



Open Access & Data Management in
Horizon 2020
18 & 19 October 2018



How to manage your data to make them open and FAIR

Iryna Kuchma, EIFL Open Access Programme Manager, @irynakuchma
Health NCP Net 2.0 Training Open Access & Data Management in Horizon 2020, EKT

19th October 2018

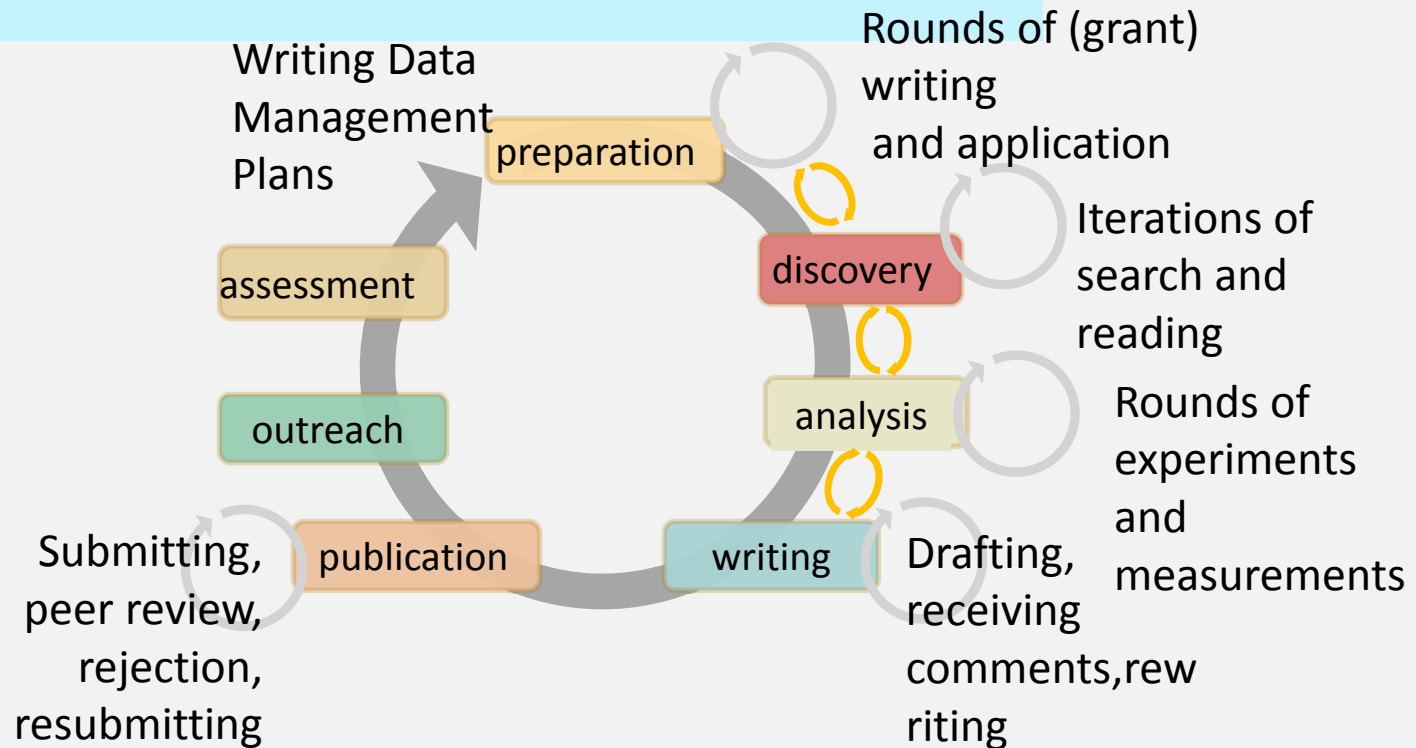


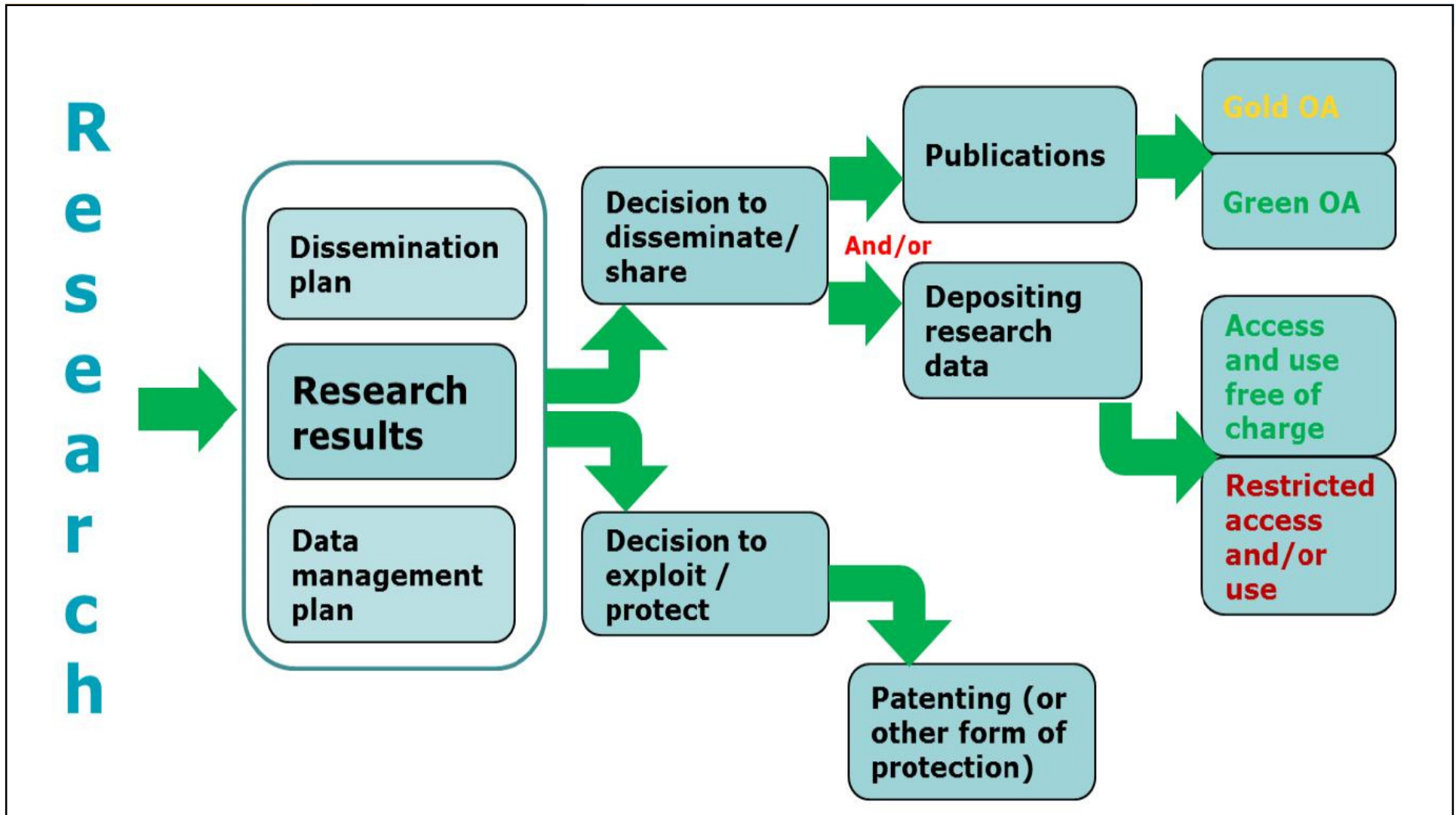
**What makes a good Data
Management Plan (DMP)?
Common themes to address in
DMPs and some sample texts**

A model of the research workflow



A model of the research workflow



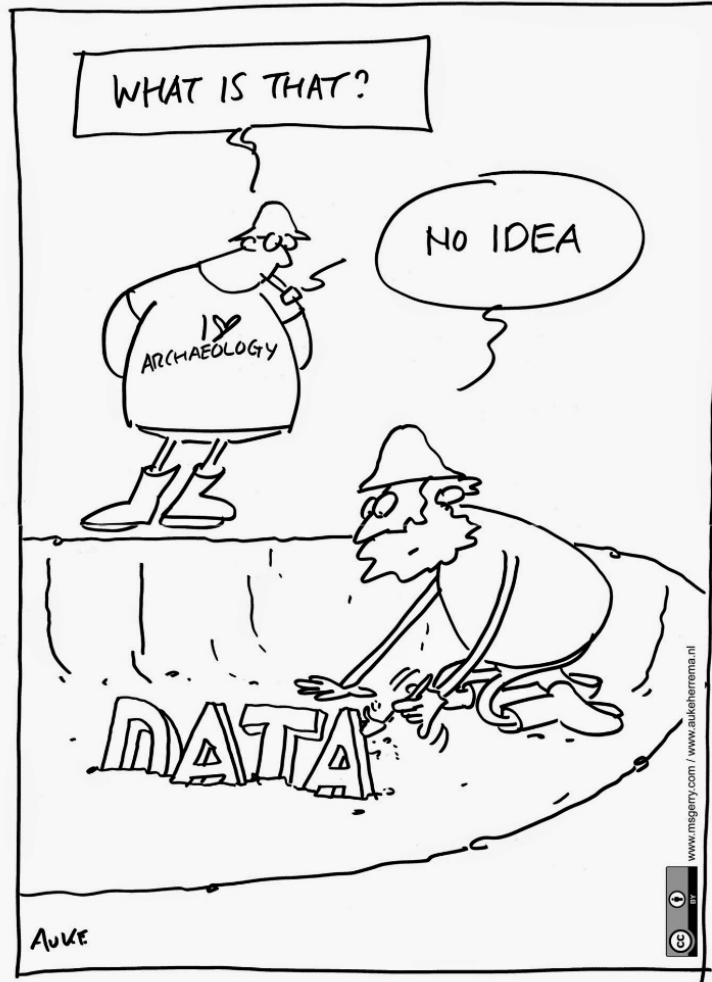




Defining your data

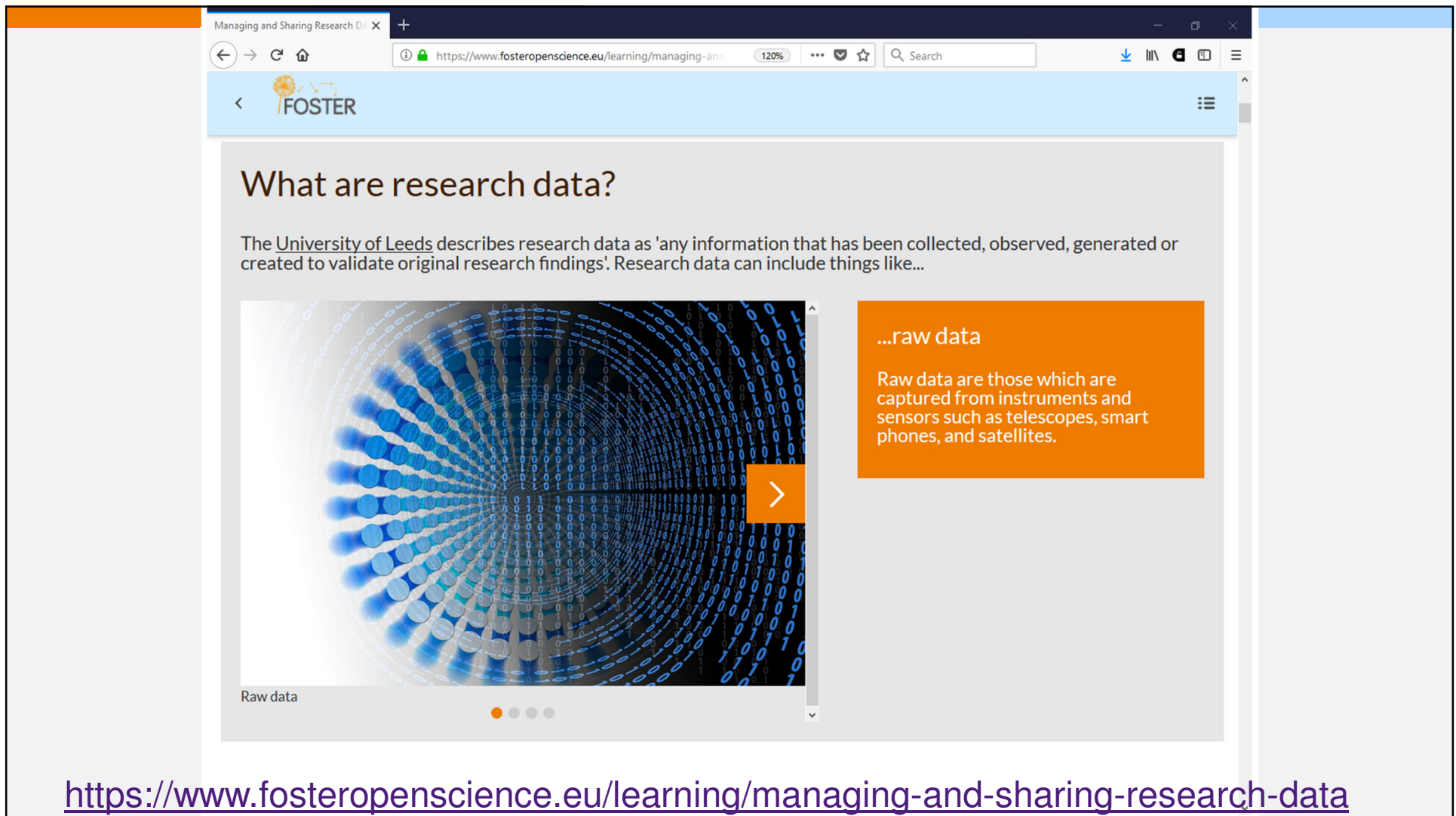
Describe your data (e.g. type, format, volume)





DATA FOR FUTURE GENERATIONS

Image courtesy of <http://aukeherrema.nl> CC-BY



Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data

What are research data?

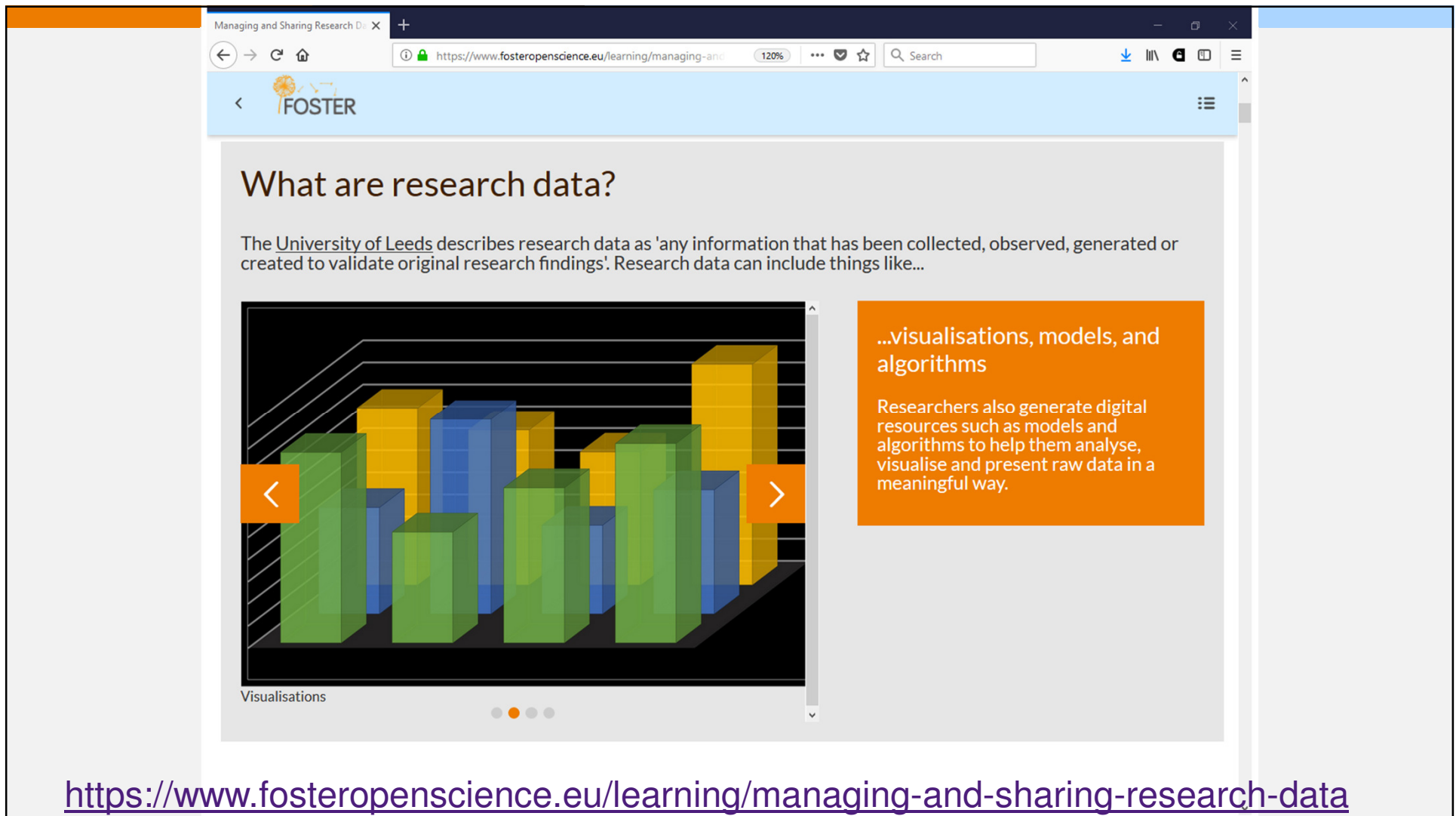
The University of Leeds describes research data as 'any information that has been collected, observed, generated or created to validate original research findings'. Research data can include things like...

Raw data

...raw data

Raw data are those which are captured from instruments and sensors such as telescopes, smart phones, and satellites.

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>



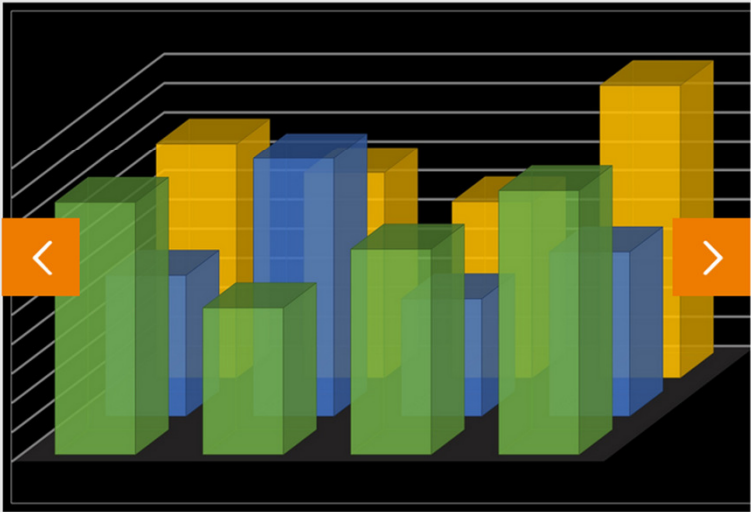
Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data

FOSTER

What are research data?

The [University of Leeds](#) describes research data as 'any information that has been collected, observed, generated or created to validate original research findings'. Research data can include things like...

