Support & Training Materials

Fostering the practical implementation of Open Science in Horizon 2020 and beyond

Iryna Kuchma, EIFL Open Access Programme Manager, @irynakuchma

Seminar for young researchers in SSH on Open Science and Data Management, EKT

5th June 2018

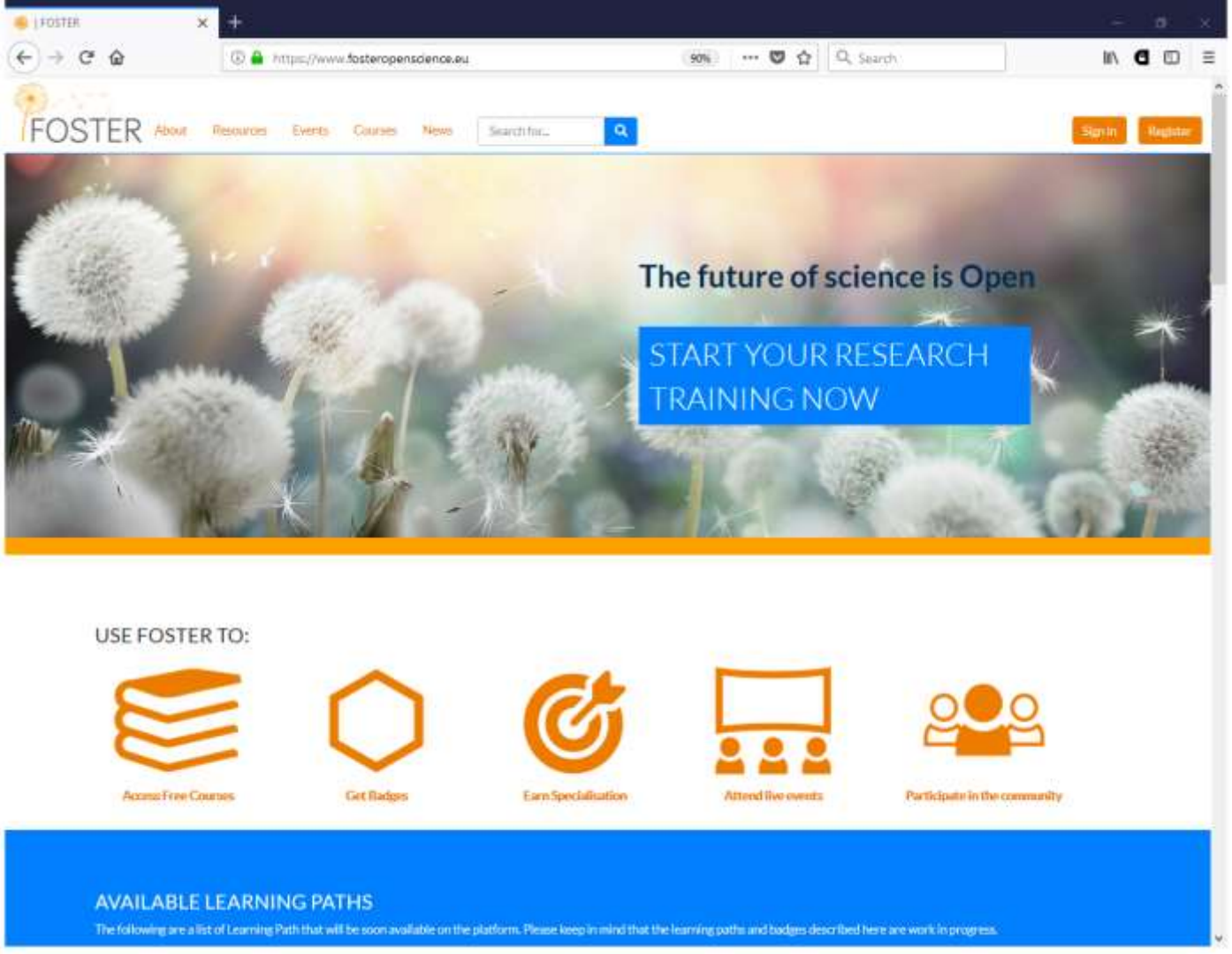
CC BY

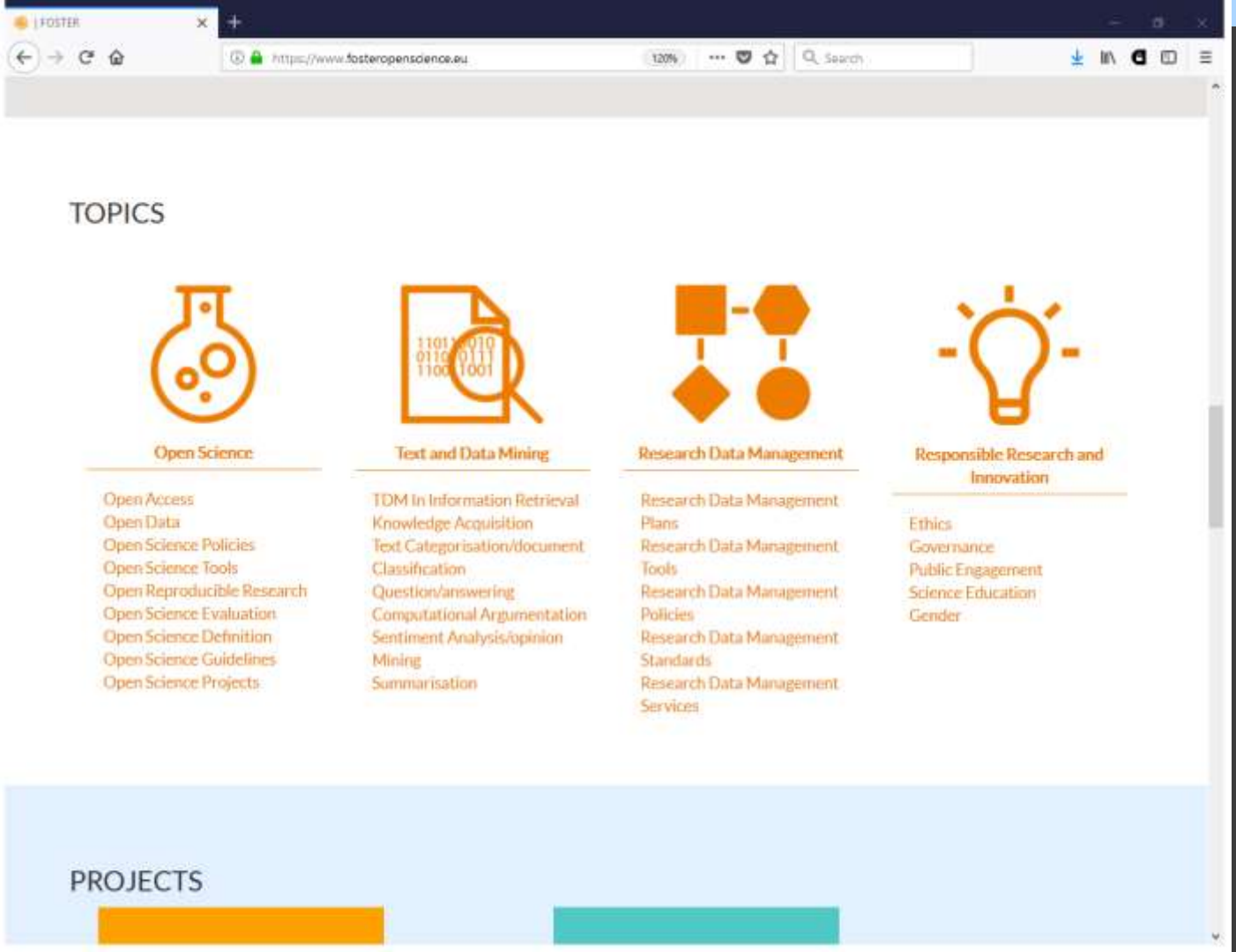
FOSTER Objectives & project activities 2017-2019

- Strengthening Open Science training capacity in ERA
- Focusing on practical implementation of Open Science & ‘training the trainers’
- Training resources: new topics RDM & Open Data + intermediate & advanced level, and discipline specific
- Involving disciplines:
