Bibliographic resources and research tools for PHD students in Industrial Engineering

Silvia Tapparo - Biblioteca di Ingegneria - Sede di Tecnica e Gestione dei Sistemi Industriali
Michela Zorzi – Centro di Ateneo per le Biblioteche
26/01/2023
What we’ll talk about

- Scholarly publishing
- Tips on authors’ rights
- Bibliometrics and bibliometric indicators
- Open Access
- Who requires Open Access
- Facilitations for authors
- Repositories + Padua Research Archive (IRIS)
- Research data management + Research Data Unipd
- Unipd Library System services for authors
Workshop materials

Workshop slides are available here:

http://biblioingegneriacentrale.cab.unipd.it/usa/laboratori/materiali
Scholarly publishing
Versioning

AUTHOR MANUSCRIPT submitted

PREPRINT - AM

> peer review >

POSTPRINT - AAM

> edit + peer review >

PROOF

> edit + correction

UNCORRECTED/ CORRECTED PROOF

Version in which changes were applied as a result of peer review and was re-submitted to the publisher.

Version of Record “PUBLISHED”

FINAL published VERSION - Version of Record

The final published version. Online, in print or both.

AUTHOR ACCEPTED MANUSCRIPT

This is the version accepted for publication after it has been peer-reviewed

EARLY ACCESS VoR

Editorial version without definitive page numbers, issue or volume. Also known as Epub, Early view, article in press, ...
Authors’ rights
Many Scholarly or Academic Journals require authors to sign away their rights to their own articles!

Intellectual property

**Intellectual Property (IP)** protects different types of intellectual creations. It is usually divided into two parts:

1. **Industrial property**
   - Designs and models protects the **new design** of an utilitarian object
   - Patent law protects **new inventions**
   - Trademark law protects **signs** used in commerce to distinguish the origin of goods or services

2. **Literary and artistic property**
   - Copyright / author’s rights
What is meant by copyright, publishing rights, related rights…

**Moral Rights**

Author’s right allows the author to be the one entitled to dispose of his/her works in an exclusive way.

**Economic Rights**
Moral Rights

- authorship (paternity)
- repentance
- on unpublished works
- to the integrity of the work

NON-TRANSFERABLE
Economic Rights

- publication
- reproduction
- transcription
- diffusion
- public representation
- communication to the public
- creation of derivative works

Secondary rights: rental and lending, resale
Amendment to the contract (addendum)

SPARC addendum and Science Commons SCAE are legal tools that authors can use to safeguard their editorial rights by amending standard publisher agreements.

In the guidelines for authors, Sparc illustrates how to use the addendum to the editorial contract to guarantee authors' rights. SPARC addendum is also available in Italian.

For authors participating in EU-funded projects who need to publish in non-Open Access journals, Horizon 2020 has proposed a specific model of publishing agreement.

In the revised editorial agreement, the author agrees with the publisher on the possibility of:

- maintain the right to reproduce, distribute and publicly display the article, not for commercial, but for teaching and research use
- self-archive the article on the personal or institutional website and deposit it in open archives
- prepare works derived from the article: for example, use the contents of the publication to create didactic material.

https://bibliotecadigitale.cab.unipd.it/en/digital-library/about-publishing/while-publishing
Italian law on author’s rights

- Civil Code (art. 2575-2583)

- Legge 22 aprile 1941 n. 633, "Protezione del diritto d'autore e di altri diritti connessi al suo esercizio"

Amended by
- Legge 18 agosto 2000 n. 248, "Nuove norme di tutela del diritto di autore"
Copyrights on doctoral thesis

Your rights:
Moral rights
Property rights

Rights to respect:
3rd party copyright material
Submission and publication of PhD theses

Regolamento per i corsi di dottorato di ricerca

ART. 32.2: La domanda di ammissione all’esame finale deve essere corredata: a) da un esemplare della tesi in formato digitale per il deposito presso l’archivio dell’Ateneo che ne garantisce la conservazione e la pubblica consultabilità. Lo stesso deposito è funzionale anche ai fini di adempiere all’obbligo di deposito della tesi presso le biblioteche nazionali di Roma e Firenze. In caso di richiesta di revisioni da parte dei valutatori la tesi definitiva deve essere depositata secondo i tempi e le modalità annualmente stabilite dall’Ateneo.