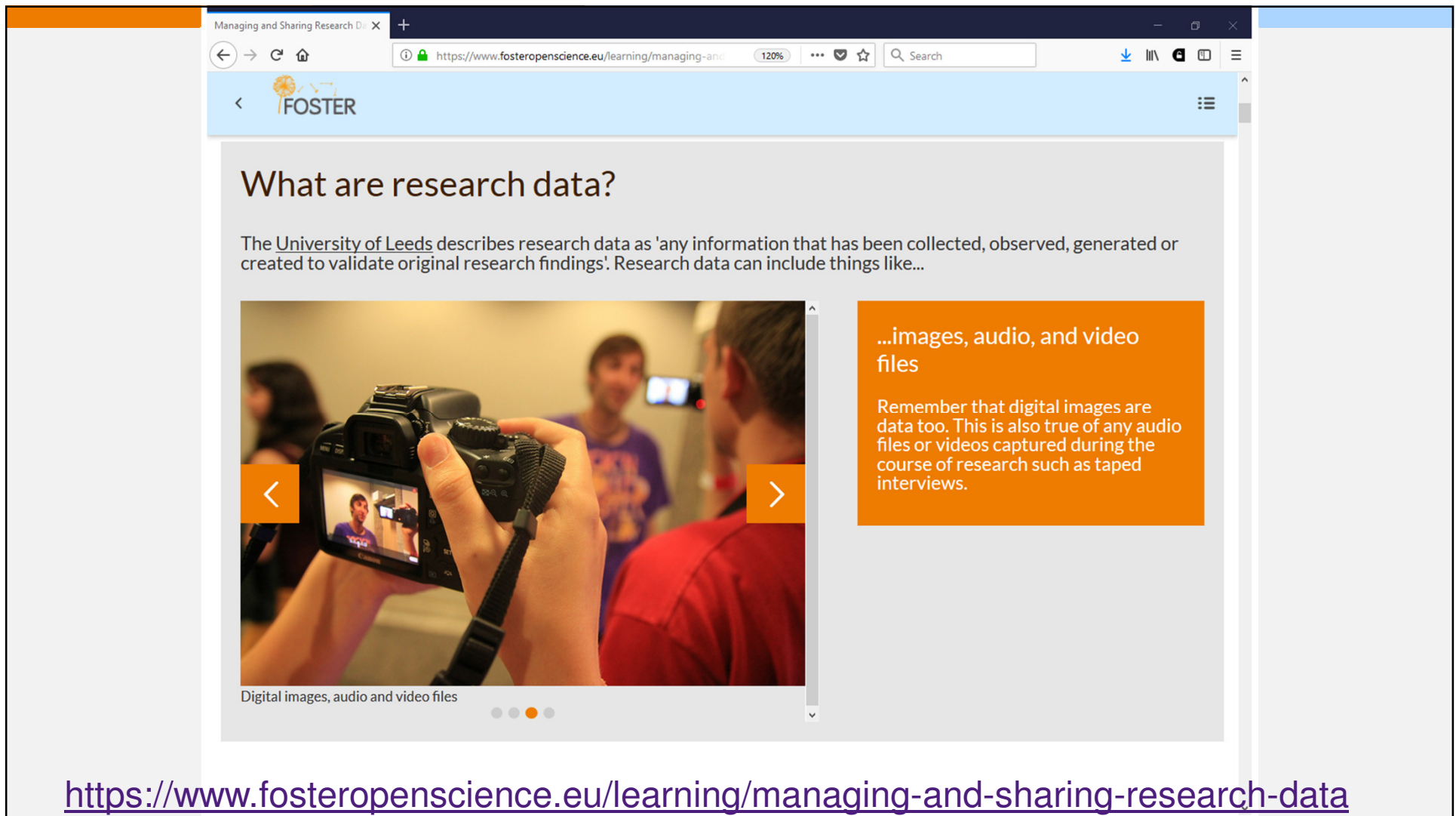


Visualisations

...visualisations, models, and algorithms

Researchers also generate digital resources such as models and algorithms to help them analyse, visualise and present raw data in a meaningful way.

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>




Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data

What are research data?

The [University of Leeds](#) describes research data as 'any information that has been collected, observed, generated or created to validate original research findings'. Research data can include things like...

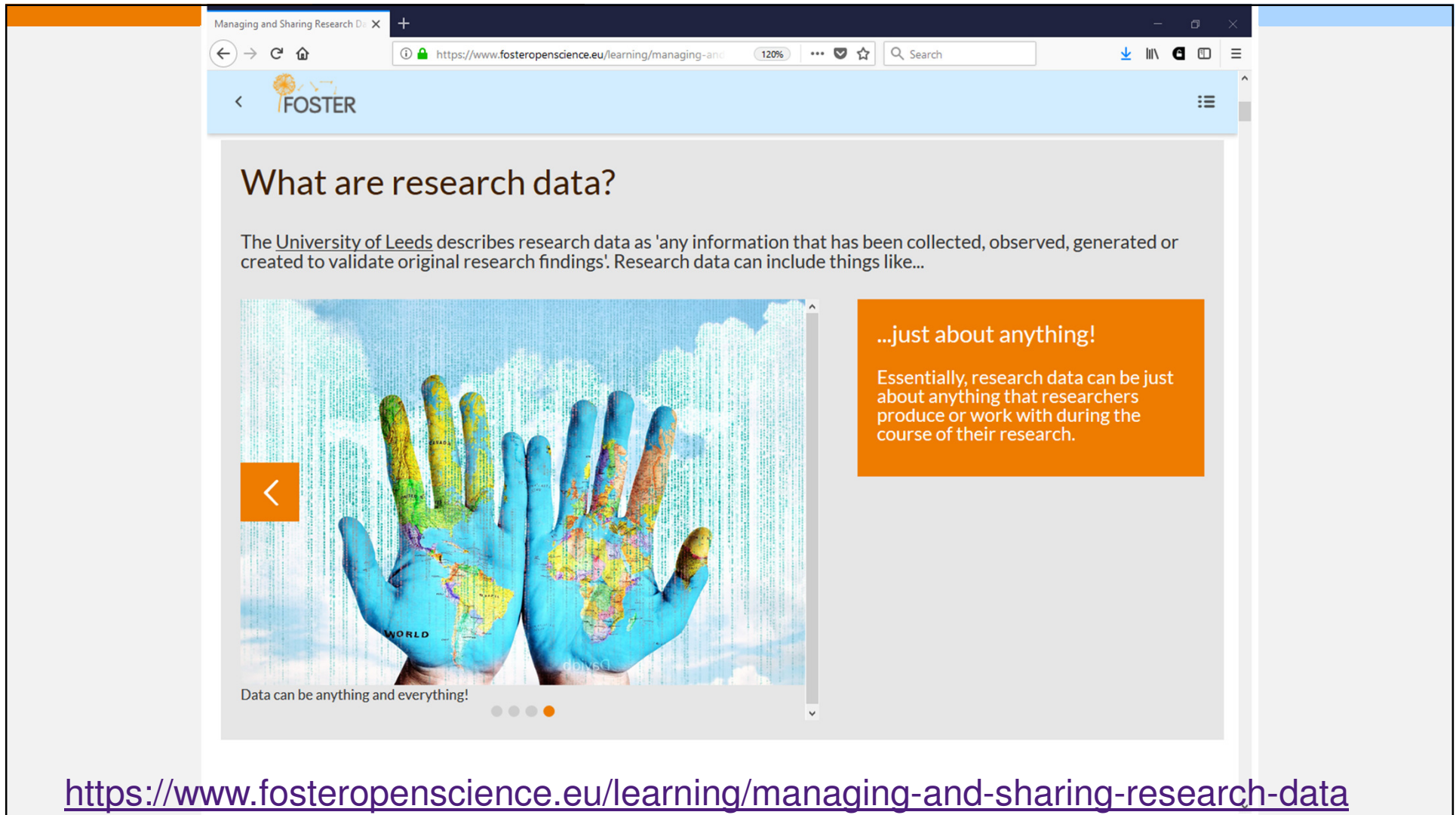


...images, audio, and video files

Remember that digital images are data too. This is also true of any audio files or videos captured during the course of research such as taped interviews.

Digital images, audio and video files

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>




Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data

FOSTER

What are research data?

The University of Leeds describes research data as 'any information that has been collected, observed, generated or created to validate original research findings'. Research data can include things like...



...just about anything!

Essentially, research data can be just about anything that researchers produce or work with during the course of their research.






Data can be anything and everything!

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>

Definition of Open Data

Open Data are online, free of cost, accessible data that can be used, reused and distributed provided that the data source is attributed.

Tip - use 5 Star Open Data Model to explain FAIR

-  make your stuff available on the Web (whatever format) under an open license
-  make it available as structured data (e.g., Excel instead of image scan of a table)
-  make it available in a non-proprietary open format (e.g., CSV instead of Excel)
-  use URIs to denote things, so that people can point at your stuff
-  link your data to other data to provide context

Tim Berners-Lee's proposal for five star open data - <http://5stardata.info>



Looking after your data

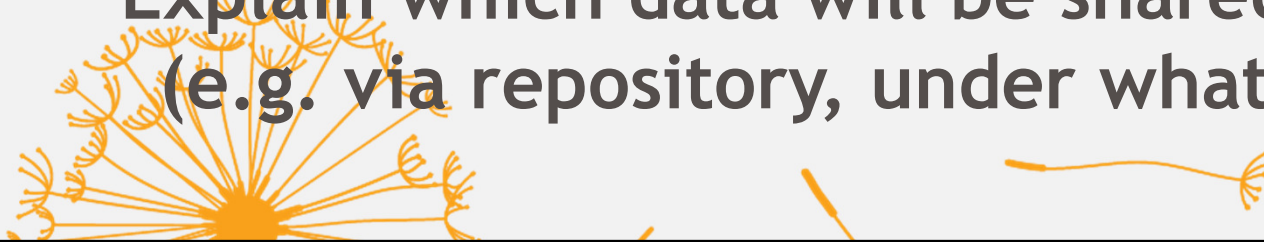
Explain how you will manage your data, noting particular concerns or issues (e.g. storage and backup, data structuring, versioning, documentation)





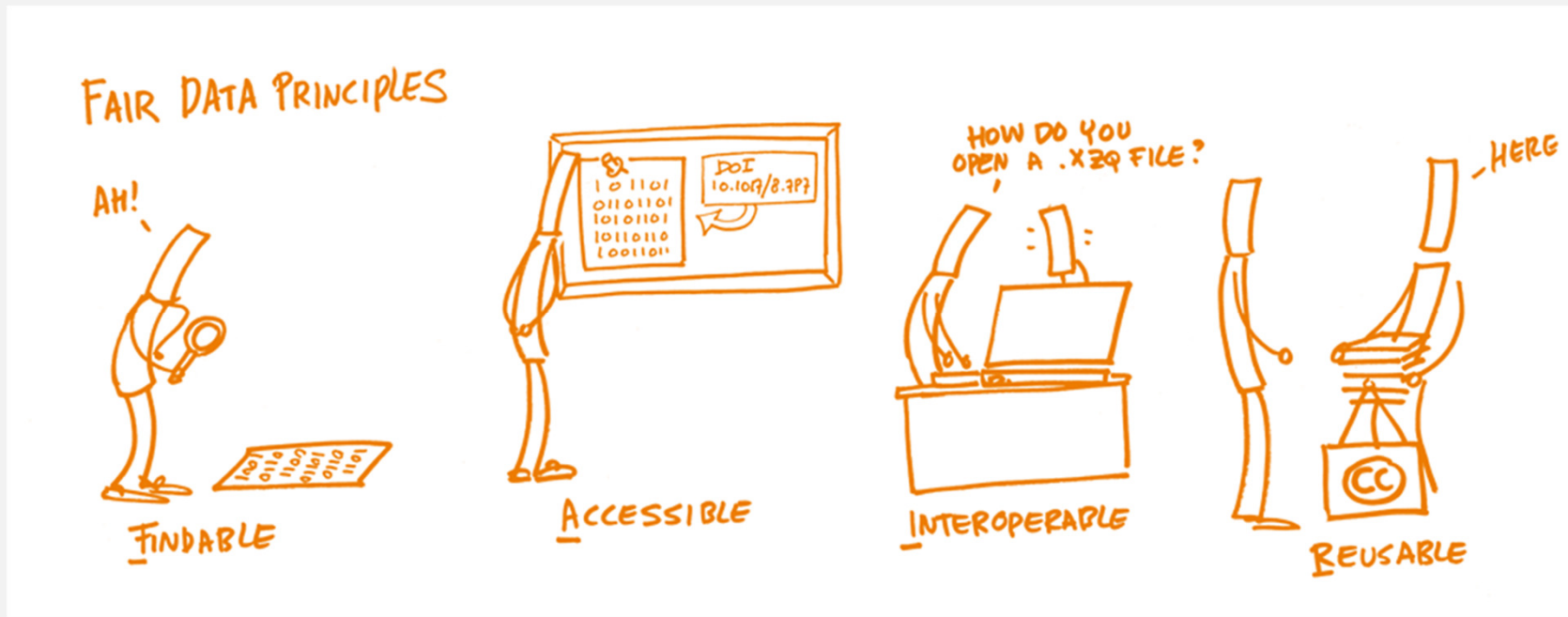
Sharing your data

Explain which data will be shared and how
(e.g. via repository, under what licence)



Misconception #1:

My web page is a FAIR way to share my data.



Better options for open data

- Domain repository (first choice)
- General repository (Figshare, Zenodo, Dryad)
- Institutional repository
- Data journal
- Journal supplementary material



The image shows a browser window displaying the homepage of re3data.org. The browser's address bar shows the URL www.re3data.org. The website header includes navigation links for Search, Browse, Suggest, Resources, and Contact, along with the DataCite logo. The main content area features the re3data.org logo and the tagline 'REGISTRY OF RESEARCH DATA REPOSITORIES'. Below the logo is a search bar with the placeholder text 'Search...' and a 'Search' button. The background of the page is a blue-tinted image of mountains. At the bottom center, there is a circular icon containing a document symbol.

re Browse by subject | re3data.org

www.re3data.org/browse/by-subject/

re3data.org Search Browse Suggest Resources Contact DataCite

Browse by subject

Graphical Text

click to zoom into subjects or to select a bottommost subject in the hierarchy as filter for the re3data search page
ctrl + click on a top subject to select it as filter

The sunburst chart displays a hierarchical classification of subjects. The inner ring shows major categories: Engineering Sciences (blue), Humanities (yellow), and Natural Sciences (green). The middle ring shows sub-categories like Computer Science, Materials, and Geosciences. The outer ring lists specific sub-fields such as Robotics, Quantum Computing, and Nanotechnology. The chart is semi-circular and uses different colors for each major category.

re3data.org

Search Browse Suggest Resources Contact DataCite

Browse by subject

Graphical Text

- A. Humanities and Social Sciences
 - a. Humanities
 - I. Ancient Cultures
 - 1. Prehistory
 - 2. Classical Philology
 - 3. Ancient History
 - 4. Classical Archaeology
 - 5. Egyptology and Ancient Near Eastern Studies
 - II. History
 - 1. Medieval History
 - 2. Early Modern History
 - 3. Modern and Current History
 - 4. History of Science
 - III. Fine Arts, Music, Theatre and Media Studies
 - 1. Art History
 - 2. Musicology
 - 3. Theatre and Media Studies
 - IV. Linguistics
 - 1. General and Applied Linguistics
 - 2. Individual Linguistics
 - 3. Typology, Non-European Languages, Historical Linguistics
 - V. Literary Studies
 - 1. Medieval German Literature
 - 2. Modern German Literature
 - 3. European and American Literature
 - 4. General and Comparative Literature and Cultural Studies

re3data.org

Search Browse ▾ Suggest Resources ▾ Contact DataCite

Browse by content type

- Archived data
- Audiovisual data
- Configuration data
- Databases
- Images
- Networkbased data
- Plain text
- Raw data
- Scientific and statistical data formats
- Software applications
- Source code
- Standard office documents
- Structured graphics
- Structured text
- other

Legal notice / Impressum DataCite

To the extent possible under law, re3data.org has waived all copyright and related or neighboring rights to the database entries of re3data.org.

Except where otherwise noted, content on this site is licensed under a [Creative Commons Attribution 4.0 International License](#).

The screenshot displays the 'Browse by country' page on re3data.org. The browser's address bar shows the URL 'www.re3data.org/browse/by-country/'. The page header includes the re3data.org logo, navigation links for 'Search', 'Browse', 'Suggest', 'Resources', and 'Contact', and the DataCite logo. The main content area is titled 'Browse by country' and contains two tabs: 'Graphical' (selected) and 'Text'. Below the tabs is a world map where countries are colored in green and blue. The map shows that most of North America, South America, and Europe are colored green, while many countries in Africa, Asia, and Oceania are colored blue. The browser's status bar at the bottom indicates a zoom level of 133%.

re3data.org

Search Browse Suggest Resources Contact DataCite

Browse by country

Graphical Text

- International
- Afghanistan
- Australia
- Austria
- Azerbaijan
- Belgium
- Benin
- Burkina Faso
- Bosnia and Herzegovina
- Brazil
- Canada
- Switzerland
- China
- Cote d'Ivoire
- Cameroon
- Colombia
- Costa Rica
- Cyprus
- Czech Republic
- Germany
- Denmark
- European Union
- Egypt
- Spain
- Estonia

There are ways to share sensitive data too

- Open metadata
- Data brokers and data access committees
- Safe havens
- Institutional data archive/vault



UK Data Service » Access control

https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control

Site search Accessibility FAQ Help Contact New users Sitemap

UK Data Service

About us Get data Use data **Manage data** Deposit data News and events

Home > Manage data > Legal and ethical > Access control

Regulating access to data

Sensitive and confidential data can be safeguarded by regulating or restricting access to and use of the data. Access controls should always be proportionate to the kind of data and level of confidentiality involved.

When regulating access, consider who would be able to access your data, what they are able to do with it, whether any specific use restrictions are required, and for how long you want the data to be available.

Advice for depositors

Researchers wishing to deposit confidential research data should [get in touch](#) if they think additional access restrictions to the data they are depositing are required.

Three tiers of access | Five Safes

Three tiers of access

The UK Data Service facilitates three levels of access for data:

- open data: for data that contain no personal or disclosive information
- safeguarded data: for data that contain no personal information, but the data owner considers there to be a risk of disclosure resulting from linkage to other data
- controlled data: for data that may be disclosive

Open data are licensed under an open licence such as Open Government Licence or Creative Commons Licence and user do not need to register to access the data.

Safeguarded data are licensed under the [End User Licence](#) and users need to be registered. Users agree to certain conditions, such as not to disseminate any identifying or confidential information on individuals, households or organisations, and not to use the data to attempt to obtain information relating specifically to an identifiable individual. Safeguarded data may have additional conditions such as requiring data owner permission or prohibiting commercial use.

Controlled data are only available to users who have been trained and accredited and

Legal and ethical

- Obligations
- Disclosure assessment
- Consent for sharing
- Anonymisation
- Access control**
- Resources

DATA CATALOGUE

GUIDE TO GOOD PRACTICE

Get the handbook on Managing and Sharing Research Data: a Guide to Good Practice from Sage publications

QUICK ACCESS TO

- Data Access Policy
- Five Safes for Secure Lab

<https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control>

Three tiers of access

open data: for data that contain no personal or disclosive information (licensed under an open licence)

safeguarded data: for data that contain no personal information, but the data owner considers there to be a risk of disclosure resulting from linkage to other data (licensed under the End User Licence and users need to be registered)

controlled data: for data that may be disclosive (<http://blog.ukdataservice.ac.uk/access-to-sensitive-data-for-research-the-5-safes/>)

Tip - some repository decisions are tricky

- There may be a preferred repository that the funder expects
- Data from multidisciplinary studies may not have an obvious home
- Data types and volumes will also need to be taken into account

Misconception #2:

I don't need to decide now if I want to share.
I can wait and see what I want to do at the
end of my project.



Open Data doesn't just happen - data management planning helps!

- What data will be created (format, types, volume...)
- Standards and methodologies to be used (incl. metadata)
- How ethics and Intellectual Property will be addressed
- Plans for data sharing and access
- Strategy for long-term preservation

A DMP is a plan to share!



Tip - use existing tools and guidance to help write their plans



<https://dmponline.dcc.ac.uk>

Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data

FOSTER

Data management planning tools - DMPonline

DMPonline is a freely available tool that helps research teams to write data management plans that meet funding body requirements. DMPonline was jointly developed by the Digital Curation Centre (DCC) and the University of California Curation Center (UC3). The tool contains a number of templates that represent the requirements of different funding bodies across Europe. Users are asked three questions at the outset to determine the appropriate template to display (e.g. the Economic and Social Research Council (ESRC) template when applying for an ESRC grant). Using tools like DMPonline takes the guesswork out of writing your data management plan by providing you with the specific set of questions that individual funding bodies want you to answer. The tool also provides users with general guidance - and where provided, institutional guidance - to make sure that your answers are realistic and implementable.

