

# FAIR principles and Open Data

*Emma Lazzeri*

ORCID iD: <https://orcid.org/0000-0003-0506-046X>

Istituto di Scienza e Tecnologie dell'Informazione  
Consiglio Nazionale delle Ricerche

# Practicalities

During the course, we will use different tools and services to

- Share
- Collaborate
- Interact



# Virtual Research Environment

- We set up a Virtual Research Environment (VRE) for this course. By entering the VRE you will find:
  - A dedicated forum (social networking) where you can ask questions after the lessons, discuss, share experiences. Trainers will use the forum to share important information about the course.
  - A dedicated workspace where trainers will share course material and other useful documentation
- To join the VRE, you can use your institutional, google or LinkedIn account:

<https://eosc-pillar.d4science.org/group/eosc-pillar-gateway/explore?siteId=273133421>



# Interactive Zoom buttons



Chat



Raise Hand



Q&A

## CHAT

Use the Chat for technical/ practical messages. Useful links will be shared here during the lessons

## RAISE HAND

If you wish to speak during the discussion sections, please raise your hand

## Q&A

Please use the Q&A button to pose questions anytime during the course. Questions will be answered in the discussion sessions, after the presentations.

# Mentimeter

- Mentimeter allows for a quick **interaction** with the audience.
- You will be able to **post anonymous comments**
- You will be asked to **answer questions anonymously**
- Results of the interactions will be **available live**
- You can **access** mentimeter from any device (mobile pc, tablet...)
  - Go to [www.menti.com](https://www.menti.com) and enter code: **17 62 71**
  - Click on the direct link: <https://www.menti.com/663okc6efu>
  - Scan the QR code





# What we will learn today

---

FAIR refers to principles, not to standards: no one size fits all!

---

We will see how to generally apply FAIR principles to your data

---

We will learn how you can apply FAIR principles in a structured community (institutional or disciplinary)

---

We will learn a good recipe for Data Management Plans





# Some news from the publishers' side...



Daniel Graeber  
@StreamDaniel

At first I thought it is a joke, but Nature journals are serious: up to 9500 EUR for open-access!



Melike Dönertaş  
@melikedonertas

9,500€ - that's the annual income of professors in Turkey. + publication fees are rarely covered by universities or institutes. Then people ask why they don't hear much from Turkey even though scientists from Turkey excel in abroad. One of many reasons 🙄



Nature News & Comment · Nov 24, 2020

The journals will charge authors up to €9,500 to make research papers free to read, in a long-awaited alternative to subscription-only publishing.



Nature journals reveal terms of landmark open-access option  
The journals will charge authors up to €9,5...  
nature.com



José Manuel García-Fernández  
@jmgarciafandez

I guess the Nature group will announce soon how much they are going to pay to experts for reviewing manuscripts, since their "production costs" are so high due to reviewing so many papers. Certainly this is not what open access is about. This is strictly business, not science.

9:03 AM · Nov 24, 2020



Mike Boylan-Kolchin  
@MBKplus

Taking a break from #JWSTDeadlineMemes to read about Nature's new Open Access policy and now I want to go back to bed.



4:11 PM · Nov 24, 2020

101 18 people are Tweeting about this

<http://bjoern.brembs.net/2020/11/are-natures-apcs-outrageous-or-very-attractive/>



Asmeret Asefaw Berhe  
@aaberhe

A charge of €9,500, US\$11,390 or £8,290 to make a single paper open access (OA) in Nature journals is so far beyond outrageous, I don't have a word for it. ... FTR, a life-altering summer research internship for an undergrad costs about US\$5,000!



Dave Lunt  
@davelunt

I've no problem with Nature's pricing, they charge what they can, just like every other company. The problem isn't them it's us. Every time you talk about "a Nature paper" as something special, you enable this. Stop



Nature journals reveal terms of landmark open-access option  
The journals will charge authors up to €9,500 to make research papers free to read, in a long-awaited alternative to subscription-...  
nature.com

8:04 PM · Nov 24, 2020





Definitions we will need in  
today's lecture



# Payload

**In the context of this course:** it is referred to any **research result** in its digital form that can be uploaded (and eventually openly shared) in a repository.

Examples: articles, dataset, software, images, videos, reports, conference poster or presentations, lectures, etc

# A Record deposited in a Repository

## Metadata

A set of data describing the digital object(s) you are depositing

The screenshot shows a Zenodo record page. The title is 'MOD01 - Research Data Management & Open Science: Introduction and Motivations - including EC policies and mandates' by Emma Lazzeri. The description states it is Module 1 of a 4-module course. A table lists a file 'INSrLazzeri\_MOD1\_20200706.pptx' (52.9 MB). The 'Citations' section shows 'No citations'. The 'Publication date' is July 6, 2020, and the DOI is 10.5281/zenodo.3939213. Keywords include 'Open Science', 'research data management', and 'open access'. Grants from the European Commission are listed. The license is Creative Commons Attribution 4.0 International. The version history shows 'Version 1' on July 6, 2020.

