.shp and associated files) with the shape of Antkythera's coastline.
- *geology* –a vector polygon dataset (.shp and associated files) with the main bedrock units on Antkythera.
- *tracts* – a vector polygon dataset (.shp and associated files) with the main stage-one survey units.
- *grids* – a vector polygon dataset (.shp and associated files) with the main stage-two survey units.
- *terraces* – vector line dataset (.shp and associated files) with all observable agricultural terraces (i.e. the location

- *other* – primarily Andrew Bevan (UCL), with further assistance from James Conolly (Trent)
- *geology* – a combination of fieldwork by Ruth Siddall (UCL) and remote sensing by Andrew Bevan (UCL)

Repository Location
UK Archaeology Data Service Collection 1115 (doi: 10.5284/1012484)

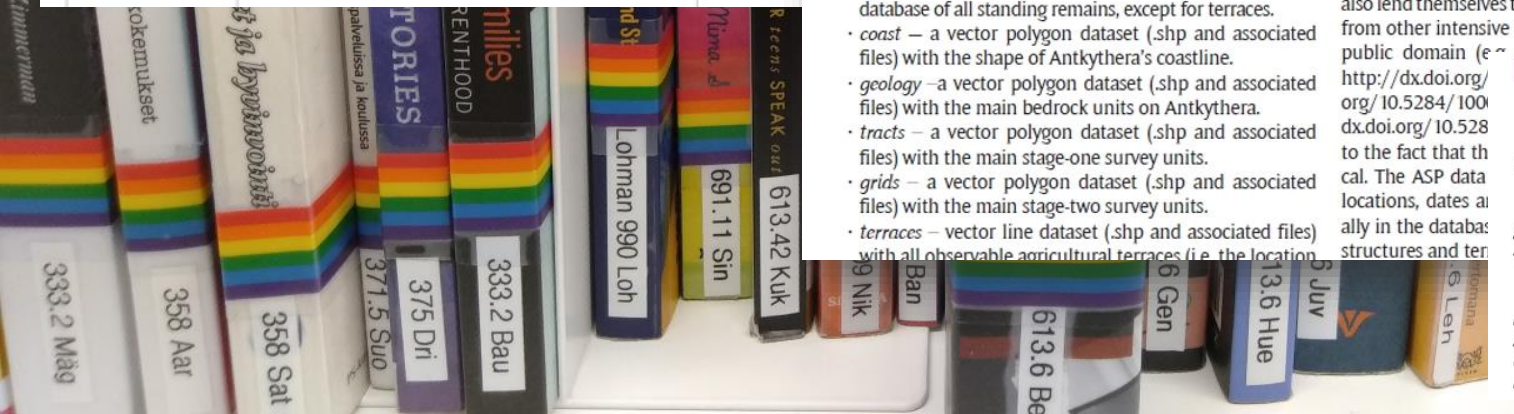
Publication Date
05/02/2012

Language
English (a Greek language summary of the project methods and results can be found at www.ucl.ac.uk/asp/ or www.tuarc.trentu.ca/asp/).

License
Creative Commons CC-BY 3.0

Reuse Potential

Due to their unusual coverage of an entire landscape, these datasets would provided a good basis for developing a tutorial on survey, GIS and/or spatial analysis in archaeology. They also lend themselves to the comparative analysis of evidence from other intensive Mediterranean surveys that are in the public domain (e.g. http://dx.doi.org/ org/10.5284/100 dx.doi.org/10.528 to the fact that th cal. The ASP data locations, dates a ally in the databas structures and ter



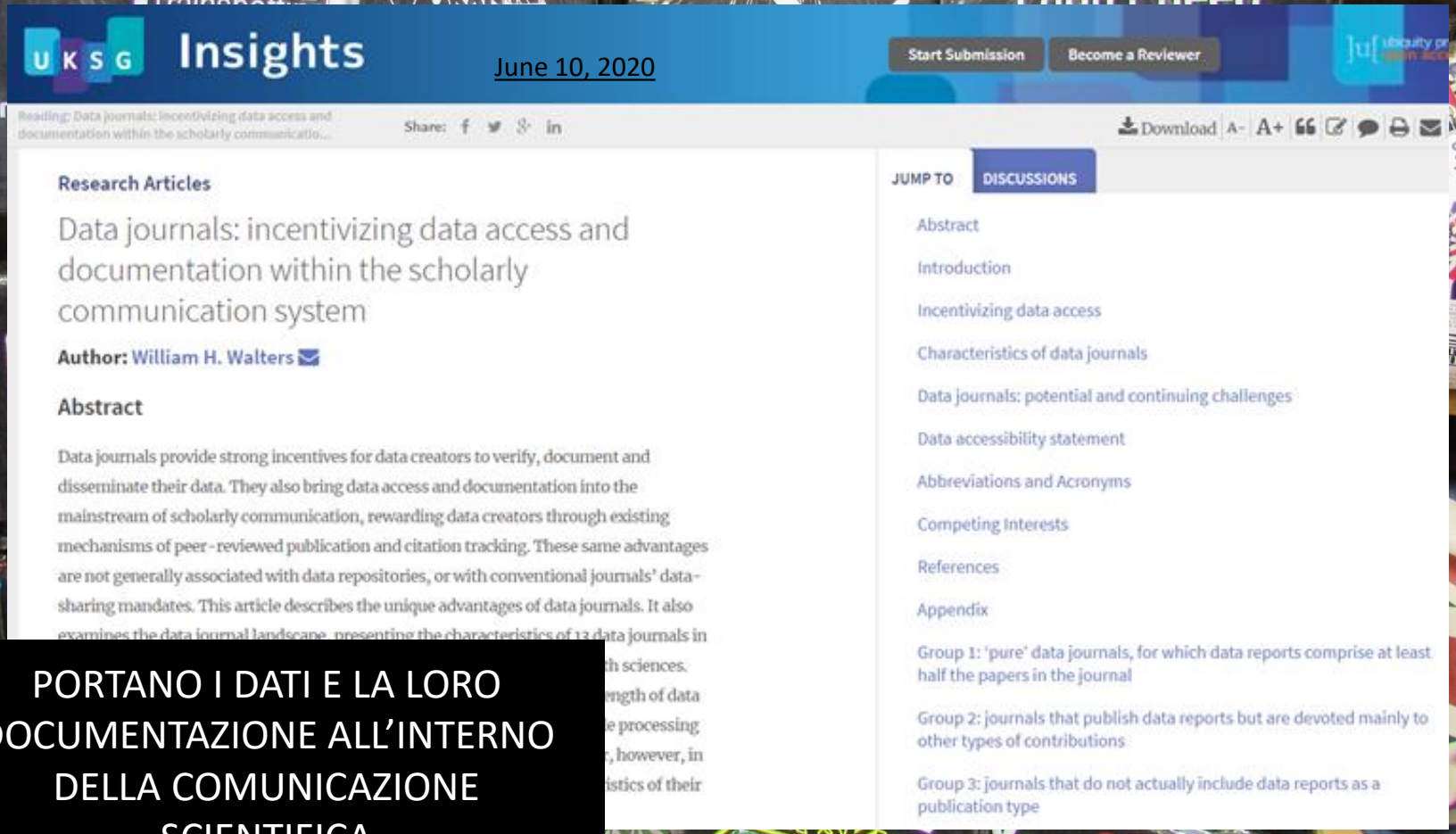
Data journals

Panayiota Polydoratou

Alexander Technological Educational Institute of Thessaloniki

European Commission Workshop
Alternative Open Access Publishing Models: Exploring New Territories in Communication
Brussels, 12 October 2015

A = Accessible. Data journals



UKSG Insights

June 10, 2020

Start Submission | Become a Reviewer

Share:

Download A- A+

Research Articles

Data journals: incentivizing data access and documentation within the scholarly communication system

Author: William H. Walters

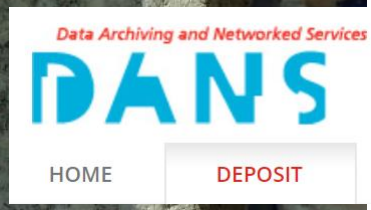
Abstract

Data journals provide strong incentives for data creators to verify, document and disseminate their data. They also bring data access and documentation into the mainstream of scholarly communication, rewarding data creators through existing mechanisms of peer-reviewed publication and citation tracking. These same advantages are not generally associated with data repositories, or with conventional journals' data-sharing mandates. This article describes the unique advantages of data journals. It also examines the data journal landscape, presenting the characteristics of 13 data journals in

JUMP TO DISCUSSIONS

PORTANO I DATI E LA LORO DOCUMENTAZIONE ALL'INTERNO DELLA COMUNICAZIONE SCIENTIFICA

A = Accessible. Formati



Type	Preferred format(s)	Non-preferred format(s)
Text documents	• PDF/A (.pdf)	• ODT (.odt) • MS Word (.doc, .docx) • RTF (.rtf) • PDF (.pdf)
Plain text	• Unicode text (.txt)	• Non-Unicode text (.txt)
Markup language	• XML (.xml) • HTML (.html) • Related files: .css, .xslt, .js, .es	• SGML (.sgml)
Spreadsheets	• ODS (.ods) • CSV (.csv)	• MS Excel (.xls, .xlsx) • PDF/A (.pdf) • OOXML (.docx, .docm)
Databases	• SQL (.sql) • SIARD (.siard) • DB tables (.csv)	• MS Access (.mdb, .accdb) (v. 2000 or later) • dBase (.dbf) • HDF5 (.hdf5, .he5, .h5)
Statistical data	• SPSS Portable (.por) • SPSS (.sav) • STATA (.dta) • DDI (.xml) • data (.csv) + setup (.txt)	• SAS (.7dat; .sd2; .tpt) • R (* under examination)
Raster images	• JPEG (.jpg, .jpeg) • TIFF (.tif, .tiff) • PNG (.png) • JPEG 2000 (.jp2)	• DICOM (.dcm) (by mutual agreement)

https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats

A – Accessible - Formati

NATIONAL ARCHIVES

Appendix A: Tables of File Formats

Quick Links

Computer Aided Design	Digital Audio	Digital Moving Images
Digital Cinema	Digital Video	Digital Still Images
Digital Photographs	Scanned Text	Digital Posters
Geospatial Formats	Presentation Formats	Textual Data
Structured Data Formats	Email	Web Records
Calendars	Navigational Charts	