  - Humanities
  - Social sciences
  - Life sciences

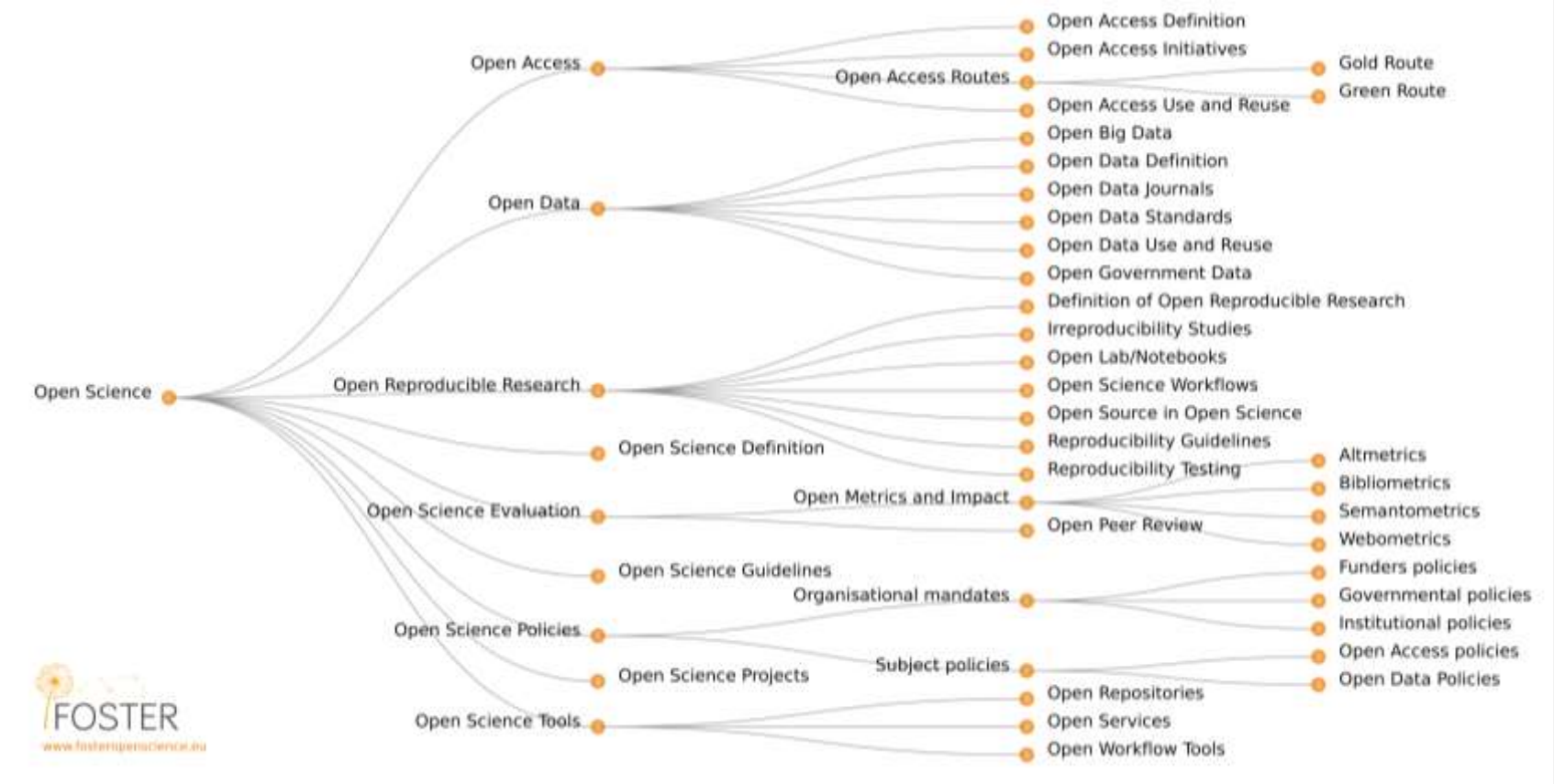
DARIAH-EU cessda elixir

https://www.fosteropenscience.eu


FOSTER
https://www.fosteropenscience.eu
FOSTER
About
Resources
Events
Courses
News
Search for...
Sign in
Register
The future of science is Open
START YOUR RESEARCH TRAINING NOW
USE FOSTER TO:
Access Free Courses
Get Badges
Earn Specialisation
Attend live events
Participate in the community
AVAILABLE LEARNING PATHS
The following are a list of Learning Path that will be soon available on the platform. Please keep in mind that the learning paths and badges described here are work in progress.


