Europa Media Trainings

Open Science and Data Management

Krisztina Toth
Managing Director, Europa Media Innovations
EMG Group

Horizon Europe Academy Part II
Horizon Europe and H2020 Project Management, Financial Rules & Reporting
15-17 June 2022
Barcelona

Open Science
The origins

Open Science

= approach based on cooperative work and systematic sharing of knowledge and tools **as early and widely as possible**

Responsible Research and Innovation (RRI)

= societal actors (researchers, citizens, policy makers, business, third sector organisations, etc.) work together during the whole R&I process to better align both the process and its outcomes with the values, needs and expectations of society.

- Public engagement

- Gender equality

- Ethics

- **Open Science**

- Science education

- Governance

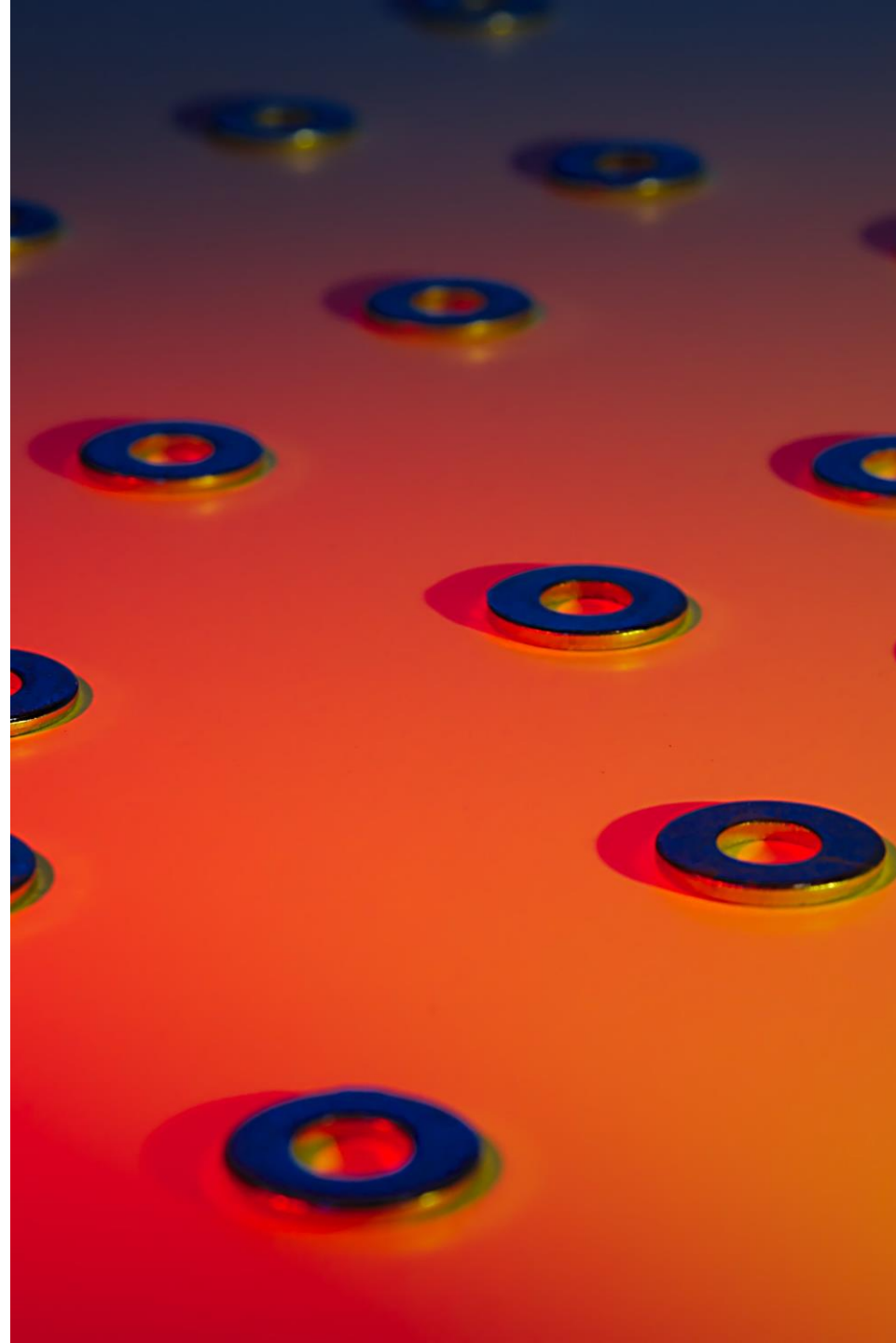# Open Science
The 3 Os

Open Innovation

✓ in your methodology for collaboration with stakeholders you highlight how that leads to open innovation
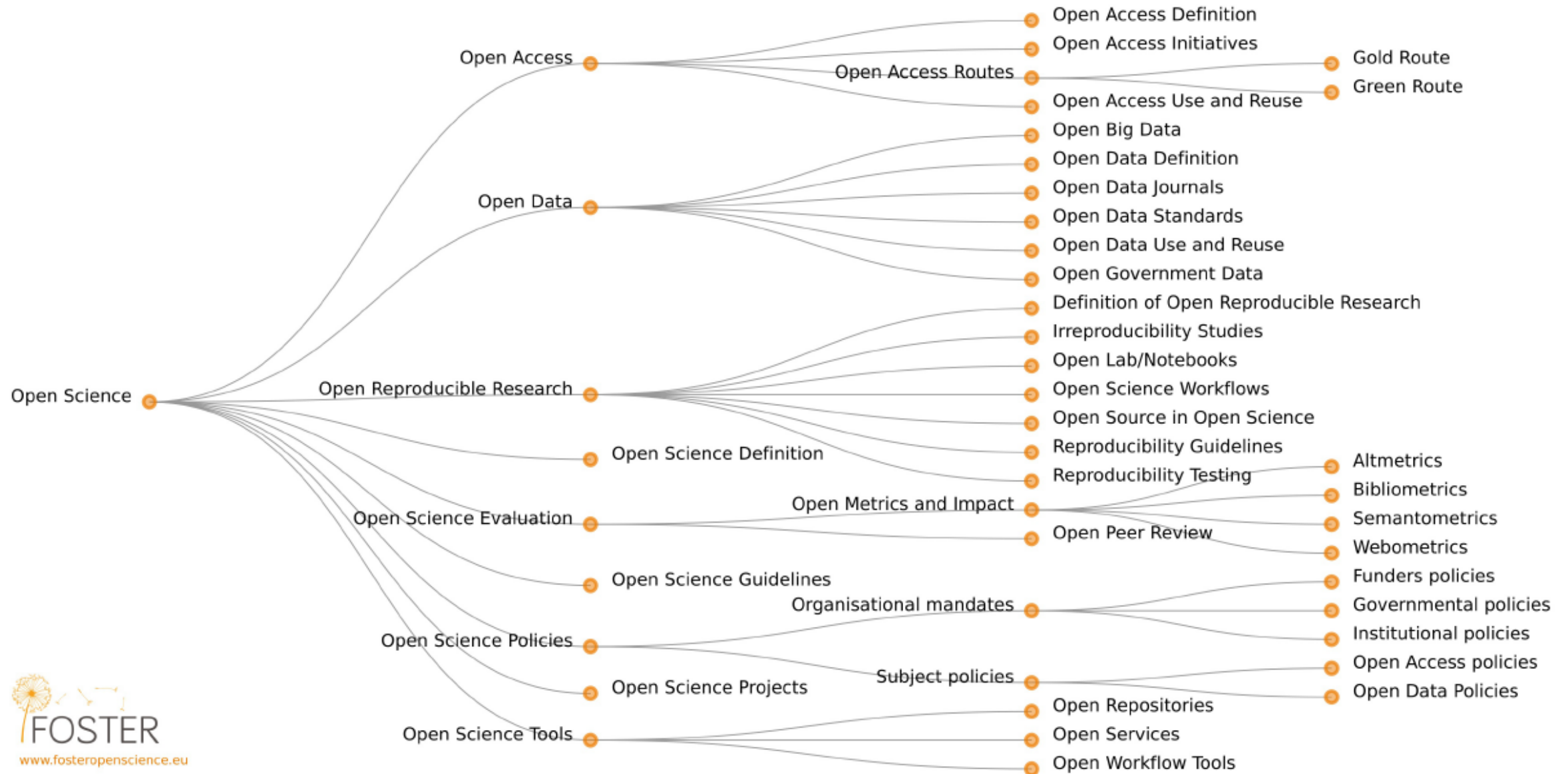
Open Science