Symbol Key

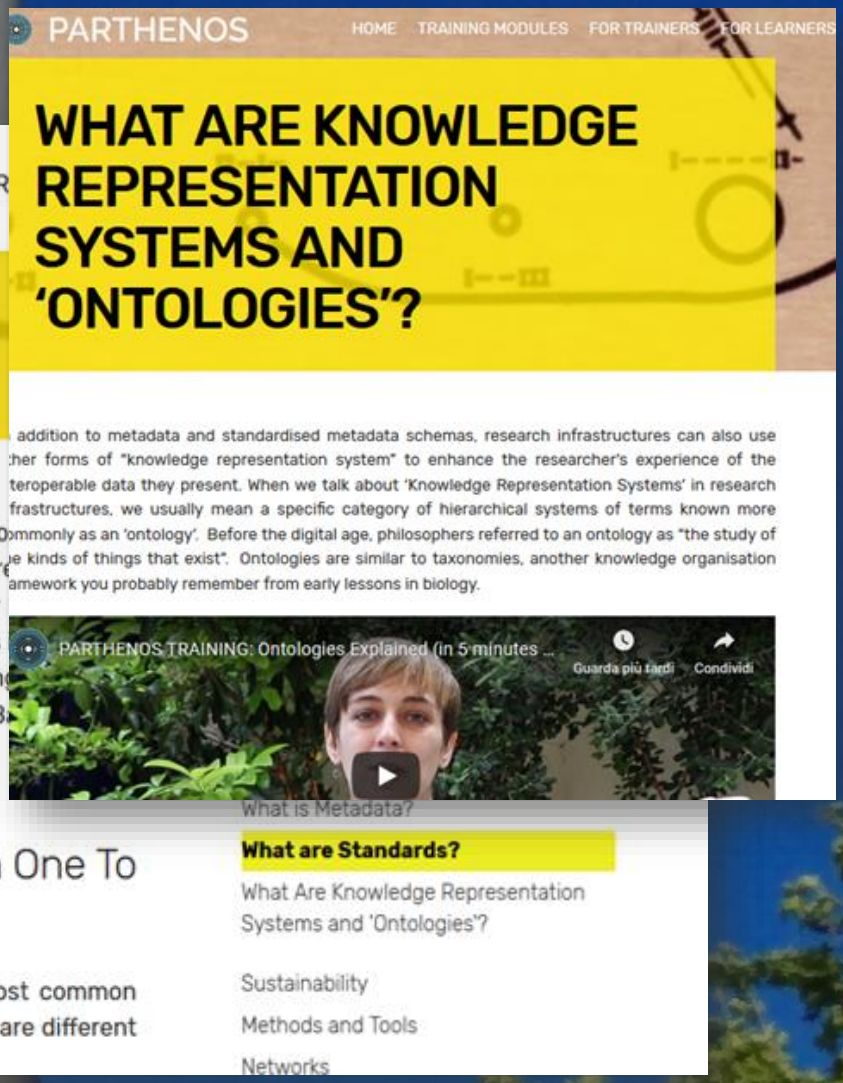
Preferred Formats

Acceptable Formats

NationalArchives

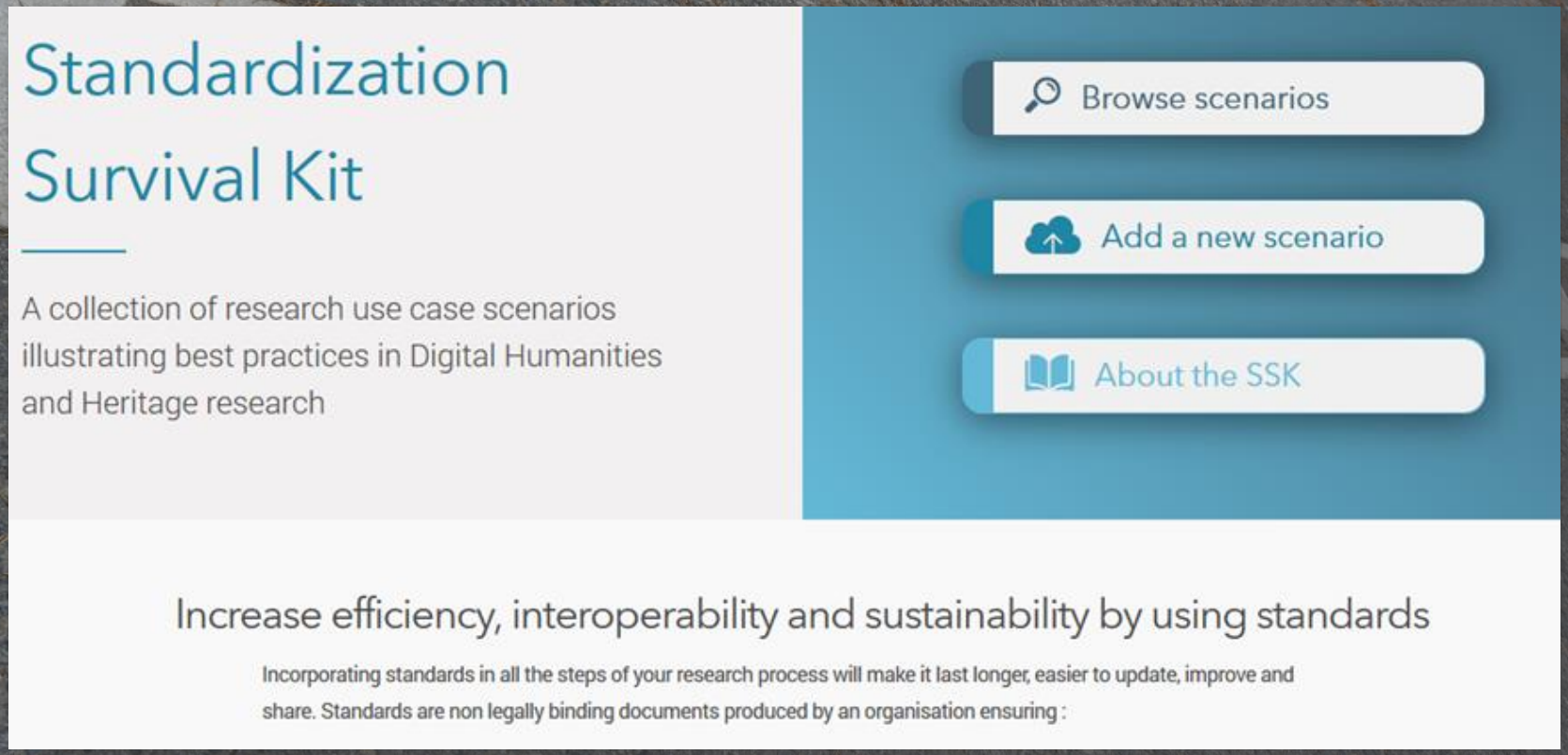
INTEROPERABLE

I = Interoperable. Standards



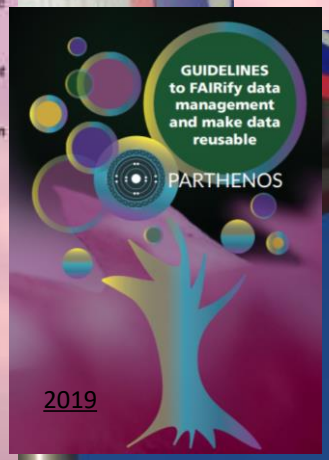
http://training.parthenos-project.eu/sample-page/intro-to-ri/interoperability/what-are-standards/

I = Interoperable. Standards

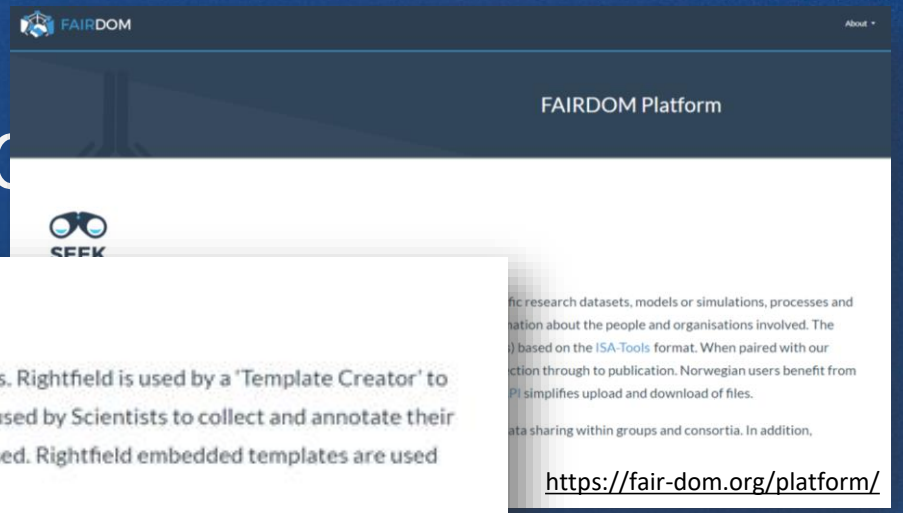


http://ssk.huma-num.fr/#/

I = Interoperable



I= Interoperable – ac



https://fair-dom.org/platform/

RightField

Rightfield is an open-source tool for adding ontology term selection to Excel spreadsheets. Rightfield is used by a 'Template Creator' to create semantically aware Excel spreadsheet templates. The Excel templates are then reused by Scientists to collect and annotate their data; without any need to understand, or even be aware of, Rightfield or the ontologies used. Rightfield embedded templates are used within the Samples framework of the SEEK.

MORE INFORMATION

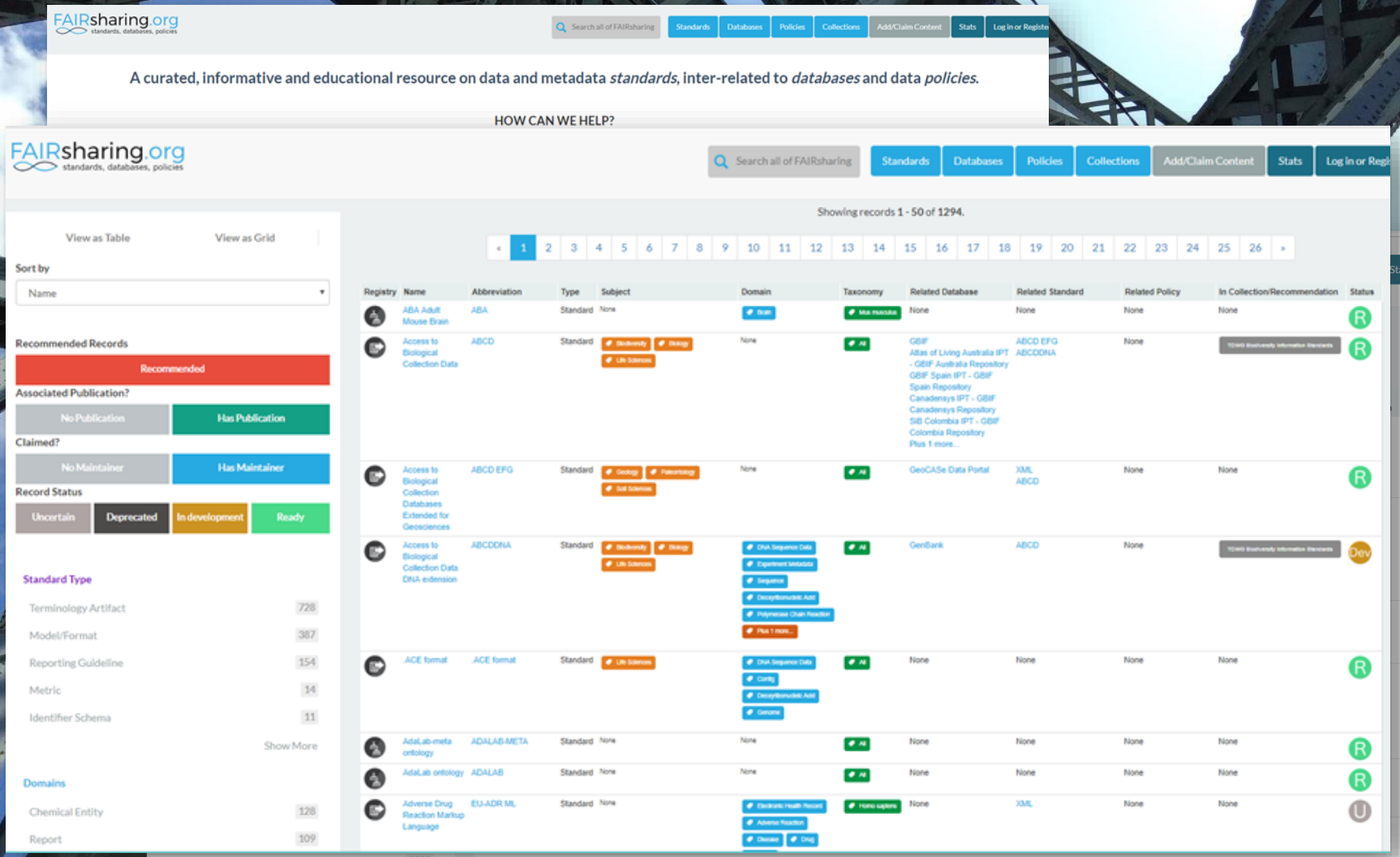
https://fair-dom.org/platform/rightfield/

Metadata	Values (examples)		Notes
Asset Title			The name of the data file
Uploader			The person submitting the asset to SEEK
Uploader SEEK ID			If you add your own SEEK ID, this will help us link this asset with your profile
Project	Project		The project that the asset belongs to
ASSAY			
Assay SEEK ID			If referring to an exisitng Assay, you can link to it via the Assay SEEK ID.
Assay Title			The title of an exisitng assay
Assay_type	ExperimentalAssayType		The assay_type describes the type of experiment you are performing
Technology_type			Describes the type of instruments and/or equipment used for the experiment
Description			A brief, human readable description.
Experimentalist			The names of the people who carried out the experiments. These can either be SEEK members or external scientists
Date			The start date for the experiment if different from the upload date
SOP			Links to SOPs and protocols used to carry out the experiment. If they are already in SEEK, you can refer to them by their SEEK ID
Publication (optional)			If this data appears in a publication, you can link it directly, or via the assay or study. If it is already registered in SEEK, you can use the PubMed ID or DOI as a reference.
Experimental_conditions			
Item	ExperimentalConditions	ExperimentalConditions	The name of the experimental condition you are fixing in your experiment (e.g. temperature, concentration, pH etc). If there is more than 1, please list them in columns across the spreadsheet
Compound (if concentration)			The compound name is only required if the item is concentration.
Unit			The SI units of the experimental conditions measurements.
Start_value (optional)			This field is used for recording changes throughout the experiment to measure different conditions (e.g. pH or dilutions)
End_value (optional)			This field is used for recording changes throughout the experiment to measure different conditions (e.g. pH or dilutions)
Comments			Additional information that would be useful for people reading this data file

amplification
by-product_formation
catabolic_response
cell_growth_optimisation
cell_size
Comparative_genomic_hybridization
comparative_genomics
continuousEnzymatic

Metadata | Organism_Sample | Instrument_(optional) | Data_ | Result_

[FAIRsharing. Per essere interoperabili]



https://fairsharing.org

REUSABLE

R = Reusable. Documentazione

DOCUMENTAZIONE (README FILE) PER
- EVITARE USO SCORRETTO/CATTIVE INTERPRETAZIONI DEI VOSTRI DATI
- MANTENERE INTEGRITÀ

cessda TRAINING

Project-level documentation

Project-level documentation explains the aims of the study, what the research questions/hypotheses are, what methodologies were being used, what instruments and measures were being used, etc. In the accordion the questions which your project-level documentation should answer are stated in more detail:

- ⊕ 1. For what purpose was data created
- ⊕ 2. What does the dataset contain
- ⊕ 3. How was data collected
- ⊕ 4. Who collected the data and when
- ⊕ 5. How was the data processed
- ⊕ 6. What possible manipulations were done to the data
- ⊕ 7. What were the quality assurance procedures
- ⊕ 8. How can data be accessed

Data-level documentation

Data-level or object-level documentation provides information at the level of individual objects such as pictures or interview transcripts or variables in a database. You can embed data-level information in data files. For example, in interviews, it is best to write down the contextual and descriptive information about each interview at the beginning of each file. And for quantitative data variable and value names can be embedded within the data file itself.

⊖ Quantitative data

Variable-level annotation should be embedded within a data file itself. If you need to compile an extensive variable level documentation that can be created by using a structured metadata format.

Data-level documentation for quantitative data

For quantitative data document the following:

- **Information about the data file**
  Data type, file type and format, size, data processing scripts.
- **Information about the variables in the file**
  The names, labels and descriptions of variables, their values, a description of derived

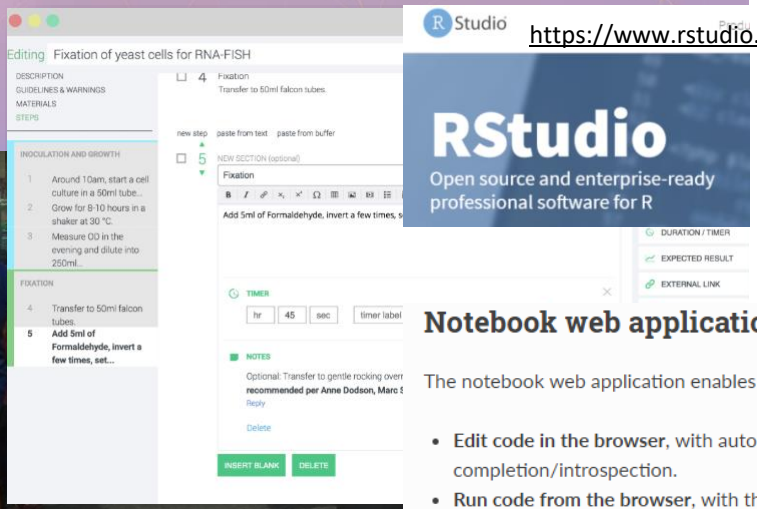
https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/Documentation-and-metadata

R = Reusable. Documentazione

protocols.io

Make your science more reproducible

protocols.io is the #1 open access repository for science methods



Studio

https://www.rstudio.com/

RStudio

Open source and enterprise-ready professional software for R

http://jupyter.org/index.h

What is an Open Notebook?

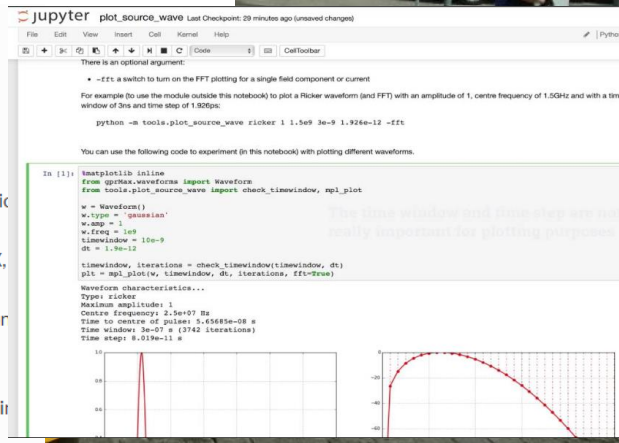
Open Notebooks are documents that contain equations, visualisations, narrative text and live code that can be executed independently and interactively, with output visible immediately beneath the input.

They bring together analysis descriptions and results, which can be executed to perform the data analysis in real time.

Notebook web application

The notebook web application enables users to:

- **Edit code in the browser**, with automatic syntax highlighting, indentation, and tab completion/introspection.
- **Run code from the browser**, with the results of computations attached to the code whic generated them.
- See the results of computations with **rich media representations**, such as HTML, LaTeX, SVG, PDF, etc.
- Create and use **interactive JavaScript widgets**, which bind interactive user interface cor and visualizations to reactive kernel side computations.
- Author **narrative text** using the Markdown markup language.
- Include mathematical equations using **LaTeX syntax in Markdown**, which are rendered i browser by MathJax.



