Research (FAIR) data management

| Mariarita de Luca |

https://orcid.org/0000-0002-5507-968X
mariarita.deluca@areasciencepart.it

Institute for Research and Innovative Technologies (RIT)
AREA SCIENCE PARK

1° Workshop for National PhD in "Theoretical and Applied Neuroscience", Bertinoro 18.10.2023

Scuola Internazionale Superiore di Studi Avanzati SISSA

Research office

(https://www.sissa.it/research-funding)

SISSA

research@sissa.it
Research Office SISSA – Scuola Internazionale Superiore di Studi Avanzati

Work experience and education

Research Manager and Administration (RMA) at Research Funding and International Relations Office

Secondary school teacher (mathematics and physics)

Senior Post-doctoral Fellowships

PhD in Applied Mathematics

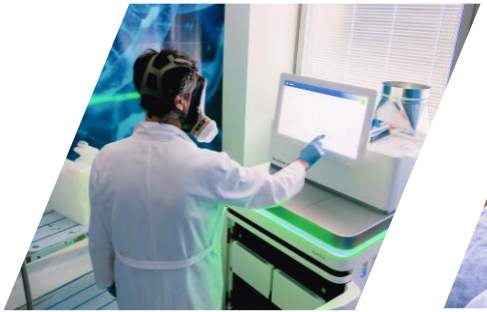
Visiting student(during PhD)

AREA SCIENCE PARK (TRIESTE)

Research Institute for Technological Innovation

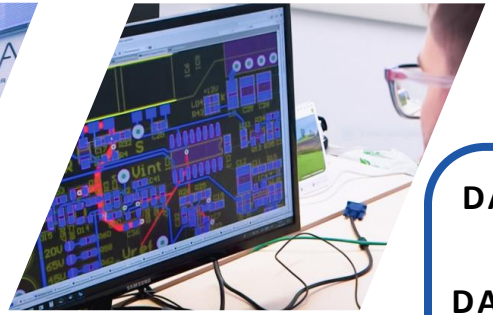
Three laboratories active in creating an **integrated system** of research infrastructures and platforms.

They provide, in an **open access mode**, knowledge and services aimed at carrying out experimental testing and applied and industrial research projects.



GENOMICS AND EPIGENOMICS LABORATORY
LAGE

ELECTRON MICROSCOPY LABORATORY
LAME



DATA ENGINEERING LABORATORY
LADE

DATA INFRASTRUCTURE

DATA MANGEMENT AND CURATION

DATA SCIENCE

LIFE SCIENCES PLATFORM

The **Life Sciences Platform** integrates the instrumentation and expertise of big labs located in Area Science Park campuses dedicated to **structural, cellular and molecular biology, biophysics and microscopy**, with applications in the **pharmaceutical, biotechnology, nutraceutical and diagnostics sectors**.

AREA SCIENCE PARK

INNOVATIVE MATERIALS PLATFORM

The **Innovative Materials Platform** integrates the instrumentation and expertise of big labs located in Area Science Park Campuses dedicated to application of **imaging, structural-analysis and chemical-characterization techniques for the study and development of advanced surfaces and materials**.

AREA SCIENCE PARK

Outline

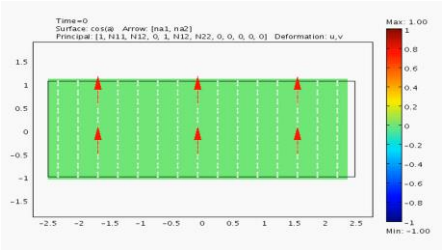
- Research data
- Research data lifecycle
- FAIR principles and tools
- Data management tools and practices
- Data management Plan example

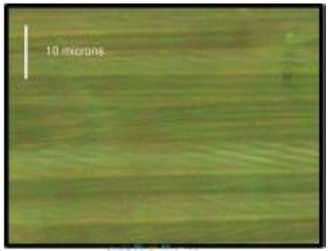
AREA SCIENCE PARK

Past scientific interests

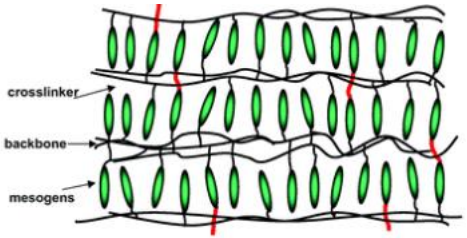
Mathematical models for soft-active materials

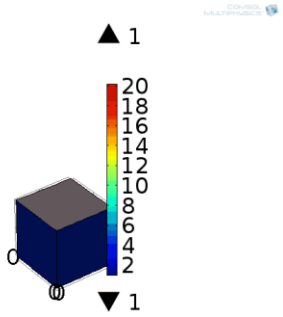
- Elasticity within large deformation framework (non-linear models)
- Deformation of active-smart materials (swelling materials, nematic elastomers, …)

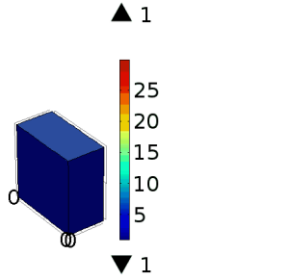




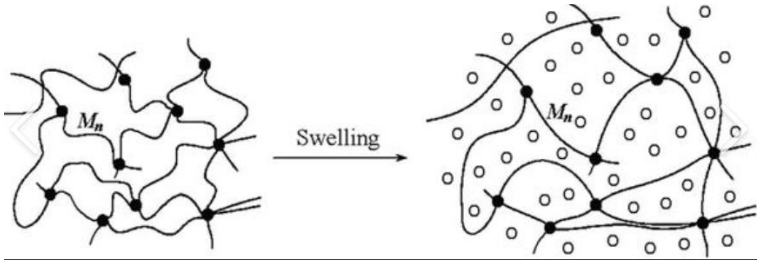
M. de Luca, A. Petelin, M. Copic and A. DeSimone, "Sub-stripe pattern formation in liquid crystal elastomers: Experimental observations and numerical simulations", JMPS, 61 (2013) 2161 – 2177
https://doi.org/10.1016/j.jmps.2013.07.002







M. de Luca, A. DeSimone. Elastomeric Gels: A Model and First Results. Innovative Numerical Approaches for Multi-Field and Multi-Scale Problems.
Lecture Notes in Applied and Computational Mechanics, vol 81. Springer, Cham. (2016)
https://doi.org/10.1007/978-3-319-39022-2_4



AREA SCIENCE PARK

What about my data and my publications?

- Do I have access to my publications?
- Where are my data?
- Can I reproduce my numerical simulations?

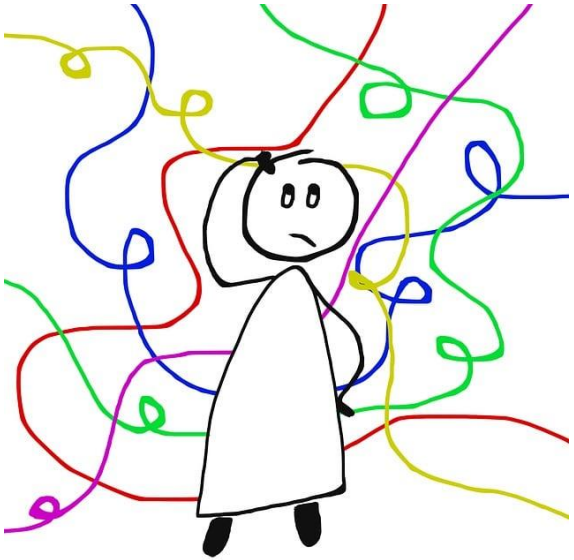


Image by Elisa from Pixabay

AREA SCIENCE PARK

Research Data

Research data are the **raw materials** collected, processed and studied in the undertaking of research. They are the evidential basis that substantiates published research findings.

They may be **primary data** generated or collected by the researcher, or **secondary data** collected from existing sources and processed as part of the research activity.

In addition to the 'raw' data, **research data include information about the means necessary to generate data or replicate results**, such as computer code, experimental methods and instruments used, and essential interpretive and contextual information, e.g. specifications of variables.