✓ new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools

✓ practices like data management

Open to the World

✓ consider the existing international collaborations, the EU's regional and bilateral agreements

Open Science Taxonomy

- **Open Science**
  - **Open Access**
    - Open Access Definition
    - Open Access Initiatives
    - Open Access Routes
      - Gold Route
      - Green Route
    - Open Access Use and Reuse
  - **Open Data**
    - Open Big Data
    - Open Data Definition
    - Open Data Journals
    - Open Data Standards
    - Open Data Use and Reuse
    - Open Government Data
  - **Open Reproducible Research**
    - Definition of Open Reproducible Research
    - Irreproducibility Studies
    - Open Lab/Notebooks
    - Open Science Workflows
    - Open Source in Open Science
    - Reproducibility Guidelines
    - Reproducibility Testing
  - **Open Science Definition**
  - **Open Science Evaluation**
    - Open Metrics and Impact
      - Altmetrics
      - Bibliometrics
      - Semantometrics
      - Webometrics
    - Open Peer Review
  - **Open Science Guidelines**
  - **Open Science Policies**
    - Organisational mandates
      - Funders policies
      - Governmental policies
      - Institutional policies
      - Open Access policies
      - Open Data Policies
    - Subject policies
  - **Open Science Projects**
  - **Open Science Tools**
    - Open Repositories
    - Open Services
    - Open Workflow Tools

FOSTER
www.fosteropenscience.eu

FOSTER Open Science taxonomy

Zenodo http://doi.org/10.5281/zenodo.2595951

Legal obligations
GA Art 17 + Annex 5

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 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 the latest available version of the 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 deposited publications must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles.

Legal obligations
GA Art 17 + Annex 5

Research Data Management

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

- **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**

- as soon as possible and within the deadlines set out in the DMP, **ensure open access — via the repository — to the deposited data**, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights

- **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.

Legal obligations
GA Art 17 + Annex 5

Additional practices

- **Where the call conditions impose additional obligations** regarding open science practices

- **Where the call conditions impose additional obligations regarding the validation of scientific publications**, the beneficiaries must provide (digital or physical) access to data or other results needed for validation of the conclusions of scientific publications

- **Where the call conditions impose additional open science obligations in case of a public emergency**, the beneficiaries must (if requested by the granting authority) immediately deposit any research output in a repository and provide open access to it under a CC BY licence, a Public Domain Dedication (CC 0) or equivalent. As an exception, if the access would be against the beneficiaries' legitimate interests, the beneficiaries must grant non-exclusive licenses —under fair and reasonable condition. **This provision applies up to 4 years after the end of the action.**

Open Science
Mandatory practices (in line with GA)

- **Open access to scientific publications**

- **Management of research data** in line with FAIR principles

- **Information about research outputs/tools/instruments** needed to validate conclusions of scientific publications or to validate/re-use research data

- **Digital or physical access to the results** needed to validate the conclusions of scientific publications

- **Public emergency:** immediate open access to all research outputs under open licenses / access under fair and reasonable conditions to legal entities that need the output to address the emergency

Open Science
'Optional' practices (recommended)

- Store or give access to research data on the **European Open Science Cloud (EOSC)**

- **Early and open sharing of research** (via preregistration, registered reports, preprints)

- **Involving all relevant knowledge actors** including citizens, civil society and end users in **co-creation**, **co-design** and **co-assessment** activities

- **Output management** beyond research data

- Participation in **open-peer review**

Open Science
Evaluation

Excellence

- Methodology: how open science practices are implemented

Capacity of participants and consortium as a whole

- How the consortium brings together the necessary disciplinary and interdisciplinary knowledge

Part A

- List up to 5 relevant publications, widely used datasets or other achievements

- Open access expected for publications

- Datasets are expected to be FAIR and 'as open as possible, as closed as necessary'

Open access

= **online access at no cost for the end user** of research outputs (e.g. scientific publications, data, software, algorithms, electronic notebooks etc.)