…WHY NOT?

- PROTOCOLS.IO PER I PROTOCOLLI
- OPEN LAB NOTEBOOK TIENE TRACCIA DI TUTTO IL LAVORO (DIFFICILE LA PRIMA VOLTA, POI…)

R = Reusable. Licenze

R= Reusable. Licenze

Copyright: protects the STRUCTURE, selection or arrangement of their contents" (Art. 3) NOT THE DATA

Sui generis database right: protects the «substantial effort» in OBTAINING data [NOT «CREATING»]… the right owner often is the institution

Database=a collection of independent works, data or other materials arranged in a systematic or methodical way (Art.1)

KEEP CALM AND

RICORDA: NESSUN COPYRIGHT SUI DATI (NON CREATIVI)

DIRECTIVE 96/9/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 11 March 1996

on the legal protection of databases

NCIL OF THE EUROPEAN UNION,

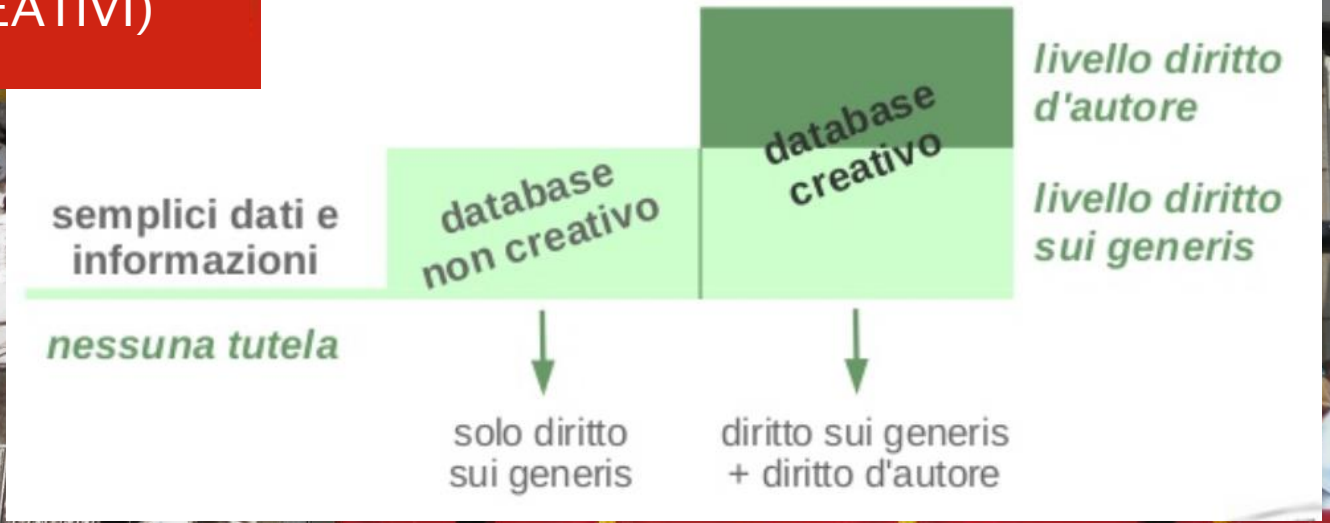
n Community, and in particular Article 57 (2), 66 and 100a thereof,

(1),

Simone Aliprandi

2014

la QUALI DIRITTI SUI DATI?



[webinar]

2013

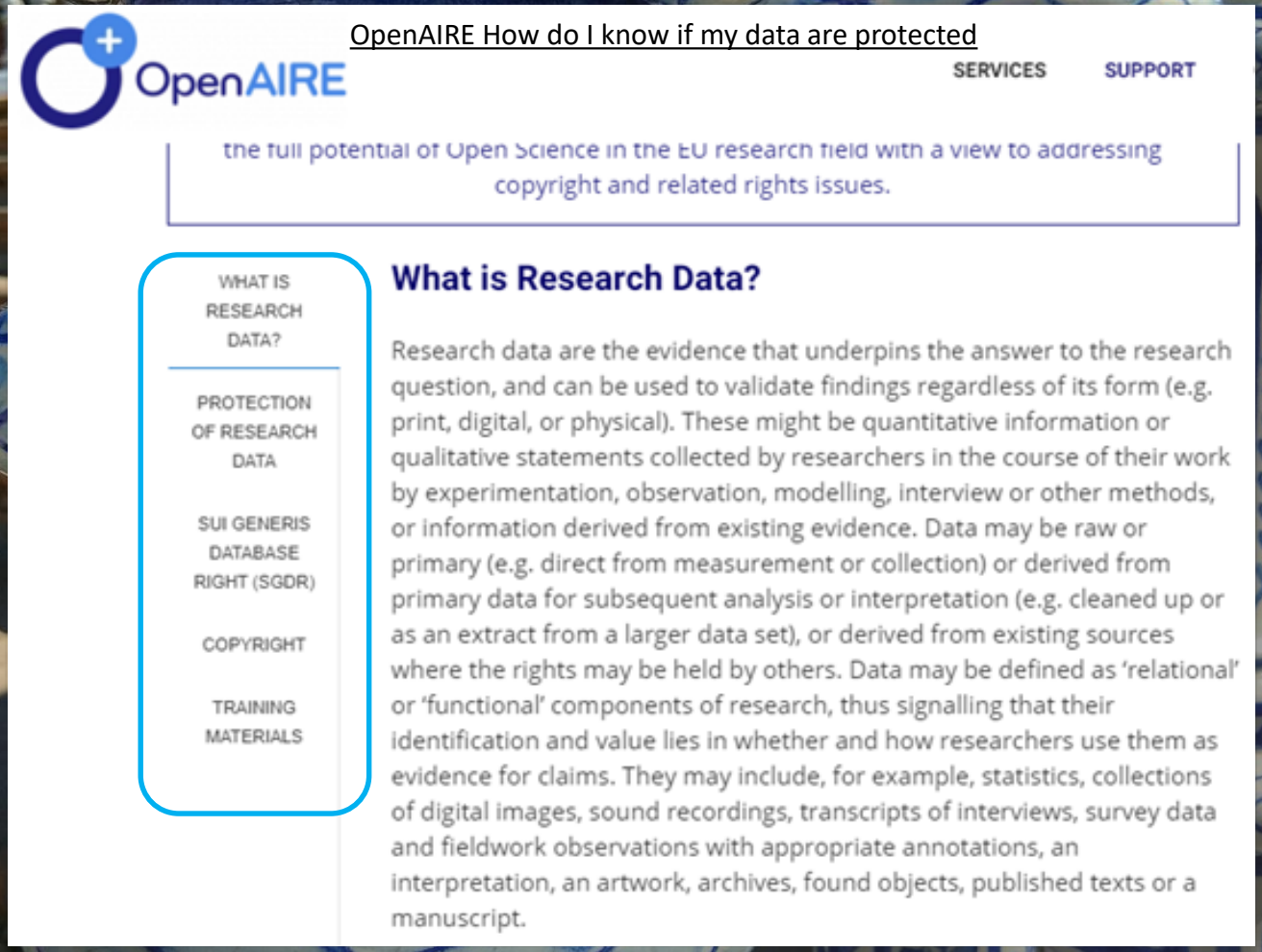
2020

- POSSONO ESSERCI ALTRE FORME DI PROTEZIONE DEI DATI (ES. CONTRATTI)
- PER DATI CHE RICADONO SOTTO GDPR VA SEMPRE ESPLICITATA LA BASE LEGALE SULLA QUALE SI CONDUCE LA RICERCA

2020

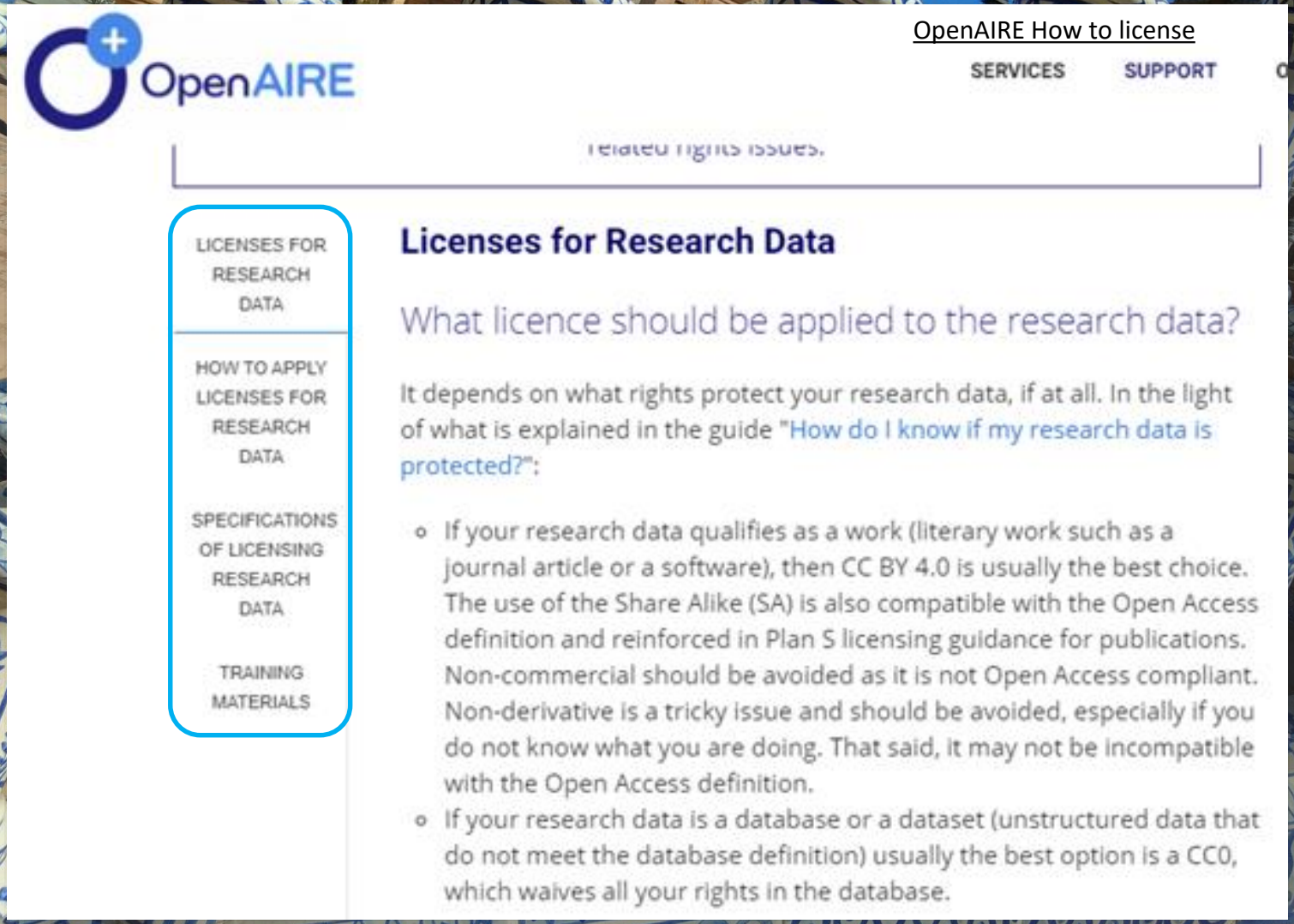
Aspetti legali

OpenAIRE How do I know if my data are protected



Aspetti legali

OpenAIRE How to license

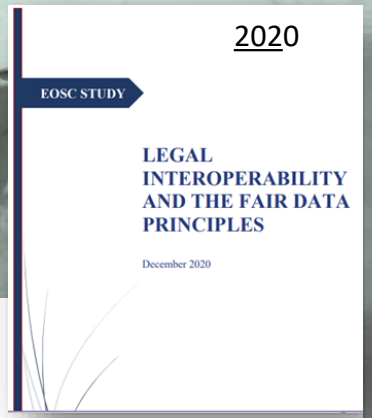


Aspetti legali

OpenAIRE Can I reuse data

- How can a protected dataset be used? +
- Where are licences found? +
- Interoperability and stacking +
- What happens if I use 'Share Alike' (SA) licensed material in my work? Does that mean I have to make my work available under the same SA licence? +
- Can a dataset be used if there is no licence? +
- What are the risks of using a dataset without a licence? +
- Training materials +

Legal interoperability



Chapter two - addresses key thematic legal issues, including:

- Copyright (including database rights) and licences;
- Other intellectual property rights (e.g. patents, trade secrets, neighbouring rights);
- Privacy and data protection (GDPR);
- Other restrictions and legitimate reasons (e.g., protection of sovereign genetic resources and traditional knowledge); and
- Private law considerations.

1. Open access to research data is an enabler of legal interoperability. The promotion of FAIR Principles should go hand-in-hand with efforts to make data open in accordance with the principle that data must be "as open as possible and as closed as necessary".

2. Regardless of whether the data is Open or not, all new data made available through the EOSC should be FAIR by design.

9. Copyrightable **data** should be FAIR and, to the greatest extent possible, be made part of the public domain or assigned a permissive licence, unless legal or legitimate reasons apply. The Creative Commons No Rights Reserved (CC0) or the Public Domain Dedication and Licence (PDDL) or an equivalent statement of rights should be preferred. In cases where liability is a concern that cannot be addressed by other means, the CC BY 4.0 licence is an appropriate alternative.

11. From a licence compatibility perspective, attribution should be pursued by means of moral and ethical obligations e.g., the European Code of Conduct for Research Integrity or the development of Persistent Identifiers, or by way of a standard form of acknowledgement, rather than by means of a licence such as the CC BY 4.0. The CC0 is, in general, preferred over the CC BY 4.0, although both are generally permissive.

Creative Com

CC Factsheet

FACT SHEET ON CREATIVE COMMONS & OPEN SCIENCE V.0.1

This information guide contains questions and responses to common concerns surrounding open science and the implications of licensing data under Creative Commons licences. It is intended to aid researchers, teachers, librarians, administrators and many others using and encountering Creative Commons licences in their work.

What is Open Science?

Open Science is the movement to make scientific research and data accessible to all for knowledge dissemination and public reuse.

How should I licence my data for the purposes of Open Science?

We recommend you use the CC0 Public Domain Dedication, which is first and foremost a waiver, but can act as a licence when a waiver is not possible.

CC ZERO LICENCE, 'NO RIGHTS RESERVED' LOGO

PUBLIC DOMAIN MARK LOGO

By applying CC0 to your data you enable everyone to freely reuse your data as they see fit by waiving (giving up) your copyright and related rights in that data.

You should keep in mind that there are many situations in which data is *not* protected as a matter of law. Such data can include facts, names, numbers – things that are considered 'non-original' and part of the public domain thus not subject to copyright protections. Similarly, your database (which is a structured collection of data) might be considered 'non-original' and thus ineligible for copyright, and it might additionally be excluded from other forms of protection (like the EU *sui generis* database right, also known as the 'SGDR', for non-original databases).