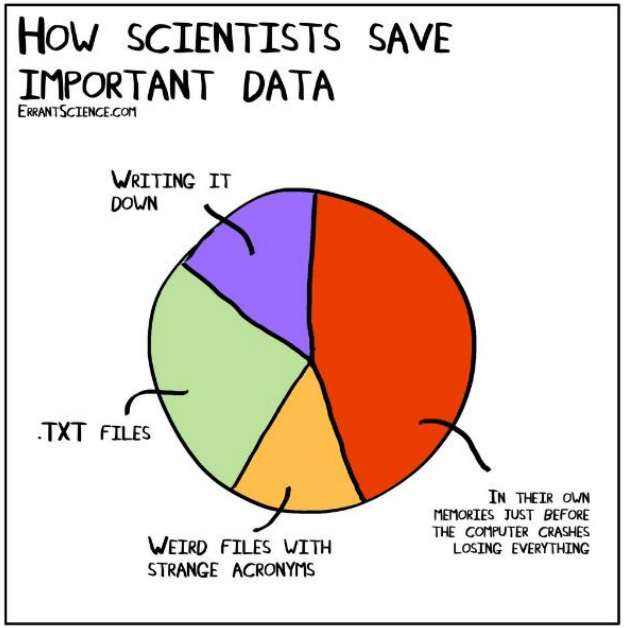


Image by Errant Science

[Research data defined (reading.ac.uk)]

Raw research data may exist in digital and non-digital formats

Observational
Facts recorded directly in real time from the physical and social environment, e.g. measurements collected by weather sensors, species abundance surveys, archaeological samples, brain scan images, experience and opinion surveys in the social sciences. These data are often unique to time and place and by definition cannot be reproduced.

Experimental
Data collected as the outputs of field or laboratory experiments and complex analytical processes, e.g. clinical trial data, chemical analyses of physical samples, DNA sequencing of organic material, field trial results. These data are generally in principle reproducible, assuming the experimental conditions can be replicated.

AREA SCIENCE PARK

Research data defined (reading.ac.uk)

Raw research data may exist in digital and non-digital formats

Simulation
Data generated by means of computational 'virtual experiments', often used to model complex systems and processes, e.g. climate and weather simulations, models of market processes. These data are usually reproducible, given information about the model, the code and computing environment used to execute the model, and any input conditions. This information may in fact be more important that the output data.

Derived or compiled
Datasets produced by processing or combining source data, e.g. databases compiled by extraction of information from multiple secondary sources, collections of digitized materials, corpora collected by means of text mining.

Reference
Published and curated data, usually existing as part of managed collections, e.g. national statistics archives, crystallographic databases, gene banks.

AREA SCIENCE PARK

Research data defined (reading.ac.uk)

Research data lifecycle

The Biomedical Data Lifecycle is a representation of stages in your research regarding data:

- **plan and design**
- collection and creation
- analysis and collaboration
- evaluation and archive
- sharing and dissemination
- publication and reuse

Research Data Management (RDM) is an essential element in any project that generates, collects or re-uses data.
Planning ahead to data needs that proposers are likely to encounter during the project is a best practice.

Harvard Biomedical Research Data Lifecycle, 10.5281/zenodo.8075933

Data Management Plan - DMP

DMPs are formal documents that outline from the start of the project all aspects of the **research data lifecycle** both during and after a project.

Writing a DMP is **part of the methodology of the project**, since good data management makes the work more efficient, saves time, contributes to safeguarding information and to increasing the value of the data during and after the research.

DMPs are thus a key means of support **when planning and conducting a research project**, and, ideally, filling in a DMP should be started prior to the beginning of the project.

A DMP should be a **living document**, which is updated and enriched as the project evolves.

programme-guide_horizon_en.pdf (europa.eu)

Image by Memed_Nurrohmad from Pixabay

AREA SCIENCE PARK

DMP

A good practice regarding DMPs is to register them as a non-restricted public deliverables to make them openly accessible, **unless legitimate reasons exist to keep them confidential.**

The DMPs should reflect **common disciplinary practices.**

The DMP must describe the research methodology, in line with the **FAIR principles for data management.**

Picture from «The Turing Way» 10.5281/zenodo.3332807

programme-guide_horizon_en.pdf (europa.eu)

F A I R principles

There is an urgent need to improve the infrastructure **supporting the reuse of scholarly data**.

A diverse set of stakeholders—representing **academia, industry, funding agencies, and scholarly publishers**—have come together to design and jointly endorse a concise and measureable **set of principles that we refer to as the FAIR Data Principles**.

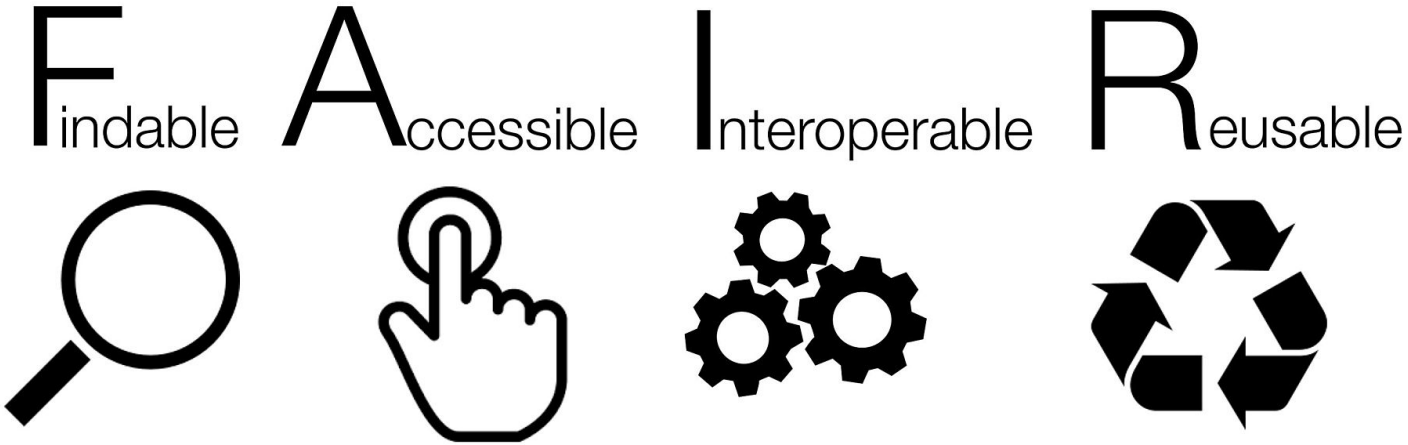
The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings.

Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on **enhancing the ability of machines to automatically find and use the data**, in addition to supporting its reuse by individuals.

This Comment is the **first formal publication of the FAIR Principles**, and includes the rationale behind them, and some exemplar implementations in the community.

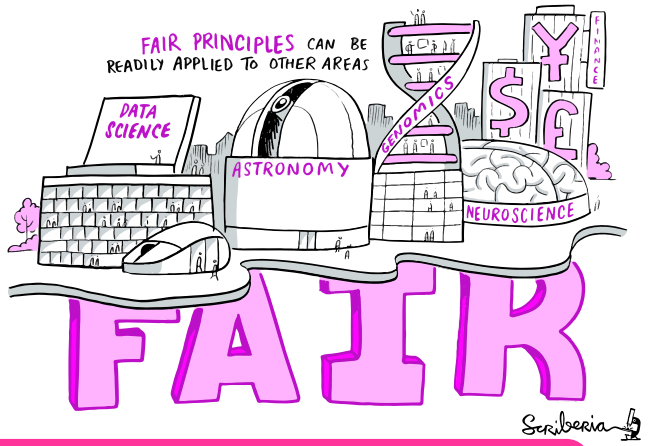
Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

F A I R principles



AREA SCIENCE PARK

F A I R principles



Picture from «The Turing Way» 10.5281/zenodo.3332807

Good data management is not a goal in itself, but rather **is the key conduit leading to knowledge discovery and innovation**, and to subsequent data and knowledge integration and reuse by the community after the data publication process.

The emphasis placed on FAIRness being applied to **both human-driven and machine-driven activities**, is a specific focus of the FAIR Guiding Principles that distinguishes them from many peer initiatives.

These high-level FAIR Guiding Principles **precede implementation choices**, and do not suggest any specific technology, standard, or implementation-solution;

Moreover, the Principles are not, themselves, a standard or a specification.

They **act as a guide** to data publishers and stewards to assist them in evaluating whether their particular implementation choices **are rendering their digital research artefacts Findable, Accessible, Interoperable, and Reusable.**

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

AREA SCIENCE PARK

FAIR principles apply during all the phases of the Research data lifecycle

Harvard Biomedical Research Data Lifecycle, 10.5281/zenodo.8075933

Box 2 | The FAIR Guiding Principles

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

FAIR principles

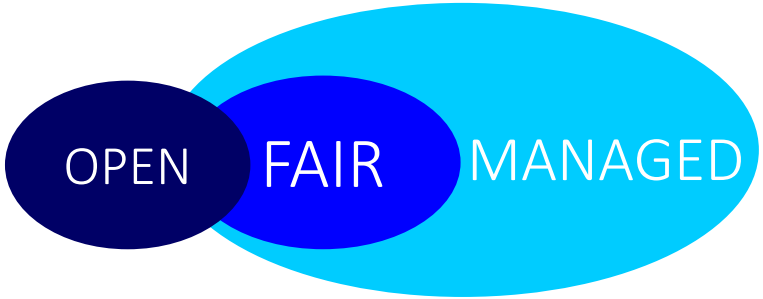
AREA SCIENCE PARK

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

FAIRness VS OPENness

Managing data properly is in the primary interest of any researcher, as the whole research process results streamlined and more effective

FAIR ≠ OPEN



1. DATA SHOULD BE AS OPEN AS POSSIBLE

2. BUT IF DATA ARE NOT «FAIR», OPENING IS RISKY (MISUSE, MISINTERPRETATION, …)

3. IF DATA ARE NOT PROPERLY MANAGED FROM THE BEGINNING, IT'S ALMOST IMPOSSIBLE TO MAKE THEM «FAIR» [WITH EOSC MANAGED/FAIR INCREASINGLY OVERLAPPING, «FAIR BY DESIGN»]

In Horizon Europe data should be «as open as possible and as closed as necessary»

AREA SCIENCE PARK

Slides by Elena Giglia DOI 10.5281/zenodo.6761944

Metadata

Metadata: Data about data, infomation to describe how the data were created/collected, by whom, with which instruments, with what methodology, ….