For more information on data management plans and tips on writing them, check out the [DCC website](#).

DMPonline Home Public DMPs Funder requirements Help Language

Welcome

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

Join the growing international community that have adopted DMPonline:

- 17,622 Users
- 203 Organisations
- 23,083 Plans
- 89 Countries

Some funders mandate the use of DMPonline, while others point to it as a useful option. You can [download funder templates](#) without logging in, but the tool provides tailored guidance and example answers from the DCC and many research organisations. [Do not sign up for an account and still not?](#)

Sign in Create account

* Email

* Password

Forgot password?

Remember email

Sign in

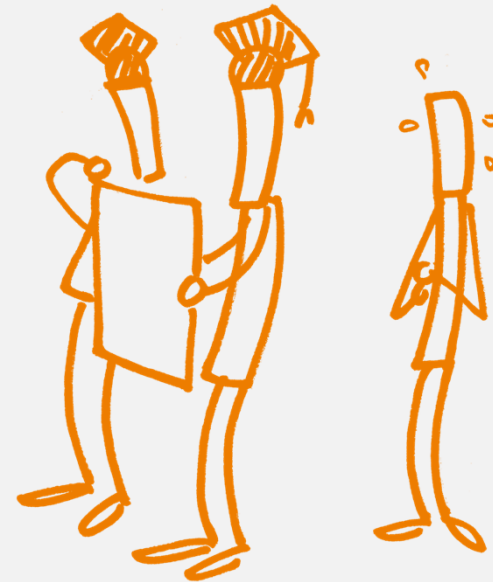
- or -

Sign in with institutional credentials (UK only)

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>

Misconception #3:

If I share my data early, I'll be scooped!



Pre-registration timestamps your work

Register Your Project



Open Science Framework

A registration on OSF creates a frozen, time-stamped version of a project that cannot be edited or deleted. The *original project* can still be edited, while the registered version cannot. You might create a registration to capture a snapshot of your project at certain points in time - such as right before data collection begins, when you submit a manuscript for peer review, or upon completion of a project.

Registrations can be made public immediately or embargoed for up to 4 years. Registrations cannot be deleted, but they can be withdrawn. [Withdrawing a registration](#) removes the content of the registration but leaves behind basic metadata, like registration title, contributors, and a reason for the withdrawal (not required).

<http://help.osf.io/m/registrations/l/524205-register-your-project>

Tips - share preprints too

- Early feedback on methods and initial findings
- Time to correct and mistakes before publishing
- Recognition for your ideas by peers

Misconception #4:

I have to keep and share everything.



Image: 'Balancing rocks' by Viewminder CC-BY-SA-ND www.flickr.com/photos/light_seeker/7780857224

Deciding which data need to be kept after the project ends

Five steps to follow

- ① **Could** this data be re-used
- ② **Must** it be kept as evidence or for legal reasons
- ③ **Should** it be kept for its potential value
- ④ **Consider costs** – do benefits outweigh cost?
- ⑤ **Evaluate criteria** to decide what to keep

5 steps to decide what data to keep

www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep

Tip - link data to other outputs for context (reuse)

Open Data



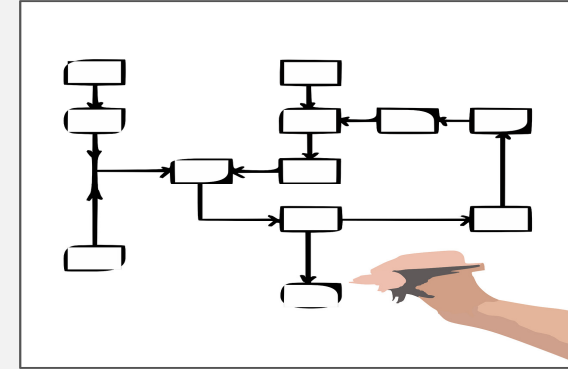
To support validation and facilitate reuse

Open Code



Software created to analyse and/or visualise the data

Open Workflows



What steps were taken and in what order?

Consider who else has a say about sharing data

- Collaborators
- Research participants
- Commercial partners
- Data repository
- Publishers
- Institutions, funders



Exercise: barriers to data sharing

In groups of 2-3, consider any barriers to sharing data.

If there are any specific issues for your discipline please feel free to note these.

10 minutes plus feedback

How to make data open?



<https://okfn.org>

1. Choose your dataset(s)
 - What can you open? You may need to revisit this step if you encounter problems later.
2. Apply an open license
 - Determine what IP exists. Apply a suitable licence e.g. CC-BY
3. Make the data available
 - Provide the data in a suitable format. Use repositories.
4. Make it discoverable
 - Post on the web, register in catalogues...





WHAT IS A DMP & WHY WRITE ONE?

Image CC-BY-NC-SA by Leo Reynolds www.flickr.com/photos/lwr/13442910354

Data Management Plans (DMP)

A DMP is a brief plan to define:

- how the data will be created
- how it will be documented
- who will be able to access it
- where it will be stored
- who will back it up
- whether (and how) it will be shared & preserved

DMPs are often submitted as part of grant applications, but are useful whenever researchers are creating data.



OPEN RESEARCH DATA

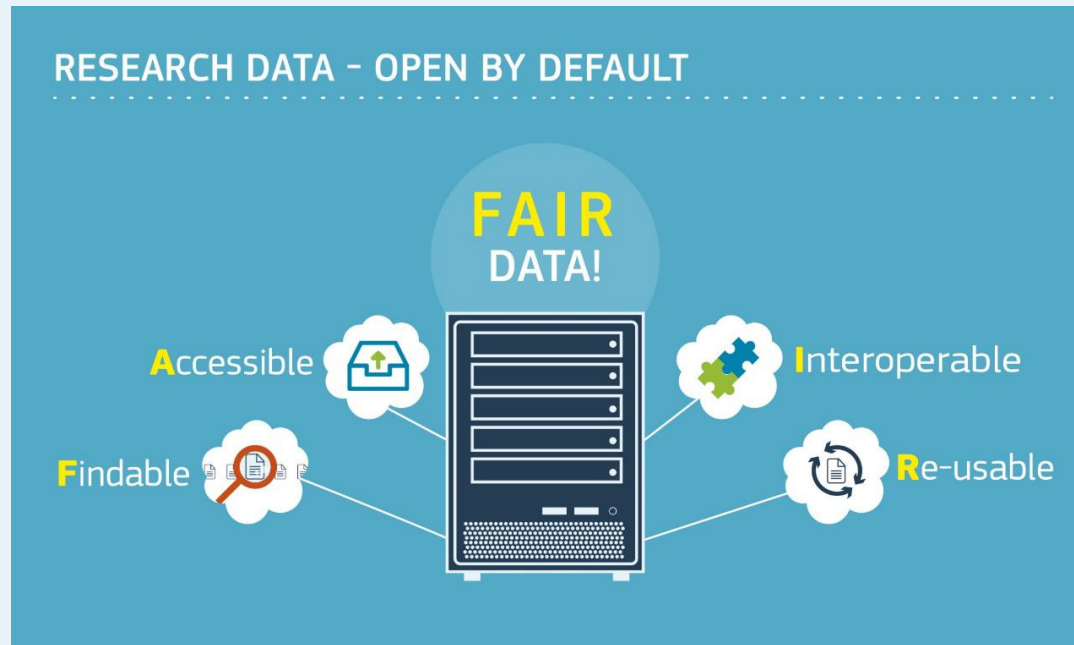
IN HORIZON 2020

**Jean-François Dechamp
& Daniel Spichtinger**

European Commission
Directorate-General for Research &
Innovation

Research and
Innovation

RESEARCH DATA - OPEN BY DEFAULT



Research and Innovation



Making data FAIR

Findable - Assign persistent IDs, provide rich metadata, register in a searchable resource,...

Accessible - Retrievable by their ID using a standard protocol, metadata remain accessible even if data aren't...

Interoperable - Use formal, broadly applicable languages, use standard vocabularies, qualified references...

Reusable - Rich, accurate metadata, clear licences, provenance, use of community standards

www.force11.org/group/fairgroup/fairprinciples

Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data


FOSTER

Assessing the FAIRness of your data

FAIR self-assessment tool

This self-assessment tool was developed by the Australian ANDS-Nectar-RDS initiative and enables you to assess the 'FAIRness' of a dataset and determine how to enhance its FAIRness. It was developed primarily for research support staff but can be used by anyone.

ANDS, Nectar and RDS are supported by the Australian Government through the National Collaborative Research Infrastructure Strategy program.



FAIR data self-assessment tool

The new FAIR data self-assessment tool enables you to assess the 'FAIRness' of a dataset and determine how to enhance its FAIRness (where applicable). This self-assessment tool has been designed predominantly for data librarians and IT staff, but could be used by anyone.

[Check out the tool](#)

Self-Assessment Tool

ANDS-Nectar-RDS FAIR self-assessment tool

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>

RESEARCH DATA - OPEN BY DEFAULT

Projects must have



Provides information on:



the data the research will generate



how to ensure its curation, preservation and sustainability



what parts of that data will be open (and how)

Research data lifecycle

RE-USING DATA:

follow-up research,
new research,
undertake research
reviews, scrutinising
findings, teaching &
learning

CREATING DATA: designing research,
DMPs, planning consent, locate existing
data, data collection and management,
capturing and creating metadata

PROCESSING DATA:

entering, transcribing,
checking, validating and
cleaning data,
anonymising data,
describing data, manage
and store data

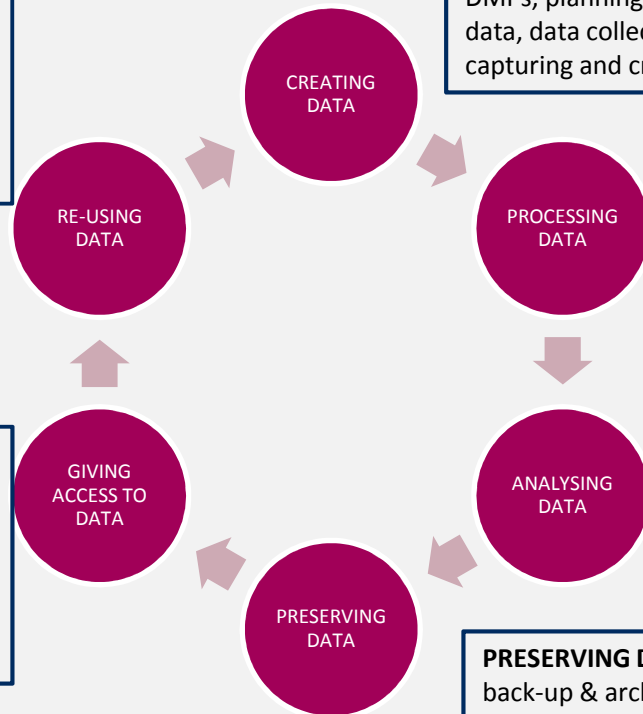
ANALYSING DATA:

interpreting, & deriving
data, producing outputs,
authoring publications,
preparing for sharing

PRESERVING DATA: data storage,
back-up & archiving, migrating to best
format & medium, creating metadata
and documentation

ACCESS TO DATA:

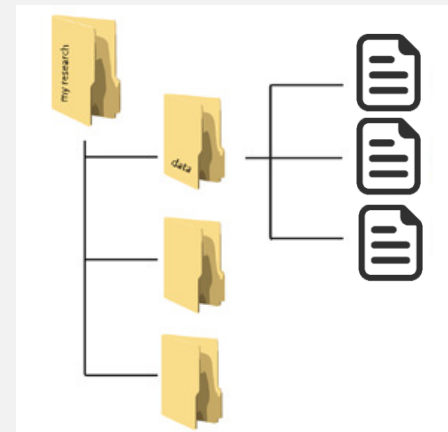
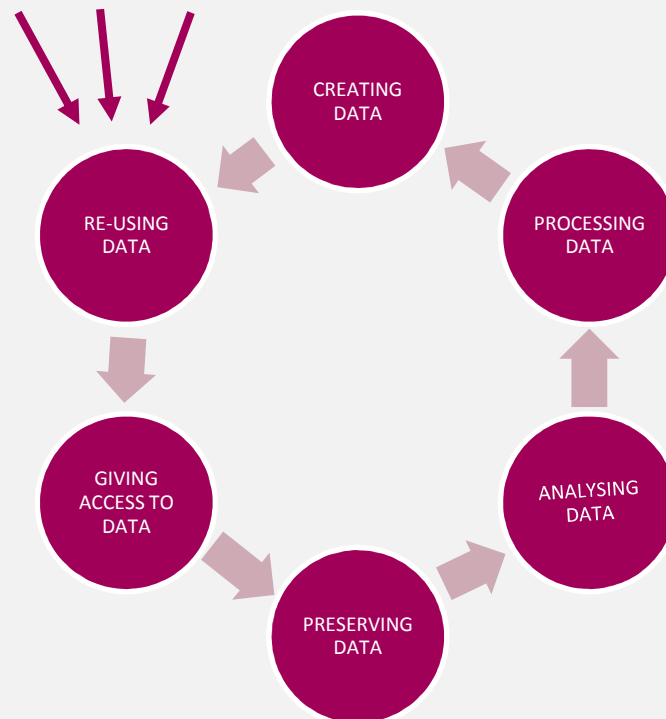
distributing data,
sharing data,
controlling access,
establishing
copyright, promoting
data



Ref: UK Data Archive: <http://www.data-archive.ac.uk/create-manage/life-cycle>

Planning trick 1: think backwards

What data organisation would a re-user like?



Data organisation

Meaningful file names

Below are tips on meaningful and consistent file names. Read more in '[Choosing a file name](#)'.⁽²⁾

- ❑ Make sure to use consistent file names. When you use a date in the file name, choose a notation (for instance, YYYYMMDD or yymmdd).
- ❑ Do not use strange characters like ?\!@*%{[<> in the file name.
- ❑ Use traceable file names, such as Project_Instrument_locatie_YYYYMMDD.ext.
- ❑ Make sure to only use each file once in the folder structure. If you store a file in more than one place, several versions of the same file can unwillingly be created.
- ❑ See also [version management](#).

It is good practice to note the file naming and its meaning in a readme.txt.

Even if a researcher is well underway with his project consistent file naming is still an option by using a [bulk file rename utility](#).⁽³⁾ It is important, however, to check if this bulk renamer delivers on its promises.



white_data_20140708.csv



blue_data_20140708.docx



red_data_20140708.R



red_data_20140708_v02.R

File naming and version management

<http://datasupport.researchdata.nl/en/start-de-cursus/iii-onderzoeksfase/organising-data>

File naming and folder structure | +

https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/2.-Organise-Document/File-naming-and-folder-structure

120%

Search

cessda eric

About | Consortium | Projects | Research Infrastructure | Contact

Expert Tour Guide on Data Management

1. Plan

2. Organise & Document

Designing a data file structure

Organisation of variables

File naming and folder structure

Documentation and metadata

Adapt your DMP: part 2

Sources and further reading


3. Process

4. Store

5. Protect

6. Archive & Publish

TIP: Batch renaming of automatically generated files



Batch renaming is organising research data files and folders in a consistent and automated way with software tools (also known as mass file renaming, bulk renaming).

Batch renaming software exists for most operating systems. See the accordion for examples.

+ Batch renaming tools

It may be useful to rename files in a batch when:

- Images from digital cameras are automatically assigned base filenames consisting of sequential numbers;
- Proprietary software or instrumentation generate crude, default or multiple filenames;
- Files are transferred from a system that supports spaces and/or non-English characters in filenames to one that doesn't (or vice versa). Batch renaming software can be used to substitute such characters with acceptable ones.

<https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/2.-Organise-Document/File-naming-and-folder-structure>

How to ... use Bulk Rename Utility

File naming and folder structure

https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/2.-Organise-Document/File-naming-and-folder-structure

cessda eric

About Consortium Projects Research Infrastructure Contact

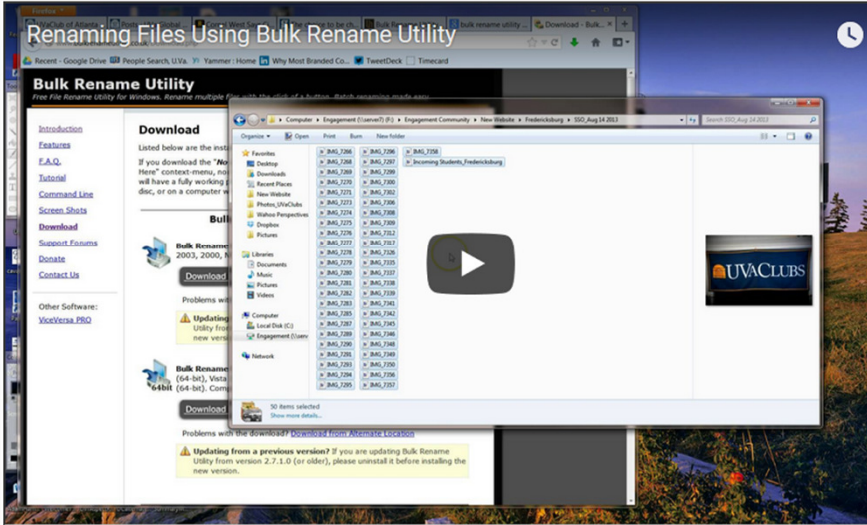
How to ... use Bulk Rename Utility

Follow the steps in the video to use Bulk Rename Utility to batch rename your files.

Expert Tour Guide on Data Management

1. Plan
2. Organise & Document
 - Designing a data file structure
 - Organisation of variables
 - File naming and folder structure**
 - Documentation and metadata
 - Adapt your DMP: part 2
 - Sources and further reading
3. Process
4. Store
5. Protect
6. Archive & Publish

Renaming Files Using Bulk Rename Utility



« Previous Next »

<https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/2.-Organise-Document/File-naming-and-folder-structure>

How to avoid a data management nightmare

A few tips for avoiding a data management nightmare. A video created by NYU Health Sciences Library's Karen Hanson, Kevin Read, and Alisa Surkis

<https://www.youtube.com/watch?v=nNBiCcBlwRA>

Why manage data?

NON PECUNIAE INVESTIGATIONIS CURATORE

SED VITAE FACIMUS PROGRAMMAS DATORUM

(Not for the research funder, but for life we make data
management plans) PROCURATIONIS

- Make your research easier
- Stop yourself drowning in irrelevant stuff
- Save data for later
- Avoid accusations of fraud or bad science
- Write a data paper
- Share your data for re-use
- Get credit for it

How to deal with data and context?

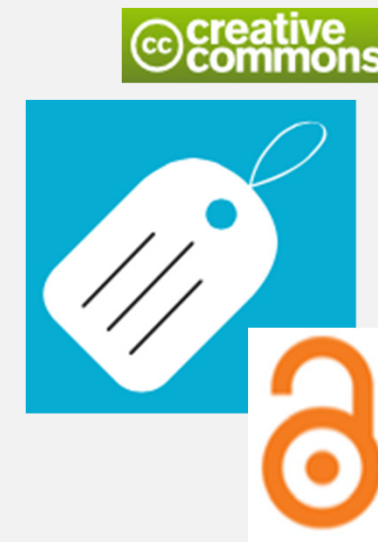
Versioning, back-up, storage and archiving

– During the project and in the long term

Ethics, consent forms, legal access

Security and technical access

Usage licences



What should be preserved and shared?

- The **data** needed to validate results in scientific publications (minimally!).
- The associated **metadata**: the dataset's creator, title, year of publication, repository, identifier etc.
 - Follow a metadata standard in your line of work, or a generic standard, e.g. Dublin Core or DataCite, and be FAIR.
 - The repository will assign a persistent ID to the dataset: important for discovering and citing the data.