## Payload

The digital object(s) you are uploading to be stored (and eventually shared). Payload includes attached files such as the file containing the data and the accompanying material(s), readmefile, etc

# FAIR Principles

These slides summarise the contents of Martínez-Lavanchy, P.M., Hüser, F.J., Buss, M.C.H., Andersen, J.J., Begtrup, J.W. (2019). 'FAIR Principles'. In: Holmstrand, K.F., den Boer, S.P.A., Vlachos, E., Martínez-Lavanchy, P.M., Hansen, K.K. (Eds.), Research Data Management (eLearning course). doi: 10.11581/dtu:00000049

Find the video to this link: <https://vidensportal.deic.dk/RDMelearn>



# FAIR Principles

- FAIR indicate a list of principles that can help you in making your data ready for Open Science
- They are **principles**, not standards!
- They were designed to enable optimal use of research data and methods
- A group of different experts designed the FAIR principles between 2014 and 2016
- They identified a set of 15 principles

# FAIR: What does it mean?



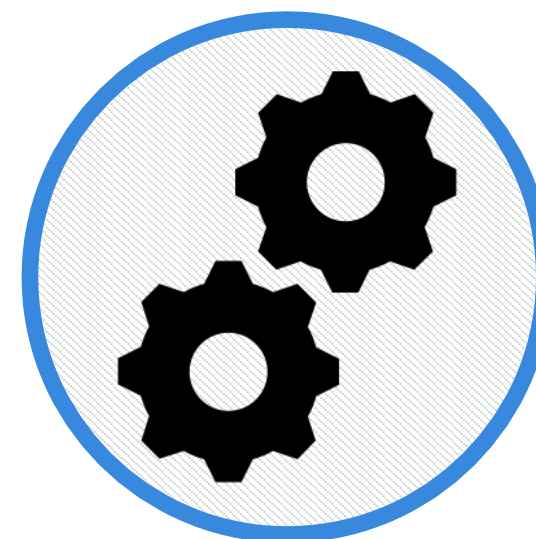
**Findable**

Other can find your data



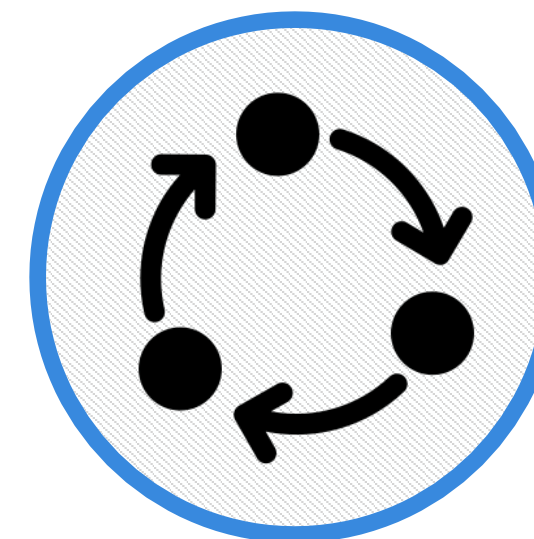
**Accessible**

Your data is accessible  
to others



**Interoperable**

Your data can be  
integrated with other data  
and/or they can be easily  
used and read by  
machines.



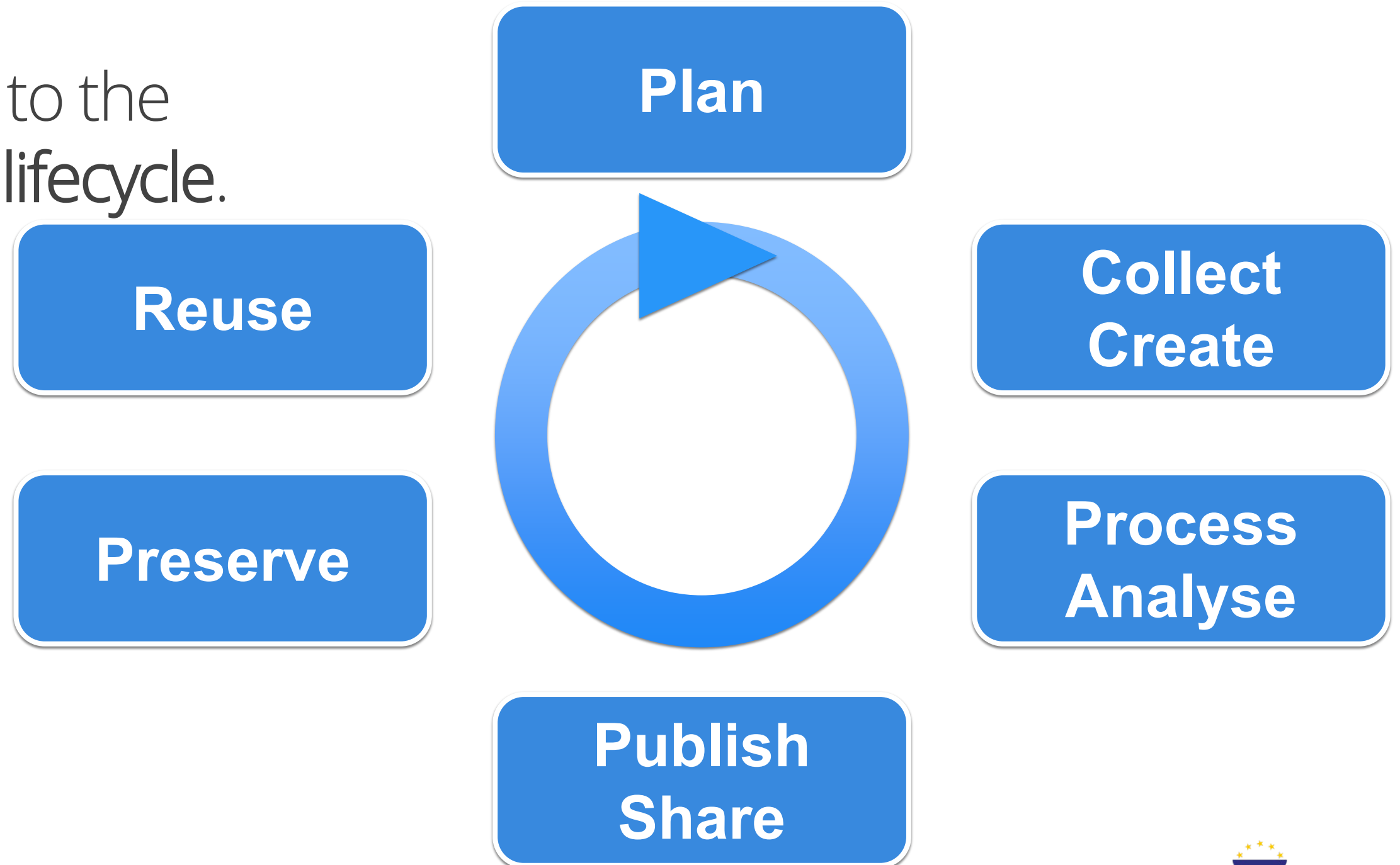
**Reusable**

Your data can be  
reused by others in new  
research

# FAIR principles

FAIR principles apply to the entire **research data lifecycle**.