For each specific discipline, we need **proper metadata** to correctly describe the data.

Metadata **must be machine readable,** then they must be written using standards, ontologies, taxonomies and semantic metadata.

Semantic metadata is data that describes the meaning of that data or content. In other words, it is making explicit the meaning of the metadata so that machines, and not just humans, can infer or interpret information about that metadata.

An **ontology** is a description of data structure–of classes, properties, and relationships in a domain of knowledge. It is meant to serve as a basis for instances of knowledge graphs, ensuring data consistency and understanding of the data model

Data taxonomy is the classification of **data into hierarchical groups to create structure**, standardize terminology, and popularize a dataset within an organization.

AREA SCIENCE PARK

FAIRsharing.org
standards, databases, policies

Digital object

information in binary form and its associated metadata

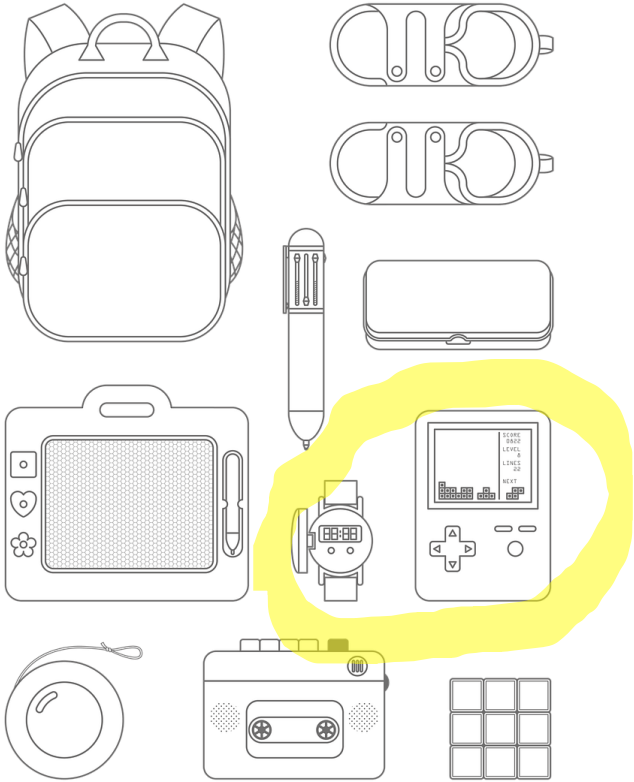


Image by yuelanliu from Pixabay

Persistent Identifiers (PIDs)

Globally, unique and long-lasting references to **digital objects** (such as data, publications and other research outputs) or **non-digital objects** such as researchers, research institutions, grants, etc.
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.

programme-guide_horizon_en.pdf (europa.eu)

The Open Researcher and Contributor ID (ORCID) is a non-proprietary alphanumeric identifier for the unique identification of scientists and other authors of scientific literature.

Definition: Wikipedia https://orcid.org/

AREA SCIENCE PARK

doi Foundation

A DOI is a digital identifier of an object, any object — physical, digital, or abstract.
Designed to be used by humans as well as machines, DOIs identify objects persistently.
They allow things to be uniquely identified and accessed reliably.
You know what you have, where it is, and others can track it too.

A digital indentifier of any object

Physical, digital, or abstract

Designed to be used by machines

(As well as humans)

Persistent identifiers

A long-lasting reference to an object or to information about an object

Here's an example

https://doi.org/10.1000/182

AREA SCIENCE PARK

Box 2 | The FAIR Guiding Principles

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

FAIR principles

AREA SCIENCE PARK

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

FINDABLE

To be findable, data must be:

- uploaded on a suitable repository
- well described with rich metadata
- have assigned a PID

How to find a repository?

OpenDOAR

OSFHOME

zenodo

AREA SCIENCE PARK

elixir

figshare

ACCESSIBLE

Foto di Arek Socha da Pixabay

Through the PID data are retrievable, with authentication and authorization protocols when possible.

To be accessible, data must define under wich condition the access is granted:

- open,
- restricted,
- closed,
- embargoed

The data repository has contingency plan to ensure that the metadata remain accessible even when the data is no longer available.

ACCESSIBLE ≠ OPEN

AREA SCIENCE PARK

An important difference

FINDABLE

- **Deposit:** upload a digital object (data, articles, …) on a platform (repository) that allows to correctly describe the object through medatada 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).

ACCESSIBLE

AREA SCIENCE PARK

Source: slide n. 49 , E. Lazzeri, F. Di Donato, FAIR principles and Open data, 10.5281/zenodo.4450515

INTEROPERABLE

Data interoperability is connected to the need of having **machine readable data and metadata.**

- The repository on which the data are deposited is connected or federated (interoperates) with other platforms/systems.
- The repository metadata schema is (machine) actionable or can be mapped in other metadata schema describing the same type of data.
- Metadata contains the links to other metadata (for example the ORCID of the researcher or the DOI of the publication).

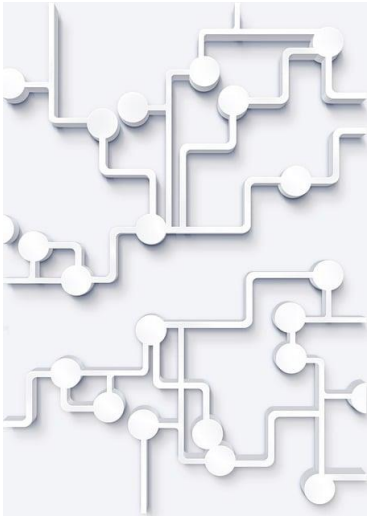


Foto di Gerd Altmann da Pixabay

AREA SCIENCE PARK

REUSABLE

Licensing FAIR Data for Reuse

DOI:10.1162/dint_a_00042

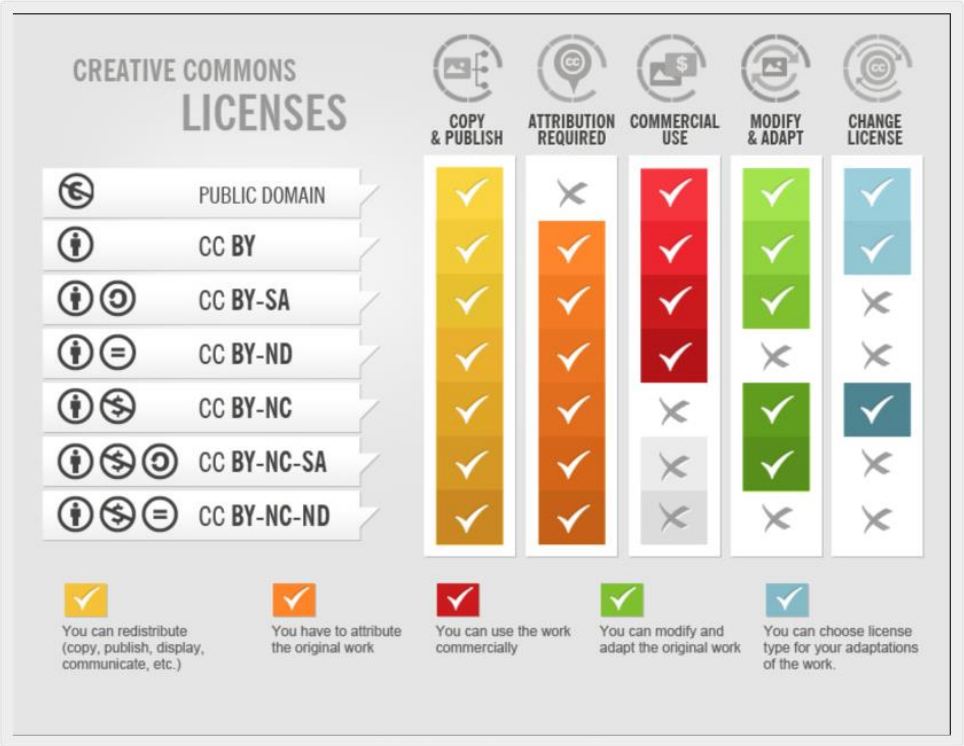
Define under which condition data can be reused (as they are, or modified) through an appropriate **license** (Creative Commons licenses).