In these cases, using a Creative Commons licence such as a CC BY could signal to users that you claim a copyright in the non-original data despite the law, and perhaps despite your real intention.

Finally, if your data is in the public domain worldwide, you might state simply and obviously on the material that no restrictions attach to the reuse of your data and apply a Public Domain Mark.

When in doubt, consider which use may be appropriate according to the chart below:

CC0 & PUBLIC DOMAIN LICENCES WHICH LICENSE TO USE AND WHEN

CC ZERO	PUBLIC DOMAIN
'Creative arrangement' of data is original, but any copyright has been waived and content is made available copyright-free	'Creative arrangement' of data is not original; the author acknowledges this and communicates the data is in the public domain

What is Open Science?

Open Science is the movement to make scientific research and data accessible to all for knowledge dissemination and public reuse.

How should I licence my data for the purposes of Open Science?

We recommend you use the CC0 Public Domain Dedication, which is first and foremost a waiver, but can act as a licence when a waiver is not possible.

CC ZERO LICENCE, 'NO RIGHTS RESERVED' LOGO

By applying CC0 to your data you enable everyone to freely reuse your data as they see fit by waiving (giving up) your copyright and related rights in that data.

You should keep in mind that there are many situations in which data is *not* protected as a matter of law. Such data can include facts, names, numbers – things that are considered 'non-original' and part of the public domain thus not subject to copyright protections. Similarly, your database (which is a structured collection of data) might be considered 'non-original' and thus ineligible for copyright, and it might additionally be excluded from other forms of protection (like the EU *sui generis* database right, also known as the 'SGDR', for non-original databases).

In these cases, using a Creative Commons licence such as a CC BY could signal to users that you claim a copyright in the non-original data despite the law, and perhaps despite your real intention.

Finally, if your data is in the public domain worldwide, you might state simply and obviously on the material that no restrictions attach to the reuse of your data and apply a Public Domain Mark.

PUBLIC DOMAIN MARK LOGO

When in doubt, consider which use may be appropriate according to the chart below:

CC0 & PUBLIC DOMAIN LICENCES WHICH LICENSE TO USE AND WHEN

CC ZERO	PUBLIC DOMAIN
'Creative arrangement' of data is original, but any copyright has been waived and content is made available copyright-free	'Creative arrangement' of data is not original; the author acknowledges this and communicates the data is in the public domain

nmons e Oper

USARE UNA CC0

- NON SIGNIFICA DIVENTARE ACCADEMICAMENTE MALEDUCATI
- **LA FONTE VA CITATA SEMPRE**
- USATE LA CC0 E ASSOCIATE UNA FORMULAZIONE DELLA CITAZIONE CHE RICHIEDETE (DA COPIARE/INCOLLARE)

But I would like attribution when others use my dataset. In that case, shouldn't I use a CC BY licence?

We recommend that you avoid using a CC BY licence. Here's why:

While attribution is a genuine, recognisable concern, not only might using a CC BY licence be legally unenforceable when no underlying copyright or SGDR protects the work, but it may also communicate the wrong message to the world. A better solution is to use CC0 and simply ask for credit (rather than require attribution), and provide a citation for the dataset that others can copy and paste with ease. Such requests are consistent with scholarly norms for citing source materials.

Legally speaking, datasets that are ***not*** subject to copyright or related rights (and are thus in the public domain) cannot be the object of a copyright licence. Despite this, agreements based in contract law may be enforceable. Creative Commons licences, however, are copyright licences. Therefore, where the conditions for a copyright or related right are not triggered, copyright licences, such as the CC BY licence, are unenforceable.

In some cases, however, rights may exist (like the *sui generis* database right previously mentioned), and permission for others to use your dataset may be legally required. These rights are meant to protect the maker's investment, rather than originality. As such, database rights do not include the moral right of attribution. So by using a CC BY licence, you signal to users that you restrict access to your dataset beyond the protections provided by the law. We are not saying that this cannot be done, we are just saying that if you choose to do this, you should make sure you fully understand what it entails.

It sounds like you're really pushing for the use of CC0 for open science datasets.

Exactly. Data is only open if anyone is free to use, reuse, and distribute it. This means it must be made available for both commercial and non-commercial purposes under non-discriminatory conditions that allow for it to be modified.

When data is made available for all reuse, others can create new knowledge from combining it. This leads to the enrichment of open datasets and further dissemination of knowledge. Accordingly, CC0 is ideal for open science as it both protects and promotes the unrestricted circulation of data.

And remember, it's bad science not to cite the source of data you use. To help others cite your data include a citation that users can copy and paste to give you credit for your hard work.

cannot be done, we are just saying that if you choose to do this, you should make sure you fully understand what it entails.

I'm uncomfortable with others using my research for commercial purposes. Should I use a non-commercial licence for my dataset?

We recommend you avoid using a non-commercial licence. Here's why:

For legal purposes, drawing a line between what is and is not 'commercial' can be tricky; it's not as black and white as you might think. For example, if you release a dataset under a non-commercial licence, it would clearly prohibit an organisation

I'm uncomfortable permitting use of my research for any and all purposes. Should I use a 'No Derivatives' (ND) licence for my dataset?

We recommend you avoid using a 'No Derivatives' licence. Here's why:

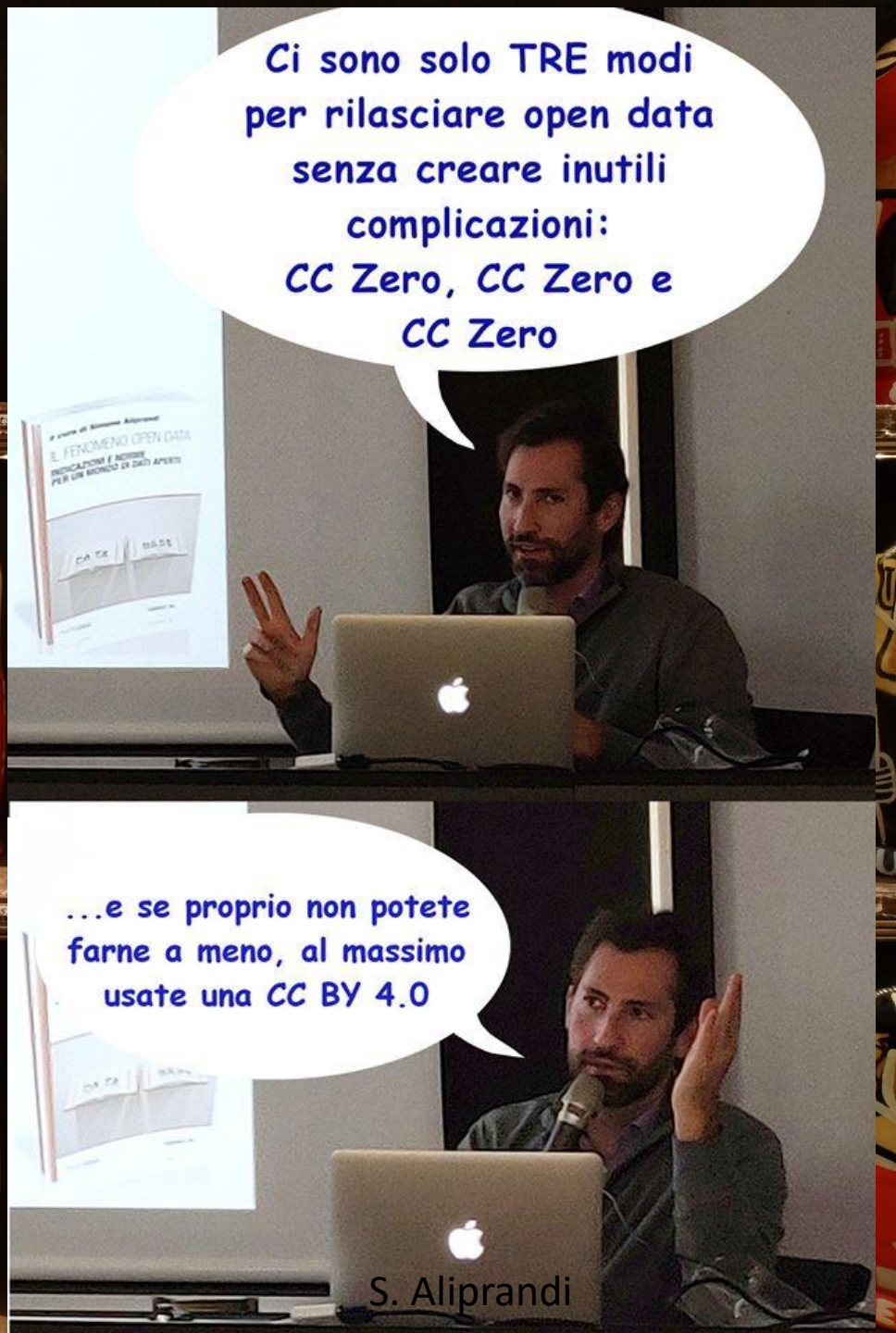
Similar to how a non-commercial licence might restrict meaningful reuse of your dataset, a ND licence can have the same effect: it may prevent someone from recombining and reusing your data for new research. For data to be truly Open Access, it must permit these important types of reuse.

...in sintesi...


12
PATRIOTS
12
9
Green Bay Packers
San Francisco 49ers


2
AUBURN
21
TEXAS A&M
5
TEXAS
25
Ohio State Buckeyes
Auburn Tigers
Texas A&M
Texas Longhorns


SPURS
BULLS
GIANTS


Ci sono solo TRE modi
per rilasciare open data
senza creare inutili
complicazioni:
CC Zero, CC Zero e
CC Zero
...e se proprio non potete
farne a meno, al massimo
usate una CC BY 4.0
S. Aliprandi

[non suoniamo tutti la stessa musica]

Obstacles to the trans-European archiving and sharing of research data

Making research data as openly available as possible is a widely recognised goal. For researchers working on an interdisciplinary project involving several countries, it can be difficult to fully comprehend in which ways open access to research data can be legally obtained. European national laws still diverge.

A report from Knowledge Exchange (Knowledege Exchange, 2011) concludes that it will remain difficult to predict when particular files of research data are protected because of:

- **Diversity in copyright protection**
  Even though most research data will fail to meet the criteria for copyright protection because they are not likely to be considered as "works" (they mainly concern facts), the lack of harmonisation of the criteria for copyright protection in Europe is tricky. E.g., whereas Germany, Denmark and the Netherlands have a relatively similar (higher) originality standard, the UK has a very low standard (skill, judgment and labour) making

- **Diversity in copyright owner**
  If protection applies, the right holder's consent is required for sharing the data. However, the designation of the copyright owner is also different in different jurisdictions. Althoug in many cases the maker of the work will be considered to be the author and therefore th right holder, only Dutch and UK law designate the employer as the right holder if the wor was made in the course of employment.

CESSDA guide

⊕ Czech Republic

⊕ Finland

⊕ Switzerland

CHIARITE SUBITO CON I PARTNER (SOPRATTUTTO ESTERI) SE CI SONO OBBLIGHI DIVERSI DI LEGGE O REGOLAMENTI INTERNI DA RISPETTARE

[regole chiare dall’inizio]

… FISSATE REGOLE CHIARE DA SUBITO

- CHI HA I DIRITTI SUI DATI (SE ESISTE)
- CHI DEVE PROVVEDERE A CONSERVARE
- CHI PUÒ SFRUTTARE

Licenze per software

ATTENZIONE A POSSIBILI RESTRIZIONI

Open Software Licenses

Using these preferred licenses
saves you a lot of extra work.
Also, legal restrictions may apply.
First,
software that is published under an open license
is actually exported to all countries in the world.
So, by sharing your software openly,
you might be in conflict with an export restriction
to a specific country.
Second,
if your software
could also be applicable
to the production of weapons of mass destruction,
you are not allowed to publish it.
Third,

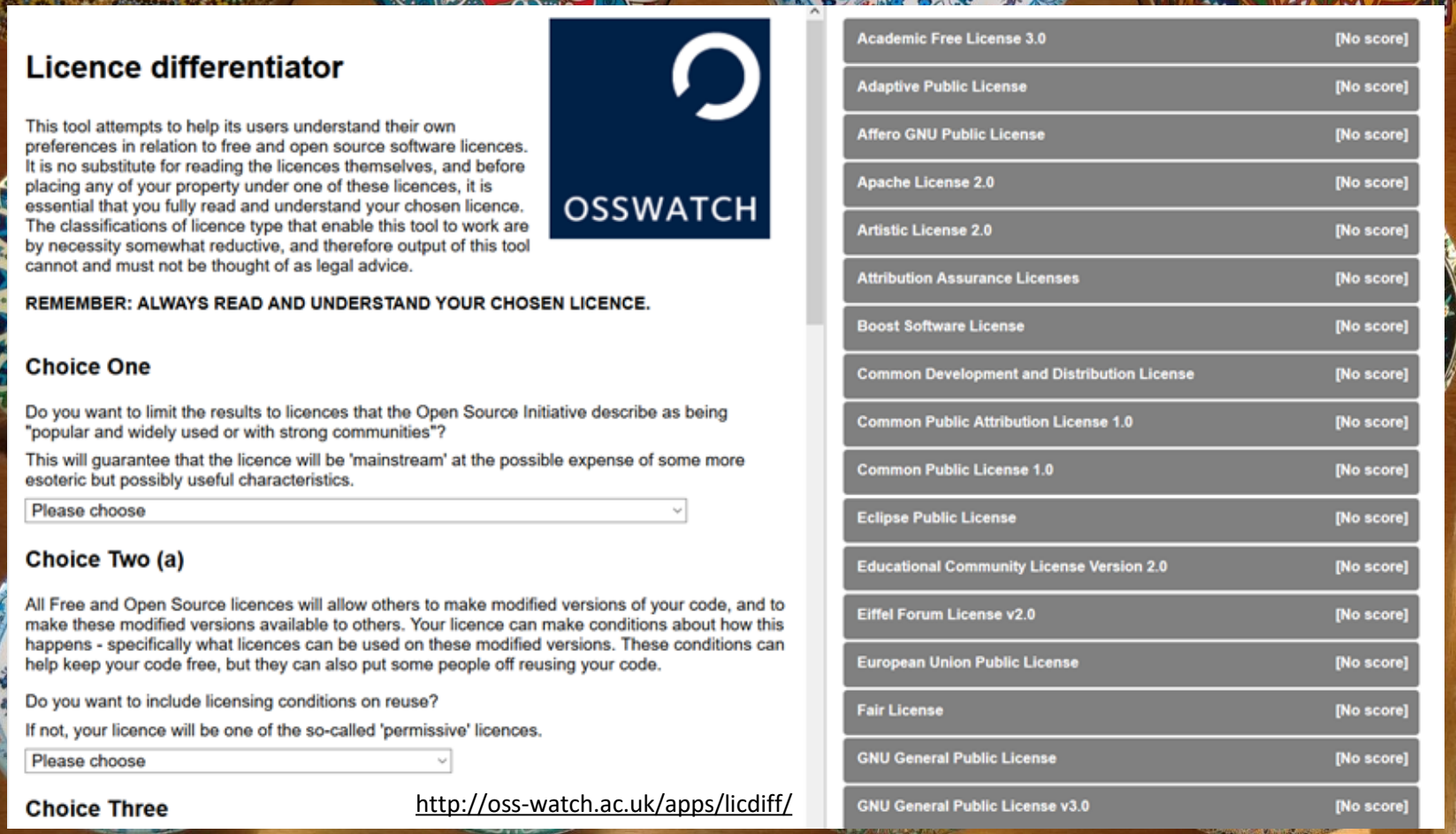
Meta Keijzer - de Ruijter, Project Manager at TUDelft – EdX MOOC on Open Science