FOSTER
https://www.fosteropenscience.eu
120%
Search
TOPICS
Open Science
Open Access
Open Data
Open Science Policies
Open Science Tools
Open Reproducible Research
Open Science Evaluation
Open Science Definition
Open Science Guidelines
Open Science Projects
Text and Data Mining
TDM In Information Retrieval
Knowledge Acquisition
Text Categorisation/document Classification
Question/answering
Computational Argumentation
Sentiment Analysis/opinion Mining
Summarisation
Research Data Management
Research Data Management Plans
Research Data Management Tools
Research Data Management Policies
Research Data Management Standards
Research Data Management Services
Responsible Research and Innovation
Ethics
Governance
Public Engagement
Science Education
Gender
PROJECTS

Open Science taxonomy



Paper available at http://oro.open.ac.uk/44719/. Image available at http://oro.open.ac.uk/47806/

USE FOSTER TO:

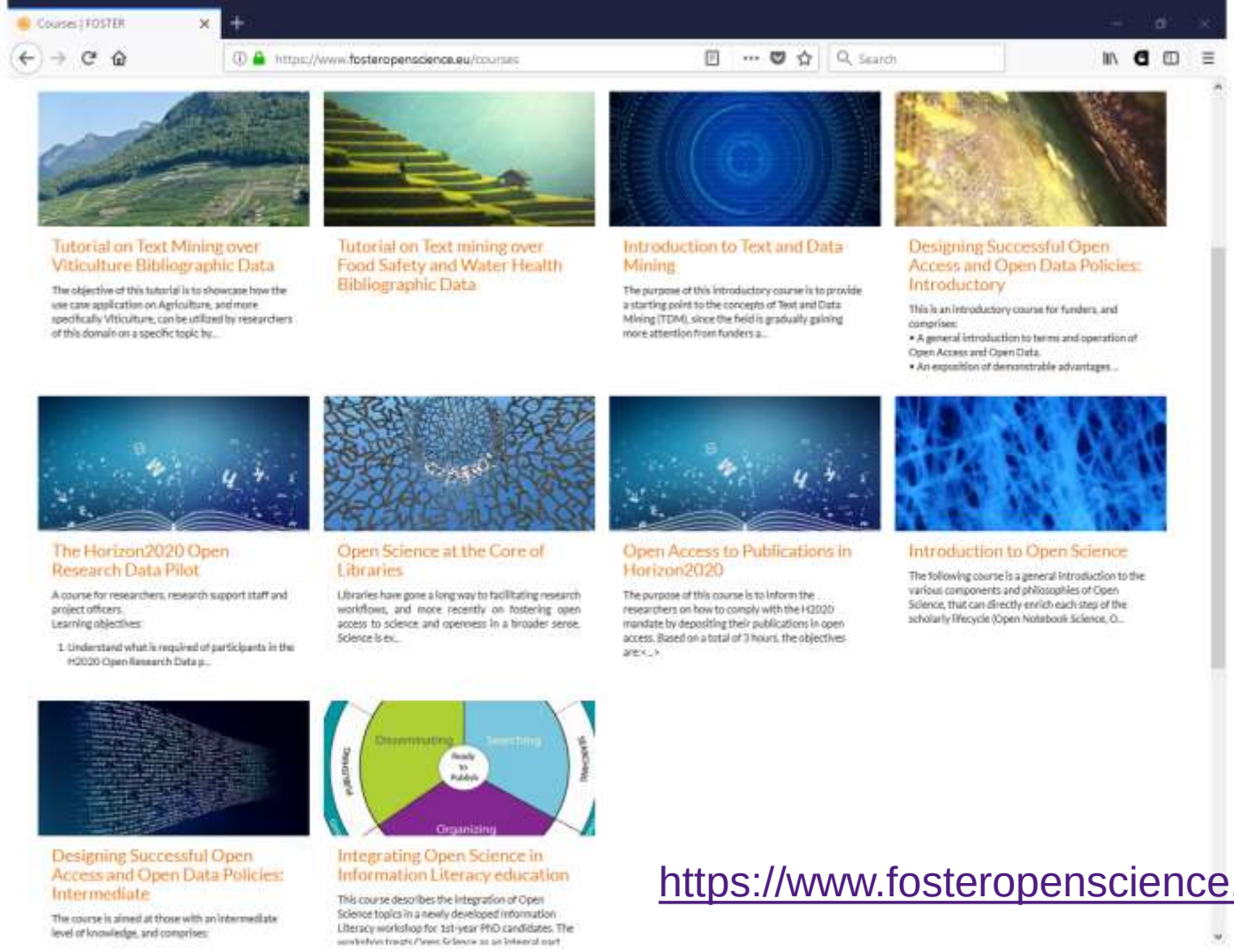
Access Free Courses

Get Badges

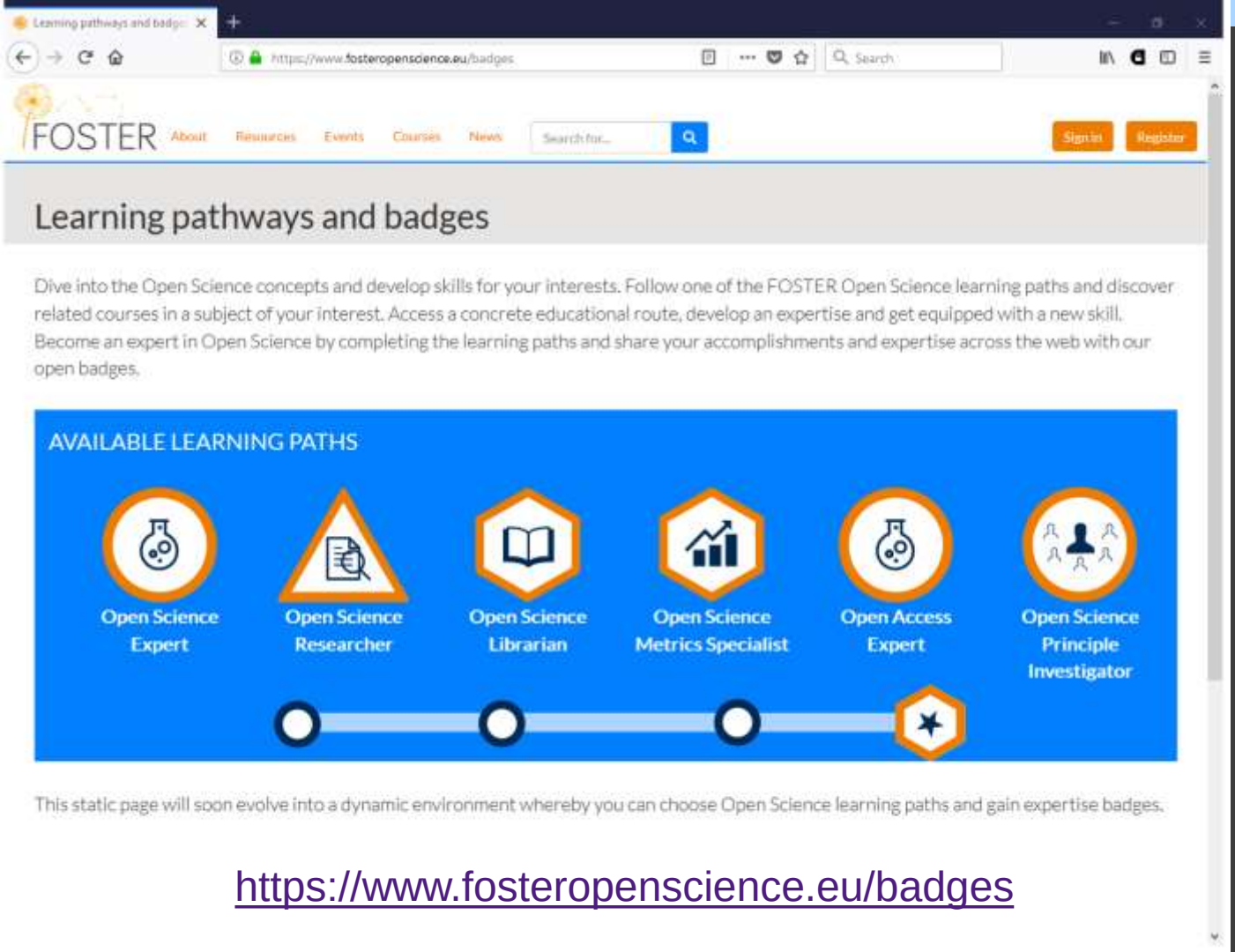
Earn Specialisation

Attend live events

Participate in the community



https://www.fosteropenscience.eu/courses



https://www.fosteropenscience.eu/badges

Learning Paths | FOSTER

https://www.fosteropenscience.eu/learning-paths

Search

The open peer reviewer

The practice of peer review is evolving to become more open. This pathway will help you to understand the move towards Open Peer Review and equip you to be able to participate - both as an author and as a reviewer.

Effort: 2-3 hours

Level: intermediate and advanced

The learning path will be completed by finishing the following courses:

- Open peer review
- Open research data
- Open source software and workflows

The responsible data sharer

Sharing data is great but you need to know what data you can share, with whom you can share, when to share, and how best to share. This pathway will equip you with the knowledge you need to make informed decisions about sharing your data responsibly.

Effort: 3-4 hours

Level: intermediate and advanced

The learning path will be completed by finishing the following courses:

- What is open science?
- Open research data
- Ethics and data protection
- Licensing

The reproducible research practitioner

Recent studies have revealed a reproducibility crisis for published findings. This pathway will help you to make sure that your published findings can stand up to scrutiny and serve as a solid foundation for others to build upon.