What should be preserved and shared? (2)

- **Documentation**: code books, lab journals, informed consent forms - domain-dependent, and important for understanding the data and combining them with other data sources.
- **Software**, hardware, tools, syntax queries, machine configurations - domain-dependent, and important for using the data. (Alternative: information about the software etc.)

Basically, everything that is needed to replicate a study should be available. Plus everything that is potentially useful for others.

Data description examples

The final dataset will include self-reported **demographic and behavioural data** from **interviews** with the subjects and **laboratory data** from urine specimens provided.

From [NIH data sharing statements](#)

Metadata examples

Metadata will be tagged in XML using the **Data Documentation Initiative (DDI) format**. The codebook will contain information on study design, sampling methodology, fieldwork, variable-level detail, and **all information necessary for a secondary analyst** to use the data accurately and effectively.

From [ICPSR Framework for Creating a DMP](#)



REUSABLE DATA

Image courtesy of <http://aukeherrema.nl> CC-BY

The screenshot shows a web browser window with the URL www.dcc.ac.uk/resources/subject-areas/biology. The page features a red and orange header with the DCC logo and the tagline "because good research needs good data". A search bar is located in the top right of the header. Below the header is a navigation menu with links for Home, Digital curation, About us, News, Events, Resources, Training, Projects, Community, and Tailored support. The breadcrumb trail reads "Home > Resources > Subject Areas > Biology".

The main content area is titled "Biology" and displays a list of subject areas in various sizes and colors, with "Biodiversity Ecology" and "Genetics Zoology" being the most prominent. The list includes:

- Entomology
- Biomaterials
- Geography
- Toxicology
- Physiology
- Proteomics
- Bioengineering
- Environmental Science
- Botany
- Neuroscience
- Molecular biology
- Palaeontology
- Multi-disciplinary
- Hydrology
- Biodiversity Ecology
- Biomechanics
- Cartography
- Geoscience
- Marine Zoology
- Topography
- Clinical Medicine
- Bioinformatics
- Plant pathology
- Metabolic biochemistry
- Biology
- Marine Biology
- Crystallography
- Biogeography
- Meteorology
- Animal physiology
- Genomics
- Chemistry
- Genetics
- Zoology
- Soil Science
- Biochemistry
- Plant physiology
- Oceanography
- Glaciology
- Maritime Geography
- Marine Science
- Animal pathology
- Hydrogeology
- Climatology
- Geology
- Cell Biology
- Hydrography

Metadata Standards

ABCD - Access to Biological Collection Data
A standard for the access to and exchange of primary biodiversity data, including specimens and observations.

Darwin Core
A body of standards, including a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biological diversity by providing reference definitions, examples, and commentaries.

EML - Ecological Metadata Language
Ecological Metadata Language (EML) is a metadata specification particularly developed for the ecology discipline.

Use Cases

CESSDA Catalogue [ⓘ]
Provides a seamless interface to datasets from social science data archives across Europe using the CESSDA MLI profile of DDI.

DDI Projects [ⓘ]
The Data Documentation Initiative website's list of projects adopting or encouraging DDI as a standard.

DDI Use Case Literature [ⓘ]
Links to DDI 3 use case papers, which were the outcomes of a workshop held at the Schloss Dagstuhl - Leibniz Center for Informatics in Wadern, Germany, November 2-6, 2009.

English Heritage Listed Buildings System [ⓘ]
A case study of the use of the MIDAS XML Monument schema as a vehicle for storing data exported from a major heritage sector information system, the English Heritage Listed Building System (LBS).

Eurostat [ⓘ]
The statistical office of the European Union, which implements SDMX in a number of projects.

ICPSR - Inter-university Consortium for Political and Social Research [ⓘ]
A data archive providing leadership and training in data access, curation, and methods of analysis for the social science research community; all metadata conforms to the DDI standard.

Open Archives Initiative [ⓘ]
Develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content.

SDMX Implementations [ⓘ]
The SDMX website's list of current and planned practical implementations from national and international organisations.

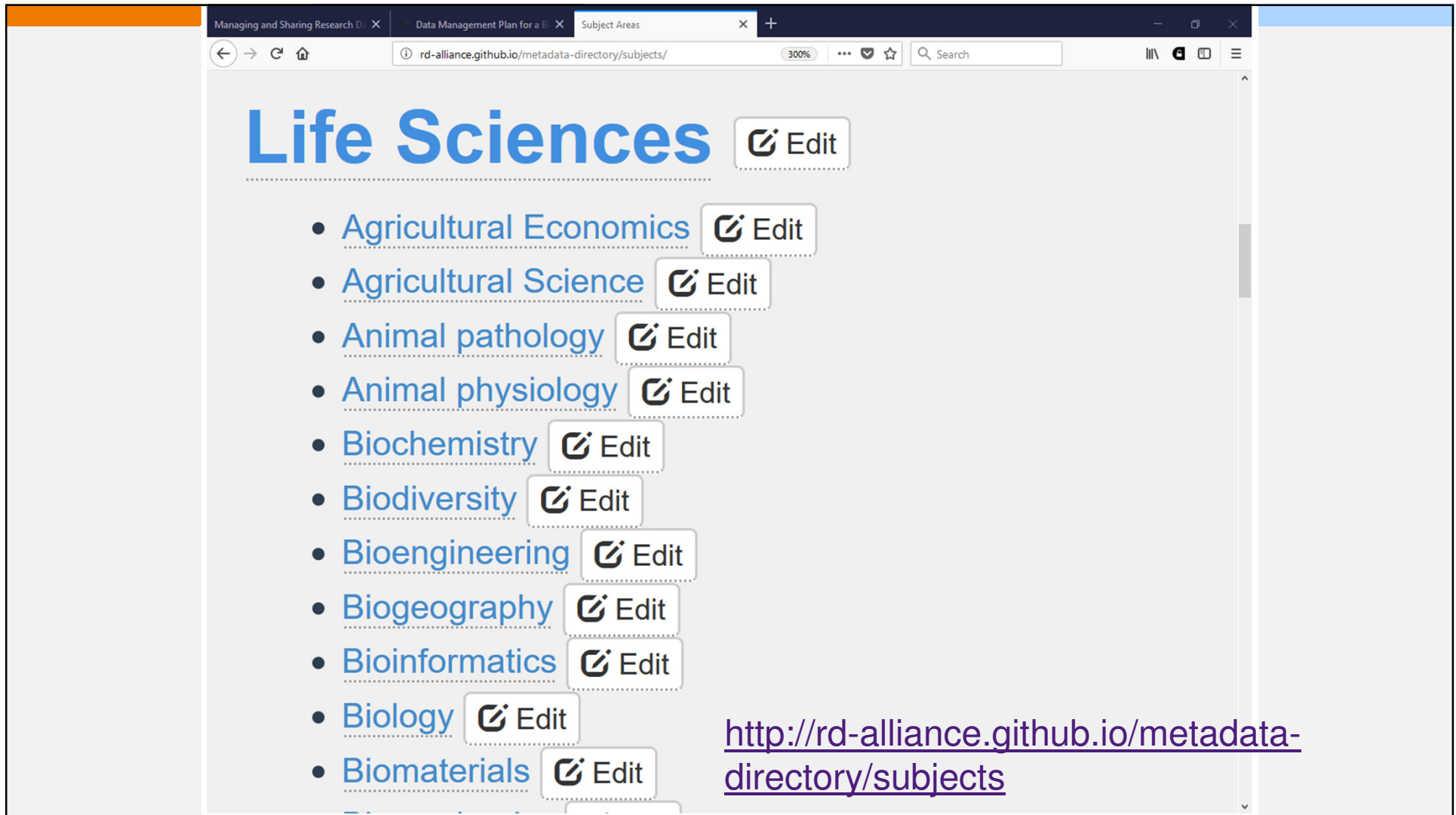
The Centre for Digital Music Research Data Repository [ⓘ]
Used by researchers at C4DM to share their research data with their colleagues and others in the digital music research community, this repository uses the DataCite metadata schema to describe its holdings.

The Institution for Social and Policy Studies (ISPS) Data Archive [ⓘ]
Provides members of the scholarly community with access to files associated with scholarly studies for the purpose of replication, for all studies conducted by ISPS-affiliated researchers. ISPS metadata records conform to DDI requirements and include a minimal set of Dublin Core metadata elements.

UK ADS - UK Archeology Data Service [ⓘ]
The ADS collects, catalogues, manages, preserves, and encourages re-use of digital resources created by archaeologists. It promotes good practice in the use of digital data in archaeology, provides technical advice to the research community, and supports the deployment of digital technologies. Its catalogue records are based on Dublin Core.

UKDA - UK Data Archive [ⓘ]
Curator of the largest collection of digital data in the social sciences and humanities in the United Kingdom, the archive uses DDI as the basis for its catalogue records.

<http://www.dcc.ac.uk/resources/subject-areas/social-science-humanities>



The image shows a web browser window with three tabs: "Managing and Sharing Research D...", "Data Management Plan for a B...", and "Subject Areas". The address bar shows the URL "rd-alliance.github.io/metadata-directory/subjects/". The main content area features a large blue heading "Life Sciences" with an "Edit" button to its right. Below this heading is a list of subject areas, each with an "Edit" button:

- [Agricultural Economics](#) Edit
- [Agricultural Science](#) Edit
- [Animal pathology](#) Edit
- [Animal physiology](#) Edit
- [Biochemistry](#) Edit
- [Biodiversity](#) Edit
- [Bioengineering](#) Edit
- [Biogeography](#) Edit
- [Bioinformatics](#) Edit
- [Biology](#) Edit
- [Biomaterials](#) Edit

At the bottom right of the page, the URL <http://rd-alliance.github.io/metadata-directory/subjects> is displayed.

FAIRsharing.org standards, databases, policies


Search all of FAIRsharing

Standards Databases Policies Collections Add/Claim Content Stats Log in or Register

A curated, informative and educational resource on data and metadata *standards*, inter-related to *databases* and data *policies*.

HOW CAN WE HELP?

We guide consumers to discover, select and use these resources with confidence, and producers to make their resource more discoverable, more widely adopted and cited.



Research data facilitators, librarians, trainers

Use FAIRsharing to provide a foundation on which to create or enrich educational lectures, training and teaching material, and to plug into data management planning tools...
[\[read more\]](#)

Researchers Developers & Curators Journal Publishers **Librarians & Trainers** Societies & Alliances Funders

[Find](#) [Discover](#) [Learn](#)

<https://fairsharing.org>

Data sharing examples

The videos will be made available [via the bristol.ac.uk website](#) (both as streaming media and downloads) HD and SD versions will be provided to accommodate those with lower bandwidth. Videos will also be made available [via Vimeo](#), a platform that is already well used by research students at Bristol. [Appropriate metadata will also be provided](#) to the existing Vimeo standard.

All video will also be available [for download and re-editing by third parties](#). To facilitate this [Creative Commons](#) licenses will be assigned to each item. In order to ensure this usage is possible, the [required permissions will be gathered](#) from participants (using a suitable release form) before recording commences.

From [University of Bristol Kitchen Cosmology DMP](#)

We will make the data and associated documentation available to users under a [data-sharing agreement](#) that provides for: (1) a commitment to using the data [only for research purposes](#) and not to identify any individual participant; (2) a commitment to [securing the data](#) using appropriate computer technology; and (3) a commitment to [destroying or returning the data after analyses](#) are completed.

From [NIH data sharing statements](#)

Examples restrictions

Because the STDs being studied are reportable diseases, we will be **collecting identifying information**. Even though the final dataset will be stripped of identifiers prior to release for sharing, we believe that there **remains the possibility of deductive disclosure of subjects** with unusual characteristics. Thus, we will make the data and associated documentation available to users **only under a data-sharing agreement**.

From [NIH data sharing statements](#)

Examples restrictions (2)

1. Share data **privately within 1 year**.
Data will be held in Private Repository, but metadata will be public
2. Release data to **public within 2 years**.
Encouraged after one year to release data for public access.
3. **Request, in writing, data privacy up to 4 years**.
Extensions beyond 3 years will only be granted for compelling cases.
4. Consult with creators of private CZO datasets prior to use.
*Pis required to **seek consent before using private data** they can access*

From [Boulder Creek Critical Zone Observatory DMP](#)

Archiving examples

The investigators will **work with staff at the UKDA** to determine **what to archive and how long** the deposited data should be retained. Future long-term use of the data will be ensured by **placing a copy of the data into the repository**.

From [ICPSR Framework for Creating a DMP](#)

Data will be provided in **file formats considered appropriate for long-term access**, as recommended by the UK Data Service. For example, SPSS Portal format and tab-delimited text for qualitative tabular data and RTF and PDF/A for interview transcripts. Appropriate **documentation necessary** to understand the data will also be provided. Anonymised data will be held for **a minimum of 10 years** following project completion, in compliance with LSHTM's Records Retention and Disposal Schedule. Biological samples (output 3) will be **deposited with the UK BioBank** for future use.

From [Writing a Wellcome Trust Data Management and Sharing Plan](#)

Sharing data: what is meant?

With collaborators
while research is active



Data are mutable

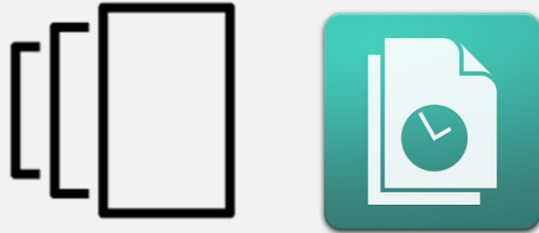
(Open) data sharing



Data are stable,
searchable, citable,
clearly licensed

Storing data: what is meant?

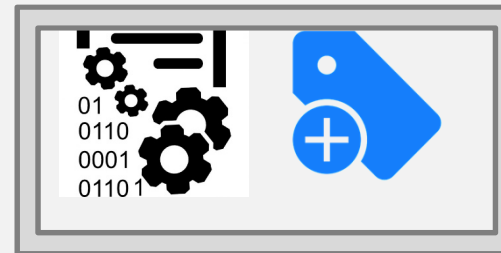
Storing and backing up
files while research is
active



Likely to be on a
networked filestore or
hard drive
Easy to change or
delete



Archiving or
preserving data in
the long-term



Likely to be
deposited in a digital
repository
safeguarded and
preserved

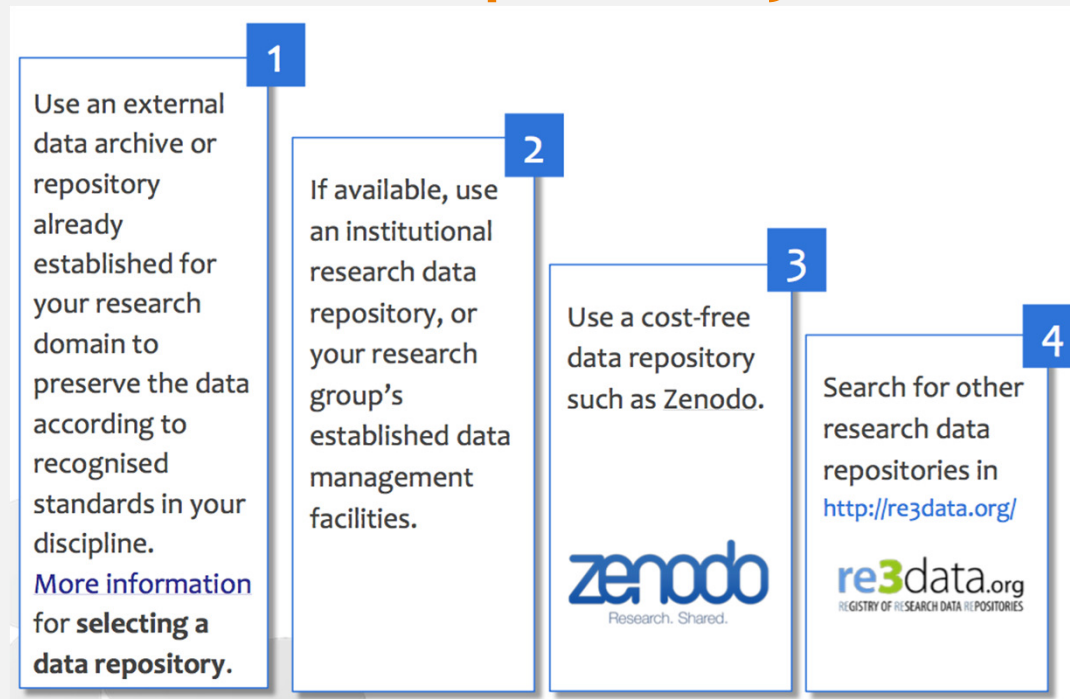
Archiving, repositories, ehm?

Select a data repository that will preserve your data, metadata and possibly tools in the long term.

It is advisable to **contact the repository of your choice when writing the first version of your DMP.**

Repositories may offer guidelines for sustainable data formats and metadata standards, as well as support for dealing with sensitive data and licensing.

Where to find a repository?



More information: <https://www.openaire.eu/opendatapilot-repository>

Zenodo: <http://www.zenodo.org>

Re3data.org: <http://www.re3data.org>

How to select a repository?

Main criteria for choosing a data repository:

Certification as a 'Trustworthy Digital Repository', with an explicit ambition to keep the data available in the long term.

- Three common certification standards for TDRs:



Data Seal of Approval: <http://datasealofapproval.org/en>

nestor seal: http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html

ISO 16363: <http://www.iso16363.org>

How to select a repository? (2)

- Matches your particular data needs: e.g. formats accepted; mixture of Open and Restricted Access.
- Provides guidance on how to cite the data that has been deposited.
- Gives your submitted dataset a persistent and globally unique identifier: for sustainable citations - both for data and publications - and to link back to particular researchers and grants.

(All) Research. Shared.

— your one stop research shop!

All research outputs from across all fields of science are welcome! Zenodo accept any file format as well as both positive and negative results. However, we do promote peer-reviewed openly accessible research, and we do curate your upload before putting it on the front-page.



Citeable. Discoverable.

— be found!

Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citeable. Zenodo further supports harvesting of all content via the OAI-PMH protocol.

www.zenodo.org

Community Collections

— create your own repository

Zenodo allows you to create your own collection and accept or reject all uploads to it. Creating a space for your next workshop or project have never been easier. Plus, everything is citeable and discoverable.

Safe

— more than just a drop box!

Your research output is stored safely for the future in same cloud infrastructure as research data from CERN's [Large Hadron Collider](#) using a CERN's battle-tested repository software [INVENIO](#) used by some of the world's largest repositories such as [INSPIRE HEP](#) and [CERN Document Server](#).

Reporting

— tell your funding agency!

Zenodo is integrated into reporting lines for research funded by the European Commission via [OpenAIRE](#). Just upload your research on Zenodo and we will take care of the reporting for you. We plan to extend with further funding agencies in the future so stay tuned!

Flexible Licensing

— not everything is under Creative Commons

Zenodo encourage you to share your research as openly as possible to maximize use and re-use of your research results. However, we also acknowledge that one size does not fit all, and therefore allow for uploading under a multitude of different licenses and access levels*.

* You are responsible for respecting applicable copyright and license conditions for the files you upload.

The screenshot shows the Zenodo website interface. At the top, there is a blue navigation bar with the Zenodo logo on the left, a search bar, and links for 'Upload' and 'Communities'. On the right side of the navigation bar, a user profile is displayed with the email 'pedroprincipe@sdum.uminho.pt'. Below the navigation bar, there is a secondary search bar labeled 'Search uploads...' and a green 'New Upload' button. The main content area features a large heading 'Get started!' followed by the text 'Make your first upload - all research outputs from across all fields of research are welcome.' and another green 'New Upload' button. At the top of this content area, there are tabs for 'Drafts 0' and 'Published 0', and sorting options for 'Most recent' and 'asc.'. The footer is a dark blue bar containing several columns of links: 'About' (with sub-links: About, Contact, Policies), 'Resources' (with sub-links: Features, FAQ), 'Developers' (with sub-links: REST API, OAI-PMH), 'Contribute' (with sub-links: GitHub, Donate), and 'Funded by' (with logos for CERN, OpenAIRE, and the European Union). The bottom-most bar contains the text 'Powered by CERN Data Centre & Invenio' on the left and 'Privacy policy Terms of Use Support' on the right.

zenodo

Search

Upload

Communities

pedroprincipe@s dum.uminho.pt

Delete

Save

Publish

New upload

Instructions: (i) Upload minimum one file or fill-in required fields (marked with a red star). (ii) Press "Save" to save your upload for editing later. (iii) When ready, press "Publish" to finalize and make your upload public.