Licenze per software 2

	Free			Non-free		
	Public domain & equivalents	Permissive license	Copyleft (protective license)	Noncommercial license	Proprietary license	Trade secret
Description	Grants all rights	Grants use rights, including right to relicense (allows proprietization, license compatibility)	Grants use rights, forbids proprietization	Grants rights for noncommercial use only. May be combined with share-alike.	Traditional use of copyright; no rights need be granted	No information made public
Software	PD, Unlicense CC0	BSD, MIT, Apache, MPL	GPL, AGPL	JRL, AFPL	Proprietary software, no public license	Private, internal software
Other creative works	PD, CC0	CC-BY	CC-BY-SA	CC-BY-NC	Copyright, no public license	Unpublished

https://en.wikipedia.org/wiki/Copyleft

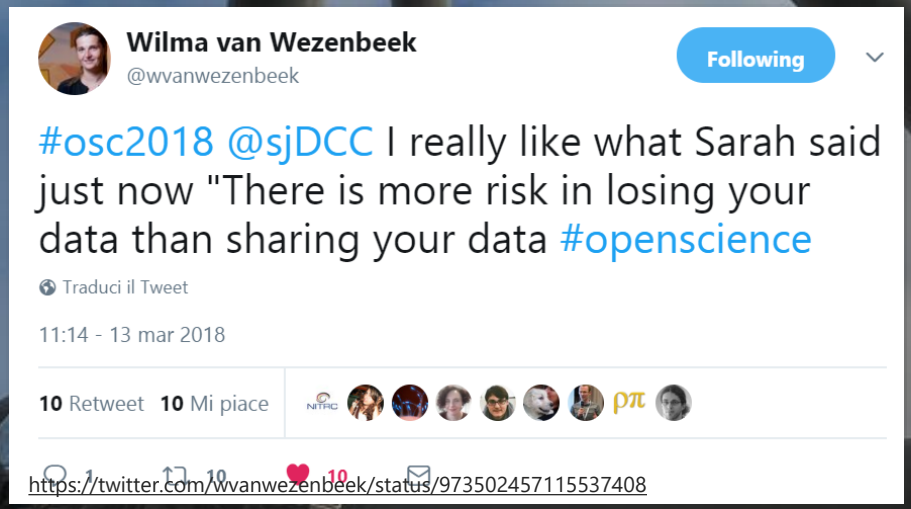
Licenze per software 3



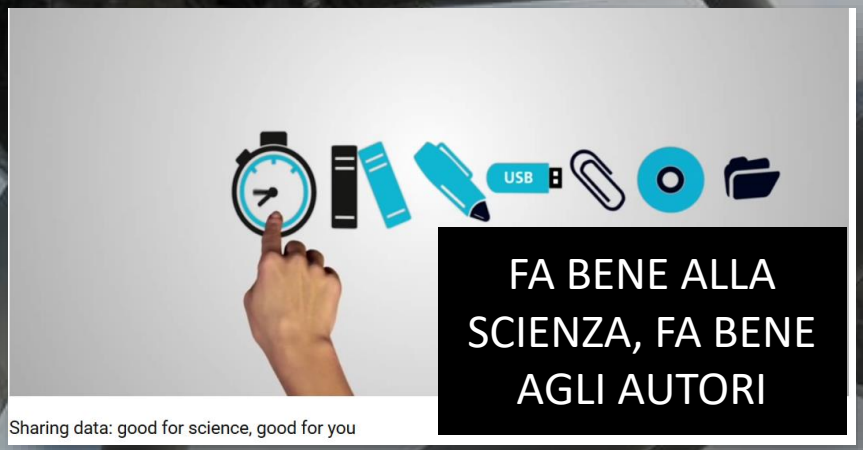
http://oss-watch.ac.uk/apps/licdiff/

3. DATI APERTI

Perché i dati aperti?

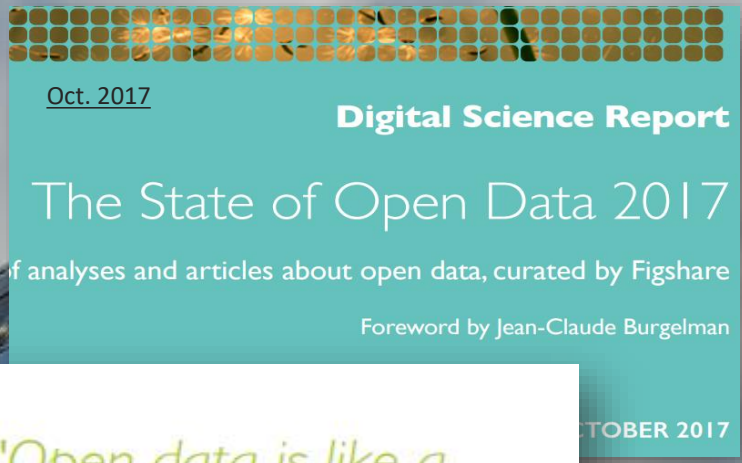


https://twitter.com/wvanwezenbeek/status/973502457115537408



FA BENE ALLA SCIENZA, FA BENE AGLI AUTORI

https://www.youtube.com/watch?v=[illegible]&feature=youtu.be



"Open data is like a renewable energy source: it can be reused without diminishing its original value, and reuse creates new value."

People will contact me to ask about stuff

Christopher and Alex (C&A) say: "This is usually
objection of people who feel overworked and th
[data sharing] isn't part of their job…" I would
learning from each other – if a researcher is opp
datasets, collaborating with others, and genera
they should be outed by their community as a p

COMUNICAZIONE SCIENTIFICA= GRANDE CONVERSAZIONE

People will misinterpret the data

C&A suggest this: "Document how it should be interpreted. Be prepared to help and correct such people; those that mis the help." From the UK Data Archive: "P contextual information for your res correctly use and understand your

IMPOSSIBILE, SE NELLA «R» DI FAIR AVETE DOCUMENTATO

It's worth mentioning, however, a s be useful to counter willful misrepr of Information legislation), as one refute the wrong interpretation."

My data is not very interesting

C&A: "Let others judge how intere people that care about them." I'd dataset has value to future researc "climate change" was a research t documenting and understanding

EHM… ALLORA PERCHÉ USA FONDI PUBBLICI?

I might want to use it in a research paper

Anyone who's discussed data sharing with a researcher is familiar with this excuse. The operative word here is *might*. How many papers have we all considered writing, only to have them shift to the back bur real concern.

EMBARGO PERFETTAMENTE COMPATIBILE CON FAIR

C&A suggest the embargo route: "One embargo; require people to archive the public after X months. You could even things that are no longer cared about eventually everything can become ope their datasets, but I would caution to have any restrictions default to sharing. That is, after X months the data are automatically made open by the repository.

I would also add that, as the original collector of the data, you are at a huge advantage compared to others that might want to use your dataset. You have knowledge about your system, the conditions during collection, the nuances of your methods, et cetera that could never be fully described in the best metadata.

I'm not sure I own the data

My data is too complicated.

C&A: "Don't be too smug. If it turns out it's not that comp professional [standing]." I would add that if it's too comp complicated to reproduce, which means it's arguably not can be solved by more documentation.

IMPOSSIBILE, SE NELLA «R» DI FAIR AVETE DOCUMENTATO

My data is embarrassingly bad

C&A: "Many eyes will help you imp accept your data for what it is." I a making the sausage. We know it's Plus it helps you strive will be at m collection phase.

MA COME SI FA A FARE RICERCA CON DATI «BRUTTI»???

It's not a priority and I'm busy

CARLY STRASSER
Open Science + Data Sharing Advocate
Home
Blog Posts
Previous Research

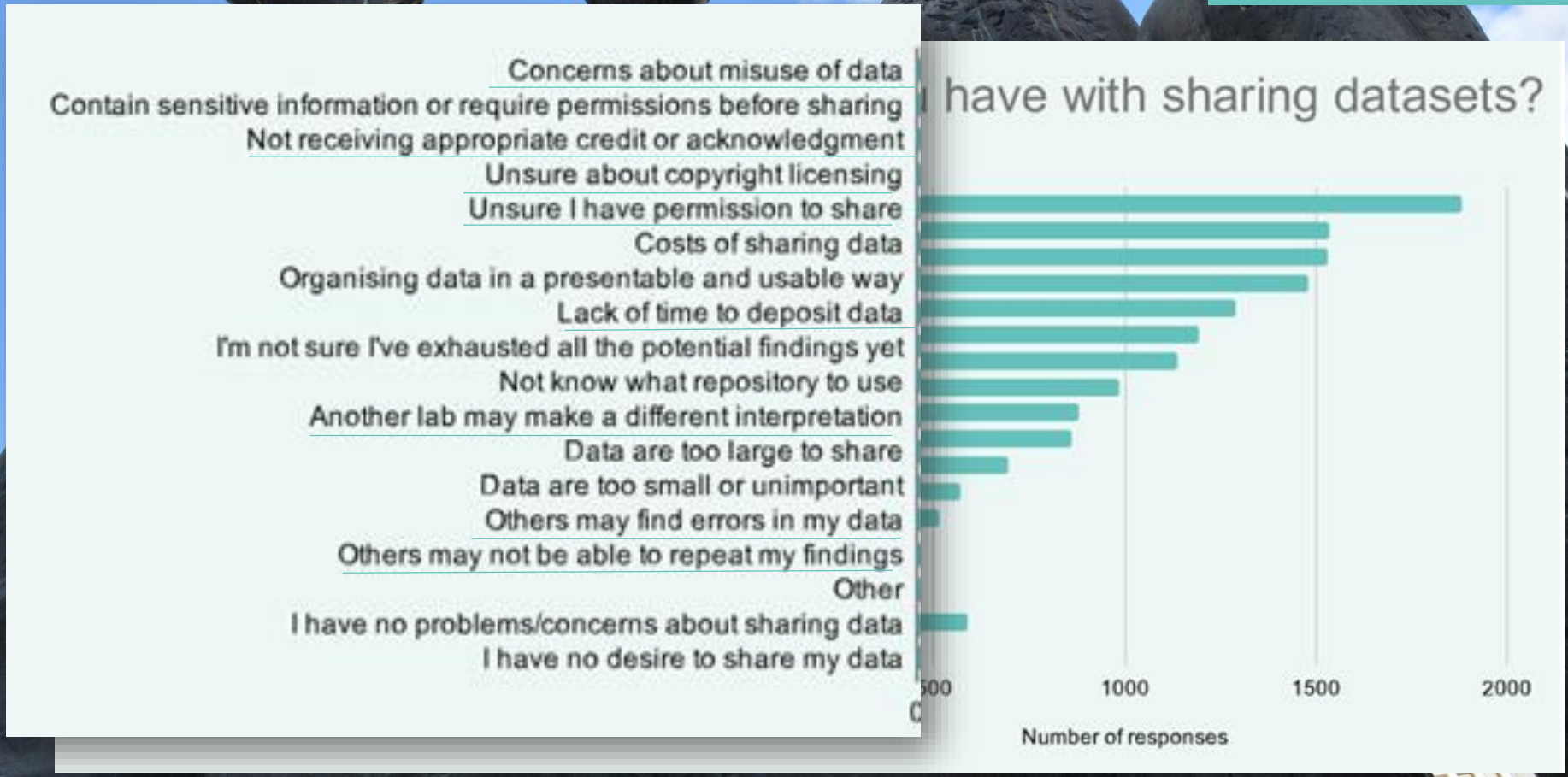
http://carlystrasser.net/closed-data-excuses-excuses/

Closed Data… Excuses, Excuses

SEMPRE MAGGIOR NUMERO DI FINANZIATORI E DI RIVISTE LO CHIEDE…
STA DIVENTANDO UNA PRIORITÀ

[scuse, scuse… aggiornato]

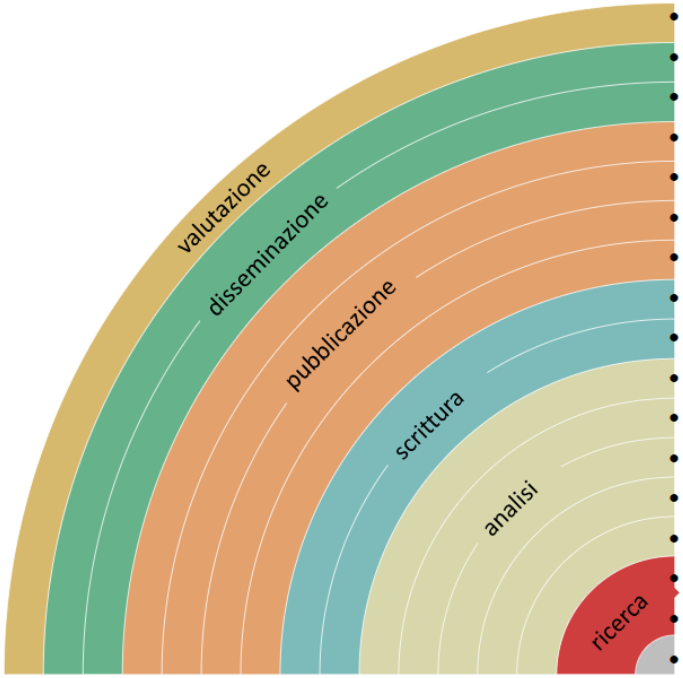
Dec.2020



… «as open as possible»…

Open science: perché solo i dati?