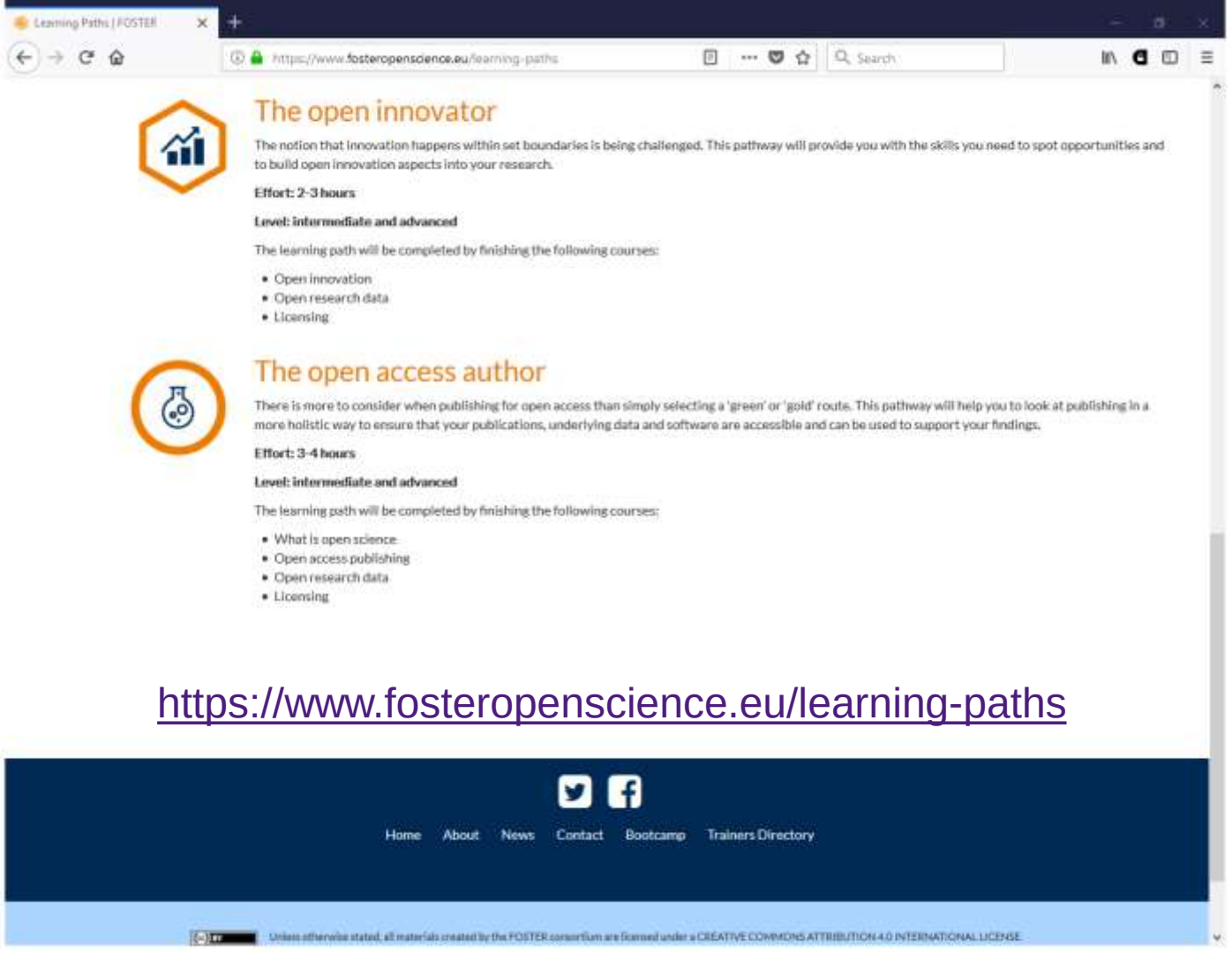
Effort: 3-4 hours

Level: intermediate and advanced

The learning path will be completed by finishing the following courses:

- What is Open Science?
- Best practices in open research
- Open Access publishing
- Open research data
- Open source software and workflows

https://www.fosteropenscience.eu/learning-paths



https://www.fosteropenscience.eu/learning-paths

Events | FOSTER

https://www.fosteropenscience.eu/events

Search

Upcoming events

https://www.fosteropenscience.eu/events

Go to past events

Seminar for young researchers in SSH on Open Science and Data Management

05.06.2018 - 01.06.2018

EKT

National Hellenic Research Foundation 48, Vassileos Konstantinou , Athens

DMOS18 - Data Management and Open Science

12.06.2018 - 15.06.2018

Pedro Fernandes

Instituto Gulbenkian de Ciência, Oeiras, Portugal

Open Science Funder Training Workshop

18.06.2018

ORION Project, www.orion-openscience.eu

Instituto de Salud Carlos III Monforte de Lemos, 5, E-28029, Madrid

Make peer review transparent again

20.06.2018

OpenUP, FOSTER

Göttingen

ESOF2018 Session: "Open science: from concept to implementation"

12.07.2018 - 12.09.2018

Marie Curie Alumni (Fernanda Bajanca, Maja Mise), EC MSCA Office & FOSTER+ (Ivo Grigorov , Reme Melero, Sarah Jones)

Compans-Caffarelli, Toulouse, France

Open Science Toolkit

"Move from being aware of open science to being able to put open science into practice in their daily workflows"

- Targeted towards researchers
- Focus on intermediate level, practical content
- Disciplinary examples via CRG, GESIS, DARIAH
- Quizzes will assess competence
- Badges will be issued on successful completion

Modules and ‘specialism’ pathways

1 hr each. Standalone topics.

- What is open science?
- Best practices in open research
- Data protection and ethics
- Licensing
- Managing and sharing research data
- Open access publishing
- Open peer review
- Open science and innovation
- Open source software and workflows
- Sharing preprints

2-4 hours. Combines content

- The reproducible research practitioner
- The responsible data sharer
- The Open access author
- The open peer reviewer
- The open innovator

For more information, see www.fosteropenscience.eu/learning-paths

Open Peer Review module example

Open Peer Review

This module will introduce you to Open Peer Reviewing and let you know how you can get started with it.

Introduction

This module introduces you to open peer review (OPR), an eme
Science.

Upon completing this module, you will:

- understand what OPR means and how it supports Open S
- be aware of OPR workflows and which aspects of the revi
- know how to write a constructive and responsible open p
- know about useful tools and services that can support yo

CC-BY-SA AJ Cann

OPR in three minutes

In this short video, Tony Ross-Hellauer introduces the concept of open peer
strongly needed in the peer review process.

What does OPR mean?

Definition of OPR

Click the forward arrow to see more.

Transparent & accountable

Open peer review is an umbrella term for various alternative review methods that seek to make classical peer review more transparent and accountable (cf. Ross-Hellauer, 2016).

Quiz - Are you an Open Peer Reviewer?

Transparency can be added to peer review through:

- Accessible evaluation reports
- Platforms that allow interaction
- Revealed identities of reviewers

Submit | Show feedback

What are the benefits of open peer review?

- It is not biased
- My results can be published more quickly
- My review is a citable research output

Submit | Show feedback

Case study approach

Using the EC Open Science Monitor approach to share practical examples of activity from the Life Sciences, Social Sciences and Humanities.

Open Access

eLIFE

Open Source Licensing

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Open Peer Review

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

Ethics

EUROPEAN GENOME-PHENOME ARCHIVE
10TH ANNIVERSARY

Life Sciences: Nextflow for reproducible in silico genomics

Why?