They are strongly interconnected.

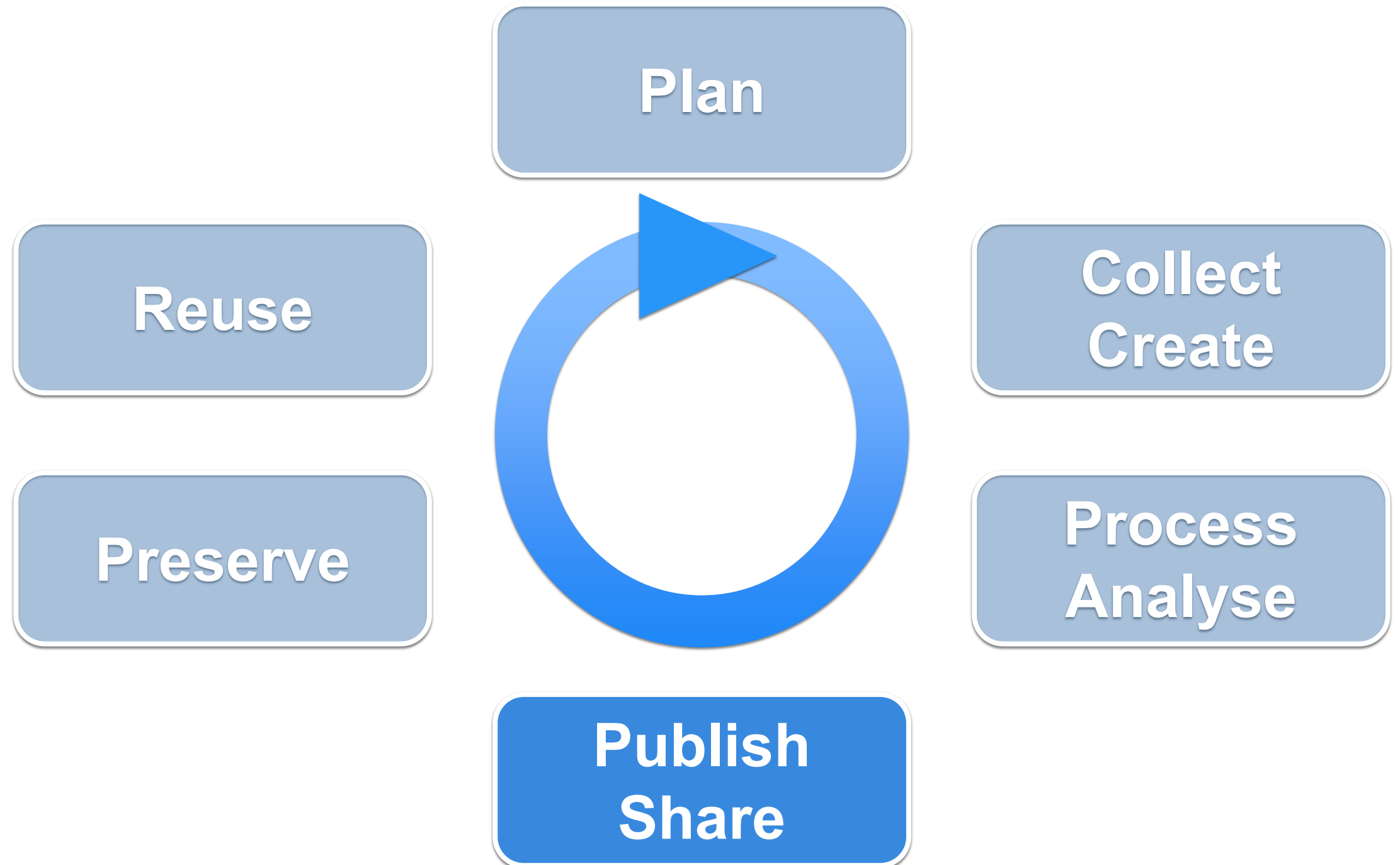




# FAIR principles

Please note:

Apply FAIR principles  
**does not** mean  
to **openly** share  
research data





FAIR Data

≠

Open Data

“

### Open Data

Data can be freely used, shared, enriched by anyone, anywhere for any purpose.

”

### FAIR Data

Data follow a series of good practices to allow data access, still respecting any ethical, legal and contractual restriction.



# Your research data could

- Contain personal information (privacy e GDPR)
- Fall under copyright (in the case of a database with creative structure)
- Fall under the Sui Generis right (database obtained thanks to a substantial investment)
- Be protected by patent or industrial secret

**Data sharing needs to respect the specific law.  
Data needs to be protected against non authorised access.**

# How can you adhere to FAIR principles if your data cannot be opened?

## Create and share a description of your data

- This way other researchers may ask for permission to access your data for reuse purposes, by giving a specific aim and following the rules defined by the law.
- **Restrict access** to the record payload (attachment, files,...)