Describe in the metadata all the additional information needed to reproduce your data (if possible).

Describe in the metadata the **provenance** of your data (sources, methodologies, ...).

AREA SCIENCE PARK

Creative Commons licenses



Does the law protect my data?

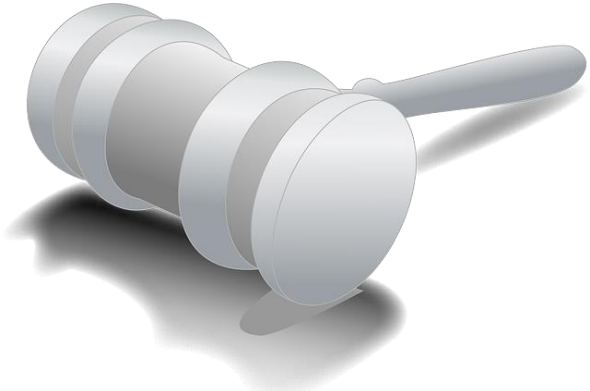


Foto di OpenClipart-Vectors da Pixabay

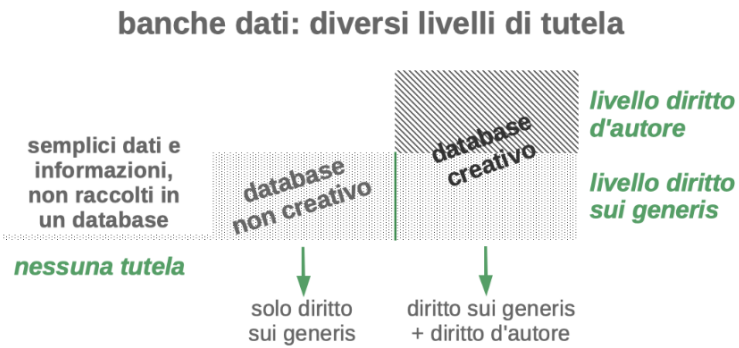
- **Primary (or raw) Data is not intellectual work, no copyright applies!**
- Copyright protection covers expressions and not ideas, procedures, operating methods or mathematical concepts as such.
- **Protection is on databases and not on data.**

How do I know if my research data is protected? (openaire.eu)

Data and law protection

Raw or primary 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)



Raw data are not YOUR data, the owner depends on who is funding your research (Institutional policies). In Horizon Europe, research data must be public (With public domain license or equivalent - C00).

ARE YOUR DATA FAIR?

FAIR | Aware

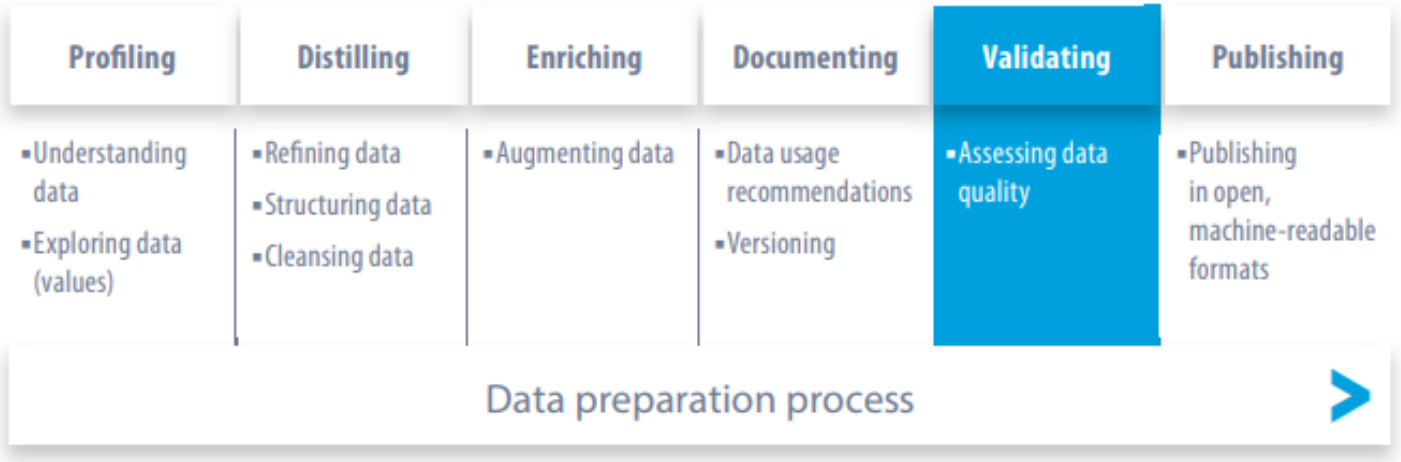
Your first step towards your FAIR data(set)

FAIR data infographic (CC-BY except F.A.I.R logos CC-BY-SA by Sangya Pundir

IF YOUR DATA ARE FAIR, YOU CAN DECIDE TO GO OPEN!!!

AREA SCIENCE PARK

Recommendations for providing high-quality data

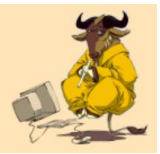


[Data.europa.eu data quality guidelines - Publications Office of the EU]

All your research output can be FAIR (and OPEN)

Repositories for Software:

- GitHub is development platform to host and review code, manage projects, and build software
- Savannah hosts free projects that run on free operating systems, with a focus on GNU software
- SourceForge is an Open Source software community and hosting platform
- Launchpad is a software collaboration and hosting platform



programme-guide_horizon_en.pdf (europa.eu)

All your research output can be FAIR (and OPEN)

Repositories for experimental workflows and protocols:

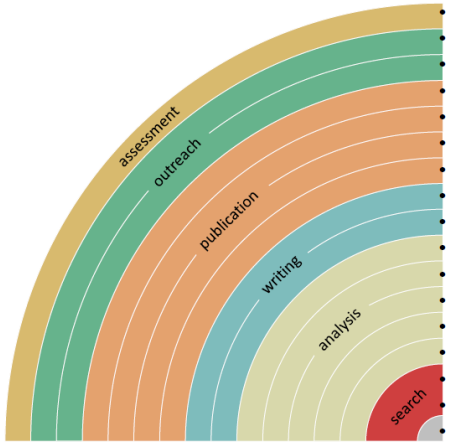
- Protocol Exchange (open repository for sharing scientific research protocols)

- Protocols (Platform for data management and protocol sharing

programme-guide_horizon_en.pdf (europa.eu)

OPEN SCIENCE tools

You can make your workflow more open by …



- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO

DOI: 10.5281/zenodo.1147025

Open science practices include

- early and open sharing of research (for example through preregistration, registered reports, pre-prints, or crowd-sourcing);
- research output management (DMP);
- measures to ensure reproducibility of research outputs;
- providing open access to research outputs (such as publications, data, software, models, algorithms, and workflows);
- participation in open peer-review;
- and involving all relevant knowledge actors including citizens, civil society and end users in the cocreation of R&I agendas and contents (such as citizen science).

AREA SCIENCE PARK

DO YOU NEED HELP?
ASK FOR SUPPORT

OpenAIRE

Get support on general open science practices
noad-it@openaire.eu
www.openaire.eu

Research Infrastructures

Get support from your specific domain
Research Infrastructure
www.esfri.eu

Competence Center

Be supported at national level from a team of experts in Open Science and EOSC related fields
Coming soon:
https://www.icdi.it/it/attivita/tf-cc

Your Institution

In case your institution has a structured open science strategy and dedicated support!

OpenAIRE

Slide from Emma Lazzeri, 10.5281/zenodo.4326562.

FAIR – by – design DATA

Data management fair-by-design means that the research data and their management are designed **from the beginning** of the project and in the most automated way in accordance with the **FAIR** principles.

AREA SCIENCE PARK

HOW TO WRITE A Data Management Plan (DMP)?

Common aspects that need to be addressed in all DMPs include:

- Data set description
- Standards and metadata
- Name and persistent identifier for the data-sets
- Curation and preservation methodology
- Data sharing methodology
- Output management, for research outputs other than data and publications
- **Costs and personnel related to RDM**

AREA SCIENCE PARK

DSW

DMPONLINE

argos

DMP – Summary section

- Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

AREA SCIENCE PARK

DMP – FINDABILITY section

- Will data be identified by a persistent identifier?
- Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
- Will metadata be offered in such a way that it can be harvested and indexed?