[You should attach to the application to the final exam: a) a digital copy of your thesis to be stored in the Unipd institutional repository, which guarantees conservation and public accessibility. The storage meets the requirement to deposit the thesis in the Italian national libraries in Rome and Florence. In case the evaluator requires a revision of the thesis, the revised final version must be stored accordingly to the calendar yearly established by Unipd.]
Publication workflow

- Soon after the PhD student career closure: migration of theses from Uniweb to Padua Research Archive (PRA/IRIS)
- Once a year: harvesting of theses toward National Libraries of Florence and Rome, according to legal deposit

Check deadlines for submission in Doctoral program [webpages](#)
Embargo  
*(consent form)*

**[temporary embargo]**

Richiede che il testo completo della tesi depositata nell’archivio Istituzionale *Padua Research Archive (IRIS)* rimanga chiuso per (selezionare):

- 12 mesi
- 18 mesi
- 36 mesi

con la seguente motivazione* (selezionare):

- Brevetto
- Motivi editoriali (allegare una copia del contratto sottoscritto con l’editore o altra attestazione)
- Protezione dei diritti economici dell’Ateneo
- Segreto industriale (se è stato firmato un accordo di non divulgazione, allegare una copia dell’accordo)

**[permanent embargo]**

È possibile chiedere che venga applicata una chiusura permanente nel caso in cui intervenga una delle motivazioni seguenti (selezionare):

- Inserimento nella tesi di testi e/o materiale altrui non autorizzato, in violazione dei diritti di terze parti, d’autore o editoriali, fatta eccezione per l’uso di immagini a bassa risoluzione
- Motivi di sicurezza pubblica o nazionale: indicare obbligatoriamente nel campo NOTE tutti i dati necessari per la verifica
- Motivi di privacy o presenza di dati sensibili: assicurarsi che il motivo non rientri nell’embargo temporaneo.

In questo caso è necessario selezionare in Uniweb l’opzione di 36 mesi di embargo e successivamente contattare l’Ufficio Dottorato e Post Lauream - Settore Dottorato di Ricerca al seguente indirizzo e-mail phd@unipd.it specificando nell’oggetto della mail: “Richiesta embargo PERMANENTE”.

**FILES IN THIS ITEM:**

There are no files associated with this item.
Editorial reasons

Theses including articles (already published or not)

- pay attention to publishers’ policies concerning the use of articles in PhD dissertations
- ask for permission through an Addendum
- choose the pre-print or the Accepted version

https://www.flickr.com/photos/21133841@N03/3542221586/in/photostream
Sherpa Romeo

Bibliometrics and bibliometric indicators
Visibility, prestige and citation impact of a journal

Before choosing a publisher or journal to submit a work for publication, it is a good idea to evaluate what opportunities for visibility in the reference scientific community that the journal offers. The more disciplinary and interdisciplinary databases that index the journal, the greater the possibility of disseminating the results of research.

The “Impact of research” section collects information on Web of Science and Scopus, databases that calculate the most popular citation indexes (citation count, Impact Factor, SJR, h-index and others).

It is important to verify the citation impact (Impact Factor) of the journal, as the number of citations that a publication receives is of great importance both for defining the impact potential of the research presented within it, and for the attribution of scores in the publications, competitions, and the recurring ANVUR assessment exercises (e.g. VQR).

Bibliometrics and...

**BIBLIOMETRICS** is a set of mathematical and statistical methods used to analyze and measure the quantity and quality of books, articles, and other forms of publications.

**Bibliometrics**
- identifies the best journals of a specific discipline
- defines the prestige of a specific journal
- determines the impact of published research

**Bibliometrics evaluates:**
- scientific journals
- single researchers
- research groups

… bibliometric indicators

**Bibliometric indicators** are very important for researchers and organizations, as these measurements are often used in funding decisions and promotions of researchers. They are becoming increasingly important since published research results are read and then quoted by other researchers.

- **quantity indicators**: measure the productivity of a particular researcher (**Impact Factor**; **SNIP**, **SCImago**)

- **quality indicators**: measure the quality or performance of a researcher's output; corresponds to the so called “peer-review”, a review by colleague-scientists (**h-index**)
Impact Factor

The **impact factor (IF)** is a measure of the frequency with which the average article in a journal has been cited in a particular year. It is used to measure the importance or rank of a journal by calculating the times its articles are cited.