Come puoi rendere Open ogni passo della ricerca…



- aggiungendo misure di impatto alternative, es. altmetrics
- comunicando sui social media, es. Twitter
- condividendo poster e presentazioni, es. su FigShare
- utilizzando licenze aperte, es. Creative Commons BY
- depositando in archivi o pubblicando su riviste Open
- provando la open peer review, es. PubPeer o F1000
- condividendo preprints, es. su OSF, arXiv o bioRxiv
- con formati leggibili dalle macchine, es. Jupyter o CoCalc
- con la scrittura collaborativa, es. Overleaf o Authorea
- condividendo protocolli e workflow, es. su Protocols.io
- condividendo note di laboratorio, es. OpenNotebookScience
- condividendo software, es. su GitHub con licenza GNU/MIT
- condividendo i dati, es. su Dryad, Zenodo o Dataverse
- pre-registrando esperimenti, es. su OSF o AsPredicted
- commentando pagine web, es. su Hypothes.is o Pund.it
- usando bibliografie condivise, es. su Zotero
- condividendo progetti di ricerca, es. su RIO Journal

ORIZZONTE DI RIFERIMENTO RESTA LA OPEN SCIENCE…
TUTTO IL CICLO VA APERTO, NON SOLO I DATI

UN MODO STRUTTURATO DI PENSARE AI DATI

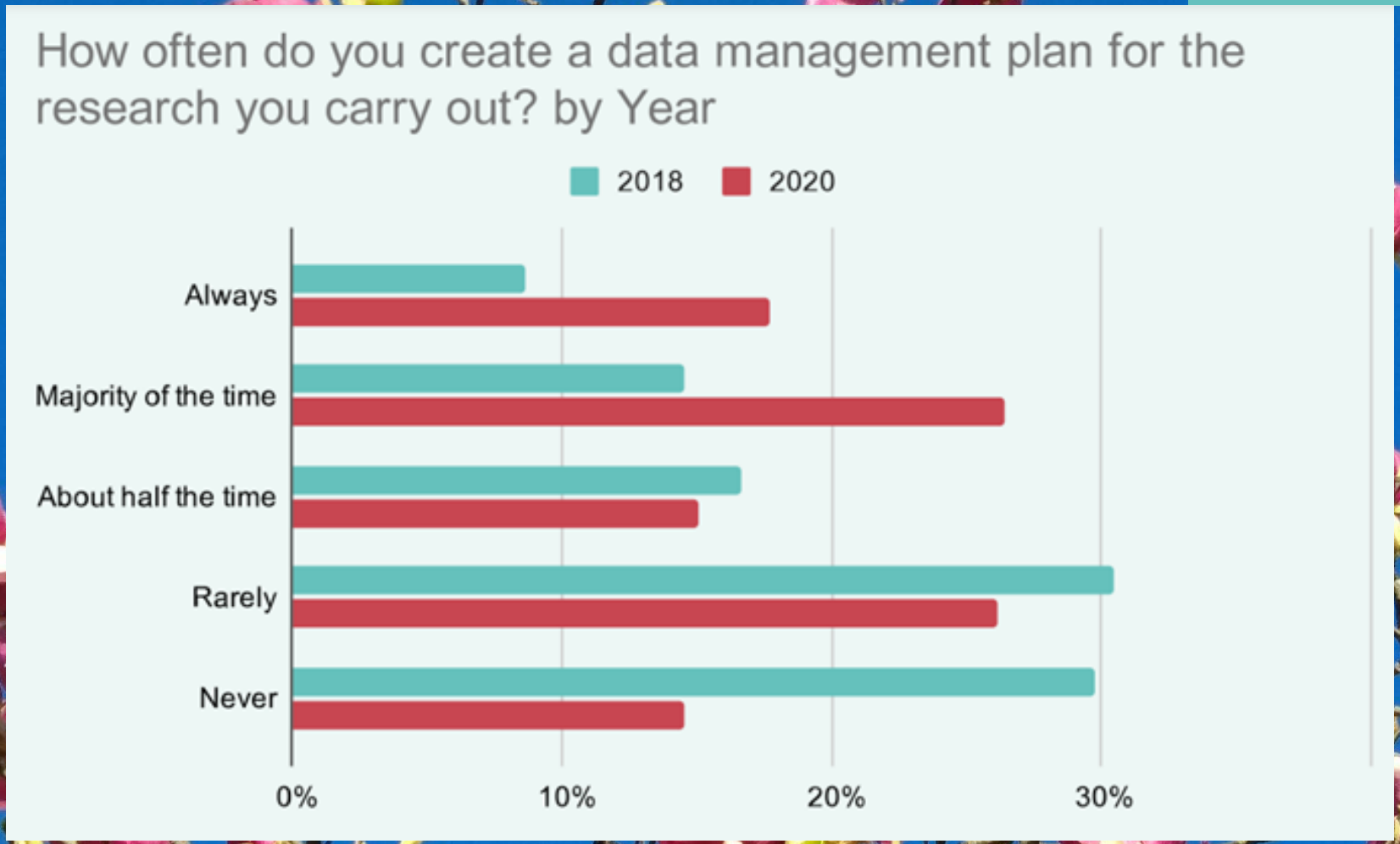
REGOLE CHIARE=MENO ERRORI DA SUBITO

È UN «LIVING DOCUMENT», CRESCE COL PROGETTO

DOVE METTERE TUTTE QUESTE INFORMAZIONI? NEL DATA MANAGEMENT PLAN

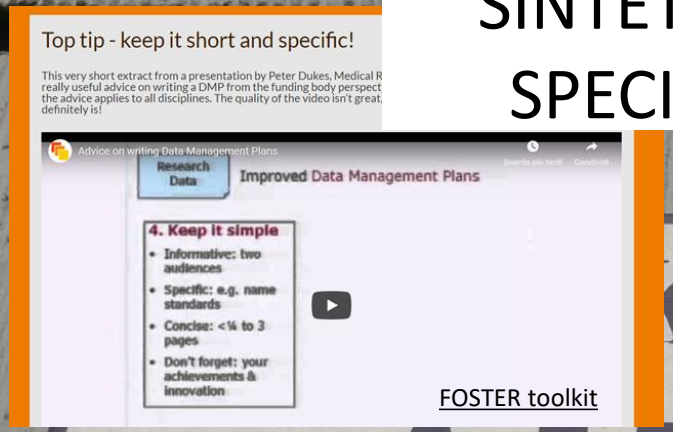
DMP in crescita

Dec.2020



Trucchi e suggerimenti

SINTETICO E SPECIFICO



NON COPIATE

OGNI DATASET È UNICO, OGNI INFRASTRUTTURA È UNICA, OGNI RICERCA HA LA SUA IMPOSTAZIONE

- USATE TABELLE, ELENCHI PUNTATI
- SIATE SCHEMATICI E NON DILUNGATEVI

ESSERE GENERICI NON SERVE A NULLA
[we expect a huge size of data; data will be available]

CIÒ CHE DICHIARATE NEL DMP POI VA FATTO VERAMENTE… QUINDI

A) NON FATE GLI SPLENDIDI
B) NON IMPEGNATEVI A FARE COSE CHE SAPETE IMPOSSIBILI

Es. DATI PSEUDONIMYZED, non ANONIMYZED

Tips&tricks / 2

DMP tips

Vantaggi d

A Review of the History, Advocacy and Efficacy of Data Management Plans 2017

A data planning process ensures that all aspects of data management are holistically explored at the start of a project. Short-term and long-term aims can be balanced, so that decisions made early in a project do not negatively impact on the ability to find and use the research data in future.

Effective management of data provides researchers with many benefits, including

- time saved through reduced duplication of effort
- decreased risk of loss, theft or inappropriate use of data
- good research practice ensures the integrity and quality of data
- data can be understood and used now and in the future

EVITA SFORZI DI DUPLICAZIONE
EVITA LA PERDITA
RISOLVE I CONFLITTI

- helps researchers find and gain access to data management – expertise and infrastructure offered at the University
- increased researcher profile through data dissemination and re-use.

A data planning process is particularly important in the context of collaborative research projects. Researchers may identify areas of potential difficulty or conflict, and these can be resolved with colleagues and collaborators before they escalate into issues. Clarifying ownership of data, and ensuring early agreement on technical standards and frameworks across institutions, are an important part of establishing trust and ensuring that a project runs smoothly.

Vantaggi di un DMP

CESSDA Guide

Benefit 3. Clarifies needed budget

Data management is not free. You do not want to find yourself running out of funding before the end of the project because you have ignored or underestimated the cost of structured, detailed, and safe data management. Therefore, an important aspect of a DMP is its use in calculating how much money will be required for managing your research data during your research project.

A DMP can be useful in the process of applying for funding. Grant applications should not only include time and resources for collecting, analysing, and publishing on data in their budget, time and resources for careful documentation as well as server space, backup solutions, and documentation software need to be included as well. A DMP is also useful once funding is granted to plan and manage your expenses. Many research funders require a DMP as part of the application and decision-making process. The arguments for making data available are several, the most popular being that the data produced by public funds should be used to the greatest extent possible and available to the public. Unless there are legal, ethical or commercial barriers, data should also be openly available so that research results can be verified, replicated and reused.

Examples of Data Management cost assessments are given by the University of Utrecht (n.d.) and the Dutch Landelijk Coördinatiepunt Research Data Management (LCRDM, n.d.) inspired by the 'Data management costing tool' by UK Data Service, 2013.

È FONDAMENTALE PER STIMARE I COSTI DI GESTIONE - STIMATE LA DATA STEWARDSHIP (IN-KIND?) -POSSIBILI COSTI DI STORAGE COSTI SONO RIMBORSABILI IN H2020 (6.2.D.3 AMGA)

Costs - data curation

Costs related to the implementation of the Open Research Data pilot *(e.g. costs for providing open access, related research data management costs, data curation and data storage costs)* may be eligible *(see Article 6.2.D.3)*.

ELIGIBLE COSTS:

- DATA CURATION
- DATA STORAGE
- DATA MANAGEMENT

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf#page=238

Guida al DMP

SCIENCE EUROPE

Jan. 27, 2021

Contact us | Member log in |

ABOUT US OUR PRIORITIES WHAT'S GOING ON OUR RESC

> Our resources

27.01.2021

Practical Guide to the International Alignment of Research Data Management – Extended Edition

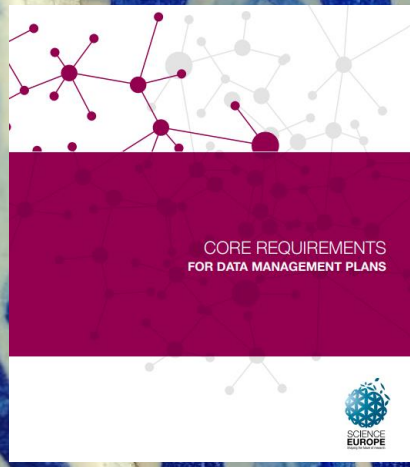
This resource offers targeted guidance for organisations, scientific communities, as well as individual researchers, to organise research data and preserve it appropriately

Originally released in 2019, and following its successful uptake by many organisations, the extended edition features a brand-new rubric to facilitate the evaluation of a data management plan (DMP). The guide also presents core requirements for DMPs, criteria for the selection of trustworthy repositories, and guidance mply with organisational requirements.

EDIZIONE AGGIORNATA DELLA GUIDA 2018

DMP Core Requirements

Jan. 27, 2021



CORE REQUIREMENTS
FOR DATA MANAGEMENT PLANS

When developing solid data management plans, researchers are required to deal with the following topics and answer the following questions:

1. Data description and collection or re-use of existing data

a. How will new data be collected or produced and/or how will existing data be re-used?

b. What data (for example the kinds, formats, and volumes) will be collected or produced?

2. Documentation and data quality

a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?

b. What data quality control measures will be used?

3. Storage and backup during the research process

a. How will data and metadata be stored and backed up during the research process?

b. How will data security and protection of sensitive data be taken care of during the research?

4. Legal and ethical requirements, codes of conduct

a. If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?

b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

c. How will possible ethical issues be taken into account, and codes of conduct followed?

SEZIONI MINIME IN UN DMP

5. Data sharing and long-term preservation

a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

b. How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?

c. What methods or software tools will be needed to access and use the data?

d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

6. Data management responsibilities and resources

a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?

b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

DMP Core Requirements

Translating the Core Requirements into a DMP template

The following example of a data management plan template is based on the core requirements for DMPs.[6] These core requirements should be considered as a minimum standard, leaving the flexibility to formulate additional guidelines according to the needs of specific domains or to national or local legislation.

The template presented below refers to the 15 questions covering six core requirements for good data management. Additional guidance and explanations are provided to help researchers fill out such a template and to assure that all relevant aspects of research data management are covered. The below table is an example of how the core requirements can be transformed into a DMP template. It will be up to the individual organisations and disciplines to develop templates that fit their needs.

GENERAL INFORMATION	
Administrative information	• Provide information such as name of applicant, project number, funding programme, version of DMP.
1 DATA DESCRIPTION AND COLLECTION OR RE-USE OF EXISTING DATA	
1a **How will new data be collected or produced and/or how will existing data be re-used?**	• Explain which methodologies or softwa will be used if new data are collected or produced. • State any constraints on re-use of existing data if there are any. • Explain how data provenance will be documented. • Briefly state the reasons if the re-use of any existing data sources has been considered but discarded.

2 DOCUMENTATION AND DATA QUALITY	
2a **What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?**	• Indicate which metadata will help others identify and discover the data. • Indicate which metadata standards (for example DDI, TEI, EML, MARC, CMDI) will be used. • Use community metadata standards where these are in place. • Indicate how the data will be organised during the project, mentioning for example conventions, version control, and folder structures. Consistent, well-ordered research data will be easier to find, understand, and re-use. • Consider what other documentation is needed to enable re-use. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, and so on. • Consider how this information will be captured and where it will be recorded for example in a database with links to each item, a 'readme' text file, file headers, code books, or lab notebooks.
2b **What data quality control measures will be used?**	• Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeated samples or measurements, standardised data capture, data entry validation, peer review of data, or representation with controlled vocabularies.

UTILI COME SPUNTO PER LE PRIME VOLTE, PER CAPIRE QUALI DOMANDE FARSI

Jan. 27, 2021

DMP – Rubric for evaluation

3 STORAGE AND BACKUP DURING THE RESEARCH PROCESS			
Guidance for Researchers		**Sufficiently Addressed The DMP…**	**Insufficiently Addressed The DMP…**
3a **How will data and metadata be stored and backed up during the research?**	• Describe where the data will be stored and backed up during research activities and how often the backup will be performed. It is recommended to store data in least at two separate locations. • Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of the home institution. Storing data on laptops, stand-alone hard drives, or external storage devices such as USB sticks is not recommended.	• Clearly (even if briefly) describes: › The location where the data and backups will be stored during the research activities. › How often backups will be performed. › The use of robust, managed storage with automatic backup (for example storage provided by the home institution). or • Explains why institutional storage will not be used (and for what part of the data) and describes the (additional) locations, storage media, and procedures that will be used for storing and backing up data during the project.	• Provides no information or very vague reference to how data will be stored and backed up during the project.

1 D			
Guidance for Researchers		**Sufficiently Addressed The DMP…**	**Insufficiently Addressed The DMP…**
1a **How will new data be collected or produced and/or how will existing data be re-used?**	• Explain which methodologies or software will be used if new data are collected or produced. • State any constraints on re-use of existing data if there are any. • Explain how data provenance will be documented. • Briefly state the reasons if the re-use of any existing data sources has been considered but discarded.	• Gives clear details of where the existing data come from and how new data will be collected or produced. It clearly explains methods and software used. • Explains, if existing data are re-used, how these data will be accessed and any constraints on their re-use.	• Provides little or no details on where the data come from and what data will be collected or re-used. • Does not, if applicable, provide sufficient rationale for generating new data.

DMP questions

CESSDA guide

Adapt your Data Management Plan

A list of Data Management Questions based on the Expert Tour Guide on Data Management

PLAN

Overview

Title of the project

Date of this plan

Description of the project
- What is the nature of the project?
- What is the research question?
- What is the project time line?