Files

Choose files

Start upload

Drag and drop files here

— or —

Choose files

(minimum 1 file required, max [redacted] per dataset - contact us for larger datasets)

Upload type

required

The screenshot shows the Zenodo website interface. At the top, there is a browser window with the URL <https://zenodo.org>. The website header features the Zenodo logo, a search bar, and links for 'Upload' and 'Communities'. On the right side of the header, there are 'Log in' and 'Sign up' buttons.

The main content area is titled 'Recent uploads' and displays a list of three items:

- sjPlot - Data Visualization for Statistics in Social Science.**
 - Uploaded on October 15, 2018 (v2.6.1). Tags: Software, Open Access.
 - Author: Daniel Lüdecke.
 - Description: General Removed defunct functions. Deprecated `sjt.lm()`, `sjt.glm()`, `sjt.lmer()` and `sjt.glmer()` are now deprecated. Please use `tab_model()` instead. Changes to functions Arguments `dot.size` and `line.size` in `plot_model()` now also apply to marginal effects and diagnostic plots. `plot_model()` now uses a...
 - Uploaded on October 15, 2018.
 - 3 more version(s) exist for this record.
- OpenAIRE's DOIBoost - Boosting CrossRef for Research**
 - Uploaded on October 1, 2018 (v2.0). Tags: Preprint, Open Access.
 - Authors: La Bruzzo, Sandro; Manghi, Paolo; Mannocci, Andrea.
 - Description: Research in information science and scholarly communication strongly relies on the availability of openly accessible datasets of scholarly entities metadata and, where possible, their relative payloads. Since such metadata information is scattered across diverse, freely accessible, online...
 - Uploaded on October 11, 2018.
 - 2 more version(s) exist for this record.
- Transfer of bacteriophages between the fingers of volunteers and water or saliva**
 - Uploaded on October 10, 2018 (v1.0). Tags: Dataset, Open Access.
 - Authors: Pitol, Ana Karina; Bischel, HN; Kohn, T; Julian, Tim.
 - Description: Data sets on the percentage of virus transferred between fingers and water or saliva. Transfer with saliva is for both wet hands (where the virus inoculum was not allowed to appreciably dry before transfer) and for dry hands (where the virus inoculum was allowed to dry before transfer). The data...
 - Uploaded on October 10, 2018.

On the right side of the page, there are three informational boxes:

- Zenodo now supports usage statistics!** (with a line graph icon). Includes a link to 'Read more about it, in our newest blog post.'
- Using GitHub?** (with the GitHub logo icon). Includes a link to 'Just Log in with your GitHub account and click here to start preserving your repositories.'
- Zenodo in a nutshell** (with a list of features):
 - Research. Shared.** – all research outputs from across all fields of research are welcome! Sciences and Humanities, really!
 - Citeable. Discoverable.** – uploads gets a Digital Object Identifier (DOI) to make them easily and uniquely citeable.
 - Communities** – create and curate your own community for a workshop, project, department, journal, into which you can accept or reject uploads. Your own complete digital repository!
 - Funding** – identify grants, integrated in reporting lines for research funded by the European Commission via OpenAIRE.
 - Flexible licensing** – because not everything is under Creative Commons.
 - Safe** – your research output is stored safely for the future in the same cloud infrastructure as CERN's own LHC research data.

At the bottom of the right sidebar, there is a link: 'Read more about Zenodo and its features'.

Managing and Sharing Research D... Data Management Plan for a B... Transfer of bacteriophages between...

https://zenodo.org/record/1454505#.W8Y1jZ0Tcs

Transfer of bacteriophages between the fingers of volunteers and water or saliva

Pitol, Ana Karina; Bischel, HN; Kohn, T; Julian, Tim

Data sets on the percentage of virus transferred between fingers and water or saliva. Transfer with saliva is for both wet hands (where the virus inoculum was not allowed to appreciably dry before transfer) and for dry hands (where the virus inoculum was allowed to dry before transfer). The data were collected, analyzed, and reported within the following publication:

Pitol, A. K., Bischel, H. N., Kohn, T., & Julian, T. R. (2017). Virus Transfer at the Skin-Liquid Interface. *Environmental Science & Technology*, 51(24), 14417-14425. <https://doi.org/10.1021/acs.est.7b04949>

140 views 0 downloads

Indexed in OpenAIRE

Publication date: October 10, 2018

DOI: [DOI 10.25678/000099](https://doi.org/10.25678/000099)

Keyword(s): bacteriophage, fingers, liquid, saliva, skin, transfer, virus

Related identifiers:
 Cites: [10.1021/acs.est.7b04949](https://doi.org/10.1021/acs.est.7b04949)
 Documented by: [10.1021/acs.est.7b04949](https://doi.org/10.1021/acs.est.7b04949)
 Supplement to: [10.1021/acs.est.7b04949](https://doi.org/10.1021/acs.est.7b04949)
 Supplementary material: [10.1021/acs.est.7b04949](https://doi.org/10.1021/acs.est.7b04949)

Communities: Eawag, Swiss Federal Institute of Aquatic Science and Technology, Zenodo

License (for files): [Creative Commons Zero - CC0 1.0](https://creativecommons.org/licenses/by/4.0/)

pt_wet	pt_dry	phage
6.47E+01	1.56E+01	MS2
5.64E+01	1.59E+01	MS2
7.10E+01	1.65E+01	MS2
7.92E+01	2.03E+01	MS2
3.95E+01	1.41E+01	MS2
2.88E+01	2.90E+01	MS2
6.88E+01	2.29E+01	MS2
6.56E+01	1.67E+01	MS2
7.33E+01	2.97E+01	MS2
6.72E+01	3.18E+01	MS2

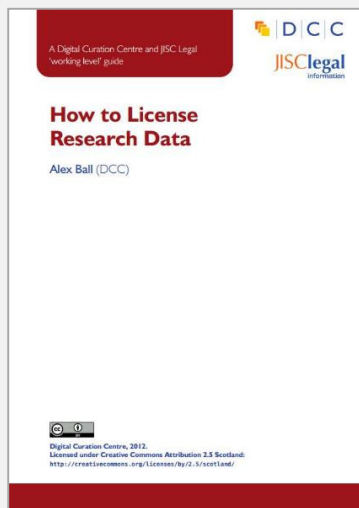
Files (8.7 kB)

Name	Size	Preview	Download
phagetransferhandtosaliva.csv	345 Bytes		
md5:a3ca4a960695603b7e680abf3fc87678			
phagetransferwatertohand.csv	7.0 kB		
md5:5b5f2b7ee28ead59c1965e35276080a2			
readme.md	1.3 kB		

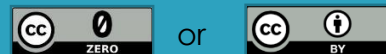
Share

Cite as

Licensing research data



Horizon 2020 Open Access guidelines point to:



This DCC guide outlines the pros and cons of each approach and gives practical advice on how to implement your licence

CREATIVE COMMONS LIMITATIONS



NC Non-Commercial
What counts as

commercial?



ND No Derivatives
Severely restricts use

These clauses are not open licenses

www.dcc.ac.uk/resources/how-guides/license-research-data

EUDAT licensing tool

Answer questions to determine which licence(s) are appropriate to use

Do you own copyright and similar rights in your dataset and all its constitutive parts?

Yes

No

Do you allow others to make commercial use of you data?

Yes

No

Creative Commons Attribution (CC-BY)

This is the standard creative commons license that gives others maximum freedom to do what they want with your work.

Public Domain Dedication (CC Zero)

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

<http://ufal.github.io/public-license-selector>

The screenshot shows a web browser window with the URL <https://riojournal.com/article/11624/>. The page features the RIO logo and navigation menus. The main content area displays the article title, author (Laurent Gatto), and sections for Abstract, Background, New information, Keywords, and Products of research. A right-hand sidebar contains a table of contents with links to Article metadata, Products of research (Software, Source code, Documentation, Reproducible framework, Data), Acknowledgements, and References. A 'Reviewable v1' badge is visible above the article content. The footer includes the text 'Powered by arpha'.

Data Management Plan for a Biotechnology and Biological Sciences Research Council (BBSRC) Tools and Resources Development Fund (TRDF) Grant

▼ Laurent Gatto

Abstract ▲

Background

This Data Management Plan (DMP) was created for Laurent Gatto's BBSRC Tools and Resources Development Fund award ([BB/N023129/1](#)).

New information

The DMP describes the management and sharing of all data and code associated with the grant, including software dissemination and release schedule, source code development and open source licensing, software documentation, reproducible framework and data annotation and dissemination.

Keywords ▲

Spatial proteomics, Bioconductor, machine learning, mass spectrometry, proteomics, software

Products of research ▲

The participants have a long history of successful collaboration and open source development and are fully committed to abiding by the BBSRC's policy on data management. Specific outputs of this

https://riojournal.com/article/11624

Powered by arpha

Managing and Sharing Research D... x Data Management Plan for a B... x +

https://riojournal.com/article/11624/

RIO Home Articles About About Pensoft Books E-Books Blog Journals Register | Login

Data Management Plan Research Ideas and Outcomes 3: e11624
https://doi.org/10.3897/rio.3.e11624 (05 Jan 2017) Reviewable v1

Data

While no new data will be generated specifically in the frame of this project, statistically sound (re-)analysis and reliable (re-)interpretation of published or private data will be produced. These data will be made available through multiple existing community resources using established standards and annotated with ample meta data. They will be distributed as dedicated R object (in well-established data structures defined in [MSnbase Gatto and Lilley 2011](#)), as used and manipulated through the [pRoloc \(Gatto et al. 2014\)](#) and [pRolocGUI](#) software and included in the open [pRolocdata \(Gatto et al. 2014\)](#) data package. All datasets will be thoroughly annotated with meta data to provide users with all necessary details on the origin or manipulation of the data in order to favour and facilitate re-use and reproducibility. Several exporters are already available, to offer these same data as spreadsheets or in the [mzTab \(Griss et al. 2014\)](#) format. When available, raw and identification data will be distributed using the [mzML](#) and [mzIdentML](#) Proteomics Standards Initiative (PSI) community formats and disseminated through the [ProteomeXchange \(PX\) project \(Vizcaíno et al. 2014\)](#) and the [Proteomics IDentifications \(PRIDE\)](#) resource. We will also distribute the data and results through the online resource [SpatialMap.org](#) that we are currently developing, which will enable users to interactively visualise, explore and search the data and annotated results stemming of our state-of-the-art statistical learning pipelines.

The refined and novel protein sub-cellular localisations will be communicated to the wider proteomics community via relevant protein databases and annotation providers like [Swiss-Prot](#), the [Gene Ontology](#) Annotation database as well as more specialised resources. The improved localisation information will be distributed with all technical details regarding the analysis and interpretation/evidence, including algorithm specifications and parameters and assignment probabilities.

Data will be made available as soon as it has been quality controlled and converted into usable computational objects. Once validated on various datasets, the algorithms will be included and distributed through the relevant software packages. The multiple sources and formats will be cross-referenced to maximise utility and availability to the research community.

Acknowledgements ▲

Contents Article info Citation Metrics Reviews Related

Refs Cited

Article metadata

Products of research

- Software
- Source code
- Documentation
- Reproducible framework
- Data

Acknowledgements

References

Powered by arpha

Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data

Top tip - keep it short and specific!

This very short extract from a presentation by Peter Dukes, Medical Research Council (MRC) back in 2012 provides really useful advice on writing a DMP from the funding body perspective. While it is an example from the Life Sciences, the advice applies to all disciplines. The quality of the video isn't great, unfortunately, but the advice provided definitely is!

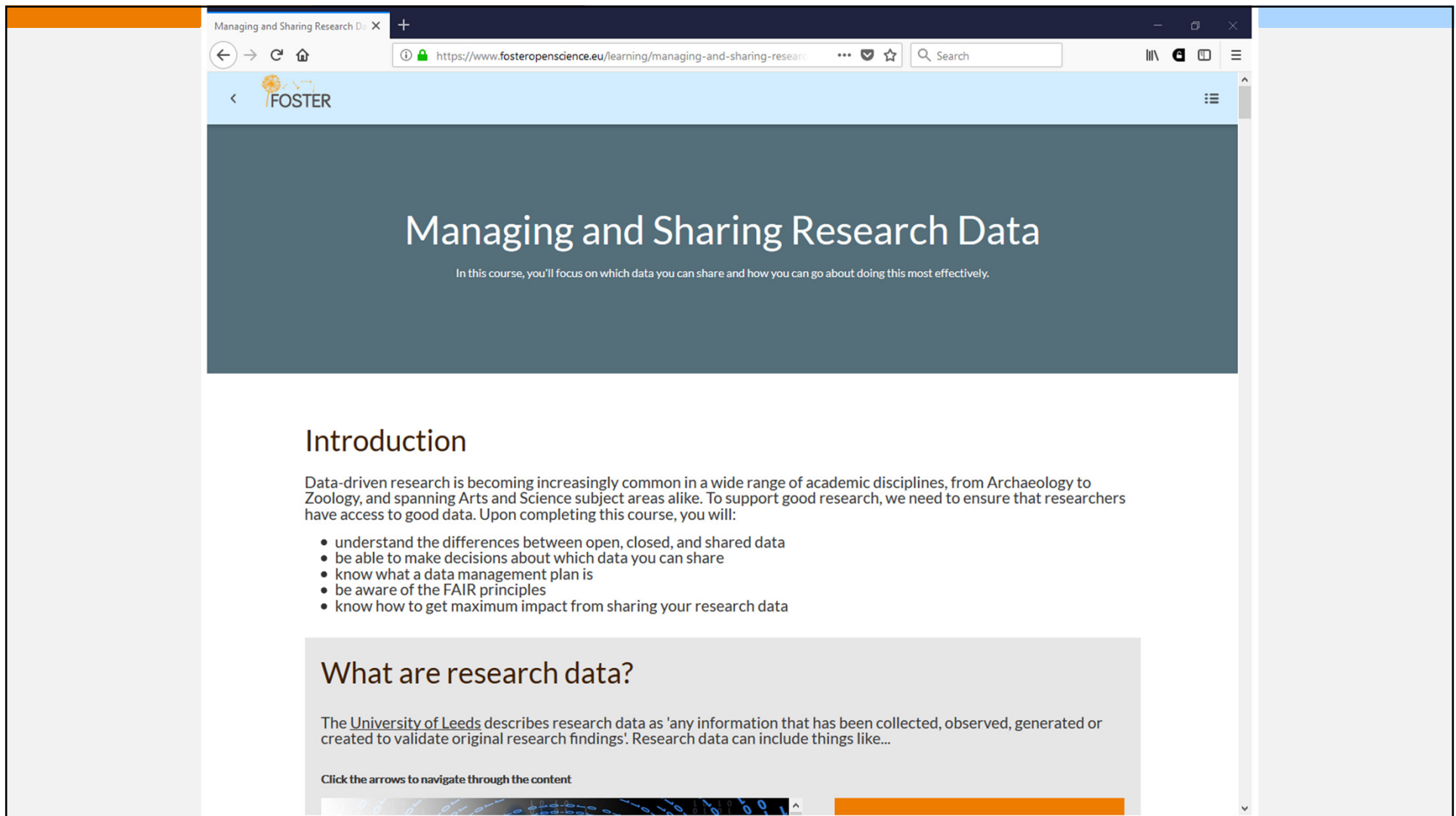
Advice on writing Data Management Plans

Research Data Improved Data Management Plans

4. Keep it simple

- Informative: two audiences
- Specific: e.g. name standards
- Concise: < 1/4 to 3 pages
- Don't forget: your achievements & innovation

<https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data>



The screenshot shows a web browser window with the following content:

- Browser Tab:** Managing and Sharing Research Data
- Address Bar:** <https://www.fosteropenscience.eu/learning/managing-and-sharing-research-data/>
- FOSTER Logo:** Located in the top left of the page header.
- Section Header:**

Managing and Sharing Research Data
- Subtext:** In this course, you'll focus on which data you can share and how you can go about doing this most effectively.
- Section Header:**

Introduction
- Text:** Data-driven research is becoming increasingly common in a wide range of academic disciplines, from Archaeology to Zoology, and spanning Arts and Science subject areas alike. To support good research, we need to ensure that researchers have access to good data. Upon completing this course, you will:
- List-Group:**
 - understand the differences between open, closed, and shared data
 - be able to make decisions about which data you can share
 - know what a data management plan is
 - be aware of the FAIR principles
 - know how to get maximum impact from sharing your research data
- Section Header:**

What are research data?
- Text:** The [University of Leeds](#) describes research data as 'any information that has been collected, observed, generated or created to validate original research findings'. Research data can include things like...
- Text:** Click the arrows to navigate through the content

The screenshot shows a web browser window with the following details:

- Browser tab: Managing and Sharing Research Data
- Address bar: <https://www.fosteropenscience.eu/learning/managing-and-sha>
- Page title: FOSTER

The main content area contains a quiz question:

Let's see how much you've learned in this course.

If I choose to share my data, then they must be available to anyone who want them.

Please select one of the options below.

True

False

Submit Show feedback

Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sha

FOSTER

Submit Show feedback

As the researcher who led the project, I decide what data can be shared.

Please select one of the options below.

True

False

Submit Show feedback

Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sha

FOSTER

Submit Show feedback

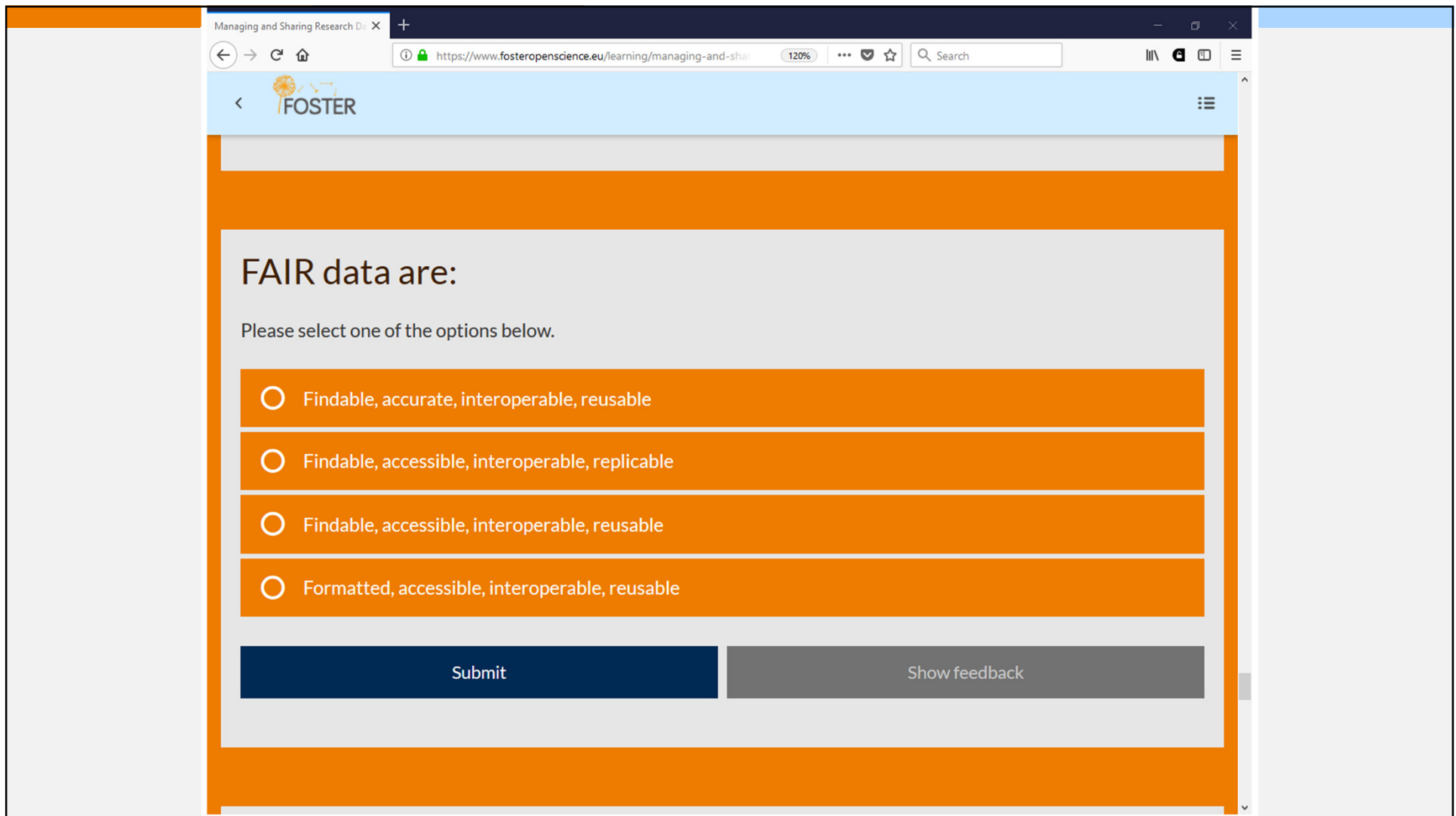
A data management plan that is submitted with a grant application cannot be changed.