Good practices to  
make your data  
FAIR



# By applying FAIR principles

- You will produce high quality data
- You will maximise the impact of your research
- You will improve the recognition within and behind your research community

**The application of FAIR  
principles strongly depends  
on the specific discipline and  
on the way the single  
researcher works**

**No one size fits all**





Why should I  
apply FAIR  
principles?



# Why do we need FAIR principles?

- The ultimate goal of FAIR principles is to make your research data (or object) reusable and safe
- You should always keep this goal in mind when trying to apply the FAIR principles to your results

**Are my results reusable by someone that was not involved in its collection/creation?**



So how can you  
make your data  
FAIR?

# FAIRification basics

- **Documentation**
  - Gives the context to make your data understandable by others
- **Metadata**
  - Make your data easy to find
- **Data formats**
  - Make your data simple to combine to other data and machine readable.
- **Access to data**
  - It means to decide who will have access to your data and how
- **Persistent identifiers**
  - Persistent links to data that allows others to find and cite (give credit to) your data.
- **Licenses**
  - Are used to tell others how they can reuse your data.

# Documentation

- Specifies the context that led to the creation/collection of your data to make them understandable
- At the beginning of a new (project) activity, you need to clearly define with your colleagues the strategy to structure and document your data.
- Document every detail of data collection/generation:
  - Methods
  - Tools
  - Software
  - Processes (who worked with the data? What did he/she did with the data? What are the relation to other data and/or publications?)
  - Metadata



# Metadata

- Data describing data
- Very important for:
  - Access
  - Comprehension
  - Process
- Use your discipline specific standards: you will spend less time **curating and interpreting data** and more time to actually make science!





# Metadata help making your data

- Findable
- Interoperable
- Reusable



Social Science and  
Humanities

innovation in metadata design, implementation & best practice

## Dublin Core™ Metadata Initiative

Home Specifications ▾ News Community ▾ Learning ▾ About ▾ Contact

Home

The Dublin Core™ Metadata Initiative supports innovation in metadata design and best practices. DCMI is supported by the National Science Foundation (NSF) and the Association of Information Science and Technology (ASIS&T).

### Stewardship



For more than twenty years, the DCMI community has developed and curated [Dublin Core Specifications](#). More recently, DCMI has become recognised as a trusted *steward* of metadata vocabularies, concept schemes and other metadata artefacts, and has taken responsibility for other [community-created specifications](#). DCMI

remains committed to this important work, and is actively developing more efficient and sustainable approaches to the stewardship of these standards, through the work of the [DCMI Usage Board](#).

### Learning



DCMI supports teachers and learners of modern metadata technologies and practices. An updated [Metadata Basics](#) page highlights current trends in descriptive metadata in the style of Dublin Core, which aims at interoperability through using globally shared vocabularies, constrained in application-specific profiles, based on principles

of Linked Data. Interested learners can also explore a [glossary](#) page, a [Linked Data Competency Index](#) that enumerates relevant skills to be learned, a [guide for users of DCMI metadata terms](#), occasional [webinars](#) and tutorials at [DCMI annual conferences](#).

### Community



DCMI responds to the evolving needs of the community, and has participated in a number of community events.

A major community event is the [DCMI Annual Conference](#). DCMI also organises regular [webinars](#), given by members of the community wishing to share their expertise with like-minded peers. Finally, DCMI [collaborates](#) with a number of other organisations.

### Development



DCMI has a long history of fostering and supporting technical development and innovation through the activities of its community, often in partnership with other organisations. Following on from the development of the ubiquitous [DCMI Metadata Terms](#), the community has in more recent years focussed on the concept of the metadata [application profile](#), developing supporting frameworks and conceptual models such as the [Singapore Framework](#). Most recently, the [Application Profiles Interest Group](#) has formed to address the next stage of development in this space.

Post-it notes image, © Bram Willemse, License: [CC BY-NC-ND 2.0](#). Other photographs, © Paul Walk, License: [CC BY 4.0](#)

Generic!

Title  
Authors  
Subject  
...

### DCMI 2020 Call for Proposals

Following on from the success of DCMI 2019 in Seoul (see Proceedings), we are pleased to announce the call for proposals in the DCMI 2020 International Conference on Metadata, Ottawa, Canada, 14-17 September 2020. We are grateful to Carleton University for offering to host us this year. This year's conference will mark the 25th anniversary of the original Dublin Core™ workshop. We will both reflect on two and a half decades of innovations while looking ahead to future developments.

[read more...](#)

ISO 15836 Part 2 is published based on a revision of DCMI Metadata Terms

# Accessibility

- Can I make my data accessible to others?
- Who will be granted access?
- How?



Findable



# Findable

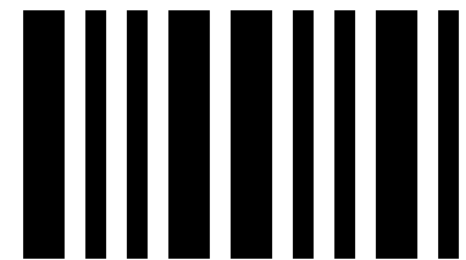
- The first step in (re)using data is to find them.
- Metadata and data should be easy to find for both humans and computers.
- Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.