**How Impact Factor is Calculated?**
The calculation is based on a two-year period and involves dividing the number of times articles were cited by the number of articles that are citable. **The Impact Factor is used to compare different journals within a specific disciplinary field.**

The [Journal of Citation Report](https://www.journalcitationreports.com) indexes more than 11,000 science and social science journals. It is important to note that Impact Factor is a journal metric and should not be used to assess individual researchers or institutions.
H-Index

The \textit{h-index} quantifies an individual’s scientific research output (cit. J.E. Hirsch).

The \textit{h-index} evaluates an author impact inside a specific scientific community on the basis of the number of his/her publications and citations obtained.

The \textit{h-index} is one of the most important function in \textit{Scopus}.
SCImago Journal Ranking

**SCImago** a database that can be accessed for free online, which allows you to obtain statistics on the citations of articles published in peer-reviewed journals. It provides statistics and compares the number of published articles and citations in each country.

**Journal ranking**

**Country rankings**
Introduction to Open Science

“Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society”

FOSTER consortium

Open Science

Open Data
Open Source (in Open Science)
Open Methodology
Open Peer Review
Open Access
Open Educational Resources

Andreas E. Neuhold – Opera propria – CC BY 3.0
Open access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions.

Open Access without barriers and restrictions to scientific knowledge

Peter Suber (2012), Open Access, MIT Press
https://cyber.harvard.edu/hoap/Open_Access_(the_book)
Open Access benefits

- Public can access your findings
- More visibility
- Funder and grant compliance
- 18% higher citations on average
- Practitioners and policymakers can apply your findings
- Researchers in developing countries can see your work

https://www.waikato.ac.nz/library/guidance/guides/open-access-information
Open licenses: Creative Commons

Choose the CC license for your work
Publishing in OA

Green OA
Self-archiving in compliance with publisher’s policy
Embargo
COST
NO
Licence
AUTHOR or PUBLISHER’s CHOICE

Diamond OA
Full, immediate OA publishing, without Article Processing Charge - APC

COST
NO
Licence
CREATIVE COMMONS CC BY CC BY SA

Gold OA
Full, immediate OA. Author pays an APC

COST
YES
Licence
OPEN or CC LICENCE

Hybrid OA
Immediate OA of a single article/chapter in a paywalled Journal/Ebook, author pays an APC

COST
YES
Licence
CC LICENCE on single item

Bronze OA
Immediate or delayed free contents, without a clear open licence or reuse permission

COST
PUBLISHER’S CHOICE
Licence
NO

https://bibliotecadigitale.cab.unipd.it/en/digital-library/about-publishing/open-access
6 Ways To Identify Predatory Publishers or vanity Press

1. The journal does not identify a formal editorial/review board.
2. The publisher has no policies or practices for digital preservation.
3. The publisher begins operations with a large fleet of journals, often using a template to quickly create each journal’s home page.
4. The publisher provides insufficient information or hides information about author fees, offering to publish an author’s paper and later sending a previously-undisclosed invoice.
5. The name of a journal does not adequately reflect its origin.
6. The journal falsely claims to have an impact factor, or uses some made up measure (e.g. view factor), feigning international standing.

https://publicdomainvectors.org/it/dominio-pubblico/
How to identify a predatory publisher

It is useful to consult the infographic on predatory publishing to evaluate better the publisher with whom to publish.

Think, Check & Submit: a tool to evaluate the publisher and verify its policies.

DOAJ - the Directory of Open Access Journal: it's the best way to find trustworthy information on Open Access journals

https://beallslist.net/: a list of potential predatory journals (updated to 2016)
Who requires Open Access?
EU funding programs

From OpenAIRE webinar: Horizon 2020 Open Science Policies and beyond, October 22nd, 2019 by Emilie Hermans (Ghent University)
https://www.slideshare.net/OpenAIRE_eu/horizon-2020-open-science-policies-and-beyond-with-emilie-hermans-openaire
Funders: OA Mandatory Policies

Projects funded with public funds

- **Italian Law 112/2013:** Publicly funded research (+50%) outputs in Scholarly Journals (18-24 month embargo, preprint, AM, VoR)

- **Plan S & cOAlition S** funders: rights retention via mandatory deposit with CC-BY licence, no embargo, AM o VoR version, Hybrid OA only under Transformative Agreement

- **EU Horizon Framework (H2020 / H Europe – MSCA fellowships),** see Plan S & Open Res. Europe

- **ERC** [Article 29.2]: mandatory deposit 6-12 M Embargo AM o VoR version. The ERC Scientific Council recommends as disciplinary repos: Europe PubMed Central for publications in the life sciences domain and arXiv for publications in the physical and engineering sciences. Recommended repository for monographs (etc.) is the OAPEN Library.