To consider:

- If scientific peer-reviewed publication are produces, they must be open access under open licenses (e.g. Creative Commons)

- Retain sufficient IPR
    - Retain the copyright on the work and grant, nonexclusive licenses to publishers
    - Put in place institutional policies to ensure copyright retention and compliance with the open access requirements

- Data should be deposited in a trusted repository as soon as possible after production

- 'As open as possible as closed as necessary'
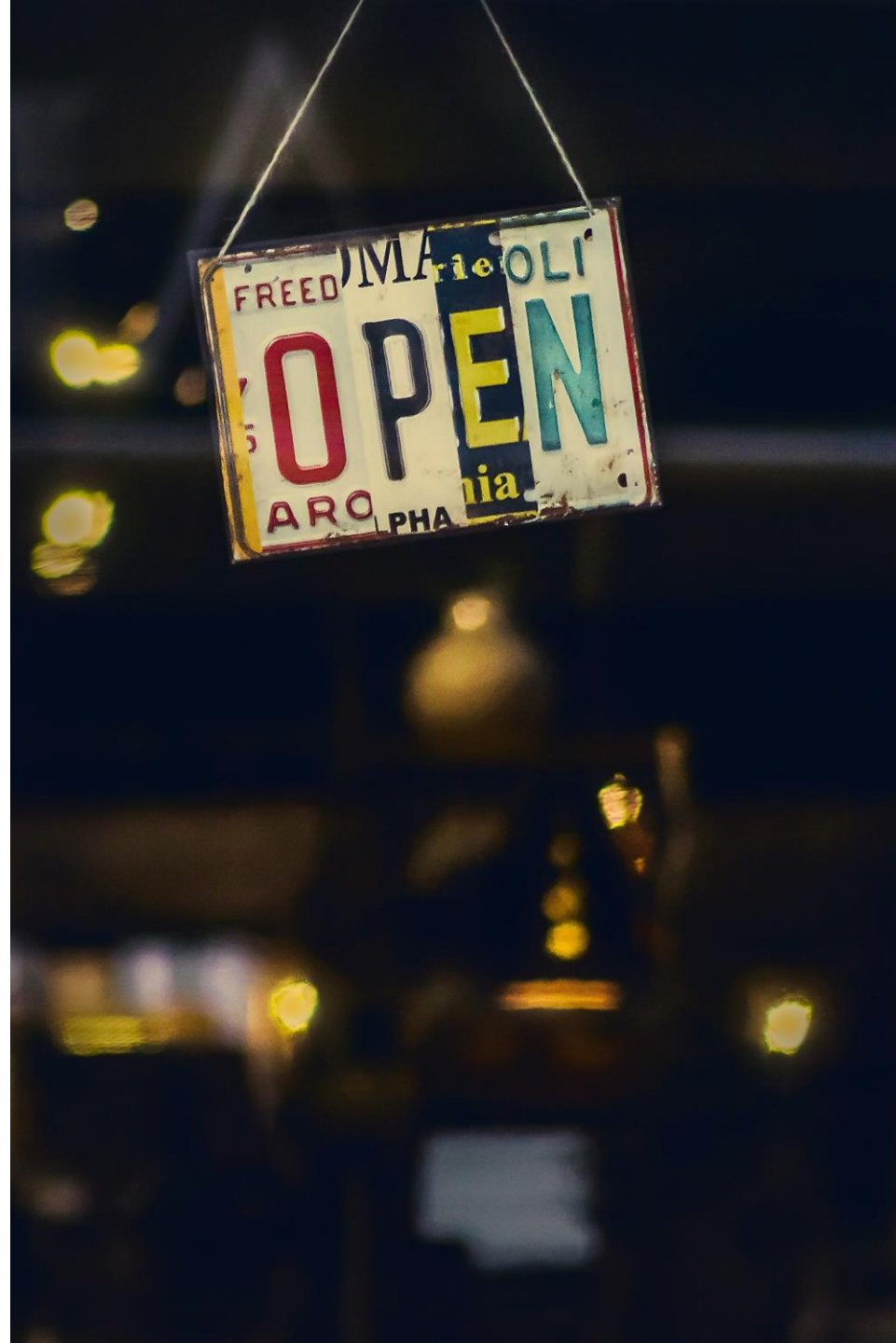
Early and open sharing

= make available research, methodologies, outputs, and findings as soon as possible in the research process.

HOW?

