E|M|G GROUP 20 YEARS ANNIVERSARY

EM TRAININGS

in collaboration with

EMDESK®

present...

Europa Media/EMDESK webinar

EM Guidelines for data management

Speakers:

Gabriella Lovász &Frosina Ilievska

14:00 – 15:30 CET // 28 May 2020

Welcome!

- 14:00– 15:00 | Presentation
- 15:00 – 15:30 | Q&A session

SPEAKERS

Gabriella Lovasz

- Contact me on LinkedIn
- On Twitter as @gabocsek
- Write to gabriella.lovasz@europamedia.org

Frosina Ilievska

- Contact me on LinkedIn
- On Twitter as @frosinailievska
- Write to frosina.ilievska@europamedia.org

CONTENT

1. Publishing an article
    1. Access and costs
2. Data Management
    1. Concept, costs, planning
3. Data for everyone
    1. FAIR principles and ethics

TRAININGS

E|M|G GROUP

CONTENT

1. Publishing an article
    1. Access and costs
2. Data Management
    1. Concept, costs, planning
3. Data for everyone
    1. FAIR principles and ethics

Publishing an article

Open, open, open!

Horizon 2020

- Open access to **research results, research data**
- **Data management plan**
- **Security** (open as much as possible, closed as much as necessary)
- See the Guidelines on Open Access to Scientific Publications and Research Data and/or the Guidelines on data management

Publish an article or paper (1)

What to give access to

Get your latest publication

- Deposit your published version or final pee-reviewed manuscript (accepted for publication) in a repository for scientific publications
  - Also deposit the research data which was needed to validate the results presented in the deposited scientific publication

Source: https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-publications

Publish an article or paper (2)
Timeframe

When should you guarantee open access to the deposited publication via the repository?

- **on publication** (gold way / open access journal)
- within **six months** of publication in any other case
- except for publications in the Social Sciences and Humanities journals, where it can be **twelve months**

Source: https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-publications

Publish an article or paper (3)
Ensuring access to metadata

You need to ensure open access to the descriptive metadata which identify the deposited publication.

This metadata must be in a standard format.

It must include all of the following:

- the words ["European Union (EU)" and "Horizon 2020"]
- the name of the project, acronym and grant number
- the publication date, and length of embargo period if applicable and
- a persistent identifier (like DOI for objects and ORCIDs for people).

Source: https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-publications

Giving access to an article or paper

Knowing your publisher's policy

How to know which version you should upload?

You need to check your publisher's policy and to see if an embargo period applies, check the Sherpa/Romeo database.

What if the publisher's policy differs from the Horizon 2020 requirements?