AREA SCIENCE PARK

DMP – ACCESSIBILITY section

Repository:

- Will the data be deposited in a trusted repository?
- Have you explored appropriate arrangements with the identified repository where your data will be deposited?
- Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

AREA SCIENCE PARK

DMP – ACCESSIBILITY section

Data:

- Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions.
- If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- Will the data be accessible through a free and standardized access protocol?
- If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
- How will the identity of the person accessing the data be ascertained?
- Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

AREA SCIENCE PARK

DMP – ACCESSIBILITY section

Metadata:

- Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?
- How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?
- Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

AREA SCIENCE PARK

DMP – INTEROPERABILITY section

- What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?
- In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?
- Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

AREA SCIENCE PARK

DMP – REUSABILITY section

- How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?
- Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?
- Will the data produced be useable by third parties, in particular after the end of the project?
- Will the provenance of the data be thoroughly documented using the appropriate standards?
- Describe all relevant data quality assurance processes.
- Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

AREA SCIENCE PARK

DMP – Other research outputs section

- In addition to the management of data, the reseracher should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).
- Researchers should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

AREA SCIENCE PARK

DMP – Allocation of resources section

- **What will the costs be for making data or other research outputs FAIR** in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?
- **How will these be covered?**
- Who will be responsible for data management in your project?
- How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

DMP – Data security section

- What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?
- Will the data be safely stored in trusted repositories for long term preservation and curation?

AREA SCIENCE PARK

DMP – Data security section

- Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review.
- Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?
- If you are dealing with personal data you must comply with the requests of GDPR.

DMP – Other issues section

- Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

AREA SCIENCE PARK

FLEXIN

Context-dependent flexibility in innate behaviours and their underlying neural circuitry

FLEXIN

Animal survival requires innate reactions to be fast, reliable, and flexible. The ERC-funded FLEXIN project will investigate the way by which brain circuits facilitate context-specific reactions to cues critical to survival such as threat and prey. Using a combination of viral tools, high-throughput neural recording techniques, and behaviour quantification in rodents, the study will uncover neural circuits which control changes in the animals' adjustment to different situations. The project will analyse how diverse contexts affect these well-defined circuits and behaviours, revealing principles of context-specific flexibility, neural circuit conservation, and brain region adaptation to environmental changes.

Courtesy of Dr. Katja Reinhard, Reinhard Lab - Team

DSW

Horizon Europe
Data Management Plan

AREA SCIENCE PARK

OPEN SCIENCE

(within the framework of European Projects)

| Mariarita de Luca |

https://orcid.org/0000-0002-5507-968X
mariarita.deluca@areasciencepart.it

Institute for Research and Innovative Technologies (RIT)
AREA SCIENCE PARK

1° Workshop for National PhD in "Theoretical and Applied Neuroscience", Bertinoro 18.10.2023

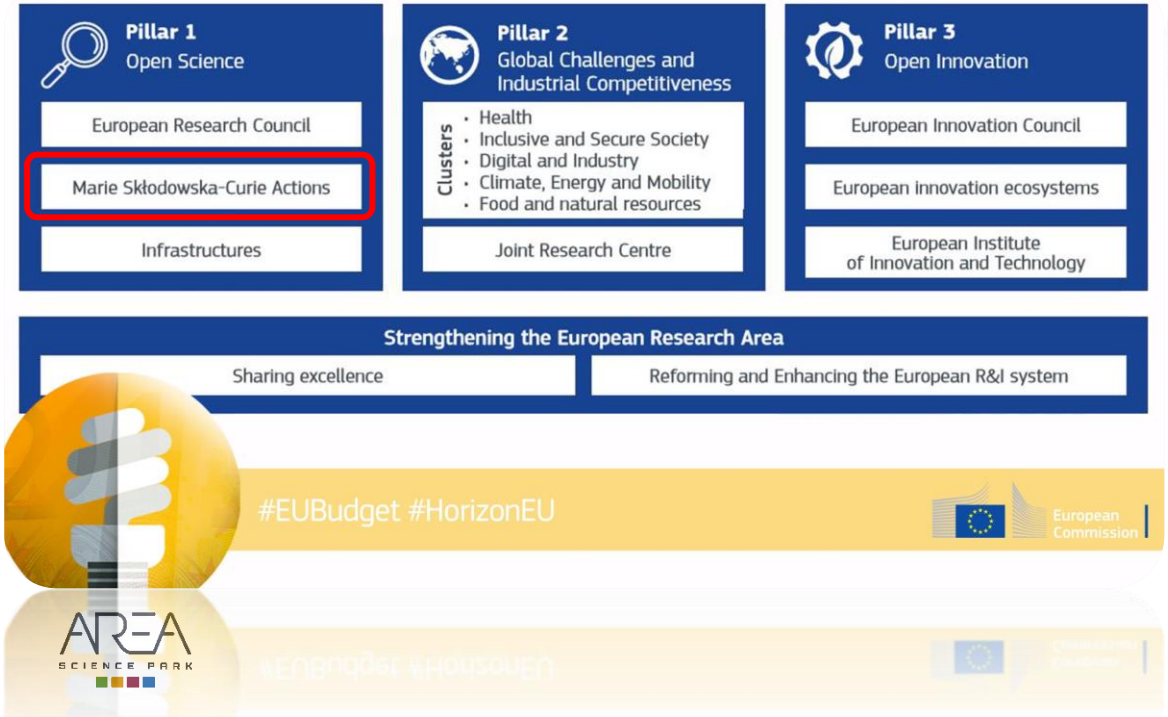
Outline

- European context (Horizon Europe - HE)
- Mandatory and suggested Open Science Practices in HE
- OS in European Projects: MSCA funding programme

AREA SCIENCE PARK

HORIZON EUROPE:

Horizon Europe is the EU's key funding programme for research and innovation with a budget of €95.5 billion



New elements in Horizon Europe

- **European Innovation Council**: Support for innovations with potential breakthrough and disruptive nature with scale-up potential that may be too risky for private investors. This is 70% of the budget earmarked for SMEs.

- **Missions**: Sets of measures to achieve bold, inspirational and measurable goals within a set timeframe. There are 5 main mission areas as part of Horizon Europe.

- **Open science policy**: Mandatory open access to publications and open science principles are applied throughout the programme.

New approach to partnerships: Objective-driven and more ambitious partnerships with industry in support of EU policy objectives.

AREA SCIENCE PARK

OPEN SCIENCE

(in Horizon Europe)

Open science is an approach based on open cooperative work and systematic sharing of knowledge and tools as early and widely as possible in the process.
It has the potential to increase the quality and efficiency of research and accelerate the advancement of knowledge and innovation by sharing results, making them more reusable and improving their reproducibility.
It entails the involvement of all relevant knowledge actors.

WHY DOES EUROPEAN COMMISSION REALLY PROMOTE OS?

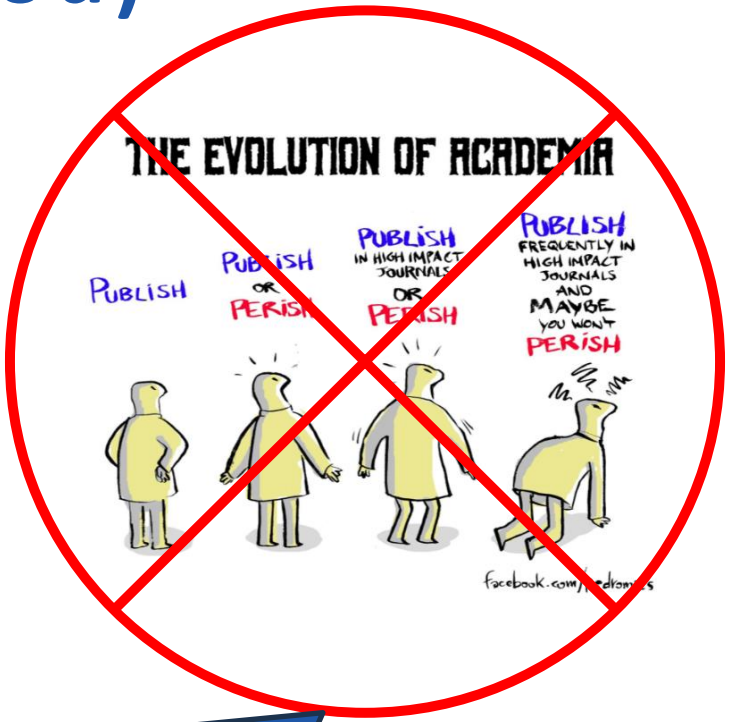
AREA SCIENCE PARK

OPEN SCIENCE (for you)

Open science is:

- a responsible research methodology
- a change of perspective, from competition to shared research
- a protected way to share your research (through CC licenses)
- an opportunity to rethink the research system (evaluation of research, publishing of research results, ….)
- ...