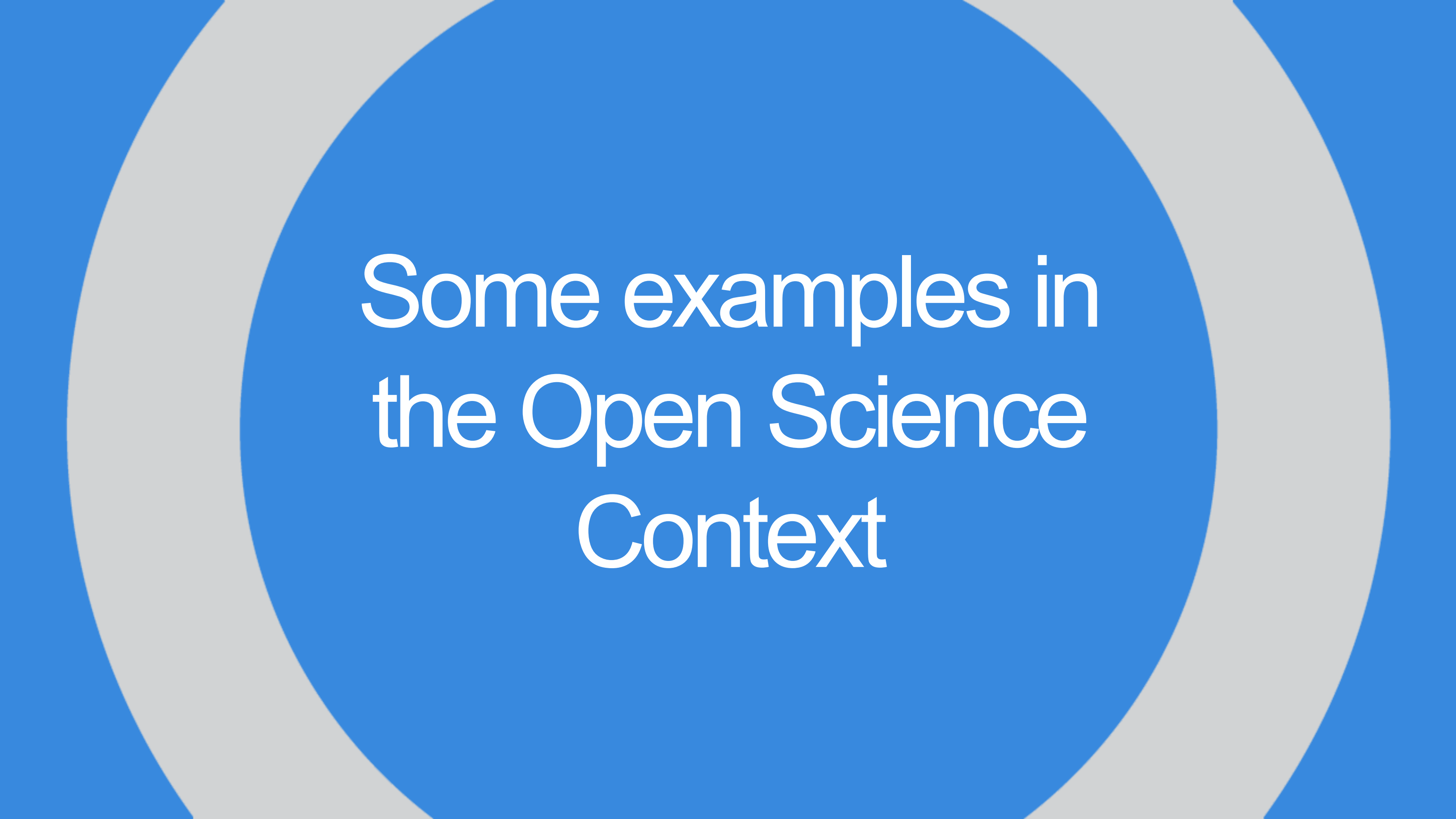


Some definitions

# Persistent Identifiers

- A **persistent identifier** (PI or PID) is a long-lasting reference to a document, file, web page, or other object.
- The term persistent identifier is usually used in the context of **digital objects** that are accessible over the Internet.
- Typically, such an identifier is not only persistent but **actionable**: you can plug it into a web browser and be taken to the identified source.
- It is like the bar code used on products...





Some examples in  
the Open Science  
Context





The Open Research and Contributor ID (ORCID) is a nonproprietary alphanumeric code to uniquely identify scientific and other academic authors and contributors

Do you have one? You should...

The screenshot shows the ORCID website interface. At the top, the ORCID logo is displayed with the tagline "Connecting Research and Researchers". To the right, there are links for "SIGN IN/REGISTER" and a language dropdown set to "English". Below this is a navigation bar with tabs for "FOR RESEARCHERS", "FOR ORGANIZATIONS", "ABOUT", and "HELP". The main content area displays the profile of "Emma Lazzeri". On the left, there is a green box for the "ORCID iD" with the URL "https://orcid.org/0000-0003-0506-046X", a "Print view" link, the "Country" (Italy), and "Other IDs" including "ResearcherID: M-1642-2014" and "Scopus Author ID: 24376669700". On the right, the "Employment (2)" section is expanded, showing two entries: "Consiglio Nazionale delle Ricerche: Pisa, IT" (2017-11-01 to present) and "Scuola Superiore Sant'Anna: Pisa, Toscana, IT" (2011-01-15 to 2017-10-31). Both entries are marked as "Preferred source".



# DOI – Digital Object Identifier

- In computing, a **digital object identifier** (DOI) is a **persistent identifier** or **handle** used to identify objects uniquely, standardized by the **International Organization for Standardization** (ISO).
- A DOI aims to be **resolvable**, usually to some form of access to the information object to which the DOI refers.
- This is achieved by **binding the DOI to metadata** about the object, such as a **URL**, **indicating where** the object can be found
- a DOI differs from identifiers such as **ISBNs** and **ISRCs** which aim only to identify their referents uniquely

# An example

3 x 3 optical switch by exploiting vortex beam emitters based on silicon microrings with superimposed gratings

Optics Letters

2017 | journal-article

DOI: [10.1364/OL.42.003749](https://doi.org/10.1364/OL.42.003749)

WOSUID: WOS:000411904500008

The screenshot shows the OSA Publishing website interface. The browser's address bar is highlighted with a red box, showing the URL: [osapublishing.org/ol/abstract.cfm?uri=ol-42-19-3749](https://osapublishing.org/ol/abstract.cfm?uri=ol-42-19-3749). The website header includes the OSA Publishing logo, a search bar, and navigation links. The article title is prominently displayed, along with a small image of the optical switch. The authors listed are: [Name], Muhammad N. Malik, Emma Lazzeri, Charalambos Klitis, Laura Meriggi, Ning [Name], Marc Sorel, and Antonella Bogoni. The article is from Optics Letters, Vol. 42, Issue 19, pp. 3749-3752 (2017). The DOI link <https://doi.org/10.1364/OL.42.003749> is provided at the bottom of the article page.