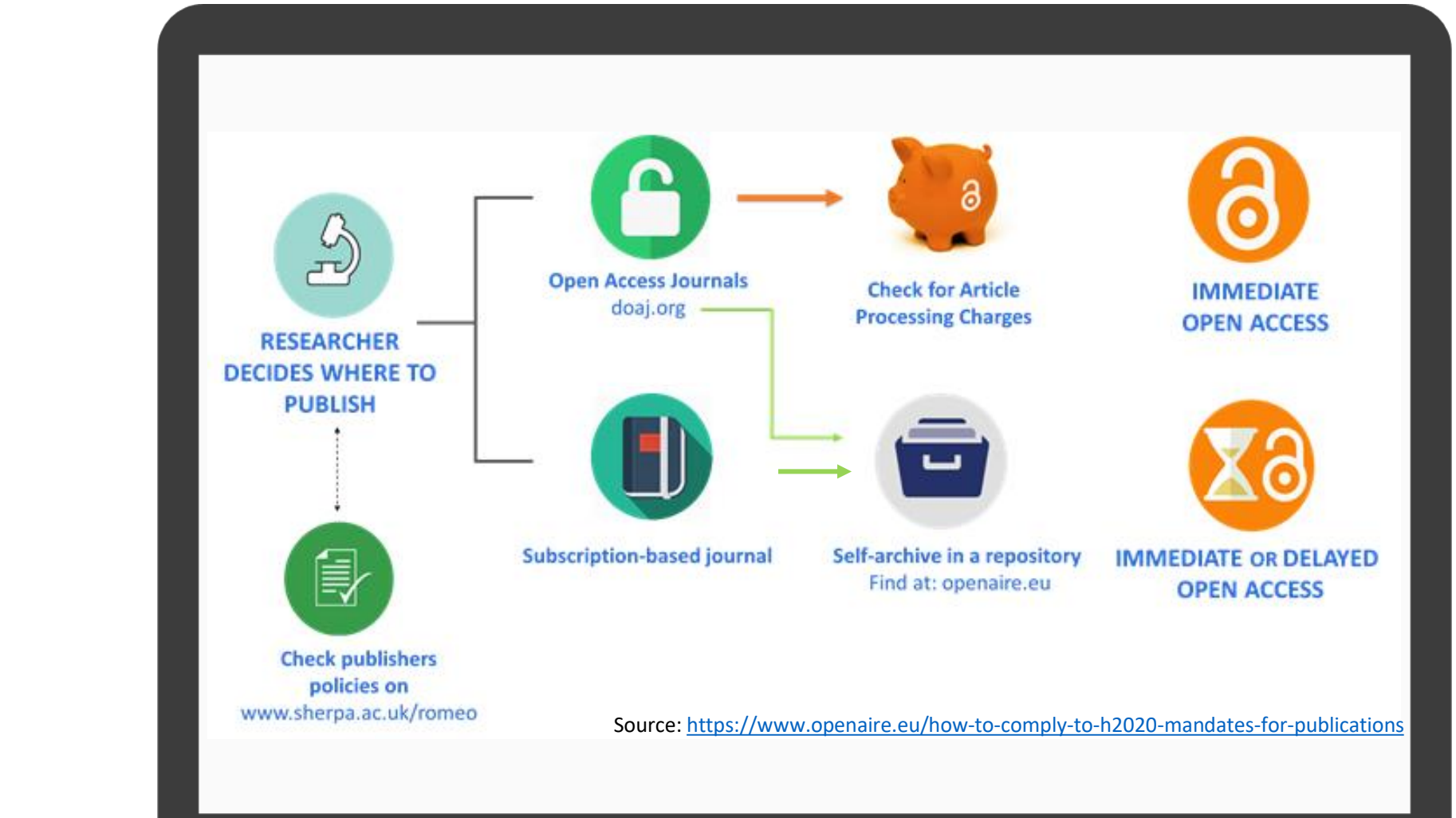
Then contact the publisher! You need to inform the publisher of the open access requirements, and ask for an exception to the publisher's policy to enable you to meet those requirements.

It is important to obtain this permission in writing.

Use a Template letter (.pdf) provided by the European Commission when writing to the publisher asking for an amendment to your publishing agreement.

Source: https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-publications

Costs of publication (1 / 2)



Source: https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-publications

Costs of publication (2 / 2)
how the budget is calculated

The Budget for Publications = Average APC x number of publications.

Methods to calculate it:

- Method 1: Average APC based on list of journals used by the consortium (look up prices at publisher website and/or consult a librarian).
- Method 2: Average APC based on general market figures.

  Björk & Solomon (2014) estimated the average price of Article Processing Charges (APC) for established OA journals at ca. 1,020 EUR and for hybrid journals (subscription journal with OA option for individual articles) at ca. 1,980 EUR. More recently, the Open APC initiative, that releases datasets on fees paid for OA journal articles by universities and research institutions under an open database license in Github, estimates the average payment for fully OA journals at 1,484 EUR, whereas for hybrid journals the average fee is 2,492 EUR.