- **MUR (ITA) SIR 2014, PRIN 2015, PRIN 2017, PRIN 2020,** see Law 112/2013

- **UniPD Supporting TAlent in ReSearch** - STARS Grants 2019, 2020, 2021

Projects funded with private funds

- Bill & Melinda Gates Foundation
- Wikimedia Foundation
- Telethon
- …
UniPD promotes the free and open distribution of the University's research results, as defined in its:

- Statute [Title III, Art. 56, Paragraph 3]
- “Policy sull’Accesso Aperto (Open Access) alla letteratura scientifica"
- "Regolamento per l'Accesso Aperto (Open Access) alla produzione scientifica dell'Università di Padova"
Facilitations for authors
Facilitations for UniPd authors

Read and Publish contracts aka Transformative agreements

https://bibliotecadigitale.cab.unipd.it/biblioteca-digitale/per-chi-pubblica/evolazioni-per-gli-autori

Read & Publish contracts
main characteristics

- they are **commercial contracts** and, just like the “classic” subscription contracts, they are negotiated between publishers and institutions that in Italy are represented by CARE-CRUI;
- in addition to historical contents with closed access, they offer an open access publishing service;
- they mainly apply to **hybrid journals**, but in some cases they also apply to full open access journals (Gold OA).

  - authors no longer pay **APCs** (paid centrally by the institution);
  - authors retain the rights to their works
How does OA Read & Publish work?

- when submitting the article, the **corresponding author** follows the instructions given by the publisher, selecting the OA option;

- **the corresponding author is identified** by the publisher as an affiliate of UniPD through these information:
  1. Membership body
  2. Institutional address
  3. ID (university network or auth-proxy);

- **after the article has been accepted**, the University Library Center is required to verify the correctness of the information provided by the corresponding author and to definitively approve the OA option.
Contracts in progress

● American Chemical Society - ACS (Hybrid OA)
● Annual Reviews (S2O)
● Cambridge University Press - CUP (Hybrid & Gold OA)
● De Gruyter (Hybrid OA)
● Emerald (Hybrid OA)
● IEEE (Hybrid & Gold OA)

● Institute of Physics - IOP (Hybrid & Gold OA)
● Lippincott (Hybrid OA)
● Royal Society of Chemistry - RSC (Hybrid OA)
● Springer (Hybrid OA)
● Wiley (Hybrid OA)
Other kinds of facilitations (discounts on APCs)

- British Medical Journal - BMJ
- Elsevier
- Taylor & Francis
- SCOAP3
- MDPI
Facilitations for Authors

Thanks to specific agreements stipulated with publishers, Unipd authors who wish to publish their work in Open Access can take advantage of facilitations on the payment of APCs (Article Processing Charge).

Transformative agreements

According to the new type of agreements defined as “transformative”, only authors belonging to the institutions adhering to the specific transformation contract can publish without additional costs. Each article managed under the agreement is defined “token” or “voucher”: the vouchers for the free publication of Open Access articles, although available in large numbers for each institution adhering to the consortium contract, are limited at the national level.

Requirements for accessing the facilitations provided for by the Transformative Agreements

The corresponding author must formally belong to the University of Padua. Furthermore, the use of the institutional e-mail @unipd.it is preferable.

Repositories
Disciplinary archives

= databases collecting the production of a specific scientific community
Interdisciplinary archives: Zenodo

For self-archiving of publications and data, open to all researchers in the world.

Managed by CERN for OpenAIRE (EU)

Search results are stored securely in the same cloud infrastructure as CERN's LHC search data.

Assignment of a digital object identifier (DOI)

Possibility of identifying any subsidies, as they are integrated into the reporting lines.