A tip for you: create a citation record from DOI with <https://citation.crosscite.org/>

# Persistent identifiers make your data:

- **Findable**
- **Accessible**



# How can you assign a Persistent Identifier to your digital object?

Persistent identifiers need to be assigned by an entity that can ensure the persistency of the link to the object

Zenodo assigns DOIs to digital objects that do not already have one

***Your repository will probably assign a persistent identifier to your digital object!***



Accessible

# Accessible

Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.



How do you give  
access to your  
data?

Through a  
**Repository**



# Open Access Repositories

Who does  
curate/deposit in the  
repository?

- **Thematic or disciplinary repositories**

Designed for specific contents, curated by specific communities: ArXiv, bioarXiv, PMC...

[http://oad.simmons.edu/oadwiki/Disciplinary\\_repositories](http://oad.simmons.edu/oadwiki/Disciplinary_repositories)

- **Institutional or national repositories**

Maintained and curated by single institutions/countries. Typically only authors based in the specific institution/country can deposit, everyone can access

- **Literature Repositories**

Reserved to text deposit (articles, reports, books, ...). Metadata reflect the repository contents.

<https://v2.sherpa.ac.uk/opensoar/>

- **Data repositories**

Designed to deposit data. They often are disciplinary and have specific metadata to describe the type of data they preserve. <https://www.re3data.org/>

- **Catch-all repositories**





All research products can be deposited (data, literature, presentations, poster, images, software, ...). Example: [Zenodo](#)

What are the  
repository contents?





# An important difference

- **Deposit:** upload a digital object (data, articles, ...) on a platform that allows to correctly describe the object through metadata and that implements long-term preservation.
- **Give access:** once the object has been deposited, the authors can choose the type of access that can be granted (open, restricted, closed, embargoed,...) and assigns a licence to reuse the contents (Creative Commons)

# Access Rights in Zenodo





-  Open Access
-  Embargoed Access
-  Restricted Access
-  Closed Access

Access right \*

- ☐  Open Access
- ☐  Embargoed Access
- ☐  Restricted Access
- ☒  Closed Access





Required. Open access uploads have considerably higher visibility on Zenodo.

Access right \*

- ☒  Open Access
- ☐  Embargoed Access
- ☐  Restricted Access
- ☐  Closed Access





Required. Open access uploads have considerably higher visibility on Zenodo.

Access right \*

- ☐  Open Access
- ☒  Embargoed Access
- ☐  Restricted Access
- ☐  Closed Access

Required. Open access uploads have considerably higher visibility on Zenodo.

Access right \*

- ☐  Open Access
- ☐  Embargoed Access
- ☒  Restricted Access
- ☐  Closed Access

Required. Open access uploads have considerably higher visibility on Zenodo.

Conditions \*

  **B** *I* ~~S~~  $x_2$   $x^2$                 Sorgente

# Warning!

**Attaching** your data to the article you published does not mean you are depositing the data.

Journals do not guarantee long term **preservation** and **curation** of the data.





Interoperable

# Interoperable

The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.

Use community standard or best practice!

# Linking Objects in Zenodo

Related/alternate identifiers

recommended

Specify identifiers of related publications and datasets. Supported identifiers include: DOI, Handle, ARK, PURL, ISSN, ISBN, PubMed ID, PubMed Central ID, ADS Bibliographic Code, arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs and URLs.

Related identifiers

e.g. 10.1234/foobar.56789

+ Add another related identifier

Optional. Resource type of the related identifier.

✓

cites this upload

is cited by this upload

is supplemented by this upload

is a supplement to this upload

is referenced by this upload

references this upload

is previous version of this upload

is new version of this upload

continues this upload

is continued by this upload

has this upload as part

is part of this upload

reviews this upload

is reviewed this upload

documents this upload

is documented by this upload

is compiled/created by this upload

compiled/created this upload

is the source this upload is derived from

has this upload as its source

is identical to this upload

is an alternate identifier of this upload

✓ N/A

Publication

Annotation collection

Book

Book section

Conference paper

Data management plan

Journal article

Other

Patent

Preprint

Project deliverable

Project milestone

Proposal

Report

Software documentation

Taxonomic treatment

Technical note

Thesis

Working paper

Dataset

Image



Reusable



# Reusable

The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings

# Licenses

- Tell others how they can reuse your data!

# Provenance

- Where is your data coming from?

# Research data may be:

- Automatically protected by the law;
- Regulated by contract;
- Subject to community norms such as academic best practices.

# Some consideration on data protection

- Copyright is a property right in certain types of original literary, artistic and scientific works.
- Copyright does not protect ideas.
- Confidentiality protects confidential information. This might be imposed by a contract or if the information is marked confidential. Use of confidential information might give rise to a claim for compensation if confidentiality is breached.
- Data Subject Rights arise in information that identifies individuals and are recognised by data protection laws in the EU.
- Patents are registered rights in novel inventions of products or processes.
- Some research data may not benefit from any legal protection, although moral and ethical considerations may apply.