Source: https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-publications

Guides

Some much needed platforms

- **OpenAIRE guide:**

Guiding you in Open Science https://www.openaire.eu/guides

- **GitHub:**

The best way to build and ship software.
Powerful collaboration, code review, and code management for open source and private projects.

https://guides.github.com/

- **Choosing a License:**

An open source license protects contributors and users. Businesses and savvy developers won't touch a project without this protection.

https://choosealicense.com/

CONTENT

1. Publishing an article
    1. Access and costs
2. Data Management
    1. Concept, costs, planning
3. Data for everyone
    1. FAIR principles and ethics

Data management

Data Management
Understanding the concept

- **Data**: content generated in your project such as, raw data, publications, documentations (reports, lab procedure), research data, events (workshops, webinars) or anything that can be used to support the integrity of the results and reproduce the research
- **Research data:** data created in a digital form (born digital) or converted to a digital form (digitized); research data are collected, observed or created for the purposes of analysis to produce and validate original research results
- **Data management** refers to all aspects of creating, housing, delivering, maintaining, and archiving and preserving data. It is one of the essential areas of responsible conduct of research - continuously maintained and kept up-to-date.

Costs

- Costs are eligible for management and publication under H2020 - they must be calculated at proposal stage.

- Consider the following when you calculate your data management costs:
  - Responsible staff for implementation of DMP
  - Hardware and software
  - Charges to data repositories
  - Journal publication

Data Management

Questions you need to ask yourself

- **What research data** you will be creating or collecting?
- **Who will be responsible** for each aspect of the management plan you are developing?
- **What policies** (funding, institutional, and legal) will apply to your data?
- **How will the data be organized** (folder structures, file naming conventions, file versioning)?
- **How will the data be documented** during the collection and analysis phase of your research?
- **What data management practices** (backups, storage, access control, archiving) will you use to store & secure your data?
- **What facilities and what equipment will be required** (hard-disk space, backup server, and repository)?
- **Who will have ownership** and access rights to your data?
- **How the data will be preserved** and made available in the long term, once your research is completed?

Useful links: DMPonline

Planning Research Data Management (RDM)

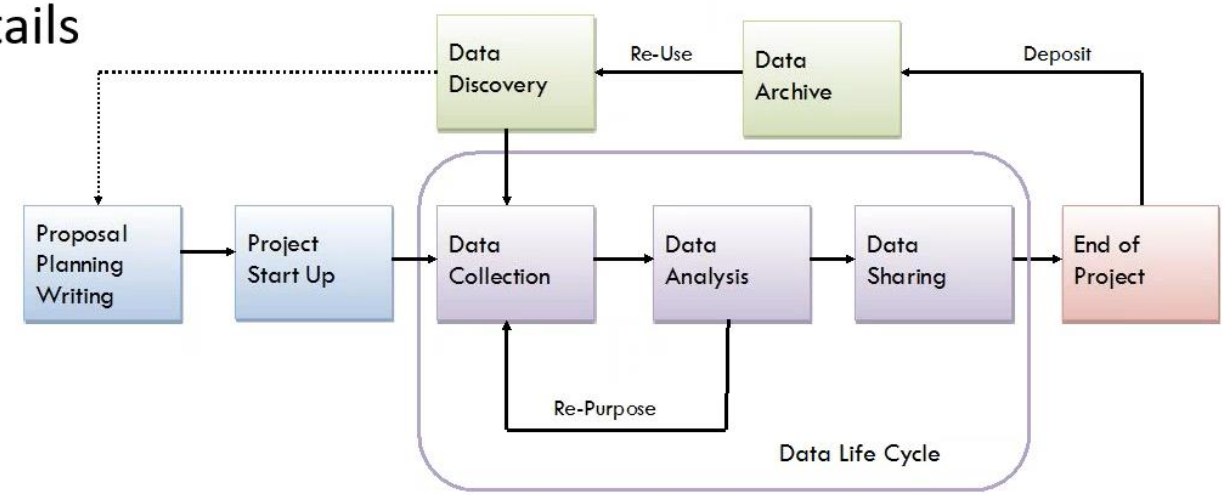
- Who plans the RDM?