- **Preregistration** in a public repository

- **Registered reports:** research articles that are peer-reviewed and published in 2 stages

- **Preprints:** scientific manuscripts that are publicly shared prior to peer-review and journal application via preprint platforms

Resources

- ORION

- The Centre for Open Science

- Sherpa Romeo

- Preregistration repositories: OSF, AsPredicted, etc.

- Preprint servers: Zenodo, Preprints, ArXiv, SocArXiv, etc.

Open peer-review

= like peer-review but more transparent and accountable

- Authors and reviewers are aware of each other's identity

- Review reports are published alongside the relevant article

- The wider community is able to contribute to the review process

- Manuscripts are made immediately available in advance of the formal peer-review procedure

- Review or commenting on the final 'version of record' is made possible

- Direct, reciprocal discussion between authors and reviewers and/or between reviewers is allowed and encouraged

- Review can be decoupled from publishing when facilitated by a different organisational entity than the venue of publication (e.g. publishing platforms)

Reproducibility of results

= possibility for the scientific community to obtain the same results as the originators of specific findings.

Practices to increase reproducibility

- Specify the research design and methodologies applied

- Specify how you deal with negative results

- Make prior searches and checks on existing results and data to avoid duplication

- Specify how you are making use or preprints, preregistration

- Detail steps you will take to make your research process and tools transparent

- Mention what steps you will take to ensure validity and quality of the project process and results (e.g. peer review)

- Plan to use the DMP and make sure your data are FAIR

Citizen, civil society and end-user engagement

= refers to **opening of R&I processes to society** to develop better, more innovative and more relevant outcomes and to increase societal trust in the processes and outcomes of R&I

Activities to consider

- **Co-design** = workshops, focus groups to develop R&I agendas, roadmaps, policies

- **Co-creation** = involve citizens or end users directly in the development of new knowledge or innovations

- **Co-assessment** = assisting in monitoring and evaluation of the project progress and ensure interaction with citizens, civil society and end users on quality, utilization and impact of project outputs
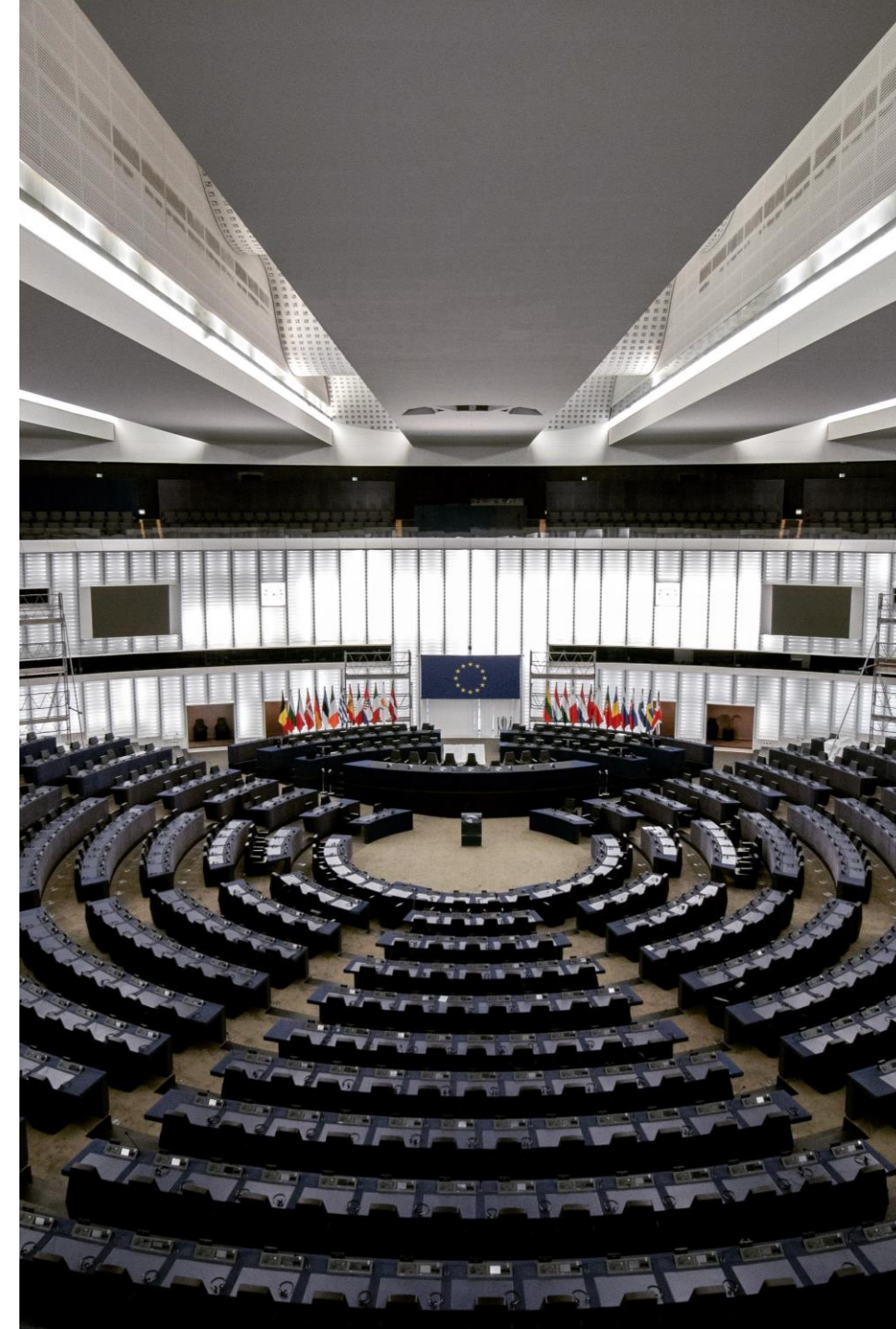
Research Data Management
The concept