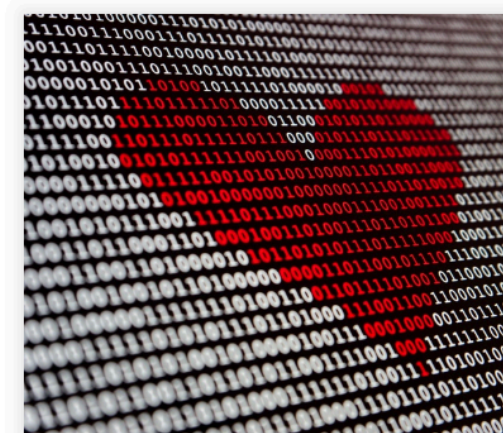
# Data is not yours!

- Data is **not** intellectual work, no copyright applies!
- **Copyright** protection covers expressions and not ideas, procedures, operating methods or mathematical concepts as such.
- Protection is on databases and not on **data**. The data are protected only and especially when they are collected and organized in a database.
- **The sui generis database right (only in Europe)** covers not only the reproduction and dissemination of the database, but also the extraction and reuse of substantial parts of the database.

Read complete article [here](#).

A similar content article in English [here](#).

OpenAIRE Guidelines on data protection: available [here](#)



11 Dicembre 2019

## Data governance: un dato non appartiene a nessuno... a meno che sia personale

di [Simone Aliprandi](#)

Quando un nostro dato è personale? Come è giusto tutelarlo? La risposta deve comprendere due punti di vista, quello della proprietà intellettuale e quello della privacy.

CONDIVIDI



### L'autore



#### [Simone Aliprandi](#)

Simone Aliprandi ha un dottorato di ricerca in Società dell'Informazione ed è un avvocato che si occupa di consulenza, ricerca e formazione nel campo del diritto d'autore e più in generale del diritto dell'ICT. È responsabile del progetto copyleft-italia.it, è membro del network Array e collabora come docente con alcuni istituti universitari; ha pubblicato articoli e libri sul mondo delle tecnologie open e della cultura libera, rilasciando tutte le sue opere con licenze di tipo copyleft.

[Sito e blog](#)



## Non c'è solo la privacy, quando si parla di dati e di diritto

Si sente spesso parlare di *tutela del dato* o *titolarità del dato*, soprattutto in questi ultimi due/tre anni in cui temi come [big data](#) e open data sono diventati di pubblico dominio e in cui l'entrata in vigore del GDPR ([il nuovo regolamento europeo sui dati personali](#)) ha portato un'ondata – per certi versi ridondante – di corsi di formazione, consulenze, articoli sul tema della protezione dei dati.

Mi occupo di consulenza e formazione proprio in quest'ambito e mi rendo conto che spesso tra gli utenti non c'è piena consapevolezza di come il diritto considera e tratta i dati. Noto soprattutto che alle parole *tutela del dato* o *titolarità del dato* **la gente pensa automaticamente all'ambito della privacy**, della tutela del dato personale.

Per inquadrare il tema correttamente e in modo completo, è tuttavia necessario tenere in debita considerazione anche il punto di vista della cosiddetta proprietà intellettuale, punto di vista che a me sta particolarmente a cuore. Anche perché quando acquisiamo, gestiamo, diffondiamo dei dati **non è detto che siano dati personali** e dunque non sempre le norme sulla privacy (GDPR e simili) entrano in gioco. Cerchiamo di capire meglio la questione.

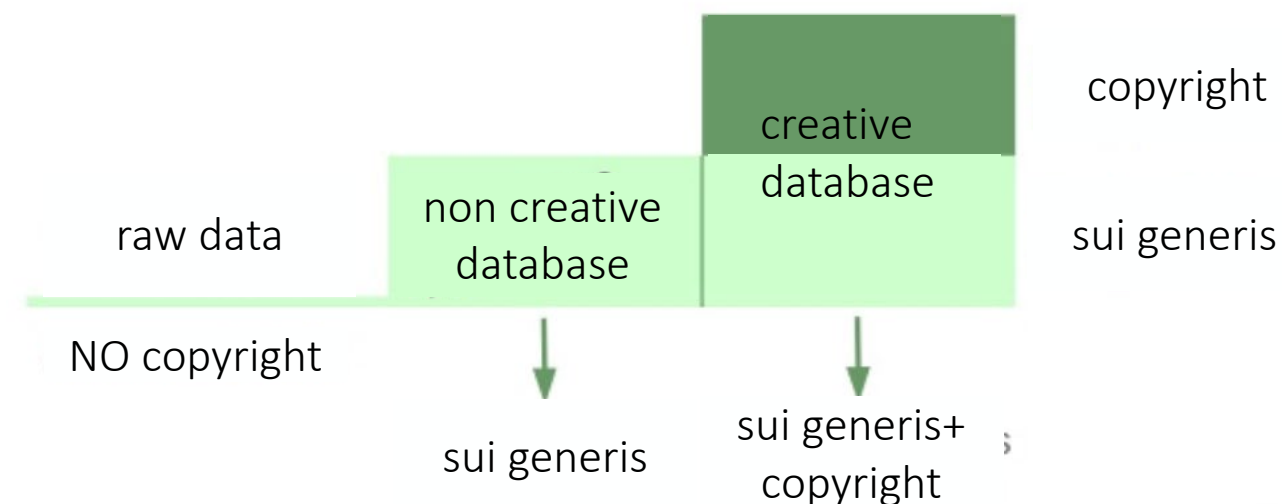
## Nessun copyright su idee e dati

Innanzitutto: **non esiste alcun diritto di proprietà intellettuale sul dato in sé**. I dati nudi e crudi e le informazioni che da essi si deducono non sono oggetto di alcun tipo di proprietà intellettuale. Questo in virtù di uno dei principi cardine del diritto d'autore secondo cui il diritto tutela non l'informazione, bensì la specifica forma espressiva con cui l'informazione è presentata. Basti leggere il testo dell'articolo 9, numero 2 dell'[Accordo TRIPS](#):



# Data and law protection

- **Raw data** are not protected by copyright
- **Database** is defined as a collection of independent works, data or other materials arranged in a **systematic or methodical way**
- **Copyright** protects the structure, selection or arrangement of the database contents, **not the data**
- **Sui generis database right**: protects the substantial effort in obtaining data (not creating)
  - Note: the right owner is often the institution.