  Researchers and project members, their DMPs and Data Protection Officers

- They should handle the following:
  - Collection/creation
  - Documentation
  - Storage
  - Security
- What additional resources are needed?
  - Training
  - Data repository fees
  - Hardware/software

Managing Data in the Research Life Cycle

- Choosing file formats
- File organization & naming conventions
- Version control
- Document all project/file details
- Access control & security
- Backup & storage
- File format conversions
- Sharing and preservation



video: DOI: 10.5281/zenodo.2564974

TRAININGS

E|M|G GROUP

CONTENT

1. Publishing an article
    1. Access and costs
2. Data Management
    1. Concept, costs, planning
3. Data for everyone
    1. FAIR principles and ethics

Data for everyone

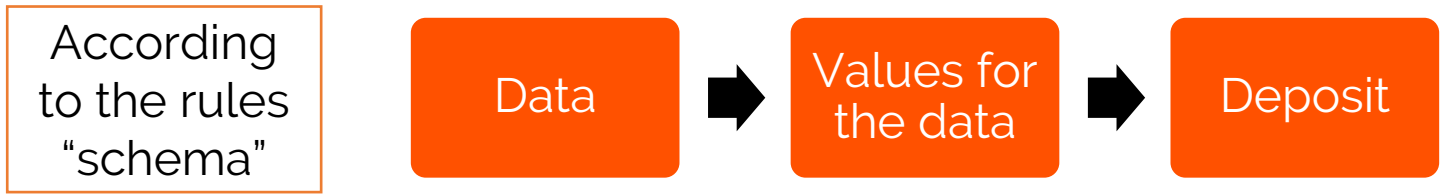
FAIR principles

- **Findable**
  - Persistent identification, registered metadata
- **Accessible**
  - deposited in publicly accessible data repositories
- **Interoperable**
  - use standardised vocabularies and formats
- **Re-usable**
  - have clear licensing and use accepted community standards

Making data findable

- **Persistent identifier** (PID) – long lasting reference to a digital source
- **Digital object identifiers** (DOIs)- digital identifiers of a digital of physical object.
- And many more
- You need to have **rich metadata** (descriptive, administrative and structural). Visit: Dublin Core metadata generator.
- **Naming conventions** according to discipline.
- Versioning

Basic structure of metadata



Making data accessible

- “As Open as Possible, As Closed as Necessary”

(H2020 AGA Annotated Model Grant Agreement).

- Data should be easily retrievable online
- Secured
- Stored safely
- It has restrictions/license for usage