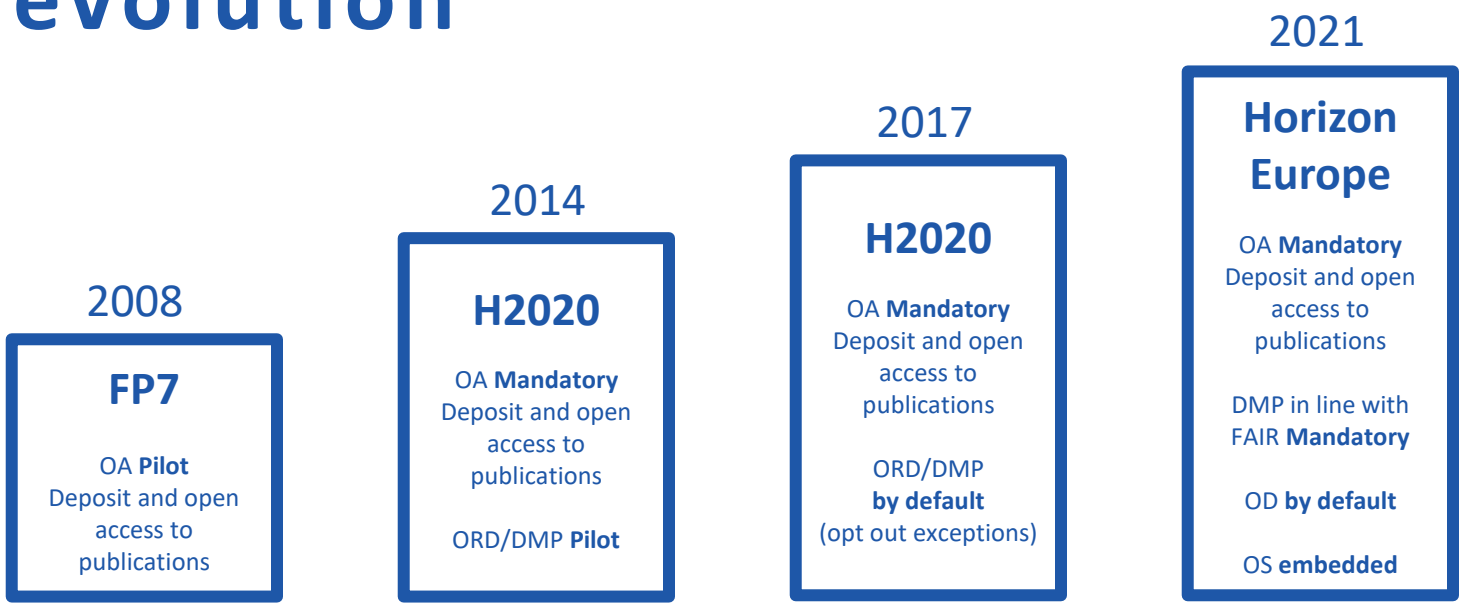
HOW DO YOU PICTURE YOURSELF IN DOING RESEARCH?



In HEU Research results and other research outputs are no more evaluated through the impact factor

AREA SCIENCE PARK

European Policies evolution



Several policy and legal developments on EU open data have taken place since 2003.

OPEN ACCESS routes

Open Access **doesn't** mean paying for publishing! Do not confuse Open Access with the business model behind Gold Open Access

Green OA:

- an author publishes in a subscription-based journal and a copy of the research (usually the author's final, peer-reviewed manuscript) is deposited in either an institutional or subject repository (such as DRO).
- There is no fee to be paid to the publisher by the researcher.
- Following any potential embargo period (set by the publisher) the manuscript is then made free to access.
- The published final version of the journal remains behind a subscription paywall on the journal website, but a version of the paper (accepted manuscript or Version of Record if permitted) is available to anyone from the chosen repository.
- **No more allowed in Horizon Europe (still allowed in H2020 with 6 months embargo period)**

Gold OA:

- a published work is free to access via the publisher's website immediately upon publication.
- It will often have clear re-use rights (perhaps detailed through a Creative Commons licence) which go beyond what is permitted by copyright legislation.
- A publisher may charge a fee for this through applying an Article Processing Charge (APC) to the individual work.
- Some journals are fully/pure open access and contain only open access articles. Others, known as hybrid journals, offer both open and subscription content.
- There are journals which do not charge authors (or their institutions) to publish open access.
- **APC covered by Horizon Europe projects only for Full OA Journals.**

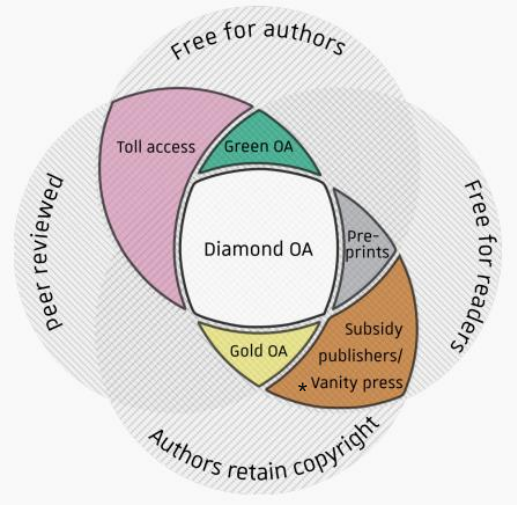
AREA SCIENCE PARK

OPEN ACCESS routes

Diamond OA: refers to academic texts (such as monographs, edited collections, and journal articles) published/distributed/preserved with no fees to either reader or author.

In 2021, it is estimated that between 17,000 and 29,000 scientific journals rely on a diamond open access model. They make up 73% ** of the journals registered in the Directory of Open Access Journals (DOAJ) and 44% of the articles, as their mean output is smaller than commercial journals.

In 2022, new national and international policies, such as the UNESCO recommendation on open science, and the *Action Plan for Diamond Open Access* promoted by the cOAlition S aim to support the development of non-commercial or community-driven forms open access publishing.



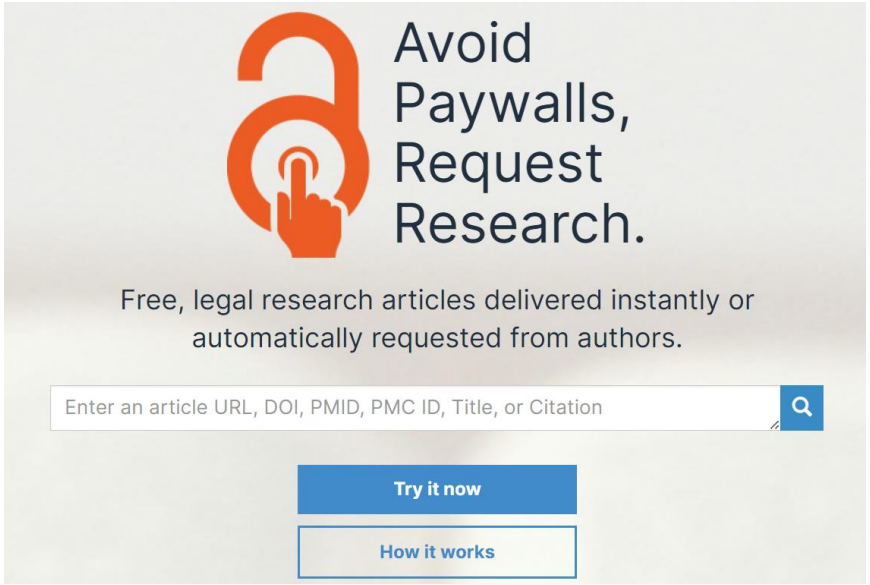
Diamond among the different open access models
Jamie-farquharson - https://doi.org/10.30909/vol.01.01.i-ix

*A vanity press or vanity publisher, sometimes also subsidy publisher, is **a publishing house where anyone can pay to have a book published**

** *Bosman, Jeroen; Frantsvåg, Jan Erik; Kramer, Bianca; Langlais, Pierre-Carl; Proudman, Vanessa (2021-03-09). OA Diamond Journals Study. Part 1: Findings (Report). doi:10.5281/zenodo.4558704. Retrieved 2021-07-10.*

How to read OA?

[Unpaywall: An open database of 20 million free scholarly articles]



[Open Access Button]: add the extension to your browser

OpenAIRE | EXPLORE

A comprehensive and open dataset of research information covering **179m publications**, **59m research data**, **380k research software items**, from **129k data sources**, linked to **3m grants** and **195k organizations**.
All linked together through citations and semantics.

How to publish OA?

Journal of High Energy Physics
JHEP

1029-8479 (ONLINE)

Website | ISSN Portal

DOAJ seal

About | Articles

PUBLISHING WITH THIS JOURNAL

There are **NO PUBLICATION FEES** (article processing charges or APCs) to publish with this journal.

Look up the journal's:

BEST PRACTICE

This journal began publishing in **open access** in 2014.

This journal uses a **CC BY** license.

→ Look up their open access statement and their license terms.

JOURNAL METADATA

Publisher
SpringerOpen, Germany

Society or institution
International School for Advanced Studies (SISSA)

Manuscripts accepted in
English

AREA SCIENCE PARK

Search Sherpa Services

Helping authors and institutions make informed and confident decisions in open access publication and compliance.