The analysis of big data in a performant and reproducible manner is an increasing pressing issue in many scientific fields including and mostly in life science disciplines. This problem has been fuelled by the combined reliance on increasingly complex data analysis methods and the exponential growth of biological datasets. When considering the installation, deployment and maintenance of bioinformatic pipelines, an even more challenging picture emerges due to the lack of community standards. Moreover, the effect of limited standards on reproducibility is amplified by the very diverse range of computational platforms and configurations on which these applications are expected to be applied (workstations, clusters, HPC, clouds, etc.). The Nextflow open source technology provides a simple but yet effective solutions to many of these problems.

Open Research Data

Example use of EBI metagenomics

Open Innovation

docker

FOSTER

Ideas for opening up during each stage of your research

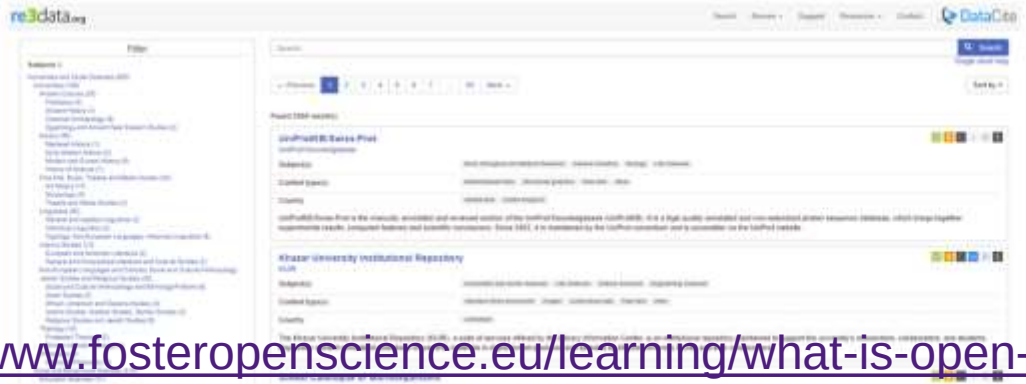
− During the planning stage

You might want to consider writing a blog post outlining your ideas and approaches early on to get community feedback.

In addition to seeking involvement from other researchers, be sure to consider involving other stakeholders too. For instance, seeking collaboration with industrial partners at the idea stage can be a great way to see your research outputs applied in a real life setting more quickly. That is great for your impact!

Be sure to check data repositories to see if there are existing data that you can reuse or build upon during your research rather than starting from scratch. re3data is a great way to find relevant repositories for your area of research. As you can see from the screenshot below, re3data provides access to repositories across a very wide range of subjects. There are more than 600 repositories listed for the Humanities and Social Sciences and more than 1000 for the Life Sciences - each of which allows users to search more specifically by sub-domain.

Reusing data means you ask for less public money in your proposal (reviewers generally like this!) and gives you the potential to identify any gaps that targeted data collection could help to fill.



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

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

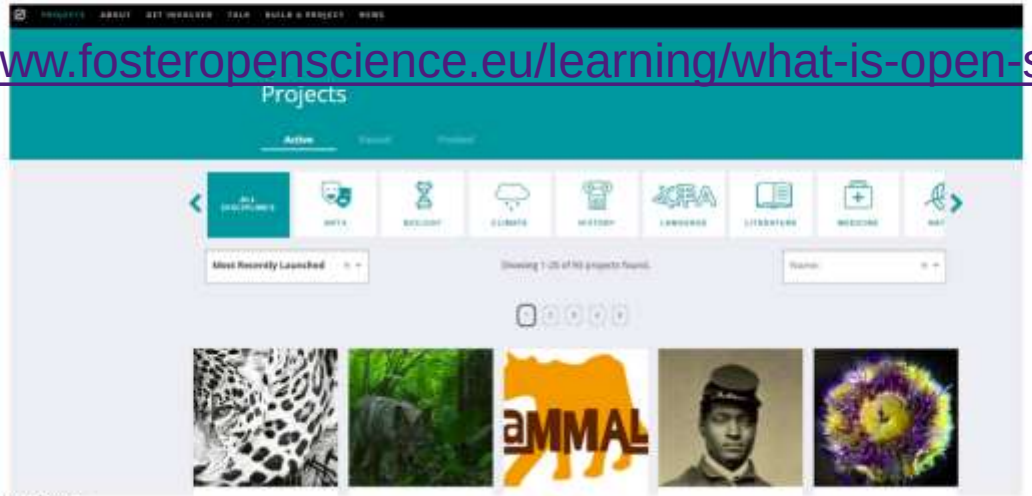
FOSTER

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.



Zooniverse

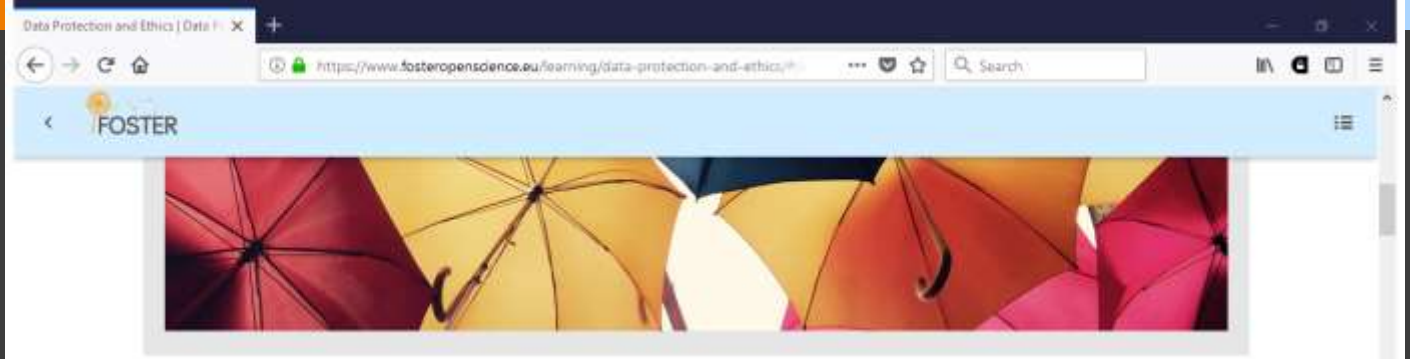
https://www.fosteropenscience.eu/learning/what-is-open-science

− Towards the end of your research

Make sure to publish with an Open Access journal and/or to deposit your publications in an Open Access repository. This means that anyone can read - and cite - your findings in the short and longer-term. That all adds up to more citations for you!