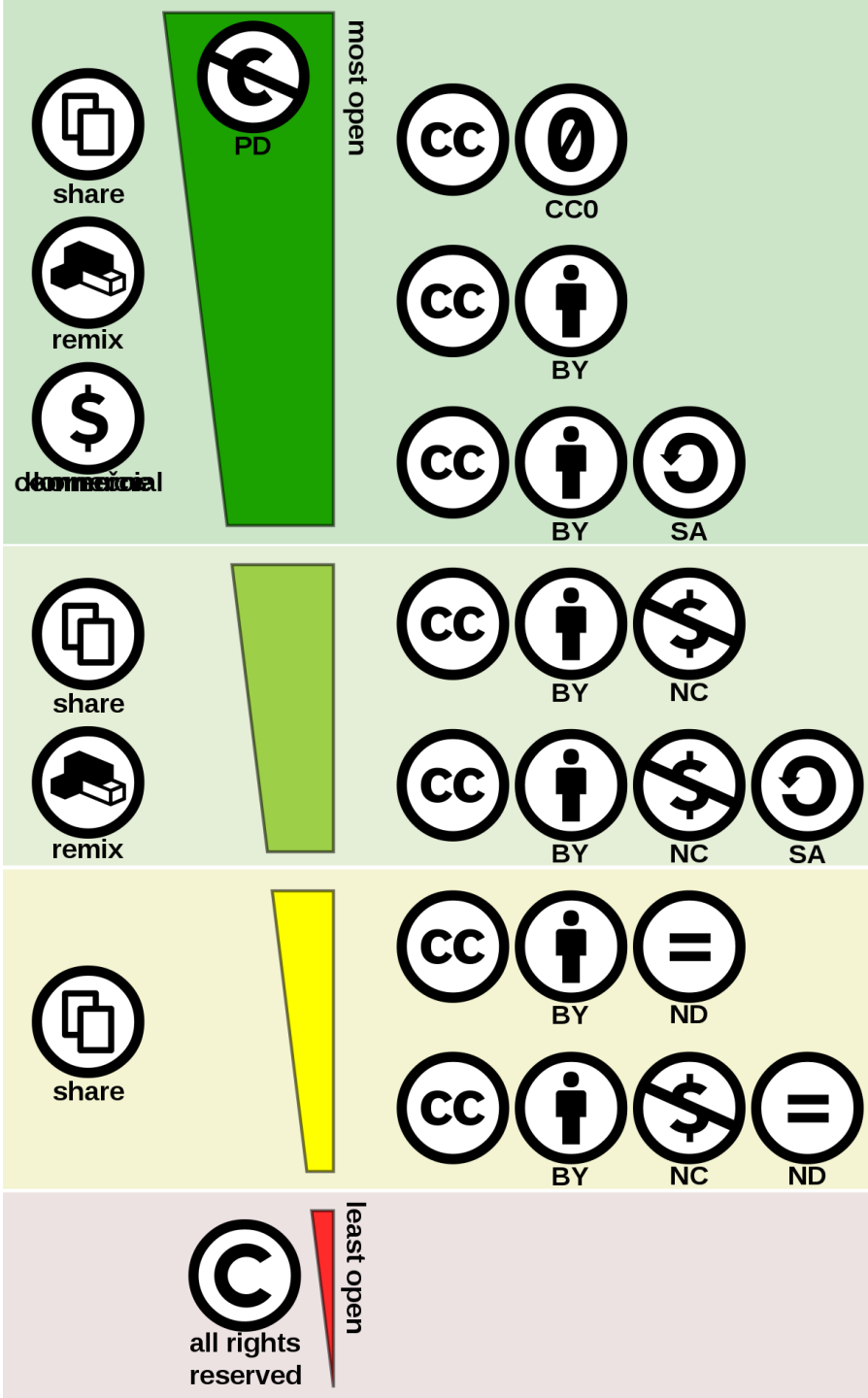
Making data interoperable

- Ensuring that **your data can be used between different systems** and technologies.

- This can be done by:
  - Common formats and standards
  - Discipline standard vocabulary

Making data re-usable

- Well-documented:
  - guidelines on types of data, definitions of categories, naming, versioning etc.
- Licensing:
  - look into Creative Commons licenses



Sharing
Data Publication and preservation

Data should be published in the following formats:

- discipline-specific data archives/centres or repositories
  - Life science Europe PMC (by EMBL-EBI)
  - ArXiv.org (physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics – by Cornell University)
  - Look at http://re3data.org/
- generic repositories such as Zenodo (by CERN and OpenAIRE)
- platforms like Open Science Framework (https://osf.io/)
- a data supplement in a journal
- a data journal

Also consider **data types and formats**

Metadata (capturing, creation, management, standards)

Access options

different levels of openness & their importance

What access levels are there?

- open to everyone (public use file)
- embargo (accessible after a certain date/period)
- limited user group (only registered users, campus use files etc.)
- on request
- closed (on-site-use/access only; data centres)

What is it important for?

- Proposals
- Consortium Agreement
- Joint ownership agreements e.g.

Ethical issues

Horizon 2020

- Management and risk assessment
- Research with humans
- Protection of Personal data
- Recruitment
- Informed consent
- Third countries involved
- Misuse