Possibility of assigning flexible guarantees of use, since not everything is under Creative Commons.

https://zenodo.org/
Institutional archives

= databases that collect the scientific production of an institution

Main types of documents:

- Scholarly articles
  - Pre print
  - Post print
  - Version of record
- Conference papers
- Book chapters
Trustworthy Digital Repositories (TDRs)

<table>
<thead>
<tr>
<th>Principle</th>
<th>Guidance for repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparency</td>
<td>To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.</td>
</tr>
<tr>
<td>Responsibility</td>
<td>To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.</td>
</tr>
<tr>
<td>User Focus</td>
<td>To ensure that the data management norms and expectations of target user communities are met.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>To sustain services and preserve data holdings for the long-term.</td>
</tr>
<tr>
<td>Technology</td>
<td>To provide infrastructure and capabilities to support secure, persistent, and reliable services.</td>
</tr>
</tbody>
</table>

in: Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). [https://doi.org/10.1038/s41597-020-0486-7](https://doi.org/10.1038/s41597-020-0486-7)


Institutional Repositories vs. Academia.edu or ResearchGate

<table>
<thead>
<tr>
<th>Feature</th>
<th>Open access repositories</th>
<th>Academia.edu</th>
<th>ResearchGate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supports export or harvesting</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Long-term preservation</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Sends you lots of e-mails (by default)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Wants your address book</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fulfills requirements of Unipd’s OA policies</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
PRA (IRIS): homepage - info and instructions

https://www.research.unipd.it/
PRA (IRIS) and PhD students

Regolamento per i corsi di dottorato di ricerca

ART. 24.9: E’ condizione necessaria per il conseguimento del titolo di dottore di ricerca che il dottorando acceda al catalogo IRIS e inserisca i dati relativi alla propria produzione scientifica, se presente.

[It is a necessary condition for the attainment of the PhD degree that students log into IRIS catalogue and submit information about their own scientific production, if produced.]
<table>
<thead>
<tr>
<th>PRA (IRIS): Research outputs archive</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Research outputs archive</td>
</tr>
<tr>
<td>• Open Access archive</td>
</tr>
<tr>
<td>(pre-print; post-print - embargo if needed; editorial version, if it is allowed)</td>
</tr>
<tr>
<td>• Platform for the assessment of research (eg. VQR 2015-2019, ASN, ...):</td>
</tr>
<tr>
<td>○ if publication rights are transferred to the publisher: the attachment will be visible only to the evaluators</td>
</tr>
<tr>
<td>○ if contribution are already published in OA: the attachment will be visible to everyone (Attachments declared completely open access by authors are still checked)</td>
</tr>
</tbody>
</table>
Padua Research Archive (PRA) as an Open Access archive: IRIS beyond Research evaluation

Once you have uploaded your work to IRIS/PRA the research support group:

- Checks whether the publisher’s policies allow you to publish open access
- Checks the embargo dates and validates the attachment
- Supports authors via SBA Help - Research Support - OA

The validation process involves a delay in the publication of the OA content in PRA, but protects the author. It is possible to report contributions that need to be displayed faster on the public portal.
<table>
<thead>
<tr>
<th>Who should you contact when uploading your documents in PRA/IRIS?</th>
<th>Who should you contact for technical problems related to PRA/IRIS?</th>
<th>Who should you contact when uploading documents in open access?</th>
</tr>
</thead>
<tbody>
<tr>
<td>People charged by departments</td>
<td>Settore supporto informativo valutazione della ricerca e qualità - Ufficio ricerca e Qualità helpdesk coda: &quot;Catalogo Padua Research Archive (IRIS)&quot;</td>
<td>Library System Helpline</td>
</tr>
</tbody>
</table>
Tips on the management of research data
What are research data?

Recorded information (regardless of the form or the media in which they may exist) necessary to support or validate a research project's observations, findings or outputs

BUT ALSO…
- Computer Aided Design (CAD)
- Waveforms
- Computer codes
- Statistics (SPSS, SAS)
- File Matlab
- Artistics products
- Web files
- …
# Data Types