Please select one of the options below.

True

False

Submit Show feedback



Managing and Sharing Research Data

https://www.fosteropenscience.eu/learning/managing-and-sha

120%

Search

FOSTER

FAIR data are:

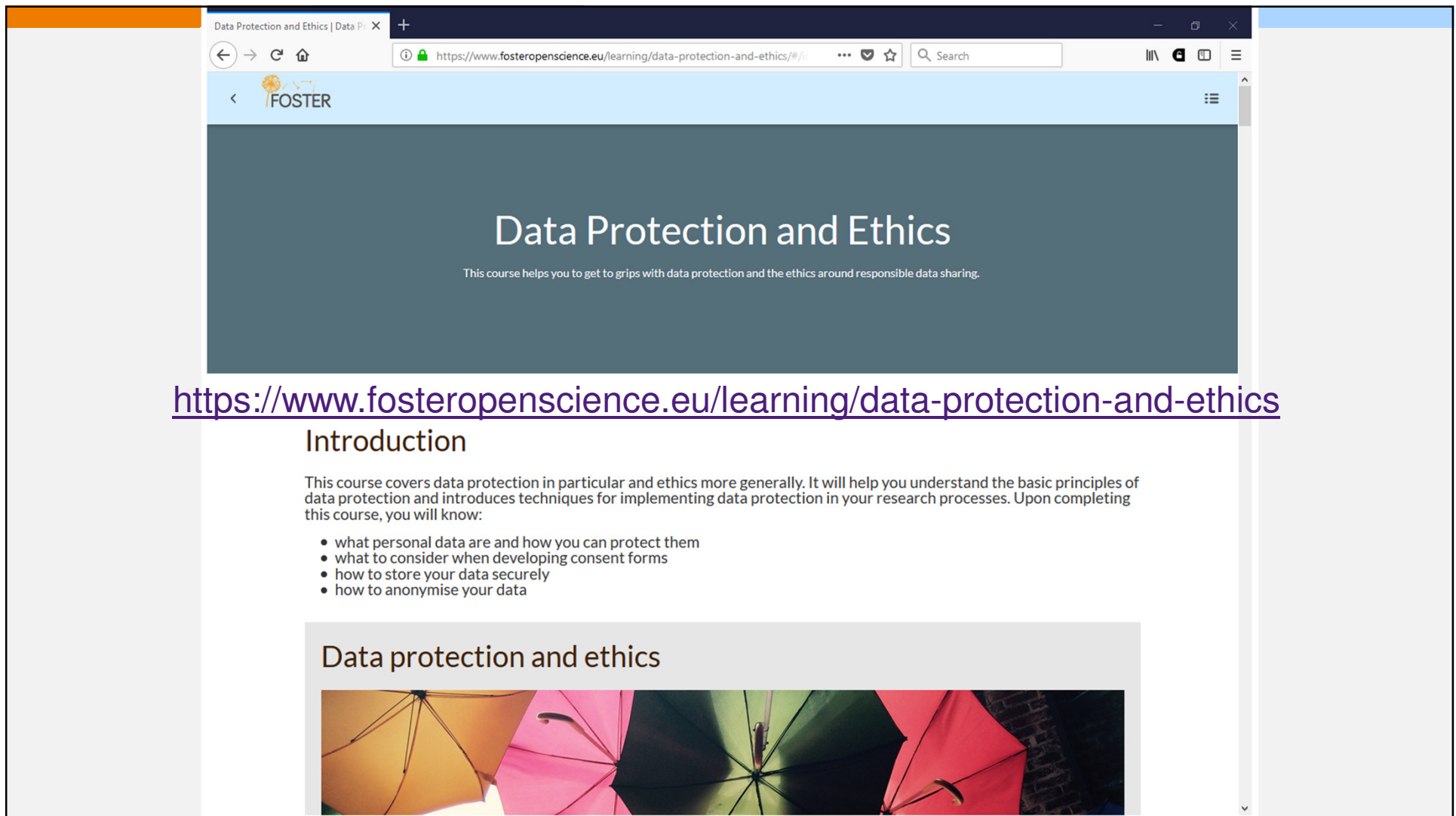
Please select one of the options below.

- Findable, accurate, interoperable, reusable
- Findable, accessible, interoperable, replicable
- Findable, accessible, interoperable, reusable
- Formatted, accessible, interoperable, reusable

Submit Show feedback

The screenshot shows a web browser window with the following elements:

- Browser Tab:** "Managing and Sharing Research D..."
- Address Bar:** "https://www.fosteropenscience.eu/learning/managing-and-sha..."
- Page Header:** "FOSTER" logo and a hamburger menu icon.
- Main Content:**
 - Section Header:** "To get the most out of the data you share, you should make it available..."
 - Text:** "Please select one of the options below."
 - Radio Buttons:**
 - from my personal web page
 - from a discipline specific repository
 - Buttons:** "Submit" (dark blue) and "Show feedback" (grey).



The screenshot shows a web browser window with the following elements:

- Browser tab: Data Protection and Ethics | Data P...
- Address bar: <https://www.fosteropenscience.eu/learning/data-protection-and-ethics/#/>
- FOSTER logo in the top left corner.
- Section title:

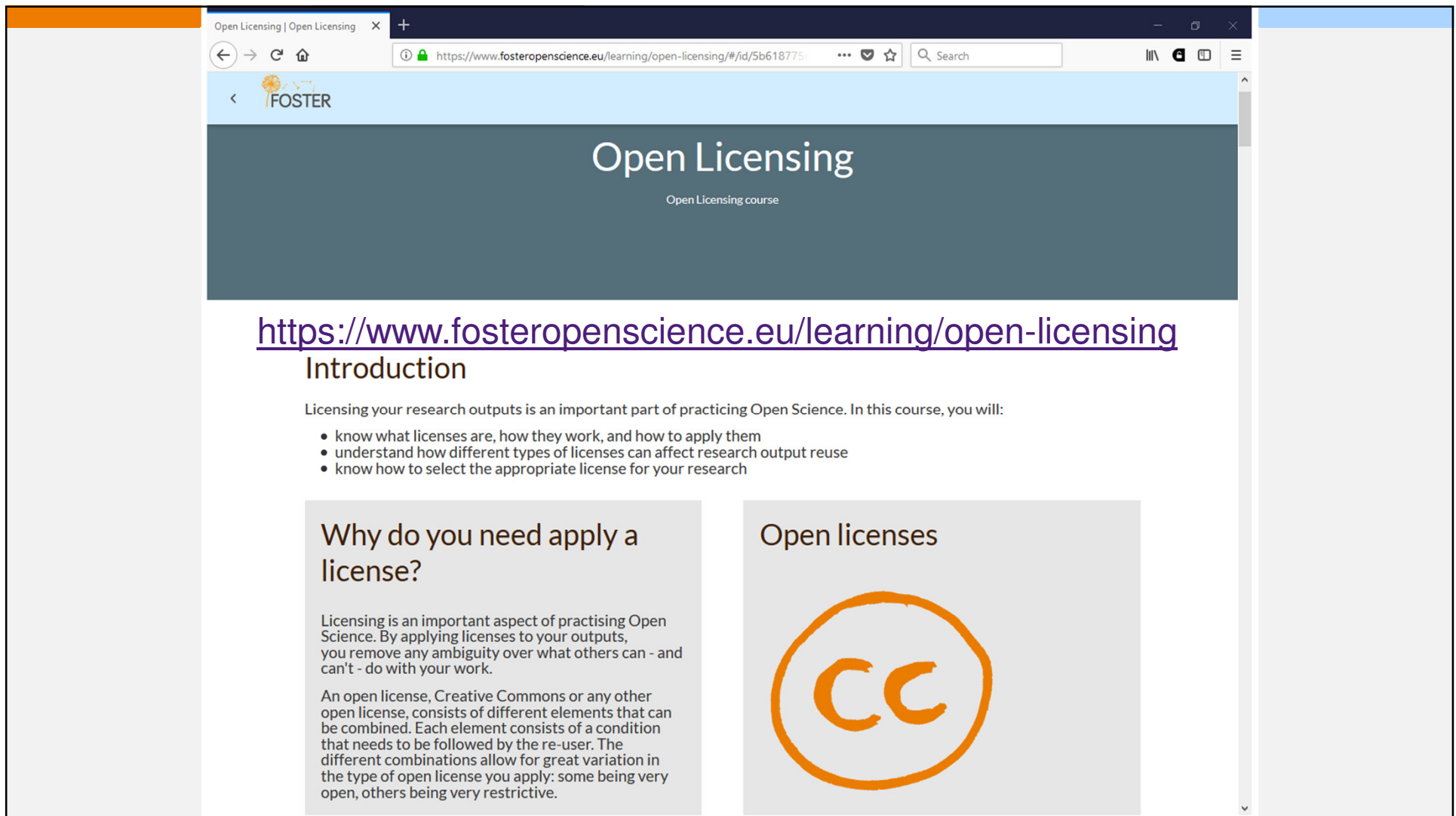
Data Protection and Ethics
- Subtext:

This course helps you to get to grips with data protection and the ethics around responsible data sharing.
- Section title:

Introduction
- Text:

This course covers data protection in particular and ethics more generally. It will help you understand the basic principles of data protection and introduces techniques for implementing data protection in your research processes. Upon completing this course, you will know:
- List of topics:
 - what personal data are and how you can protect them
 - what to consider when developing consent forms
 - how to store your data securely
 - how to anonymise your data
- Section title:

Data protection and ethics
- Image: A photograph of several colorful umbrellas (orange, pink, green, red) against a dark background.



Open Licensing | Open Licensing x +

https://www.fosteropenscience.eu/learning/open-licensing/#/id/5b618775

FOSTER

Open Licensing

Open Licensing course

<https://www.fosteropenscience.eu/learning/open-licensing>

Introduction

Licensing your research outputs is an important part of practicing Open Science. In this course, you will:


- know what licenses are, how they work, and how to apply them
- understand how different types of licenses can affect research output reuse
- know how to select the appropriate license for your research

Why do you need apply a license?

Licensing is an important aspect of practising Open Science. By applying licenses to your outputs, you remove any ambiguity over what others can - and can't - do with your work.

An open license, Creative Commons or any other open license, consists of different elements that can be combined. Each element consists of a condition that needs to be followed by the re-user. The different combinations allow for great variation in the type of open license you apply: some being very open, others being very restrictive.

Open licenses



Guidelines on DMPs

How to develop a DMP www.dcc.ac.uk/resources/how-guides/develop-data-plan

RDM brochure and template

https://dans.knaw.nl/en/about/organisation-and-policy/information-material?set_language=en

OpenAIRE RDM Handbook <https://www.openaire.eu/rdm-handbook>

ICPSR framework for a DMP

www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html

Other resources

Where to keep research data <http://www.dcc.ac.uk/resources/how-guides-checklists/where-keep-research-data/where-keep-research-data>

Five steps to decide what data to keep

<http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep>

Re3data <http://www.re3data.org/>

Figshare <https://figshare.com/>

Genbank <https://www.ncbi.nlm.nih.gov/genbank/>

How to write a lay summary <http://www.dcc.ac.uk/resources/how-guides/write-lay-summary>

Lay summaries <https://www.bhf.org.uk/research/information-for-researchers/how-to-apply/lay-summaries>

With thanks to

Joy Davidson, University of Glasgow

Bianca Kramer & Jeroen Bosman, Utrecht
University Library

Marjan Grootveld, DANS

Sarah Jones, DCC

Acknowledgements:

Jonathan Rans, DCC

Thanks to DANS and DCC for reuse of slide



Thank you! Questions?

Facebook: @fosteropenscience

Twitter: @fosterscience

Youtube: [FOSTER Open Science](#)





Why should data be open? Benefits of Open Access to research data

Iryna Kuchma, EIFL Open Access Programme Manager, @irynakuchma
Health NCP Net 2.0 Training Open Access & Data Management in Horizon 2020, EKT
19th October 2018



WHY SHOULD YOU BE OPEN?



Image by wonderwebby CC-BY-NC-SA www.flickr.com/photos/wonderwebby/2723279491



PUBLICATIONS AND DATA

It's part of good research practice

"It was **never** acceptable to publish papers without making data available."

- Ewan Birney

#OpenData
#OpenScience



Original image via doi:10.1038/461145a. "Research cannot flourish if data are not preserved and made accessible. Data management should be woven into every course in science." - *Nature* 461, 145

Science as an open enterprise

“Much of the remarkable growth of scientific understanding in recent centuries is due to open practices; open communication and deliberation sit at the heart of scientific practice.”

Royal Society report calls for ‘intelligent openness’ whereby data are accessible, intelligible, assessable and usable.



<https://royalsociety.org/policy/projects/science-public-enterprise/Report>

Cut down on academic fraud

The screenshot shows a news article on the Nature website. The article title is "Report finds massive fraud at Dutch universities". The sub-headline is "Investigation claims dozens of social-psychology papers contain faked data." The author is Diederik Stapel, a Dutch psychologist from Tilburg University. The article discusses an investigation into social-psychology papers, claiming that many contain fabricated data. It mentions that the investigation was led by a committee and that the findings were reported in a recent Science paper. The article also includes a photo of Diederik Stapel and a caption identifying him as a Dutch psychologist from Tilburg University.

nature International weekly journal of science

nature news home | news archive | specials | opinion | features | news blog | nature journal

Published online 1 November 2011 | *Nature* **479**, 15 (2011) | doi:10.1038/479015a
Updated online: 1 November 2011
Updated online: 8 December 2011

Report finds massive fraud at Dutch universities

Investigation claims dozens of social-psychology papers contain faked data.

Even Callaway

When colleagues called the work of Dutch psychologist Diederik Stapel too good to be true, they meant it as a compliment. But a preliminary investigative report (go.nature.com/tmp5c) released on 31 October gives literal meaning to the phrase, detailing years of data manipulation and blatant fabrication by the prominent Tilburg University researcher.

"We have some 30 papers in peer-reviewed journals where we are actually sure that they are fake, and there are more to come," says Pim Levelt, chair of the committee that investigated Stapel's work at the university.

Stapel's eye-catching studies on aspects of social behaviour such as power and stereotyping garnered wide press coverage. For example, in a recent *Science* paper (which the investigation has not identified as fraudulent), Stapel reported that untidy environments encouraged discrimination ([Science 332, 251-253; 2011](http://Science.332.251-253.2011)).

Dutch psychologist Diederik Stapel.
Persbureau van Eindhoven

Related stories

- Seven days: 9-15 September 2011
14 September 2011
- Chaos promotes stereotyping
07 April 2011

Naturejobs

Tenure-Track Faculty Positions (Assistant / Associate / Full Professor) Yale University, Department of Genetics
Yale University School of Medicine

Assistant Professor
Harvard Medical School

- More science jobs
- Post a job for free

Resources

- PDF Format
- Send to a Friend
- Reprints & Permissions
- RSS Feeds

external links

- Tilburg University
- Interim investigation report



www.nature.com/news/2011/111101/full/479015a.html

Validation of results

“It was a mistake in a spreadsheet that could have been easily overlooked: a few rows left out of an equation to average the values in a column.

The spreadsheet was used to draw the conclusion of an influential 2010 economics paper: that public debt of more than 90% of GDP slows down growth. This conclusion was later cited by the International Monetary Fund and the UK Treasury to justify programmes of austerity that have arguably led to riots, poverty and lost jobs.”



The error that could subvert George Osborne's austerity programme

The theories on which the chancellor based his cuts policies have been shown to be based on an embarrassing mistake

Charles Arthur and Phillip Inman
The Guardian, Thursday 18 April 2013 21.10 BST



George Osborne says that Ken Rogoff, the man whose economic error has been uncovered, has strongly influenced his thinking. Photograph: Stefan Wermuth/PA

www.guardian.co.uk/politics/2013/apr/18/uncovered-error-george-osborne-austerity

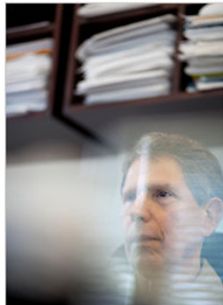
More scientific breakthroughs

Sharing of Data Leads to Progress on Alzheimer's

By GINA KOLATA
Published: August 12, 2010

In 2003, a group of scientists and executives from the [National Institutes of Health](#), the [Food and Drug Administration](#), the drug and medical-imaging industries, universities and nonprofit groups joined in a project that experts say had no precedent: a collaborative effort to find the biological markers that show the progression of [Alzheimer's disease](#) in the human brain.

[Enlarge This Image](#)



Now, the effort is bearing fruit with a wealth of recent scientific papers on the early diagnosis of Alzheimer's using methods like PET scans and tests of spinal fluid. More than 100 studies are under way to test drugs that might slow or stop the disease.

And the collaboration is already serving as a model for similar efforts against [Parkinson's disease](#). A \$40 million project to look for biomarkers for Parkinson's, sponsored by the [Michael J. Fox Foundation](#), plans to enroll 600 study subjects in the United States and Europe.

“It was unbelievable. Its not science the way most of us have practiced in our careers. But we all realised that we would never get biomarkers unless all of us parked our egos and intellectual property noses outside the door and agreed that all of our data would be public immediately.”

Dr John Trojanowski, University of Pennsylvania

www.nytimes.com/2010/08/13/health/research/13alzheimer.html?pagewanted=all&r=0



A citation advantage

A study that analysed the citation counts of 10,555 papers on gene expression studies that created microarray data, showed:

“studies that made data available in a public repository received 9% more citations than similar studies for which the data was not made available”



Data reuse and the open data citation advantage,
Piwowar, H. & Vision, T. <https://peerj.com/articles/175>



Increased use and economic benefit

The case of NASA Landsat satellite imagery of the Earth's surface:

Up to 2008

- Sold through the US Geological Survey for US\$600 per scene
- Sales of 19,000 scenes per year
- Annual revenue of \$11.4 million



Since 2009

- Freely available over the internet
- Google Earth now uses the images
- Transmission of 2,100,000 scenes per year.
- Estimated to have created value for the environmental management industry of \$935 million, with direct benefit of more than \$100 million per year to the US economy
- Has stimulated the development of applications from a large number of companies worldwide

<http://earthobservatory.nasa.gov/IOTD/view.php?id=83394&src=ve>

BE PART OF THE NEW ERA OF OPEN SCIENCE



reach more
people,
have greater
impact



avoid
duplication
of efforts



preserve data
for future
researchers



simplify final
Horizon 2020
reporting
thanks to an
up-to-date DMP

2011 Germany *E. coli* O104:H4 outbreak

From Wikipedia, the free encyclopedia



This article needs to be **updated**. Please update this article to reflect recent events or newly available information. *(July 2011)*

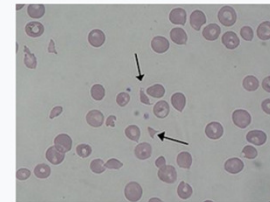
A novel strain of ***Escherichia coli* O104:H4 bacteria** caused a serious outbreak of **foodborne illness** focused in northern Germany in May through June 2011. The illness was characterized by bloody diarrhea, with a high frequency of serious complications, including **hemolytic-uremic syndrome** (HUS), a condition that requires urgent treatment. The outbreak was originally thought to have been caused by an **enterohemorrhagic (EHEC)** strain of *E. coli*, but it was later shown to have been caused by an **enteroaggregative *E. coli* (EAEC)** strain that had acquired the genes to produce **Shiga toxins**, present in **organic fenugreek sprouts**.