Be sure to deposit any data required to validate your findings as well as any software you've developed to analyse or visualise them. Link your papers, data, and code to each other through the assignment of DOIs. Link all of these back to you through your ORCID!

To make your research accessible to non-experts, consider writing a lay summary to describe your research approach and findings. Bear in mind, non-experts can be researchers in other fields as well as journalists and the general public. A bit of effort here can pay dividends if your research is picked up by the media. The Digital Curation Centre's guide can help you. INVOLVE, funded by the National Institute for Health Research provides some very helpful guidance on writing plain English summaries as well.



https://www.fosteropenscience.eu/learning/data-protection-and-ethics

Data Protection

Considering aspects of data protection is crucial for your research, particularly if you are planning to share your research data. Protecting research data means protecting the rights of humans involved in the research process.

Key elements

As a responsible researcher, you should become familiar with:

- the legal requirements that need to be respected when sharing data
- securing informed consent
- selecting appropriate anonymisation strategies
- securely storing and transferring data

Remember - data protection should start at the earliest stages of the research process!

https://www.fosteropenscience.eu/learning/data-protection-and-ethics

What are personal data?

The term personal data refers to any information that can be used to identify living (and sometimes even deceased) individuals. For example, their name or date of birth. Some personal information - such as a person's origin, political opinion, religious beliefs, health, trade union membership, or sexual orientation - is classified as sensitive personal data. If you are handling personal and/or sensitive personal data as part of your research, you need to make sure it is protected.

If you aren't sure whether your research involves collecting personal data, it is best to speak with someone in your ethics team or data protection office to get advice as early as possible so that you can build in the appropriate data protection measures.

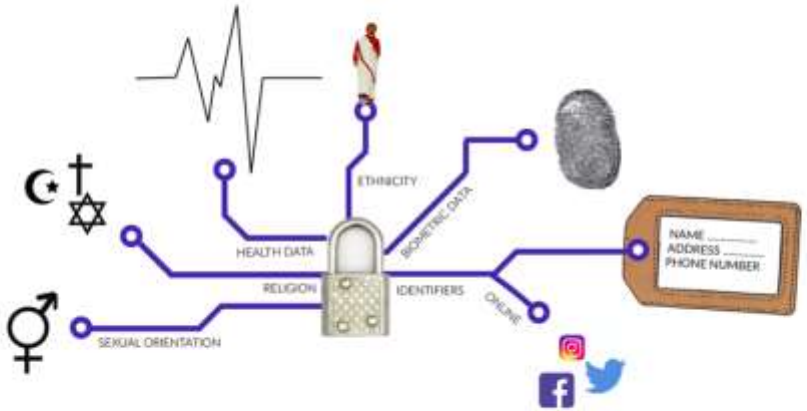


Image reused from CESSDA ERIC Expert Tour Guide on Data Management which is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

https://www.fosteropenscience.eu/learning/data-protection-and-ethics

FOSTER

Protecting personal data

From the outset of your research project, try to include measures for data protection in each step. Below are some actions you should take when starting each new project.

- Discuss whether you actually need to collect personal data to carry out your research.
- Consider collecting data anonymously if possible.
- Identify which personal data which will be included in your research.
- Include aspects of data protection in your data management plan (DMP).
- Create and use consent forms. Find more information on consent forms below.
- Find the appropriate anonymisation strategy for your research. You will find more information on anonymisation strategies below.
- Securely store, control access, and transfer your data.

CESSDA provides some useful tips on adapting your DMP to cover data protection here. Data cleaning and anonymising data can be very costly. It is a good idea to use the data management planning process as a way to identify any anonymisation activities that will incur additional costs and to request these in your grant application.

https://www.fosteropenscience.eu/learning/data-protection-and-ethics

Personal Data

− Legal requirements - EU General Data Protection Regulation (GDPR)

In April 2016 the European Union adopted its regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). The GDPR becomes enforceable in May 2018.

By collecting and/or handling personal data, researchers are thereby required to follow a number of principles including:

- Transparency - processing personal data "lawfully, fairly and in a transparent manner"
- Data Minimization - data use shall be limited to the purpose of the respective research
- Accuracy - inaccurate data must be "erased or rectified without delay"
- Integrity and Confidentiality - data must be protected by appropriate security measures (technical and organizational)

Find out more about how the EU GDPR affects processing of personal data here.

GDPR

+ Legal requirements - GDPR research exemptions

https://www.fosteropenscience.eu/learning/data-protection-and-ethics

FOSTER

− Legal requirements - GDPR research exemptions

With its General Data Protection Regulation (GDPR), the European Union provides a legal framework for data protection inside the EU, as well as for the export of these data outside its borders.