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Value</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational data captured around the time of the event</td>
<td>Usually irreplaceable</td>
<td>Sensor readings, telemetry, survey results, neuro-images</td>
</tr>
<tr>
<td>Experimental data from lab equipment</td>
<td>Often reproducible but can be expensive</td>
<td>Gene sequence, chromatograms, toroid magnetic field readings</td>
</tr>
<tr>
<td>Simulation data generated from test models</td>
<td>Model and metadata (inputs) more important than output data. Large modules can take a lot of computer time to reproduce</td>
<td>Climate models, economic models</td>
</tr>
<tr>
<td>Derived or compiled data</td>
<td>Reproducible (but very expensive)</td>
<td>Text and data mining, compiled databases, 3D models</td>
</tr>
</tbody>
</table>
Data Costs

Data management costing tool

Guides for Researchers
How to identify and assess Research Data Management (RDM) costs

Unless otherwise indicated, all materials created by OpenAIRE are licenced under CC ATTRIBUTION 4.0 INTERNATIONAL LICENSE
Research data lifecycle

From: Passport for Open Science – A Practical Guide for PhD Students
Licensed under a CC BY-SA licence
Open Data

It is the philosophy of Open Access applied to data
Data are open when anyone can access, use and share
Anyone means: academics, politicians, private citizens, economic stakeholders
Open data: a five-stars rating system

OL (On Line)  • available on the web + distributed with an open license
RE (Readable) • machine-readable structured data
OF (Open Format)  •• encoded with non-proprietary software
URL (Uniform Resource Identifier)  ••• identified by an URL
LD (Linked Data)  ••••• linked to other data sets
Funding programs requiring OA: European Union

COMMISSION RECOMMENDATION (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information

DIRECTIVE (EU) 2019/1024 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 June 2019 on open data and the re-use of public sector information

Horizon Europe Programme Guide
Managing research data: 5 steps

1. Collect research data
2. Reasonably name data
3. Structure data in hierarchical systems
4. Record data through metadata
5. Pay attention to the file format (Guide on "naming and version control")

Tips on metadata standard for different disciplines
First step: collect research data

1. Develop a clear picture of the data you need
2. Locate appropriate data resources
3. Set up a search query and search the data resource
4. Select data candidates
5. Evaluate data quality

Second step: file name strategy

A file name is a principal identifier of a file

- File name should help to identify the content of the file.
- Good file names provide useful clues to the status and version of a file, uniquely identify a file and help in classifying and sorting files.

File naming strategy should be consistent in time and among different people

- In both quantitative and qualitative research file naming should be systematic and consistent across all files in the study
- A group of cooperating researchers should follow the same file naming strategy.

Third step: structure research data

Structuring your data files in folders is important for making it easier to locate and organize files and versions.

The decision on how to organize your data files depends on the plan and organization of the study. All material relevant to the data should be entered into the data folders, including detailed information on the data collection and data processing procedures.

Fourth step: annotate using metadata

Metadata means "data about data".

It is defined as the data providing information about one or more aspects of the data and it is used to summarize basic information about data, which can make easier to track and work with specific data.

Examples of metadata standards

http://www.ucl.ac.uk/library/research-support/research-data/best-practices-guides/creating
Fifth step: file formats

When preparing to collect research data, you should choose **open, well-documented and non-proprietary formats** wherever possible. The choice of format will vary depending on how you plan to analyze, store and share your data.

It is advisable to **store your data for use in future**, which means to convert them from a current data format to a long-term preservation format. Most software applications offer export or exchange formats that allow a text-formatted file to be created for importing into another program.

Guides on **formats**

Recommended **formats**
Organize data: dataset versions

Versioning is important for long-term research data management where metadata and/or files are updated over time.

It is used to track any metadata or file changes (e.g., by uploading a new file, changing files structure, adding or editing file metadata…) once a dataset has been published.

Draft dataset

Dataset v 1
Authors, Title, Year, DOI, Repository

Dataset v 2
Big metadata change
Changes in data

Dataset v 1.1
Small metadata changes (e.g. Title editing)

Useful guides on naming and version control

Data **storage** in safe archives adhering to relevant standards.

**Preservation** actions should ensure that data remains authentic, reliable and usable while maintaining its integrity.

**Checklist for storage and preservation**

- Regular backup
- Checkup of integrity of files
- Track changes in metadata and files (versioning)
- Open, non-proprietary, well documented formats
- Multiple and different storage media
- Copy or migration of files

Checklist for [storage](#) and [preservation](#)
Since 25 May 2018, the General Data Protection Regulation (GDPR, European Union, 2016) applies to any EU researcher who collects personal data of living persons.