Epidemiological fieldwork suggested fresh vegetables were the source of infection. The agriculture minister of **Lower Saxony** identified an **organic farm**^[1] in **Bienenbüttel**, Lower Saxony, Germany, which produces a variety of **sprouted foods**, as the likely source of the *E. coli* outbreak.^[2] The farm was shut down.^[2] Although laboratories in Lower Saxony did not detect the bacterium in produce, a laboratory in **North Rhine-Westphalia** later found the outbreak strain in a discarded package of sprouts from the suspect farm.^[3] A control investigation confirmed the farm as the source of the outbreak.^[4] On 30 June 2011, the German *Bundesinstitut für Risikobewertung (BfR)* (Federal Institute for Risk Assessment), an institute of the German **Federal Ministry of Food, Agriculture and Consumer Protection**, announced that seeds of organic^[5] **fenugreek** imported from **Egypt** were likely the source of the outbreak.^[6]

In all, 3,950 people were affected and 53 died, 51 of whom were in Germany.^[7] 800 people suffered hemolytic uremic syndrome (HUS), which can lead to kidney failure.^[8] A handful of cases were reported in several other countries including **Switzerland**,^[9] **Poland**,^[9] the **Netherlands**,^[9] **Sweden**,^[9] **Denmark**,^[9] the **UK**,^[9]^[10] **Canada** and the **USA**.^[11] Essentially all affected people had been in Germany or France shortly before becoming ill.

Initially, German officials made incorrect statements on the likely origin and strain of *Escherichia coli*.^[12]^[13]^[14]^[15] The German health authorities, without results of ongoing tests, incorrectly linked the O104 serotype to **cucumbers** imported from Spain.^[16] Later, they recognised that Spanish greenhouses were not the source of the *E. coli* and cucumber samples did not contain the specific *E. coli* variant causing the outbreak.^[17]^[18] Spanish consumers expressed anger about having its produce linked with the deadly *E. coli* outbreak, which cost Spanish exporters US\$200 million per week.^[19] Russia banned the import of all fresh vegetables from the European Union from early June 2011.^[20]

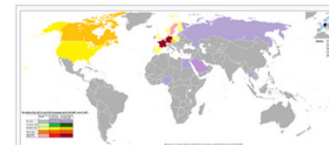
2011 German *E. coli* O104:H4 outbreak



Schistocytes seen in a person with hemolytic-uremic syndrome

Classification and external resources

[\[edit on Wikidata\]](#)



The 2011 *E. coli* O104:H4 outbreak.

LEGEND	No food restrictions or tests.	Food and/or medical tests done.	Food and/or travel restrictions.
No cases.	Grey square	Light purple square	Dark purple square
Suspected cases	Light green square	Green square	Dark green square
Known non-native cases	Yellow square	Orange square	Dark orange square
Known native cases	Orange square	Dark orange square	Dark orange square

https://en.wikipedia.org/wiki/2011_Germany_E._coli_O104:H4_outbreak

2011 Germany *E. coli* O104:H4 outbreak

The strain was analyzed by scientists at BGI-Shenzen in China working together with those in Hamburg, and **3 days later a draft genome was released under an open data license.**

This kick-started analysis by bioinformatic groups on **4 continents.**



2011 Germany *E. coli* O104:H4 outbreak (2)

24 hours after the release of genome it was assembled.

Within a week **two dozen reports have been filed on an open site.**



2011 Germany *E. coli* O104:H4 outbreak (3)

They produced results in time to help contain the outbreak and by July 2011 scientists published papers based on the analysis.

2011 Germany *E. coli* O104:H4 outbreak (4)

By opening up their early sequencing results to international collaboration, researchers in Hamburg produced results that were quickly tested by a wide range of experts, used to produce new knowledge and ultimately to control a public health emergency.

<https://www.thesgc.org>

The screenshot shows the homepage of the Structural Genomics Consortium (SGC). The browser address bar displays <https://www.thesgc.org>. The page features the SGC logo, a search bar, and a navigation menu with links for News, Careers, Publications, Open Lab Notebooks, Contact, and Log in. Below the navigation menu are dropdown menus for About, Science, Reagents & Resources, People, and News & Outreach. The main content area is dominated by a large banner for the 'ALS Reproducible Antibody Platform', which includes the text: 'Open Science to enable consistent data and accelerate reliable discoveries in Amyotrophic Lateral Sclerosis'. The banner features a 3D molecular model of a protein complex and a yellow checkmark icon. Below the banner are four columns of content: 'Latest Structures' with two entries (MLL3 and HUWE1), 'New Publications' with one entry (The dual methyltransferase METTL13), 'Chemical Probes' with three entries (VinSpinIn, TP-238, and DDR-TRK-1), and 'News from SGC' with two entries (Open Science Probe Project and Conference announcement). The footer of the page contains the SGC logo and the text 'Open Science and Innovation | Open Access'.

Open Science and Innovation | Open Access

SGC | Structural Genomics Consortium

<https://www.thesgc.org>

SGC

UNIVERSITY OF OXFORD UNICAMP GOETHE UNIVERSITÄT FRANKFURT AM MAIN

News Careers Publications Open Lab Notebooks Contact Log in

About Science Reagents & Resources People News & Outreach

ALS Reproducible Antibody Platform

Open Science to enable consistent data and accelerate reliable discoveries in Amyotrophic Lateral Sclerosis

Background image: Hermann Aberle, University of Munster

Latest Structures

MLL3: Human lysine methyltransferase 2C, 6th PHD domain, in complex with H4 peptide

HUWE1: Human HECT, UBA and WWE domain containing 1, E3

New Publications

The dual methyltransferase METTL13 targets N terminus and Lys55 of eEF1A and modulates codon-specific translation rates.
Nat Commun

Chemical Probes

VinSpinIn - A chemical probe for the Spin family proteins
12th September 2018

TP-238 - A chemical probe for CECR2/BPTF bromodomains
5th September 2018

DDR-TRK-1 - A chemical probe

News from SGC

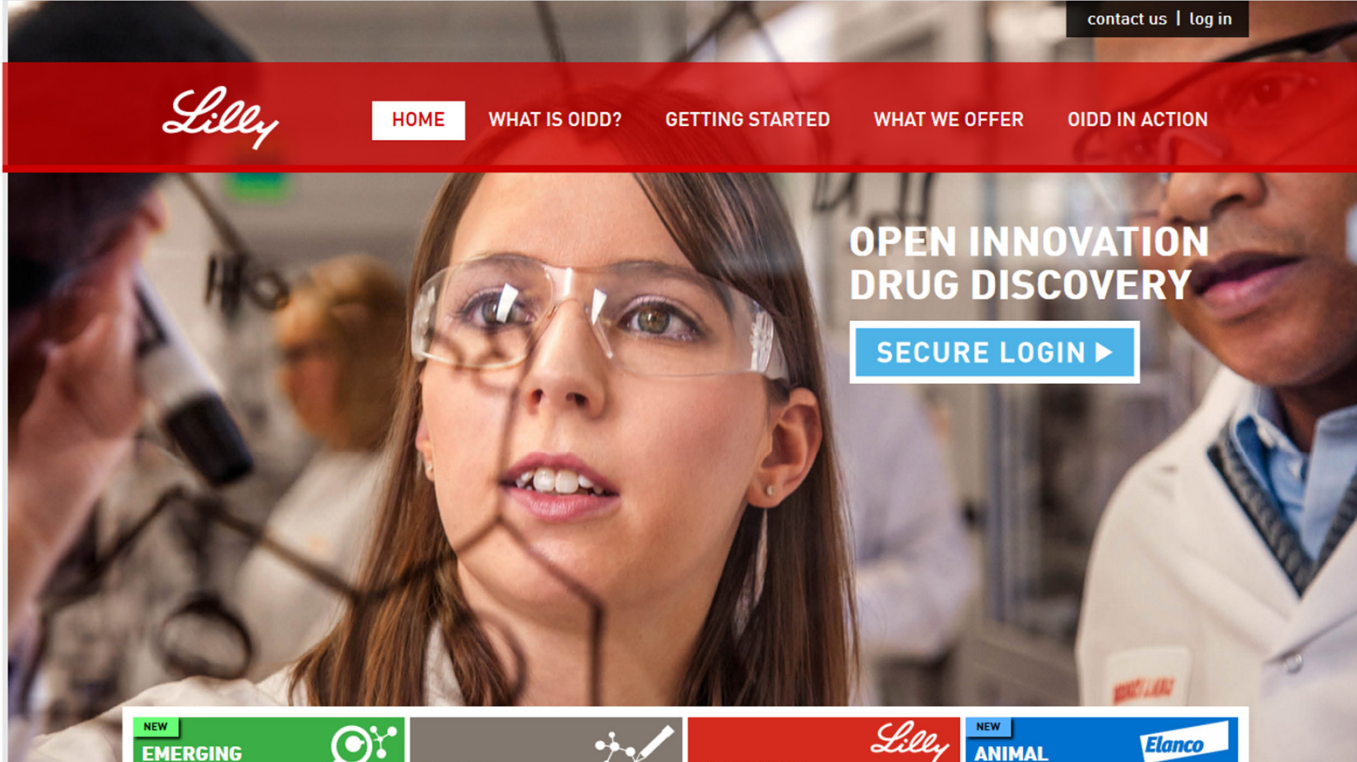
Open Science Probe Project launched
Posted on 9th May 2018

Conference announcement: BMP Signalling in Cancer II
Posted on 16th March 2018

[contact us](#) | [log in](#)



- [HOME](#)
- [WHAT IS OIDD?](#)
- [GETTING STARTED](#)
- [WHAT WE OFFER](#)
- [OIDD IN ACTION](#)



OPEN INNOVATION DRUG DISCOVERY

[SECURE LOGIN ▶](#)

<p>NEW</p> <p>EMERGING BIOLOGY</p> <p>Access to Lilly privileged compound libraries</p>	<p>DESIGN</p> <p>Cheminformatics tools for designing molecules</p>	<p>SCREENING</p> <p>Biological screening in strategic areas of interest</p>	<p>NEW</p> <p>ANIMAL HEALTH</p> <p>Biological screening in strategic areas of interest</p>
<p>COMPOUND ACQUISITION</p> <p>Diversification of Lilly's compound collection</p>	<p>SYNTHESIS</p> <p>Opportunity to access Automated Synthesis Lab</p>	<p>NEGLECTED & TROPICAL DISEASES</p> <p>Work with non-profit public-private partnerships</p>	

<https://openinnovation.lilly.com/dd>



Latest news

★ 26 July 2018
The European Lead Factory contributing to the fight against Alzheimer's
The key goal of ELF is to create validated hit series. Another successful ELF programme aimed at identifying small...

★ 10 July 2018
High number of SMEs benefit greatly from the European Lead Factory
The European Lead Factory is remarkably attractive for SMEs. The consortium consists of 30 partners, of which 10 are...

★ 07 June 2018
European Lead Factory reaches milestone of 200,000 novel compounds
Within a period of five years, the European Lead Factory (ELF) has reached one of its goals by completing the Public...

Interested?

Contact the programme office to know more how you can interact with the European Lead Factory

[Contact us!](#)

Publications

07 June 2018
Diastereoselective Synthesis of Highly Substituted, Amino-and Pyrrolidino-Tetrahydrofurans as Lead-Like Molecular Scaffolds

09 May 2018
Translation of innovative chemistry into screening libraries: an exemplar

European Lead Factory



Hitting the targets
The video report from the European Lead Factory (ELF) Stakeholder Meeting, 25-26 April, 2017 with an exciting agenda focussing on four years of successful collaboration, the impact of the European Lead Factory and the future of drug discovery.

Events

22 October 2018
IMI 10th anniversary Scientific Symposium

02 October 2018
5th Annual Drug Discovery USA Congress

<https://www.europeanleadfactory.eu>



Strengthening regional partnerships to respond to the needs of neglected patients

The DNDi Africa regional office currently works in five countries on multiple R&D projects, as well as capacity building initiatives within the Leishmaniasis East Africa Platform. Read about the office's R&D activities for neglected patients in Eastern Africa.

Neglected Patients

Neglected diseases continue to cause significant morbidity and mortality in the developing world. Yet, of the 850 new therapeutic products approved between 2000 and 2011, only 4% (and only 1% of all approved NCEs) were indicated for neglected diseases, even though these diseases account for 11% of the global disease burden.



CONTACT US



Home | About the Pathogen Box | Team | Request the Pathogen Box

What is the Pathogen Box?

The Pathogen Box contains ~400 diverse, drug-like molecules active against neglected diseases of interest and is available free of charge.

[Request the Pathogen Box](#)



<https://www.pathogenbox.org>

Updates

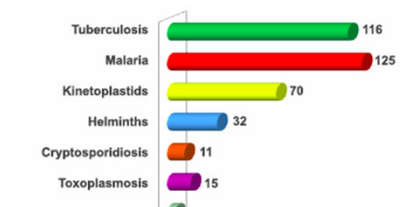
DMPK data for the Pathogen Box now available
09 May 2017

MMV awards seven endemic-region scientists Challenge Grants to 'exploit the Pathogen Box'
25 July 2016

Pathogen Box Orders



Composition





OPEN SOURCE MALARIA

Looking for New Medicines

Also Check Out



Lab Notebook



Project Wiki



Molecule Database
Chrome Only



Catch us on the Daily Show?

If you're here to see more about the **Breaking Good** project, as featured on the Daily Show

[Click here!](#)

The Open Source Malaria project is trying a different approach to curing malaria. Guided by open source principles, everything is open and anyone can contribute.

[Read More](#)

<http://openourcemalaria.org>

Recent Contributors

MFernflower	25
holeung	23
mattodd	21
drc007	19
cdsouthan	8



Activity



Join the Team



Meet Us

[View All](#) [Administration](#) [Collaboration Request](#) [Non-Science](#) [Series 3](#) [FYI](#) [Series 4](#) [Cheminformatics](#) [PR](#) [Writing](#)

Some talking points for funding.

[Administration](#) [Collaboration Request](#) [Non-Science](#)

25 Jul 2018 at 00:36 [8 Comments](#)

Following my recent [project](https://protect-au.mimecast.com/s/2ga3CwVLQmiYKv1zh9eOmp?domain=github.com), I wanted to raise some money to increase opportunities within OSM. I had the thought of starting up a charity and spent a little time drafting a [website](https://www.fundscience.co.uk/). Altho

[Read More](#)

Open Science and Innovation | Open Access | CO-ADD Community for Open Antimicrobial Drug Discovery

www.co-add.org

CO-ADD Community for Open Antimicrobial Drug Discovery

LOGIN DONATE

HOME PARTICIPATE ANTIMICROBIAL SCREENING DRUG-RESISTANT INFECTIONS NEWS & EVENTS ABOUT CO-ADD CONTACT

Open-access antimicrobial screening program

Helping Chemists Discover New Antibiotics to Fight Drug-Resistant Infections

Do you have the next antibiotic?

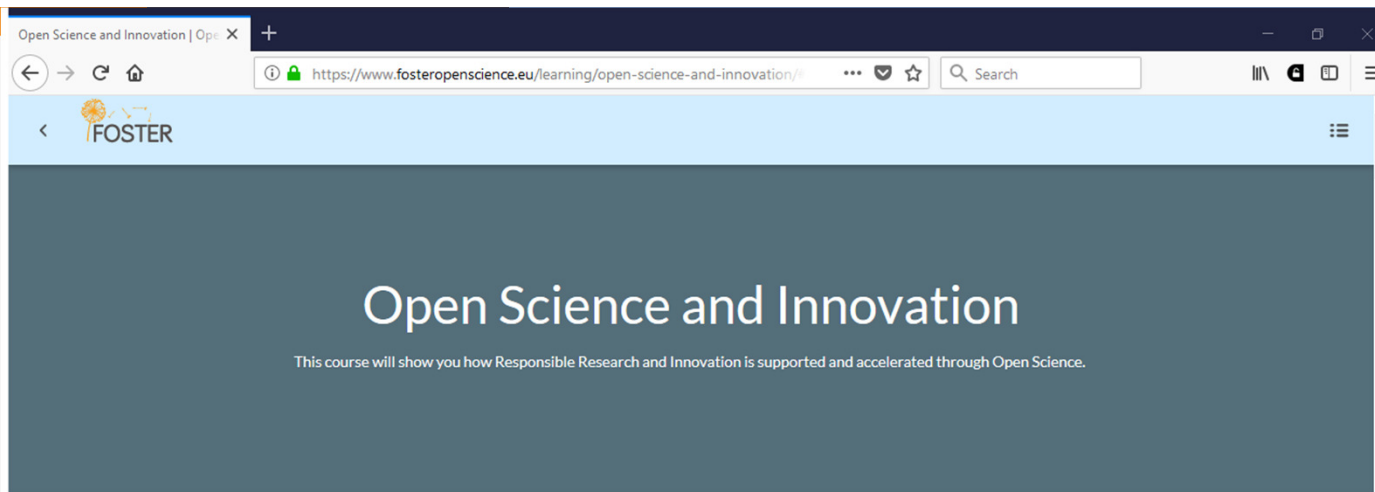
CO-ADD is a **not-for-profit** initiative led by academics at The University of Queensland.
Our goal is to **screen compounds for antimicrobial activity** for academic research groups for **free**.
We aim to help researchers worldwide to find new, diverse compounds to combat **drug-resistant infections**.

<http://www.co-add.org>

SEND COMPOUNDS

Screening for antimicrobial activity

5 Bacteria 2 Fungi



Open Science and Innovation | Opri X +

https://www.fosteropenscience.eu/learning/open-science-and-innovation/

FOSTER

Open Science and Innovation

This course will show you how Responsible Research and Innovation is supported and accelerated through Open Science.

Introduction

This course helps you to understand open business models and responsible research and innovation (RRI) and illustrates how these can foster innovation. By the end of the course, you will:

- understand key concepts and values of open business models and and responsible research and innovation
- know how to plan your innovation activities
- be able to use Creative Commons licenses in business
- understand new technology transfer policies with the ethos of Open Science
- learn how to get things to market faster

Open innovation

Open innovation is part of a broad family of concepts that often share the word “open” and the concept of “openness” and is based on the same basic ideas: by collaborating with others, by re-using (and by being allowed to re-use) the results of others and by allowing others to use and improve the results of our efforts, we all get better. See more in [Thoughts on Open Innovation: Essays on Open Innovation from leading thinkers in the field.](#)

<https://www.fosteropenscience.eu/learning/open-science-and-innovation>



With thanks to

Sarah Jones, DCC

Acknowledgements:

Jonathan Rans, DCC



Thank you! Questions?

Facebook: @fosteropenscience

Twitter: @fosterscience

Youtube: [FOSTER Open Science](#)



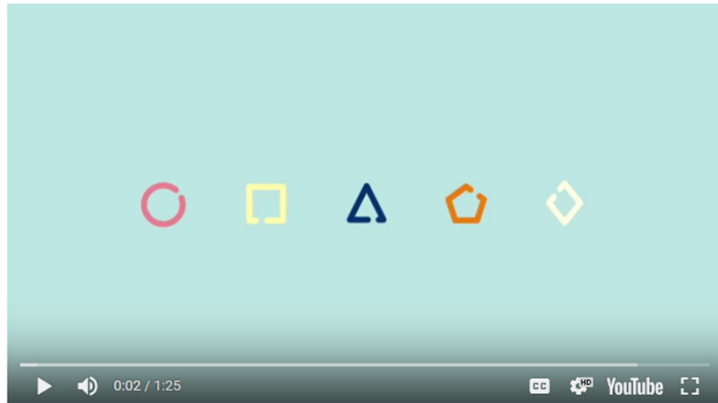


Winning H2020 with Open Science: How to gain a competitive evaluation advantage?