Origin of Data
- What kind of data will be used during the project?
- If you are reusing existing data: What is the scope, volume and format? How are different data sources integrated?
- If you are collecting new data can you clarify why this is necessary?

Principal researchers
- Who are the main researchers involved?
- What are their contact details?

Collaborating researchers (if applicable)
- What are their contact details and their roles in the project?

Funder (if applicable)
- If funding is granted, what is the reference number of the funding granted?

Data producer
- Which organisation has the administrative responsibility for the data?

Project data contact
- Who can be contacted about the project after it has finished?

Data owner(s)
- Which organisation(s) own(s) the data?
- If several organisations are involved, which organisation owns what data?

Roles
- Who is responsible for updating the DMP and making sure that it's followed?
- Do project participants have any specific roles?
- What is the project time line?

Costs
- Are there costs you need to consider to buy specific software or hardware?
- Are there costs you need to consider for storage and backup?
- Are potential expenses for (preparing the data for) archiving covered?

UTILI COME SPUNTO PER LE PRIME VOLTE, PER CAPIRE QUALI DOMANDE FARSI

ORGANISE & DOCUMENT

Organising and documenting your data

Data collection
- How will the data be collected?
- Is specific software or hardware or staff required?
- Who will be responsible for the data collection?
- During which period will the data be collected?
- Where will the data be collected?

Data organisation
- How will you organise your data?
- Will the data be organised in simple files or more complex databases?
- How will the data quality during the project be ensured?
- If data consists of many different file types (e.g. videos, text, photos), is it possible to structure the data in a logical way?

Data type and size
- What type(s) of data will be collected?
- What is the scope, quantity and format of the material?
- After the project: What is the total amount of data collected (in MB/GB)?

File format
- In what format will your data be?
- Does the format change from the original to the processed/final data?
- Will your (final) data be available in an open format?

Folder structure and names
- How will you structure and name your folders?

File structure and names
- How will you structure and name your files?

Documentation
- What documentation will be created during the different phases of the project?
- How will the documentation be structured?

Metadata
- What metadata will be provided with the collected/ generated/ reused data?
- How will metadata for each object be created?
- Is there any program that can be used to document the data?
- Can metadata be added directly into the files or will the metadata be produced in another program or document?

Metadata standard (if applicable)
- What metadata standard(s) will you use?

DMP questions

cessda
Consortium of European Social Sciences Data Archives

PROTECT

Protecting your data

Ethical review (if applicable)
- Does your project require approval by a local ethics committee?

Informed consent (if applicable)
- Do you require informed consent for your project?
- If so, how will permission be obtained?
- How are consent files organised and stored?

(sensitive) Personal data /confidential information (if applicable)
- How will access to (sensitive) personal data during the project be controlled?
- How will collaborators be granted access to the data in a secure way?
- If the research project is going to have data that includes confidential information or information that requires informed consent, is there a requirement to notify a privacy officer?
- Is there any confidential information within the material that requires special treatment and/or limits the access to it during/after the project?
- How will the material be protected during/after the project?
- How will permissions and restrictions be enforced?

Intellectual property rights (IPR)/Copyrights
- Are there IPR or copyright issues to consider?
- Will permission be needed to collect/reuse the data?
- Will these rights be transferred to another organisation for data distribution and archiving?

Agreements (if applicable)
- What are the agreements with other stakeholders?

Restrictions (if applicable)
- ny other restrictions that need to be considered?

PROCESS

Processing your data

Versioning
- What is your strategy concerning versioning your data files (and scripts) during the project?
- Will you create and/or follow a convention for versioning your data?
- Who will be responsible for securing that a "Masterfile" will be maintained, documented and versioned according to the project guidelines?
- How can different versions of a data file be distinguished?

Interoperability
- Will you make use of established software and hardware? If not, how does the software and hardware you use relate to other research?

If applicable:

- Will you make use of established terminologies/ontologies (i.e. structured controlled vocabularies) in the project? If not, how do your terminologies relate to established ones?
- Which coding is used (if any)? Will you build on established coding schemes? If not, how does your coding relate to other research?

STORE

Storing your data

Storage
- How and where will the data be stored during the project?
- For how long will the data be stored?

Backup
- How, where and at what intervals will the data be backed-up?
- How will data be recovered in the case of a data loss incident?

Security
- How will sensitive data be protected? (if applicable)
- How will data access be managed?

UTILI COME SPUNTO PER LE PRIME VOLTE, PER CAPIRE QUALI DOMANDE FARSI

Basic Information.

- State the purpose of the data collection/generation.
- Explain the relation to the objectives of the project
- Consider what data will be collected or created as part of the study (RAW data).
- Consider what data will be produced by processing the RAW data (Secondary, processed data).
- Specify if existing data is being re-used (if any)
- Specify the origin of the data
- Specify the types and formats you plan to use for the data generated/collected (raw, processed, published).
- Consider what data will be published as the result of your study (Published data).

Volume and Life Cycle of the Data.

If you are using FAIRDOM, we will look after data that will be retained and potentially exchanged by your projects. It will help with local storage for temporarily-held local data prior to processing.

For RAW data, please consider the following:

- How much RAW data you think will be produced (Estimates, per month, year, full project duration)?
- Will all of the RAW data be kept for the duration of the study or will the RAW data be deleted once it is processed?
- For large scale RAW data (images, sequence) have you planned the local storage capacity necessary for processing?
- Do you require help to organise a suitable local management system for RAW data?
- Do you have policies that govern the management and usage of RAW data?
- How long will RAW data be kept? Will there be a long-term archive?

FAIRDOM About Platform

Data Management Checklist

https://fair-dom.org/knowledgehub/data-management-checklist/

UTILI COME SPUNTO PER LE PRIME VOLTE, PER CAPIRE QUALI DOMANDE FARSI

For Secondary and Published data, please consider the following:

- What data processing is foreseen in the project?
- How much processed data will be produced, and stored (can you make estimates per month, year, full project)?
- How much of this data will be published? (Estimates per month, year, full project)?
- Does your institution, or the project funders, have policies governing the access and usage of processed data?

Additional for personally sensitive data (e.g medical data)

- When looking at the data flow through the project, define what data is:
  - aggregated (typically safe to share, if names cannot be recovered)
  - anonymized (name cannot be recovered from the data)
  - pseudonymized (name can be recovered by some)
  - non-anonymized (name linked to data)
- Determine which organisational boundaries have to be traversed by which data.
- Make sure with your *local* data protection officer and ethics commission that the data can be shared with your partners along the flow described with the anonymisation levels as described. Why local? Some laws change across surprising boundaries. E.g. in Germany Universities and other public organisations are subject to another data protection law than enterprises. Why seek advice? In some cases you may be required to be able to recover the name-data-relation, e.g. to enable study participants to *leave* a study.

DMP online

PERSONALIZZABILE

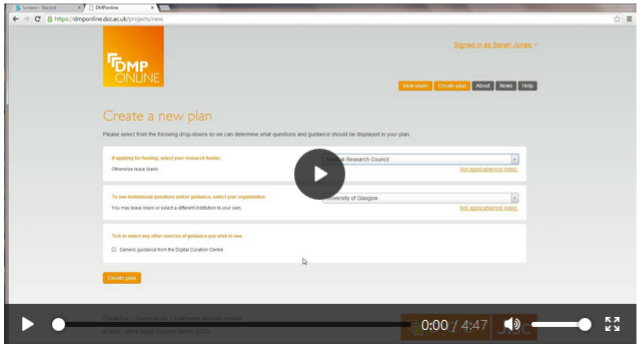
DMP ONLINE

Home | About | Future plans | Help | Change language

Welcome.

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It has been jointly developed by the Digital Curation Centre (DCC) and the University of California Curation Center (UC3).

Screencast on how to use DMPonline



Sign in

Veteran tapes

Project Details | Plan overview | Write Plan | Share | Download

expand all | collapse all

13/13 answered

Data Collection (2 / 2)

What data will you collect or create?

The "Veteran tape " project will collect and generate different types of datasets:

Type of data	Volume	Format	Storage format
Video recordings	600 x 1Gb	.mkv	.mkv
Transcriptions	600 x 1500Kb	MS Word	.txt
Structured interview text	1 x 500Kb	MS word	.txt

For the video recordings the selected format is .mkv; the same .mkv format will be used for the long-term preservation .

Transcriptions will be written in MS Word and then stored as .txt files.

We checked the format compatibility against EASY File format https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats

As the total volume of data is greater than 50Gb, DANS requires a fee for the storage. We are currently in touch with EASY to determine the costs of archiving.

Save

Guidance | Comments (1)

DCC | DCC guidance

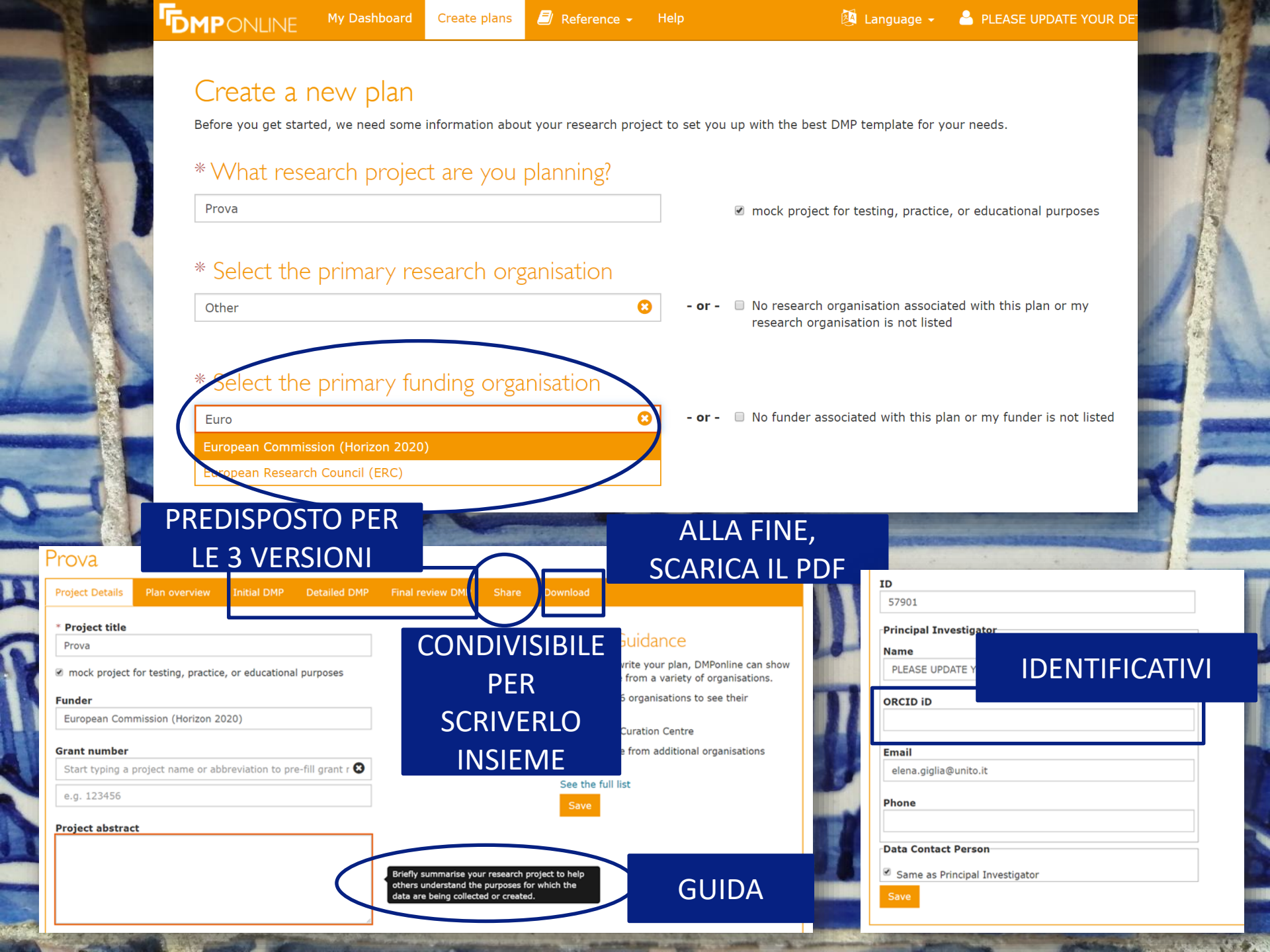
Guidance
Questions to consider:

- What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to the data?
- Are there any existing data that you can reuse?

Guidance:

Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.

GRATUITO
BASTA REGISTRARSI; POI SI ACCEDE E SI TROVANO TUTTI I PROPRI DMP NEL DESKTOP


DMP ONLINE
My Dashboard
Create plans
Reference
Help
Language
PLEASE UPDATE YOUR DE
Create a new plan
Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.
* What research project are you planning?
Prova
mock project for testing, practice, or educational purposes
* Select the primary research organisation
Other
- or -
No research organisation associated with this plan or my research organisation is not listed
* Select the primary funding organisation
Euro
European Commission (Horizon 2020)
European Research Council (ERC)
- or -
No funder associated with this plan or my funder is not listed
PREDISPOSTO PER LE 3 VERSIONI
ALLA FINE, SCARICA IL PDF
Prova
Project Details
Plan overview
Initial DMP
Detailed DMP
Final review DM
Share
Download
* Project title
Prova
mock project for testing, practice, or educational purposes
Funder
European Commission (Horizon 2020)
Grant number
Start typing a project name or abbreviation to pre-fill grant r
e.g. 123456
Project abstract
CONDIVISIBILE PER SCRIVERLO INSIEME
Guidance
rite your plan, DMPonline can show
from a variety of organisations.
organisations to see their
Curation Centre
from additional organisations
See the full list
Save
Briefly summarise your research project to help others understand the purposes for which the data are being collected or created.
GUIDA
ID
57901
Principal Investigator
Name
PLEASE UPDATE Y
ORCID iD
Email
elena.giglia@unito.it
Phone
Data Contact Person
Same as Principal Investigator
Save
IDENTIFICATIVI


Project Details
Plan overview
Initial DMP
Detailed DMP
Final review DMP
Share
Download
Set plan visibility
Public or organisational visibility is intended for finished plans. You must answer at least 50% of the questions to enable these options. Note: test plans are set to private visibility by default.
Private: visible to me, specified collaborators and administrators at my organisation
Organisation: anyone at my organisation can view
Public: anyone can view
PRIVATO O PUBBLICO?
Manage collaborators
Invite specific people to read, edit, or administer your plan. Invitees will receive an email notification that they have access to this plan.
Email address
Permissions
elena.giglia@unito.it
Owner
Invite collaborators
* Email
* Permissions
Co-owner
Editor
Read only
SI POSSONO INVITARE COLLABORATORI CON LIVELLI DI PERMESSO DIVERSI

Project Details | Plan overview | Initial DMP | Detailed DMP | Final review DMP | Share | Download

Horizon 2020 DMP

This plan is based on the "Horizon 2020 DMP" template provided by European Commission (Horizon 2020).