So, when processing personal data, researchers should adhere to the following six principles:

I. Process lawfully, fair and transparent
II. Keep to the original purpose
III. Minimise data size
IV. Personal data should be accurate and, where necessary kept up to date
V. Remove data which are not used
VI. Ensure data integrity and confidentiality

The so-called "research exemption" means that principles 2 and 5 are less stringent when applied to data collected for research projects, public interest, statistical purposes.
Privacy and personal data

Works containing sensitive data relating to identifiable persons must not be disseminated in Open Access!

BEFORE collecting data:
- Carry out a risk assessment
- Choose which data to collect + follow the minimization principle
- Prepare an informed consent document (information about the research, the subjects involved, the way data is going to be shared and stored)

AFTER collecting data:
- Protect the identities of the subjects involved (e.g. pseudonyms; keeping the information that allows identification in a separate archive)
- Anonymize + aggregate data
- Regulate access

GDPR (General Data Protection Regulation)  Information on research integrity and research ethics
DMP = Data Management Plan

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>□ What kind of data are generated</td>
<td>□ How will ethical issues be handled</td>
<td>□ How are the data stored?</td>
<td>□ How and where will the data be shared?</td>
</tr>
<tr>
<td>□ How will data be generated</td>
<td>□ How are the data accessed</td>
<td>□ Are there back up systems</td>
<td>□ How are sensitive data protected</td>
</tr>
<tr>
<td>□ What metadata are needed</td>
<td>□ Are there copyright issues</td>
<td>□ How are data safely preserved</td>
<td>□ How can data be accessed</td>
</tr>
<tr>
<td>□ Are there sensitive data</td>
<td>□ What about intellectual property rights</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


© Universität Zürich
DMP: Guidelines & tools

**DCC = Digital Curation Centre**
- [https://www.dcc.ac.uk/dmponline](https://www.dcc.ac.uk/dmponline)

**DMP Online**
DMP templates according to funders' specific requirements
- [https://dmponline.dcc.ac.uk/public_templates](https://dmponline.dcc.ac.uk/public_templates)

**Data Stewardship Wizard**
DMP which can be exported using selected template and format, including machine-actionable ones
- [https://ds-wizard.org/](https://ds-wizard.org/)

**Science Europe**
Research Data Management Guidance and Templates for DMPs

**Horizon Europe**
Funding & tender opportunities > Project reporting templates > Data management plan (HE)

**Tools for researchers**
UniPD Ufficio Ricerca Internazionale: strumenti per la progettazione e il proposal writing (with SSO)
- [https://elearning.unipd.it/ufficiserviziapplicazioni/course/view.php?id=112](https://elearning.unipd.it/ufficiserviziapplicazioni/course/view.php?id=112)
DMP: two examples

**CESSDA ( Consortium of European Social Science Data Archives)**

Link to pdf version [here](#)  
Link to editable version in this [page](#)  

**DCC (Data Curation Centre)**

Link to DMP Checklist [here](#)
Research Data Unipd
Research Data Unipd

Authentication via SSO of the University

FAIRness: self-archiving of datasets of any format with FAIR mode (Findable, Accessible, Interoperable, Reusable)

Connection between dataset and articles from the publisher's website or stored in Padua Research Archive / IRIS

DOI attribution

ERC subjects
FAIR principles

https://book.fosteropenscience.eu/
FAIR principles

Findable
Rich metadata
Persistent identifiers

Accessible
Metadata always available
Open, standard protocols
Documented formats

Reusable
Usage licenses
Info on provenance

Interoperable
Standard vocabularies
Linked metadata

https://www.fairsfair.eu/
About the Repository

Research Data Unipd supports research produced by members of the University of Padua. The service aims to facilitate data discovery, data sharing, and reuse as required by funding institutions (e.g. European Commission).

According to the University of Padua Policy on the Management of Research Data, data must be stored in a correct, complete and reliable way, respecting their integrity. They must also be accessible, identifiable, traceable, interoperable and, where possible, available for subsequent use as stated in the FAIR principles.

Quality

Datasets published in the Archive have a set of metadata that ensure that data are described and discoverable. Before publication, dataset records are checked by Editors for the presence of appropriate metadata.

Metadata Policy

All published metadata are released under a [CC0 licence](https://creativecommons.org/publicdomain/zero/1.0/).