Research Data Management (RDM)

= the **process** within the research lifecycle that includes the data collection or acquisition, organisation, curation, storage, (long-term) preservation, security, quality assurance, allocation of persistent identifiers (PIDs), provision of metadata in line with disciplinary requirements, licencing, and rules and procedures for sharing of data.

Elements to consider in your project's RDM:

- **Persistent identifiers** (PIDs) to ensure findability of research outputs and data

- **Standardised metadata frameworks** for the findability of research outputs and their potential reuse

- **Trusted repositories** for the deposition of and access to publications and research data
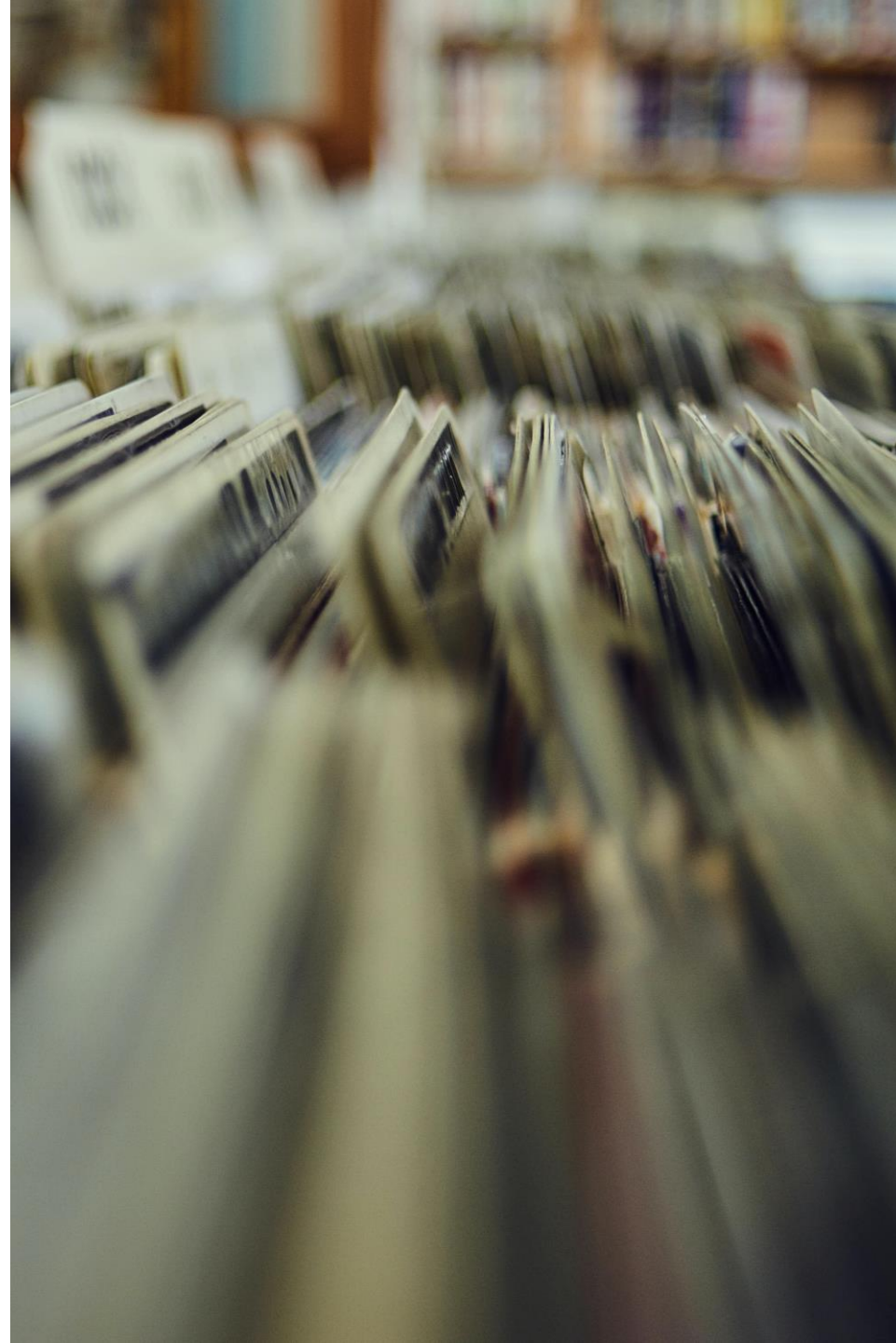
Research Data Management
Data Management Plan