The Commission is running a flexible pilot under Horizon 2020 called the Open Research Data Pilot (ORD pilot).

Projects participating in the pilot must submit a first version of the DMP (as a deliverable) within the first 6 months of the project. The DMP needs to be updated over the course of the project whenever significant changes arise.

Further details are provided in the Guidelines on FAIR Data Management in Horizon 2020 (v.3, 26 July 2016).

Template version 1, published on 16 May 2019

Initial DMP (6 sections, 9 questions) +

Detailed DMP (9 sections, 31 questions) +

Final review DMP (9 sections, 31 questions) +

Initial DMP (6 sections, 9 questions)

Write plan

Instructions

Once a project has had its funding approved and has started, you must submit a first version of your DMP (as a deliverable) within the first 6 months of the project. The Commission provides a DMP template, the use of which is recommended but voluntary. That template has been replicated by the UK Digital Curation Centre for use in DMPonline.

It is not required to provide detailed answers to all the questions in the first version of the DMP that needs to be submitted by month 6 of the project. Rather, the DMP is intended to be a *living document* in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. Therefore, DMPs should have a clear version number and include a timetable for updates.

1. Data summary
   - Provide a summary of the data addressing the following issues:
     - State the purpose of the data collection/generation
     - Explain the relation to the objectives of the project
     - Specify the types and formats of data generated/collected
     - Specify if existing data is being re-used (if any)
     - Specify the origin of the data
     - State the expected size of the data (if known)
     - Outline the data utility: to whom will it be useful

2. FAIR data
   - 2.1 Making data findable, including provisions for metadata:
     - Outline the discoverability of data (metadata provision)
     - Outline the identifiability of data and refer to standard identification mechanism. Do
     - Outline the approach for clear versioning
     - Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how
   - 2.2 Making data openly accessible:
     - Specify which data will be made openly available? If some data is kept closed provide rationale for doing so
     - Specify how the data will be made available
     - Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?
     - Specify where the data and associated metadata, documentation and code are deposited
     - Specify how access will be provided in case there are any restrictions
   - 2.3 Making data interoperable:
     - Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.
     - Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?
   - 2.4 Increase data re-use (through clarifying licenses):
     - Specify how the data will be licenced to permit the widest reuse possible
     - Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed
     - Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why
     - Describe data quality assurance processes
     - Specify the length of time for which the data will remain re-usable

3. Allocation of resources
   - Explain the allocation of resources, addressing the following issues:
     - Estimate the costs for making your data FAIR. Describe how you intend to cover these costs
     - Clearly identify responsibilities for data management in your project


Prova
Project Details
Plan overview
Initial DMP
Detailed DMP
Final review DMP
Share
Download
expand all | collapse all
0/9 answered
1. Data summary (0 / 1)
2. FAIR data (0 / 4)
3. Allocation of resources (0 / 1)
4. Data security (0 / 1)
5. Ethical aspects (0 / 1)
6. Other (0 / 1)
...E NON «A HUGE AMOUNT OF DATA»
Data volume
Note what volume of data you will create in MB/GB/TB.
Indicate the proportions of raw data, processed data, and other secondary outputs (e.g., reports).
Consider the implications of data volumes in terms of storage, access and preservation. Do you need to include additional costs?
COSTI
Consider whether the scale of the data will pose challenges when sharing or transferring data between sites; if so, how will you address these challenges?
GUIDA RAPIDA
Provide a summary of the data addressing the following issues:
State the purpose of the data collection/generation
Explain the relation to the objectives of the project
Specify the types and formats of data generated/collected
Specify if existing data is being re-used (if any)
Specify the origin of the data
State the expected size of the data (if known)
Outline the data utility: to whom will it be useful
= README FILE
Guidance
DCC
Data volume
Data format
Data description
ALTRE GUIDE SPECIFICHE
Data format
TABELLA PER DATI DI FORMATO DIVERSO
Clearly note what format(s) your data will be in, e.g., plain text (.txt), comma-separated values (.csv), geo-referenced TIFF (.tif, .tfw).
Explain why you have chosen certain formats. Decisions may be based on staff expertise, a preference for open formats, the standards accepted by data centres or widespread usage within a given community.
Using standardised, interchangeable or open formats ensures the long-term usability of data; these are recommended for sharing and archiving.
See UK Data Service guidance on recommended formats or DataONE Best Practices for file formats.
SCRIVETE DIRETTAMENTE (USATE TABELLE E PUNTI ELENCO)
Save
IN OGNI PASSAGGIO POTETE SALVARE E CONTINUARE DOPO
USATE FORMATI STANDARD (CON ELENCO)

2. FAIR data (0 / 4)

In general terms, your research data should be 'FAIR' that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation-solution.

2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

Guidance | Comments

EC | DCC

The Research Data Alliance provides a Metadata Standards Directory that can be searched for discipline-specific standards and associated tools.

APRE LA DIRECTORY DEI METADATA STANDARDS

Save

2.2 Making data openly accessible:

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

Guidance | Comments

EC | DCC

Participating in the ORD Pilot does not necessarily mean opening up all your research data. Rather, the ORD pilot follows the principle "**as open as possible, as closed as necessary**" and focuses on encouraging sound data management as an essential part of research best practice.

The Commission recognises that there are good reasons to keep some or even all research data generated in a project closed. Where data need to be shared under restrictions, explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

RICORDA GENTILMENTE CHE TUTTI I DATI DEVONO ESSERE APERTI (SE POSSIBILE)

Save

APRE REGISTRY OF REPOSITORIES

The Registry of Research Data Repositories provides a useful listing of repositories that you can search to find a place of deposit.

2.3 Making data interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

Guidance | Comments

EC

Inteoperability means allowing
exchange and re-use between
institutions, organisations, cou
(i.e. adhering to standards for
much as possible compliant wit
(open) software applications, a
particular facilitating re-combi
different datasets from differe

Guidance | Comments

Add comments to share with collaborators

Save

2.4 Increase data re-use (through clarifying licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

Guidance | Comments

EC | DCC

The EUDAT B2SHARE tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.

Reasons for embargoes may include time to publish or seek patents. If an embargo is sought, specify why and for how long,

Save

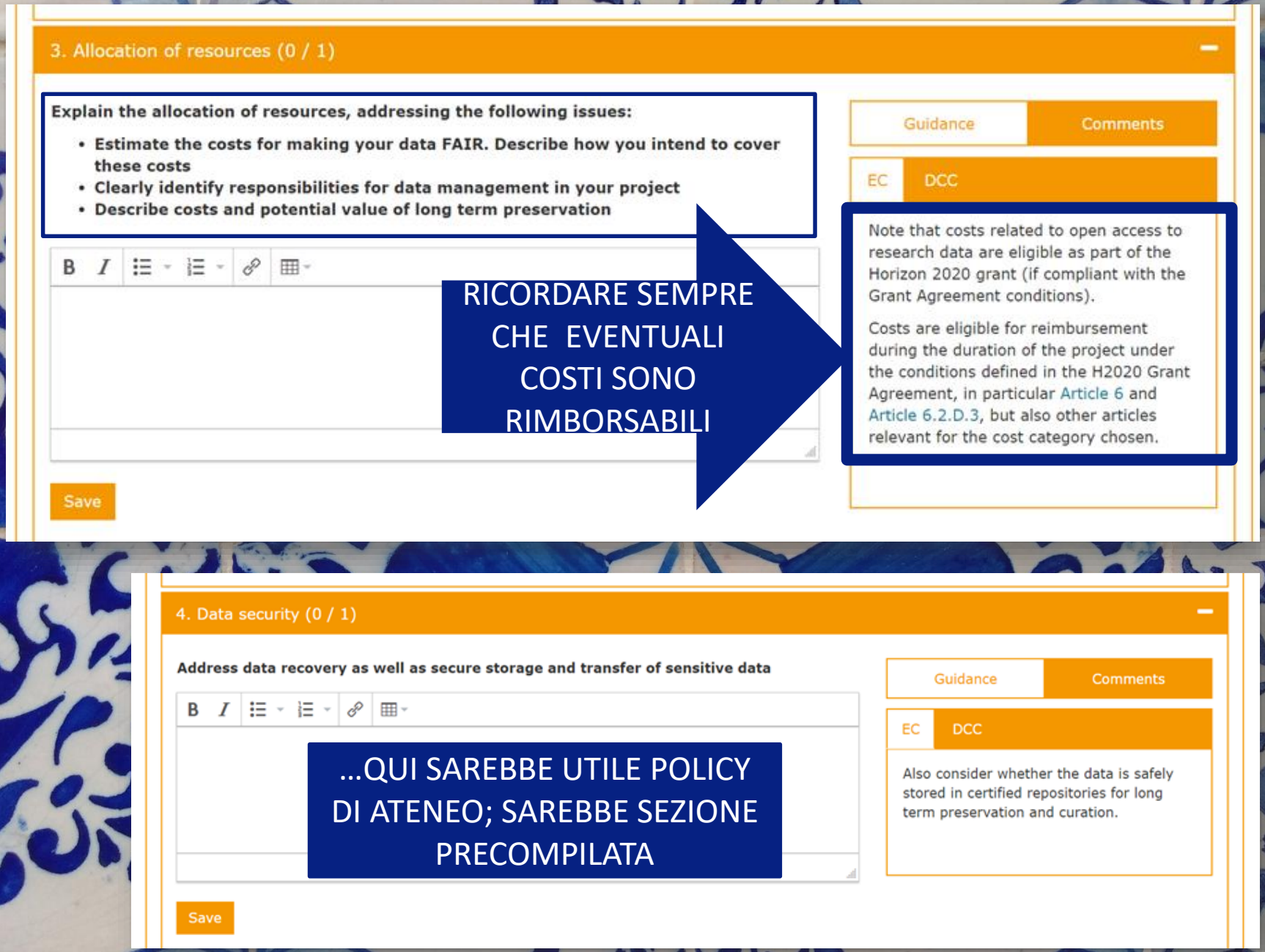
RICORDATE LORO CHE HANNO FIRMATO UN GRANT AGREEMENT…

29.3 Open access to research data

*[OPTION 1a for actions participating in the open Research Data Pilot: Regarding the digital research data generated in the action ('**data**'), the beneficiaries must:*

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

(i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;


3. Allocation of resources (0 / 1)
Explain the allocation of resources, addressing the following issues:
• Estimate the costs for making your data FAIR. Describe how you intend to cover these costs
• Clearly identify responsibilities for data management in your project
• Describe costs and potential value of long term preservation
Guidance
Comments
EC
DCC
Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).
Costs are eligible for reimbursement during the duration of the project under the conditions defined in the H2020 Grant Agreement, in particular Article 6 and Article 6.2.D.3, but also other articles relevant for the cost category chosen.
RICORDARE SEMPRE CHE EVENTUALI COSTI SONO RIMBORSABILI
Save
4. Data security (0 / 1)
Address data recovery as well as secure storage and transfer of sensitive data
Guidance
Comments
EC
DCC
Also consider whether the data is safely stored in certified repositories for long term preservation and curation.
…QUI SAREBBE UTILE POLICY DI ATENEO; SAREBBE SEZIONE PRECOMPILATA
Save


5. Ethical aspects (0 / 1)
To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former
Guidance
Comments
EC
DCC
Consider whether there are any ethical or legal issues than can have an impact on data sharing. For example, is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?
Save
BASE LEGALE PER DATI PERSONALI; CONSENSI INFORMATI VANNO IN QUESTA SEZIONE
6. Other (0 / 1)
Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)
Guidance
Comments
DCC
Related policies
Consider whether there are any existing procedures that you can base your approach on. If your group/department has local guidelines that you work to, point to them here.
List any other relevant funder, institutional, departmental or group policies on data management, data sharing and data security.
Save
QUI SAREBBE URGENTE UNA POLICY DI ATENEO


Project Details
Plan overview
Initial DMP
Detailed DMP
Final review DMP
Share
Download
Download settings
Select phase to download
Initial DMP
Optional Plan Components
project details coversheet
question text and section headings
unanswered questions
Format
pdf
PDF formatting
Font
Margin (mm)
Face
Size (pt)
Top
Bottom
Left
Right
Arial, Helvetica, Sans-Serif
10
25
20
12
12
Download Plan
AL TERMINE SI SCARICA NEL FORMATO PREFERITO (PDF, CSV, HTML, TXT) CON UN MINIMO DI FORMATTAZIONE

2. Data Management Plan – general definition

Data Management Plans (DMPs) are a *key element* of good data management. A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).

A DMP is required for all projects participating in the extended ORD pilot, unless they opt out of the ORD pilot. However, projects that opt
submit a DMP on a voluntary basis.

EUROPEAN COMMISSION
Directorate-General for Research & Innovation

H2020 Programme

Guidelines on

FAIR Data Management in Horizon 2020

Guide 2016

Version 3.0
26 July 2016

IN HORIZON 2020
DMP ENTRO M6
COME DELIVERABLE

VA FATTO SEMPRE
PERCHÉ RIGUARDA LA
CORRETTA GESTIONE
DEI DATI, NON
APERTURA/CHIUSURA

4. Research data management plans during the project life cycle

Once a project has had its funding approved and has started, you **must submit a first version of your DMP** (as a deliverable) within the first 6 months of the project. The Commission provides a DMP template in annex, the use of which is recommended but voluntary.

The DMP needs to be **updated** over the course of the project whenever significant changes arise, such as (but not limited to):

- new data
- changes in consortium policies (e.g. new innovation potential, decision to file for a patent)
- changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

The DMP should be updated as a minimum in time with the periodic evaluation/assessment of the project. If there are no other periodic reviews foreseen within the grant agreement, then such an update needs to be made in time for the final review at the latest. Furthermore, the consortium can define a timetable for review in the DMP itself.

Periodic reporting

For general information on periodic reporting please check the following sections of the online manual

- How to fill in reporting tables for publications, deliverables
- Process for continuous reporting in the grant management system.

Open data – H2020

https://goo.gl/sryNTg

How?

Open access to digital research data involves 3 steps:

Procedure for open access (research data):

Step 1 — **Deposit** the digital research data, preferably in a research data repository.

Step 2 — Provide **open access** by taking measures to enable users to access, mine, exploit, reproduce and disseminate the data free of charge *(e.g. for databases: by attaching an appropriate creative commons licence (CC-BY or CC0 tool) to the data; if the access/use is not subject to any rights: by indicating that no licence is needed).*

Open access must not be given immediately; for data needed to validate the results presented in scientific publications, as soon as possible; for other data, beneficiaries are free to specify embargo periods for their data in the data management plan (as appropriate in their scientific area).

Step 3 — Provide **information**, via the repository, about **tools and instruments** for validating the results.

Where possible, the beneficiaries should provide those tools and instruments *(e.g. specialised software or software code, algorithms, analysis protocols, etc.).*

1) DEPOSITARE IN UN DATA REPOSITORY
2) RENDERE I DATI OPEN ACCESS CON LA LICENZA PIÙ APERTA POSSIBILE
3) POSSIBILE EMBARGO
4) FORNIRE ANCHE TUTTE LE INFORMAZIONI UTILI A VALIDARE I DATI

... grazie e... ora tocca a voi!