Re-using data

We encourage researchers to use licences on their datasets to promote the reuse of the research data. The licence to be preferred is Creative Commons Attribution 4.0, but several others are used. Any re-use must acknowledge the Creators in an appropriate manner, ideally through a citation similar to that provided with the record.

Recommended formats and data files

[Formats and data files](https://data.unipd.it).
Licenses on Data

Open Data Commons Licenses

- **PDDL**
  - public domain

- **ODC-by**
  - attribution

- **ODbL**
  - attribution & share-alike

https://opendatacommons.org/index.html
Data deposit agreement

When you deposit data in the Research Data Unipd Archive, you must agree to the conditions below. This is done by clicking the "Deposit" button in the archive, before depositing the item.

This agreement confirms that you, the depositor, have the right to submit the dataset to the repository.

This agreement ensures that the archive administrators have the right to carry out activities necessary to facilitate the long-term preservation and sharing of datasets.

By submitting your dataset for a deposit, you grant a non-exclusive licence to the University of Padua to archive, publish and disseminate any material within the dataset. The licence is non-exclusive and therefore does not prevent you from exercising any rights you might have to publish and distribute any of the datasets, in its present or future versions, elsewhere.
Retractions

Retraction of articles due to lack of data, citations and methodologies are increasing.

The inability to access to data and methods prevents from reproducing experiments and validate results.
Steps dominate gas evasion from a mountain headwater stream

Data & article reported findings

The data that support the findings of this study are openly available in Botter et al. 2022 at http://researchdata.cab.unipd.it/id/eprint/619, reference number 619.
Open methodology

= the use of open methodologies throughout the entire research cycle, making it open and available to everyone online at the very moment the research is conducted.

Open Notebooks

● https://openlabnotebooks.org
● https://theopennotebook.com/
● OpenLab/Notebook % Foster

● Code Ocean
● Protocols.io
To sum up … why is it important to manage research data [properly] and make them OPEN?

- To allow the continuity of research through the use of secondary data
- To increase the efficiency of research
- To ensure compliance with the requirements set by funders
- To support the contents of a paper and improve the peer-review
- To guarantee the integrity of research and the validation of the results
- To ensure greater dissemination and greater impact
<table>
<thead>
<tr>
<th><strong>American Chemical Society Publications (ACS)</strong></th>
<th>The American Chemical Society Publications (ACS) is a non-profit scholarly publisher that provides a comprehensive collection, in any medium, of high-quality information products and services that advance the practice of the chemical and related sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U.S. Energy Information Administration (EIA)</strong></td>
<td>The U.S. Energy Information Administration (EIA) provides a wide range of information and data products covering energy production, stocks, demand, imports, exports, and prices. EIA is committed to enhancing the value of its free and open data by making it available through open data tools.</td>
</tr>
<tr>
<td><strong>Open Access Scholarly Publishers Association (OASPA)</strong></td>
<td>The Open Access Scholarly Publishers Association (OASPA) is a non-profit trade association representing the interests of open access journal publishers globally in all scientific, technical and scholarly disciplines.</td>
</tr>
<tr>
<td><strong>“S-légami! Open Access – Manuale d'uso per ricercatori”</strong></td>
<td>“S-légami! Open Access – Manuale d'uso per ricercatori” is a freely available manual that was born in the APRE Working Group dedicated to Open Science and contains the answers to the most frequent questions and concerns of researchers on open access and open data.</td>
</tr>
<tr>
<td><strong>OpenAIRE</strong></td>
<td>OpenAIRE is a pan-European research information system, which provides services for finding, storing, linking and analyzing research results from all disciplines. Its mission is to move academic communication towards openness and transparency and to facilitate innovative ways to communicate and monitor research.</td>
</tr>
</tbody>
</table>
Supporting authors
Library System support services

In the section “About publishing” of the Library System web portal, researchers will find information on Open Access, on publishing, and on the management of data.
Library System support services

Authors can submit specific requests using the Library System Help Service, choosing the address:

- Supporto alla pubblicazione accademica
- Consulenza tematiche diritto d’autore
- Supporto Open Science (Open Access, Open Data)

Engineering libraries contacts:
biblio.inge@unipd.it

Kyle James https://www.flickr.com/photos/jameskm03/2711755476