Search all | high energy | Search

Journal of High Energy Physics

Title	Journal of High Energy Physics (JHEP) [English]
ISSN	1126-6708
eISSN	1029-8479
Publisher	Springer Scuola Internazionale Superiore di Studi Avanzati (SISSA)
Publisher URL	https://www.springer.com/journal/13130
DOAJ	https://doaj.org/toc/1029-8479

Published Version

OA Publishing	This option includes Open Access publishing
Embargo	No embargo
Licence	CC BY v4.0
Copyright Owner	Authors
Locations	Any Website Named Repository (arXiv) Journal Website
Conditions	Published source must be acknowledged with citation

OPEN SCIENCE practices

Open Science practices (from HE Programme Guide):

In Horizon Europe, open science practices **are considered in the evaluation of proposals**, under 'excellence' and under the 'quality and efficiency of implementation'.
There are **mandatory open science practices**, which are required for all projects through the Model Grant Agreement and/or through the work programme or call conditions, and **recommended practices** (all open science practices that are not mandatory).

Proposers should be aware of both mandatory and recommended practices and integrate them into their proposals.

AREA SCIENCE PARK

Mandatory Open Science practices in HE:

(from Annotated Model Grant Agreement)

European Commission

open access to scientific publications:

The beneficiaries must ensure open access to peer-reviewed scientific publications relating to their results.

In particular, they must ensure that:

- at the **latest at the time of publication**, a machine-readable electronic copy of the published version or the final peer-reviewed manuscript accepted for publication, is deposited in a **trusted repository** for scientific publications
- **immediate open access is provided to the deposited publication** via the repository, under Creative Commons Attribution International Public Licence (CC BY) or a licence with equivalent rights;
- information is given via the repository about **any research output or any other tools and instruments** needed to validate the conclusions of the scientific publication (metadata of the publication).

AREA SCIENCE PARK

Mandatory Open Science practices in HE:

(from Annotated Model Grant Agreement)

open access to scientific publications:

The beneficiaries must ensure open access to peer-reviewed scientific publications relating to their results.

In particular, they must ensure that:

- Beneficiaries (or authors) must retain sufficient intellectual property rights to comply with the open access requirements.
- Metadata of deposited publications must be open under a Creative Common Public Domain Dedication (**CC 0**) or equivalent, in line with the **FAIR principles** (in particular machine-actionable) and provide information at least about the following: *publication* (author(s), title, date of publication, publication venue); *Horizon Europe or Euratom funding*; *grant project name, acronym and number*; *licensing terms*; *persistent identifiers for the publication*, *the authors involved in the action,...*
- Where applicable, the metadata must include persistent identifiers for any research output or any other tools and instruments needed to validate the conclusions of the publication.
- Only publication fees in full open access venues for peer-reviewed scientific publications are eligible for reimbursement.

AREA SCIENCE PARK

Mandatory Open Science practices in HE:

(from Annotated Model Grant Agreement)

Research data management:

The beneficiaries must manage the digital research data generated in the project ('data') responsibly, in line with the **FAIR principles** and by taking all of the following actions:

- within 6 month after the beginning of the project, establish a data management plan ('**DMP**') (and regularly update it).
- as soon as possible and within the deadlines set out in the DMP, deposit the data in a **trusted repository** and ensure open access to the deposited data, under the Creative Commons Attribution International Public License (**CC BY**) or Creative Commons Public Domain Dedication (**CC0**) or a licence with equivalent rights, following the principle **'as open as possible as closed as necessary'**.

AREA SCIENCE PARK

Mandatory Open Science practices in HE:

(from Annotated Model Grant Agreement)

Research data management:

The beneficiaries must manage the digital research data generated in the project ('data') responsibly, in line with the **FAIR principles** and by taking all of the following actions:

- If **open access is not provided** (to some or all data), this **must be justified in the DMP** provide information via the repository about any research output or any other tools and instruments needed to re-use or validate the data.
- Metadata of deposited data must be open under a Creative Common Public Domain Dedication **(CC 0)** or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the **FAIR principles** (in particular machine-actionable) and provide information at least about the following: *datasets (description, date of deposit, author(s), venue and embargo)*; *Horizon Europe funding*; *grant project name, acronym and number*; *licensing terms; persistent identifiers for the dataset, the authors involved in the action*, …
- Where applicable, the metadata must include persistent identifiers for related publications and other research outputs.

Marie Slodowska Curie Actions- MSCA


Maria Skłodowska-Curie

Marie Curie or Maria Skłodowska (7 November 1867 – 4 July 1934) was a Polish physicist, chemist and mathematician who spent most of her life in France.

In 1903 she was the first woman to be awarded the Nobel Prize. She received the **Nobel Prize in Physics**, together with her husband Pierre Curie and Antoine Henri Becquerel, for her studies on radiation.
In 1911 she received the **Nobel Prize in Chemistry** for discovering radium and polonium, whose name was chosen in honor of his native land.
She was one of five Nobel laureates to have received two and **she is the only one to have won the Prize in two distinct scientific fields.**

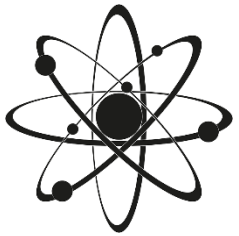


Image by Chris Raz from Pixabay

Marie Curie grew up in Russian Poland; **Since women could not be admitted to higher studies** here, then she moved to Paris and in 1891 began attending the Sorbonne, where she graduated in physics and mathematics.
In December 1897 she began to carry out studies on radioactive substances, which will remain the focus of her interests ever since.
In 1906, after the death of her husband Pierre Curie, she was allowed to teach at the Sorbonne.

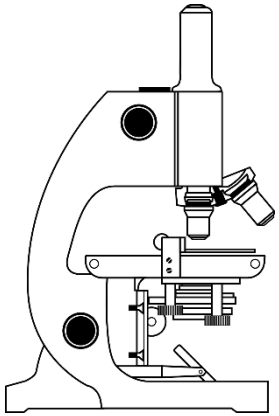


Image by Clker-Free-Vector-Images from Pixabay

Two years later she was appointed to the chair of general physics, **making her the first woman to teach at the Sorbonne.**
Wikipedia

Principles of MSCA funding programme

EXCELLENCE

- **Excellence:** The MSCA support excellent researchers. They also foster excellence in research and innovation collaborations, knowledge transfer, methodologies and content, as well as in training, supervision and career guidance.
- **Open science and responsible research and innovation:** The MSCA support Open Science and Responsible Research and Innovation.

MOBILITY

- **Mobility:** The MSCA support the mobility of researchers between countries, sectors and disciplines to acquire new knowledge, skills and competences.
- **Bottom-up and open to the world:** The MSCA are open to all domains of research and innovation and encourage international cooperation to set-up strategic collaborations.

INCLUSION

- **Excellent recruitment, working conditions and inclusiveness:** The MSCA promote the principles of the European Charter for Researchers and Code of Conduct for the Recruitment of Researchers for the recruitment, working and employment conditions of researchers.
- **Effective supervision and career guidance:** The MSCA promote effective supervision and adequate mentoring and career guidance. This contributes to creating a supportive environment for the researchers to work. The Guidelines for MSCA Supervision provide recommendations in this regard.

AREA SCIENCE PARK

The Marie Skłodowska-Curie Actions PF are the Europea Union's reference programme for **postdoctoral** training.

MSCA PF call : Expected Scope, Object and Impact

Scope: Fellowships will be provided to **excellent researchers undertaking international mobility**. Applications will be made jointly by the researcher and a beneficiary in the academic or non-academic sector.

Object: The goals of the work performed within the project, in terms of its research and innovation content. This will be translated into the **project's results**. These may range from tackling specific research questions, demonstrating the feasibility of an innovation, sharing knowledge among stakeholders on specific issues. The nature of the objectives will depend on the type of action, and the scope of the topic.

Impact: Wider **long term effects** on society (including the environment), the economy and science, enabled by the outcomes of R&I investments (long term). Impacts generally occur some time after the end of the project.
(See Template Part B in the following slides).

AREA SCIENCE PARK

Expected Outcome

Definition of Outcome: *The expected effects,* ***over the medium term****, of projects supported under a given topic.*
The results of a project should contribute to these outcomes, fostered in particular by the dissemination and exploitation measures.
This may include the uptake, diffusion, deployment, and/or use of the project's results by direct target groups.
Outcomes generally ***occur during or shortly after the end*** *of the project.*

For supported postdoctoral fellows:

- **Increased set of research and transferable skills and competences**, leading to improved employability and career prospects of MSCA postdoctoral fellows within academia and beyond;
- New mind-sets and approaches to R&I work forged through **international, inter-sectoral and interdisciplinary experience;**
- Enhanced networking and communication capacities with scientific peers, as well as with the **general public that will increase and broaden the research and innovation impact**.

For participating organisations:

- Increased alignment of working conditions for researchers in accordance with the principles set out in the European Charter for Researchers and the Code of Conduct for the Recruitment of Researchers;
- Enhanced quality and sustainability of **research training and supervision**;
- Increased global attractiveness, **visibility and reputation of the participating organisation(s)**;
- Stronger R&I capacity and output among participating organisations; **better transfer of knowledge;**
- Regular feedback of **research results into teaching and educati**on at participating organisations.