Ethical issues

Data management

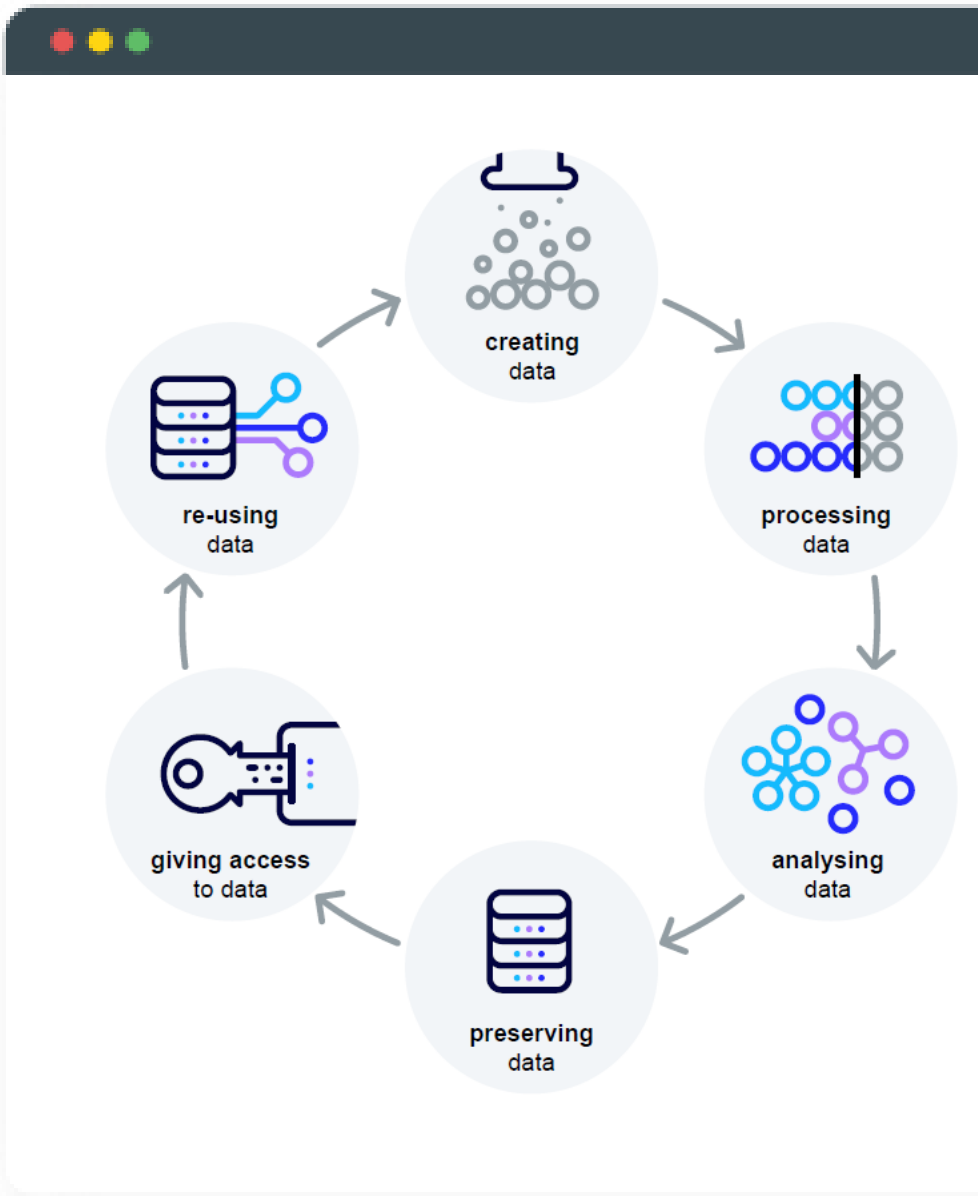
- Ethical and privacy issues regarding sharing data
- Organisation's ethical and data protection policies
- Confidential or sensitive information (written consent)
- Do data need to be anonymized?
- Secure storage and transfer of sensitive data

Research Data Management

Horizon 2020 sources, guides

- Research data management (RDM) open training materials (Zenodo)
- 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

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf



E|M|G GROUP 20 YEARS ANNIVERSARY

EM TRAININGS

VISIT

europamediatrainings.com

for more

Things you wished you knew about Horizon 2020

info@europamediatrainings.com

thank you

for your attention

Gabriella Lovasz and Frosina Ilievska
gabriella.lovasz@europamedia.org; frosina.ilievska@europamedia.org

FOLLOW US!
@EuropaMedia
@gabocsek and @FrosinaIlievska