What is a Data Management Plan (DMP)?

= your key to good data management

- Describes the data management life cycle

- The template = set of questions

- Living document **(!)**

- Deliverable – 1st version by M6

💡 Register your DMP as non-restricted public deliverable + publish it in journals, platforms or repositories (e.g. RIO, DMPOnline).

Research Data Management
To address in your DMP

- **Data set description:** what kind of data is your project generating or reusing? Estimate the size of the data set

- **Standards and metadata:** how do you structure your data and what protocols are you using?

- **Name and persistent identifier for data sets:** unique and persistent identifier and a stable link to directly access the data

- **Curation and preservation methodology:** how will you ensure the integrity of the data sets and for how much time? How will it be preserved and kept?

- **Data sharing methodology:** how can the data sets be accessed? Terms of use and license

- **Research output management** other than data and publications

- **Related costs and personnel:** data collection, documentation, storage, preservation, availability and reuse, person/team in charge

European Open Science Cloud (EOSC)

= open trusted virtual cloud to enable researcher to store, share, process, analyze, and reuse research data, publications, and software across disciplines and borders.

- Use cases on EOSC in Practice

- EOSC community and stakeholders on events and news sections

- EOSC-Hub

- Catalogue & Marketplace for services and resources for researchers

- Training

💡 Work programmes may require the use of trusted repositories federated in EOSC for depositing research data
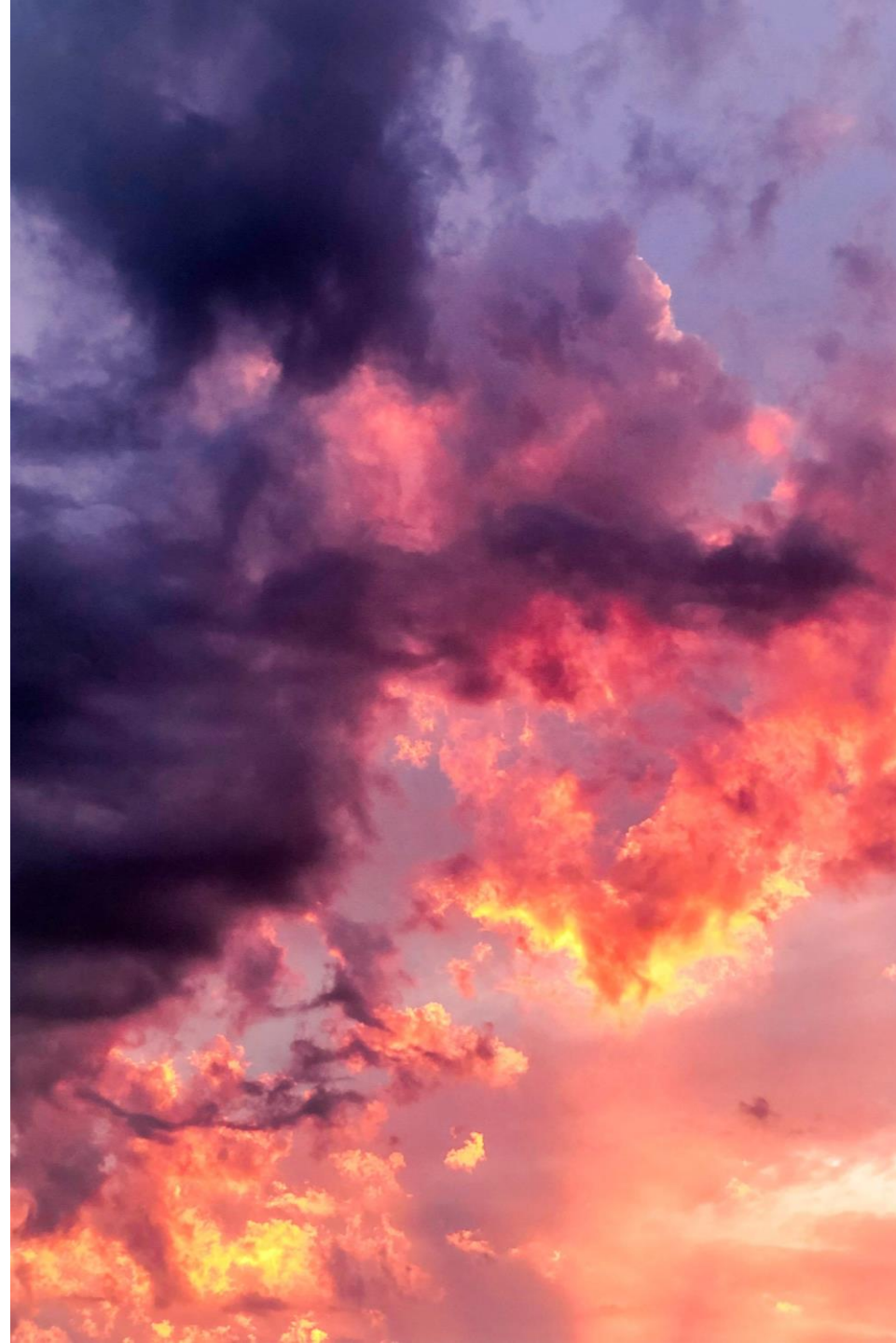
Tools and platforms

Digital profile

- ORCID (for researchers)
- CRIS (for organizations)
- ImpactStory
- Publons
- Open Science Framework (OSF)

OpenAIRE – Zenodo – Argos

- Putting your work into OpenAIRE-compliant repositories ensures that
  - you comply with H2020 mandate on Open Access
  - saves you time as you can import your project publications into the F&T Portal in one click

Open Research Europe

European Open Science Cloud (EOSC)

Open Science
Sources and guides

- [HE Programme Guide](#)

- OpenAIRE guide: *Guiding you in Open Science* https://www.openaire.eu/guides

- Research Data Alliance
  - https://www.rd-alliance.org/
  - http://rd-alliance.github.io/metadata-directory/

- Open Science Framework https://osf.io/

- Re3Data https://www.re3data.org/

- GitHub https://guides.github.com/

- Choosing a License https://choosealicense.com/

- FOSTER Open Science https://www.fosteropenscience.eu

- FIT4RRI project https://fit4rri.eu/guidelines/

Research Data Management
Sources and guides

- [Research data management (RDM) open training materials](#) (Zenodo)

- FOSTER Open Science e-learning

- Data Management Plans
    - [DMPonline](#)
    - [OneHealth EJP DMP Guide](#)
    - Webinar (video: DOI: 10.5281/zenodo.2564974; slides: DOI: 10.5281/zenodo.2565750)

- EC Guide for FAIR data management in H2020

QUESTIONS?

Thank you

for your attention

Krisztina Toth

krisztina.toth@europamedia.org

FOLLOW US!

@EuropaMedia