AREA SCIENCE PARK

Eligibility conditions

Project:

- **Bottom up** Approach for the research topic
- Proposals involving the same recruiting organisation (and for Global Postdoctoral Fellowships also the associated partner hosting the outgoing phase) and individual researcher submitted to the previous call of MSCA Postdoctoral Fellowships under Horizon Europe and having received a score of **less than 70% must not be resubmitted the following year**.
- Only **one proposal per individual** researcher can be submitted.

Participating organisations:

- Applications must be submitted by **a single independent legal entity**, established in **an EU Member State or Horizon Europe Associated Country**. This is a mono-beneficiary action.
- The associated partner hosting the outgoing phase in Global Postdoctoral Fellowships must include a letter of commitment in the proposal to ensure their active participation in the action.
- Affiliated entities are not allowed to participate as they cannot claim costs in MSCA Postdoctoral Fellowships.

AREA SCIENCE PARK

Eligibility conditions

Researcher:

- Supported fellows must be **postdoctoral researchers** at the date of the call deadline, i.e. in a possession of a doctoral degree.
- At the call deadline, supported researchers must have a **maximum of 8 years full-time equivalent experience in research**, measured from the date of award of the doctoral degree **.
- Recruited researchers must comply with the following **mobility rule**:
  they must not have resided or carried out their main activity (work, studies, etc.) in the country of the beneficiary (for European Postdoctoral Fellowships), or the host organisation for the outgoing phase (for Global Postdoctoral Fellowships) for more than 12 months in the 36 months immediately before the call deadline.
- Supported researchers can be **of any nationality.**

****Maternity**: for each child born prior to the call deadline, 18 months will be deducted from the experience in research unless the applicant can document a longer parental leave prior to the call deadline.
Paternity: for each chil dborn prior to the call deadline, the documented time of parental leave taken until the call deadline will be deducted from the experience in research.

AREA SCIENCE PARK

European and Global Fellowships

Image by GraphicMama-team from Pixabay

European Postdoctoral Fellowship (EF):

- Researcher can have any nationality and be from any country
- Host Institution must be a European Member State** or Associated Country
- **12/24 months** duration
- Mobility rule apply to the HI country
- Secondment up to 1/3 of total fellowship in any country worldwide (single period or divided in shorter mobility periods.

Placement in non-academic sector:
If requested and justified in the proposal, an **additional period of up to six months** at the end of the project can be awarded to researchers who will spend that period in a **non-academic organisation** established in an EU Member State or Horizon Europe Associated Country.
The request for such a non-academic placement must be an integral part of the proposal, explaining the added-value for the project and for the career development of the researcher, and will be subject to evaluation.

Global Postdoctoral Fellowship (GF):

- Researcher must be long term resident of a European Member State** or Associated Country
- Host Institution (ingoing phase) must be a European Member State** or Associated Country
- **12 moths** duration **ingoing** phase
- Associated Partner (outgoing phase) must be a Third Country
- **12/24 months** duration **outgoing** phase
- Mobility rule apply only to the country of the outgoing phase
- Secondment up to 1/3 of the outgoing phase.

**EU Members States: Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden

AREA SCIENCE PARK

Template Instructions - Part B

- Part B of the proposal contains the details of the proposed MSCA Postdoctoral Fellowship as well as the required supporting information.
- It will be used by the independent experts to undertake their assessment of the proposal.
- Applicants must address each of the award criteria as outlined in the relevant sections, using both descriptive text and the tables provided.
- Applicants must structure their MSCA-2023-PF proposal according to the headings indicated in the Part B proposal template.
- Please note that this call will be a **single-stage proposal** submission and evaluation procedure.
- Applicants must ensure that their proposals conform to this layout and to the instructions given.

Please be aware that proposals will be evaluated as they were submitted, rather than on their potential if certain changes were to be made.

This means that only proposals that successfully address all the required aspects will have a chance of being funded.

AREA SCIENCE PARK

Template Instructions - Part B

Part B1
Sections 1, 2 and 3 together should not be longer than **10 pages**.
The reference font for the body text of proposals is Times New Roman, the minimum font size allowed is **11 points**, ….
After the deadline, excess pages (in over-long proposals) will be automatically made invisible, and therefore will not be taken into consideration by the experts.
Note that experts will be instructed to ignore hyperlinks to information that is specifically designed to expand the proposal, thus circumventing the page limit.
Sections:
- Excellence
- Impact
- Quality and Efficiency of the implementation

Part B2
Does not have a page limit.
It includes:
- CV of the researcher (indicative length: 5 pages)
- Capacity of the participating organisation(s) (TABLE 5.2)
  - ✓ General description
  - ✓ Role and profile of supervisor
  - ✓ Key research facilities, Infrastructure and Equipment
  - ✓ Previous and current involvement in EU-funded research and training projects
- Additional ethics information (if needed)
- Additional information on security screening (if needed)
- Environmental considerations in light of the MSCA Green Charter (not mandatory but suggested)
- The commitment letter(s) of the associated partner (only for GF)

AREA SCIENCE PARK

Template Instructions - Part B1

Award criteria

- Proposals will be evaluated by experts on the basis of the award criteria **'excellence', 'impact' and 'quality and efficiency of the implementation'.**
- Evaluation scores will be awarded for each of these criteria, and not for the different aspects listed in the table beside. **Each criterion will be scored out of 5.**
- Proposals scoring equal to or above 70% will be considered for funding — within the limits of the available call budget. Other proposals will be rejected.

Excellence	Impact	Quality and efficiency of the implementation
Quality and pertinence of the project's research and innovation objectives (and the extent to which they are ambitious, and go beyond the state of the art)	Credibility of the measures to enhance the career perspectives and employability of the researcher and contribution to his/her skills development	Quality and effectiveness of the work plan, assessment of risks and appropriateness of the effort assigned to work packages
Soundness of the proposed methodology (including interdisciplinary approaches, consideration of the gender dimension and other diversity aspects if relevant for the research project, and the quality of open science practices)	Suitability and quality of the measures to maximise expected outcomes and impacts, as set out in the dissemination and exploitation plan, including communication activities	Quality and capacity of the host institutions and participating organisations, including hosting arrangements
Quality of the supervision, training and of the two-way transfer of knowledge between the researcher and the host	The magnitude and importance of the project's contribution to the expected scientific, societal and economic impacts	
Quality and appropriateness of the researcher's professional experience, competences and skills		
50%	30%	20%
Weighting		

OPEN SCIENCE in MSCA

Excellence – Template part B1

- 1.1 Quality and pertinence of the project's research and innovation objectives (and the extent to which they are ambitious, and go beyond the state of the art)
- 1.2 **Soundness of the proposed methodology (including interdisciplinary approaches, consideration of the gender dimension and other diversity aspects if relevant for the research project, and the quality of open science practices)**
- 1.3 Quality of the supervision, training and of the two-way transfer of knowledge between the researcher and the host institution
- 1.4 Quality and appropriateness of the researcher's professional experience, competences and skills

Overall Methodology: Describe and explain the overall methodology, including the concepts, models and assumptions that underpin your work.
Explain how this will enable you to deliver your project's objectives.
Refer to any important challenges you may have identified in the chosen methodology and how you intend to overcome them.

AREA SCIENCE PARK

OPEN SCIENCE in MSCA

Methodology (Interdisciplinary and Transversal aspects)

- **Integration of methods and disciplines to pursue the objectives:** Explain how expertise and methods from different disciplines will be brought together and integrated in pursuit of your objectives. If you consider that an inter-disciplinary approach is unnecessary in the context of the proposed work, please provide a justification.
- **Gender dimension and other diversity aspects:** Describe how the gender dimension and other diversity aspects are taken into account in the project's research and innovation content. If you do not consider such a gender dimension to be relevant in your project, please provide a justification.
- **Open science practices:** Describe how appropriate open science practices are implemented as an integral part of the proposed methodology. Show how the choice of practices and their implementation is adapted to the nature of your work in a way that will increase the chances of the project delivering on its objectives [e.g. up to 1/2 page, including research data management]. If you believe that none of these practices are appropriate for your project, please provide a justification here.
- **Research data management and management of other research outputs:** Applicants generating/collecting data and/or other research outputs (except for publications) during the project must explain how the data will be managed in line with the FAIR principles (Findable, Accessible, Interoperable, Reusable).

AREA SCIENCE PARK

How to apply for a MSCA PF project with an Host Institution

- **Contact a professor of the Institution you would like to apply with to have his/her scientific support and his/her committment to become your supervisor for a MSCA project.**

- **Contact the Research Office of that Institution to have all the information and the support you need.**

- **Be aware that you need to move well in advance before the deadline of the call, usually at the beginning of September.**

Foto di 200 Degrees da Pixabay

AREA SCIENCE PARK

AREA SCIENCE PARK

Thank you for your attention!