27.3.1996

EN

Official Journal of the European Communities

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

**of 11 March 1996**

**on the legal protection of databases**


















THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

# Authors and rights owners

- **Are you the author of the data you collected?**
  - Yes, in case you can prove it (deposit with clear date, DOI, ... use a data repository!)
- **Do you own any rights on the raw data you collected?**
  - No, data is facts/information and none can own rights on it!

# Creative Commons

- Not all of us are legal experts capable of writing proper licenses
- Creative Commons and Public Domain create legal certainty for everyone, who wants to use works, that are licensed respectively.
- It is important to follow and understand the different meanings of the licenses and follow the rules for using them.

CREATIVE COMMONS LICENSES		 COPY & PUBLISH	 ATTRIBUTION REQUIRED	 COMMERCIAL USE	 MODIFY & ADAPT	 CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✓
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✓
						
You can redistribute (copy, publish, display, communicate, etc.)		You have to attribute the original work				
						
You can use the work commercially		You can modify and adapt the original work				
		You can choose license type for your adaptations of the work.				

# Creative Commons

- **Public Domain**

Works are not covered by copyright

- **CC-0 (no rights reserved)**

Allows creators to give up their copyright and put their works into the worldwide public domain

- **CC-BY (Attribution)**

This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator

- **CC-BY-SA (Attribution – ShareAlike)**

This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use. If you remix, adapt, or build upon the material, you must license the modified material under identical terms.

- **CC-BY-ND (Attribution – NonDerivative)**

This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, and only so long as attribution is given to the creator. The license allows for commercial use.

- **CC-BY-NC (Attribution – NonCommercial)**

This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator.

CREATIVE COMMONS LICENSES		COPY & PUBLISH	ATTRIBUTION REQUIRED	COMMERCIAL USE	MODIFY & ADAPT	CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✓
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✓
You can redistribute (copy, publish, display, communicate, etc.)		You have to attribute the original work				
		You can use the work commercially				
		You can modify and adapt the original work				
		You can choose license type for your adaptations of the work.				



# Clear licenses help you also in combining your data:

	PUBLIC DOMAIN	PUBLIC DOMAIN	BY	BY SA	BY NC	BY ND	BY NC SA	BY NC ND
PUBLIC DOMAIN	✓	✓	✓	✓	✓	✗	✓	✗
PUBLIC DOMAIN	✓	✓	✓	✓	✓	✗	✓	✗
BY	✓	✓	✓	✓	✓	✗	✓	✗
BY SA	✓	✓	✓	✓	✗	✗	✗	✗
BY NC	✓	✓	✓	✗	✓	✗	✓	✗
BY ND	✗	✗	✗	✗	✗	✗	✗	✗
BY NC SA	✓	✓	✓	✗	✓	✗	✓	✗
BY NC ND	✗	✗	✗	✗	✗	✗	✗	✗



# Warning!

**Attaching** your data to the article you published does not mean you are depositing the data.

Journals do not guarantee long term **preservation** and **curation** of the data.

# Licensing your data: Creative Commons



## 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



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



'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



# Licensing your data: Creative Commons

- Use a CC0 or public domain
- Then ask for credit
- **Provide a citation** that researchers using your data can simply copy and paste to give you credit for your work
- Bear in mind it's bad science not to cite the source
- CC0 does not mean academic unpoliteness

**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.

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 from selling your dataset to others for a profit. However, it might also prohibit someone using the dataset in their research if they intend to eventually publish that research. This is because most academic journals are commercial businesses that charge some sort of fee for access to their content, hence, such use could qualify as 'commercial'. Consequently, using a non-commercial licence prevents researchers from using your data in work destined for publication. This can subsequently affect the dissemination, recognition, and impact of your dataset.

**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.

**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.

# To summarise...



”

Once your data is FAIR,  
decide to go Open!





**Carlos Moedas** ✓

@Moedas

 Segui

2/4 "Open as possible, as closed as necessary" is the new principle for all [#data](#) from publicly funded [#research](#) in Europe [#openaccess](#)





RETWEET  
76

MI PIACE  
32



# Open your data in practice



## Access right \*

- ☒  Open Access
- ☐  Embargoed Access
- ☐  Restricted Access
- ☐  Closed Access

Required. Open access uploads have considerably higher visibility on Zenodo.

## 🌀 License \*

Creative Commons Attribution 4.0 International

Required. Selected license applies to all of your files displayed on the top of the form. If you want to upload some of your files under different licenses, please do so in separate uploads. If you cannot find the license you're looking for, include a relevant LICENSE file in your record and choose one of the *Other* licenses available (*Other (Open)*, *Other (Attribution)*, etc.). The supported licenses in the list are harvested from [opendefinition.org](https://opendefinition.org)  and [spdx.org](https://spdx.org) . If you think that a license is missing from the list, please [contact us](#).

”

# How about Reliability? Is FAIR = Reliable

*No, it is not*

# Thank you!

**Emma Lazzeri**

[emma.lazzeri@isti.cnr.it](mailto:emma.lazzeri@isti.cnr.it)

This work was partially supported by European Union's Horizon 2020  
under projects grant Agreement numbers 831644, 857650, and

777541