Transforming human health will take longer if research outputs – like publications, data, software and biological materials such as cell lines – aren't managed, shared and used in ways that unleash their full value.

We've been leading efforts to make research more open for over 20 years, ever since we worked to make sure the results of the Human Genome Project were released immediately into the public domain.

Our short animation explains what open research is, and why we support it.



[Read the transcript](#) [PDF 111KB]

<https://wellcome.ac.uk/what-we-do/our-work/open-research>

In recent years, the research community has made significant progress. Around 80% of all Wellcome-funded research publications are now openly available. And resources like the [Protein Data Bank](#) and the [European Nucleotide Archive](#) have become the main way for data about protein structures and DNA sequences to be shared.

More related news

Contact us

If you have any questions, contact the team

@ openresearch@wellcome.ac.uk

Policies and guidance

Open access guidance

Policy on data, software and materials management and sharing

Developing an outputs management plan

Topics

Data sharing

*What if you met
the ideal
applicant*



Introduction to Open Science video

This short video by Ivo Grigorov, Technical University of Denmark (DTU) introduces the concepts of Open Science and explains how they should be applied over the entire research lifecycle.

2 What is Open Science

RESEARCH LIFECYCLE

IDEA TEST DATA MODEL CODE PUBLISH EDUCATE & TRAIN

Peer-review

Impact Factor (IF)

FOSTER

"Research Cycle" by Tenopir et al., (2011) <http://bit.ly/1fL8975>

FOSTER

<https://www.fosteropenscience.eu/learning/what-is-open-science>



Responsible Research and Innovation

Europe's ability to respond to societal challenges



*What are the
political
motivations*





[View original](#)

[Flag media](#)



EU-komissio Suomessa @EUkomissio

55d

Keskustele tiedekomissaari @Moedas kanssa 31.8 klo 17.30 Helsingissä!
Ks. europa.eu/JP44fh #EUdialogues pic.twitter.com/93i3QaelKY





“

ACKNOWLEDGES that Open Science has the potential to increase the quality, impact and benefits of science and to **accelerate advancement of knowledge** by making it more reliable, more efficient and accurate,

better understandable by society and **responsive to societal challenges**, and has the potential to **enable growth and innovation** through reuse of scientific results by all stakeholders at all levels of society, and ultimately contribute to growth and competitiveness of Europe.

Brussels 27 May 2016



Realising the European Open Science Cloud

First report and recommendations
of the Commission High Level Expert Group
on the European Open Science Cloud



“

Mostly due to current methods capture and data malpractice, approximately 50% of all research data and experiments is considered **not reproducible**, and the vast majority (likely over 80%) of data never makes it to a trusted and sustainable repository.

At an investment of Europe in data-generating research of €120 Billion between 2014-2020, the **annual capital destruction** is consequently **very substantial**.

“

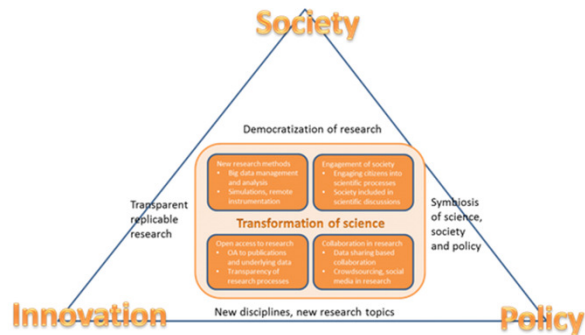
This will mean a new way of working through deep, equal partnerships between the science communities and the **ICT communities** ...

... to turn these data into knowledge as renewable, sustainable **fuel for innovation** in turn to meet global challenges.



Open Science

Open Science aims at transforming science through ICT tools, networks and media, to make research more open, global, collaborative, creative and closer to society.



About Open Science

[Policies](#) +

[Blog posts](#)

[News](#)

[Events](#)

[Projects](#)

[Funding](#)

Society

Democratization of research

New research methods

- Big data management and analysis
- Simulations, remote instrumentation

Engagement of society

- Engaging citizens into scientific processes
- Society included in scientific discussions

Transparent replicable research

Symbiosis of science, society and policy

Transformation of science

Open access to research

- OA to publications and underlying data
- Transparency of research processes

Collaboration in research

- Data sharing based collaboration
- Crowdsourcing, social media in research

Innovation

New disciplines, new research topics

Policy

Manage by stages

Starting your research project

Things to think about when you start your next research project

Click the plus sign to expand the text box

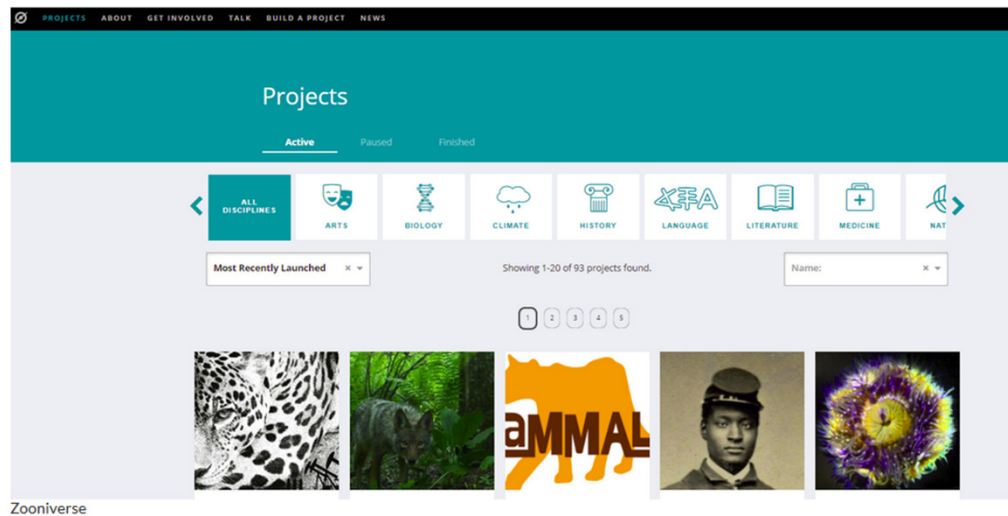
- + What will you produce during your research?
- + What tools and infrastructure will you use to carry out your work?
- + What publishing routes will you use?
- + How will you communicate your choices?
- + How will you make your data FAIR?

– During the active stage

You might share your methodologies and early findings via preprints. This is a great way to get peer feedback early on and helps you to identify any errors or problems with your approach before you publish.

Worried about getting scooped if you share early? Pre-registering your study gives you time-stamped evidence of your ideas. In addition, any peers that review your early work can vouch for you. This [Open Science Framework \(OSF\) guide](#) offers great advice on pre-registering your project.

You may want to involve citizen science in analysing the data you have generated or collected. This can dramatically increase the amount of data that you can realistically analyse in a short space of time and meaningful public engagement (i.e., not just a box-ticking exercise) is generally viewed very favourably by grant application reviewers. Check out [Zooniverse's tips](#) for setting up your own citizen science project.



During your research project

Things to think about during your research project

Click the plus sign to expand the text box

+ Handling, processing, transferring of data

+ Dealing with sensitive data

During your research project

Data management planning should take place before your research begins but you should also be updating your data management plan as you progress through your project. This video by Research Data Netherlands provides an overview of data management planning and some of the issues you may need review as you progress your research.

Wrapping up your research project

Coming to the end of your research project? Here are a few things to consider.

Click the plus sign to expand the text box

- + Have you used common standards?
- + Deposit your outputs somewhere safe
- + Have you assigned licenses and persistent identifiers?
- + Have you met your funder's expectations (e.g., is your data FAIR)?

Wrapping up your research project

*What does this mean
for Horizon2020
applicants*



*Does #OpenScience
matter at
proposal evaluation*



*Based on: Winning Horizon 2020 with Open Science,
<http://dx.doi.org/10.5281/zenodo.12247>*

www.fosteropenscience.eu

Winning Horizon2020 with Open Science?

Developed jointly between
FP7 FOSTER & FP7 OpenAIRE+
2015

DOI: 10.5281/zenodo.12247



Innovation Union & Europe 2020 Initiative

WHY Open Science in Horizon 2020?

Open Science (OS) offers researchers tools and workflows for transparency, reproducibility, dissemination and transfer of new knowledge. Ultimately, this can also have an impact on in research evaluation exercises, e.g. Research Excellence Framework (REF), set to demand greater 'societal impact' in future, rather than just research output¹. OS can also be an effective tool for research managers to transfer knowledge to society, and optimize the use and re-use by unforeseen collaborators. For funders, OS offers a better return on investment (ROI) for public funding, and underpins the EU Digital Agenda by measurably contributing to economic growth. This brief showcases why and how Open Science can optimize your Horizon 2020 proposal evaluation.

WHO is this "BRIEF" for?

This brief is developed through EC funding and specifically aimed at Horizon 2020 applicants and proposal writers seeking to comply with the Horizon 2020 Mandate (Grant Agreement article 29.1-6) and to optimize proposal evaluation and eventual societal impact of the resulting project.

HOW to use the "BRIEF"?

The text is **NOT** intended to be used verbatim as copy and paste contribution to your proposal. Instead, the brief presents suggested ways of formulating an impact section that answers the overarching political agendas and initiatives, as well as tips for ensuring that research results are effectively delivered to any users and the market place, across the various Horizon 2020 Pillars. The main text is generic, but some discipline-specific examples are included as examples, rather than covering all research fields. The footnotes also point to additional resources that will facilitate implementation to optimize project visibility and impact.

¹ Weighting of research impact confirmed for 2014 Research Excellence Framework <http://www.hefce.ac.uk/news/newsarchive/2011/mess6310.html>, 2011

HOW to write "Section 2.2 IMPACT" A generic example

The Project consortium acknowledges that the research and new knowledge generated is of societal benefit, and could potentially contribute toward solutions of societal challenges. As such, the foreground knowledge needs to be disseminated in an optimum way for impact and re-use of results, according to Responsible Research & Innovation (RRI) principles¹⁴.

Currently only 50% of research is freely accessible to the public¹⁵, resulting in measurable loss to the knowledge-based SME sector and slowing down innovation¹⁶. **The Project** consortium will thus optimize on the dissemination and impact of foreground along the full knowledge production chain, and integrate Open Science principles in its Dissemination & Communication Strategy.


In support of the EC Digital Agenda¹⁷ and the Economic Growth agenda of the Innovation Union (Green Action Plan¹⁸), the consortium will fully integrate Grant Agreement Article 29 into its workflow at task level. Foreground data (state diversity of data generated) will be permanently archived at generation in STATE REPOSITORY¹⁹ and publicly released and/or published²⁰ (with the exception of Third Party data, national security data, medical/patient data) during the lifetime of the project²¹.

Software code, tools and interfaces developed as part of the concept will be open source code and full access provided via STATE REPOSITORY²². Resulting research publications (refer to tasks/WP most likely to publish) will also be made openly available via e-infrastructure OpenAIRE²³ (DG CONNECT; request letters of support), predominantly relying on the Green Open Access strategy (self-archiving) for maximum return on investment for project and funder, and actively linked to underlying data objects, in support of the EC Open Data Pilot²⁴.


For longevity of knowledge transfer and best practice uptake beyond the project lifetime, **The Project** will cooperate with concurrent training initiatives within FP7 FOSTER²⁵ (DG Research) and OpenAIRE+, and incorporate Open Science training in any summers schools and research training workshops, to assure that the strategy is adopted by the next generation of young researchers (refer to WP/Tasks dealing with this).

Focus will be placed on demonstrating that Open Science and RRI are not only for societal and community benefit, but also directly support the career needs for impact, visibility and multiplying collaborations for individual researchers. Aiming the societal and research impact of knowledge generation can in the long-term bridge the gap between science and society.


¹⁴ EC Responsible Research & Innovation http://ec.europa.eu/research/science-society/document_library/pdf_06/responsible_research_and_innovation_brief_en.pdf
¹⁵ Archambault, E. et al. 2013. Proportion of OA Peer-Reviewed Papers at the European & World Levels 2004-2011 at http://www.science-metrics.com/pdf/SM_EOA_Availability_2004-2011.pdf
¹⁶ Houghton, J., Swan, A., Brown, S., 2011. Access to research and technical information in Denmark (WWW Document). URL: http://www.delft.nl/infoshared/Access_to_Research_and_Technical_Information_in_Denmark.pdf
¹⁷ EC Digital Agenda & Access to Knowledge <http://ec.europa.eu/digital-agenda/en/open-access-scientific-knowledge-0>
¹⁸ EC Green Action Plan for SMEs <http://ec.europa.eu/DocsRoom/documents/4790/attachments/1/translations/en/mediawatch>
¹⁹ Choose a discipline-specific general Data Repository from <http://www.europeandataportal.eu/>
²⁰ Choose likely Data Journals of relevance: e.g. Nature Scientific Data, or search <http://doi.org>
²¹ NH: embargoes can be placed to allow project to publish/exploit first, but consortium should aim for full release by end of contract, or justify why access needs to be restricted (publications may not be viewed favourably at review)
²² Choose a structured archive with minimum metadata requirements to allow maximum re-use e.g. GitHub, SourceForge, etc.
²³ EC FP7 and Horizon2020 funded e-Infrastructure <https://www.openaire.eu>, in support of EC Digital Agenda
²⁴ EC Open Data Pilot http://openaire.eu/openaire-release_EC-13-1247_en.htm
²⁵ FP7 FOSTER, Facilitating Open Science in European Research (www.fosteropen-science.eu)




“...all parameters being equal, a proposal clearly integrating Open Science into its concept, contributing to EC’s Digital Agenda and the Innovation Union’s objectives, will inherently be favoured by aware project officer/evaluator teams, over same score proposals. Supporting the principles your funders believe in, adds to the competitiveness of your proposal, even when this is not explicitly formulated in the evaluation criteria. Ultimately, a good research idea must not fail due to a lack of sufficient detail regarding openness, dissemination and broader use of the new knowledge beyond academia.”



Winning Horizon2020
with
Open Science?




“Applicants should be clear about the research outputs they anticipate producing during the life of the project along with a detailed account of how these will be stored and shared over time. Applicants should aim to provide optimum, traceable and perennial access to selected research products of the project, from *research data, software code, publications, educational resources, reports, policy briefs* etc., that can help accelerate transfer of new knowledge from academia to society, as well as improve reproducibility of public-funded research. As a minimum, applicants should retain any and all outputs that are required to verify the published research findings.”



“Open science can contribute positively throughout your proposal structure, from your attitude and mind-set when developing the concept (Section 1), to impact & dissemination (Section 2), to the actual implementation of individual workpackages and tasks.

For ease of implementation, and potentially a stronger evaluation, consider open science as part of performing transparent, rigorous and reproducible research, as described in your proposal concept.



HOW to write "Section 2.2 IMPACT" A generic example

The Project consortium acknowledges that the research and new knowledge generated is of societal benefit, and could potentially contribute toward solutions of societal challenges. As such, the foreground knowledge needs to be disseminated in an optimum way for impact and re-use of results, according to Responsible Research & Innovation (RRI) principles¹⁴.

Currently only 50% of research is freely accessible to the public¹⁵, resulting in measurable loss to the knowledge-based SME sector and slowing down innovation¹⁶. **The Project** consortium will thus optimize on the dissemination and impact of foreground along the full knowledge production chain, and integrate Open Science principles in its Dissemination & Communication Strategy.

In support of the EC Digital Agenda¹⁷ and the Economic Growth agenda of the Innovation Union (Green Action Plan¹⁸), the consortium will fully integrate Grant Agreement Article 29 into its workflow at task level. Foreground data (*state diversity of data generated*) will be permanently archived at generation in **STATE REPOSITORY**¹⁹ and publicly released and/or published²⁰ (*with the exception of Third Party data, national security data, medical/patient data*) during the lifetime of the project²¹.

Software code, tools and interfaces developed as part of the concept will be open source code and full access provided via **STATE REPOSITORY**²². Resulting research publications (*refer to tasks/WP most likely to publish*) will also be made openly available via e-Infrastructure OpenAIRE²³ (DG CONNECT; *request letters of support*), predominantly relying on the Green Open Access strategy (self-archiving) for maximum return on investment for project and funder, and actively linked to underlying data objects, in support of the EC Open Data Pilot²⁴.

For longevity of knowledge transfer and best practice uptake beyond the project lifetime, **The Project** will cooperate with concurrent training initiatives within FP7 FOSTER²⁵ (DG Research) and OpenAIRE+, and incorporate Open Science training in any summers schools and research training workshops, to assure that the strategy is adopted by the next generation of young researchers (*refer to WP/Tasks dealing with this*).

Focus will be placed on demonstrating that Open Science and RRI are not only for societal and community benefit, but also directly support the career needs for impact, visibility and multiplying collaborations for individual researchers. Alining the societal and research impact of knowledge generation can in the long-term bridge the gap between science and society.

- Export PDF
- Create PDF
- Edit PDF
- Comment
- Combine Files
- Organize Pages
- Redact
- Protect
- Optimize PDF
- Fill & Sign
- Send for Signature
- Send & Track
- More Tools

Convert and edit PDFs with Acrobat Pro DC
Start Free Trial

Excellence

Impact

“Weakness: highly focused on academic activities, and lacks an advanced communication strategy”

“data accessibility is unclear!”

“Open Access to scientific knowledge is an essential principle in the project, but there is not enough information on data management or IPR.”

“data storage & access not considered”

Implementation

Impact:

“Strengths: extensive dissemination of data to the scientific community (open access, databases)”

“outreach activities to a broad audience”

“research software is freely available”

Impact:

Societal Challenge 2, Food:

“new knowledge delivered to SMEs by depositing research outputs in OpenAIRE2020 and EMODNet”

“Data management is based on INSPIRE Directive that will favour open access for environmental data and open publications”

“web-based decision support system (with open code) to make it available to a wide, but targeted audience”

Societal Challenge 5, Climate:

“The communication plan is very effective. Training for communication and open access procedures are especially welcome.”

Benefits of Open Science Explained

Ivo Grigorov, Technical University of Denmark (DTU) runs through some of the tangible benefits you can expect from putting Open Science into practice in this short video.

5 Conclusion

RESEARCH LIFECYCLE

IDEA & PROPOSAL

TEST

DATA

MODEL CODE

PUBLISH

EDUCATE & TRAIN

Open Educational Resources

Open journals & repositories

Publish data

Publish code

Open Notebook Science


OS key part of concept

FOSTER

"Research Cycle" by Tenopir et al., (2011) <http://bit.ly/1fL8975>

What is Open Science? | What is Op: X

← → ↻ 🏠 🔒 https://www.fosteropenscience.eu/learning/what-is-open-scienc 120% 🔍 Search

<  ⋮

Are you ready to put Open Science into practice?

Open Science relates to sharing research publications only.

Select one of the options below.

True

False

Submit

Show feedback

Submit  Show feedback

I should share my research outputs only when they have been published.

Select one of the options below.

Select one of the options below.

- True
- False

Submit Show feedback

What is Open Science? | What is Op: X

https://www.fosteropenscience.eu/learning/what-is-open-scienc 120%

FOSTER

Submit Show feedback

Funding bodies expect researchers to share everything they produce in their research.

Select one of the options below.

True

False

Submit Show feedback

Practicing Open Science only benefits others more than it benefits me as a researcher.

Select one of the options below.

True

False

Submit

Show feedback

Interesting website regarding Open Data: <https://www.openaire.eu/>
with interesting factsheets: <https://www.openaire.eu/openaire-h2020-factsheets>

Contacts in Belgium of the OpenAire projects are:

Inge Van Nieuwerburgh: <https://www.openaire.eu/72-noads/47-inge-van-nieuwerburgh> and

Emilie Hermans: <https://www.openaire.eu/72-noads/129>