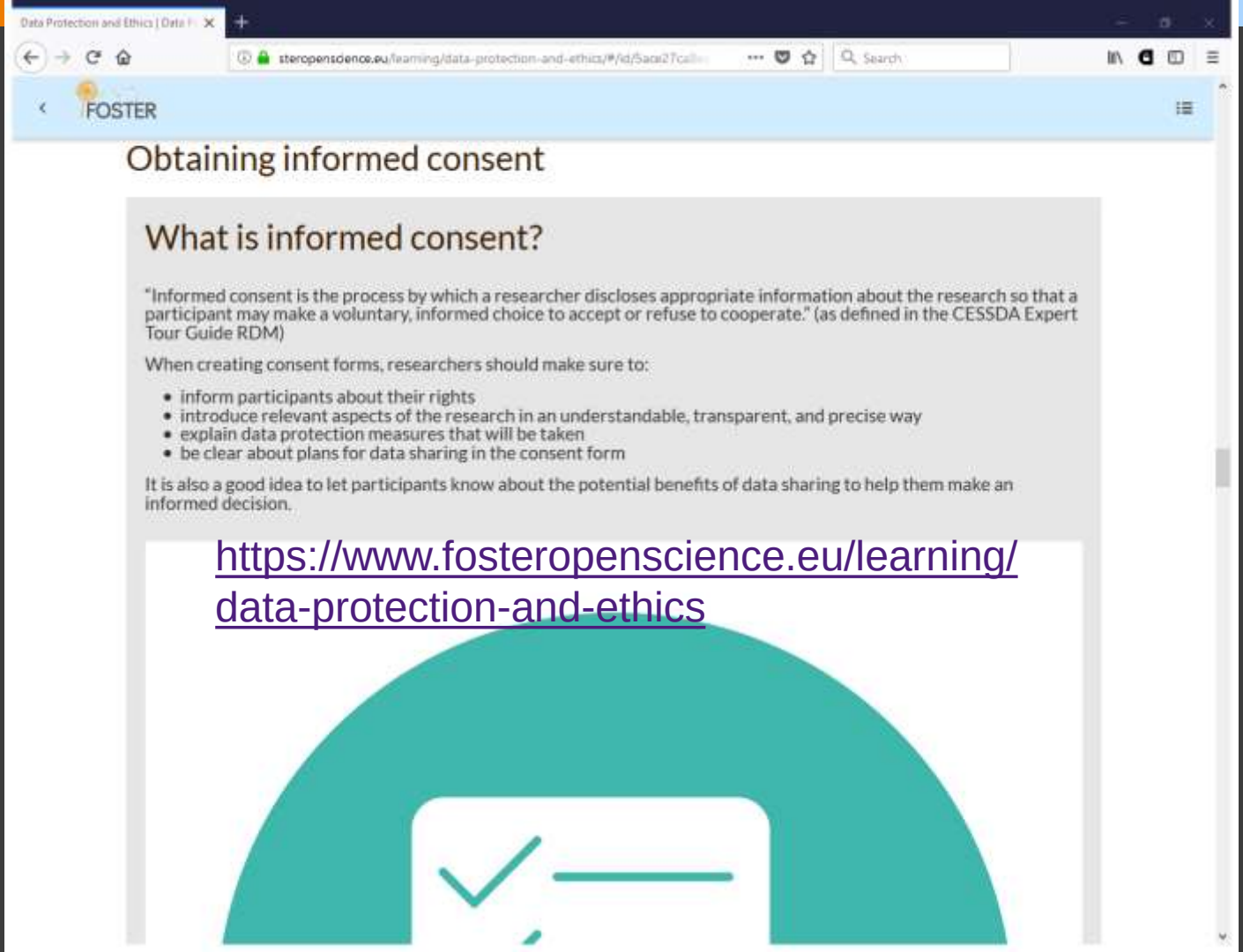
As for the collection of personal data, the GDPR includes an exemption for research:

- if it concerns "the public interest, scientific or historical research purposes or statistical purposes" (Art. 5.1 2016/679/EU); or
- if "the data subject has given consent to the processing of his or her personal data for one or more specific purposes" (Art. 6.1 2016/679/EU).

See here for the full text of the GDPR in several European languages.

Would you like to learn more about personal data and legal regulations? Check out more information from the European Union here or have a look at this Jisc webinar.

https://www.fosteropenscience.eu/learning/data-protection-and-ethics

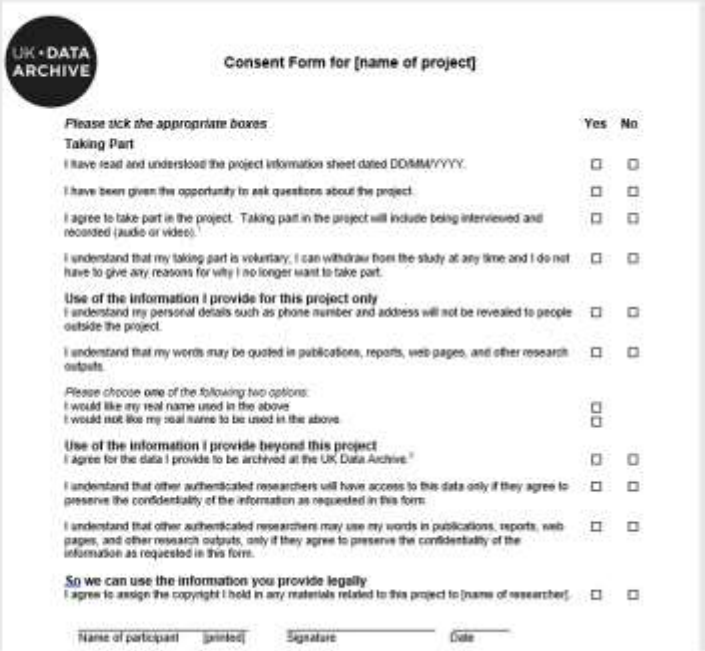

Data Protection and Ethics | Data
steropenscience.eu/learning/data-protection-and-ethics/#/id/5aca27ca8e
Search
FOSTER
Obtaining informed consent
What is informed consent?
"Informed consent is the process by which a researcher discloses appropriate information about the research so that a participant may make a voluntary, informed choice to accept or refuse to cooperate." (as defined in the CESSDA Expert Tour Guide RDM)
When creating consent forms, researchers should make sure to:
• inform participants about their rights
• introduce relevant aspects of the research in an understandable, transparent, and precise way
• explain data protection measures that will be taken
• be clear about plans for data sharing in the consent form
It is also a good idea to let participants know about the potential benefits of data sharing to help them make an informed decision.
https://www.fosteropenscience.eu/learning/data-protection-and-ethics

https://www.fosteropenscience.eu/learning/data-protection-and-ethics

Sample consent forms

If you collect personal data in your research, you are required to obtain informed consent for data collection, processing, sharing, and preservation. The CESSDA Expert Tour Guide on Data Management includes useful guidance on developing consent forms. Check out the sample forms they provide - like the one below - which can help you create your own consent forms.

For more information on ethical aspects of research including the creation of consent forms, check out this presentation by Libby Bishop, UK Data Service in the FOSTER portal.

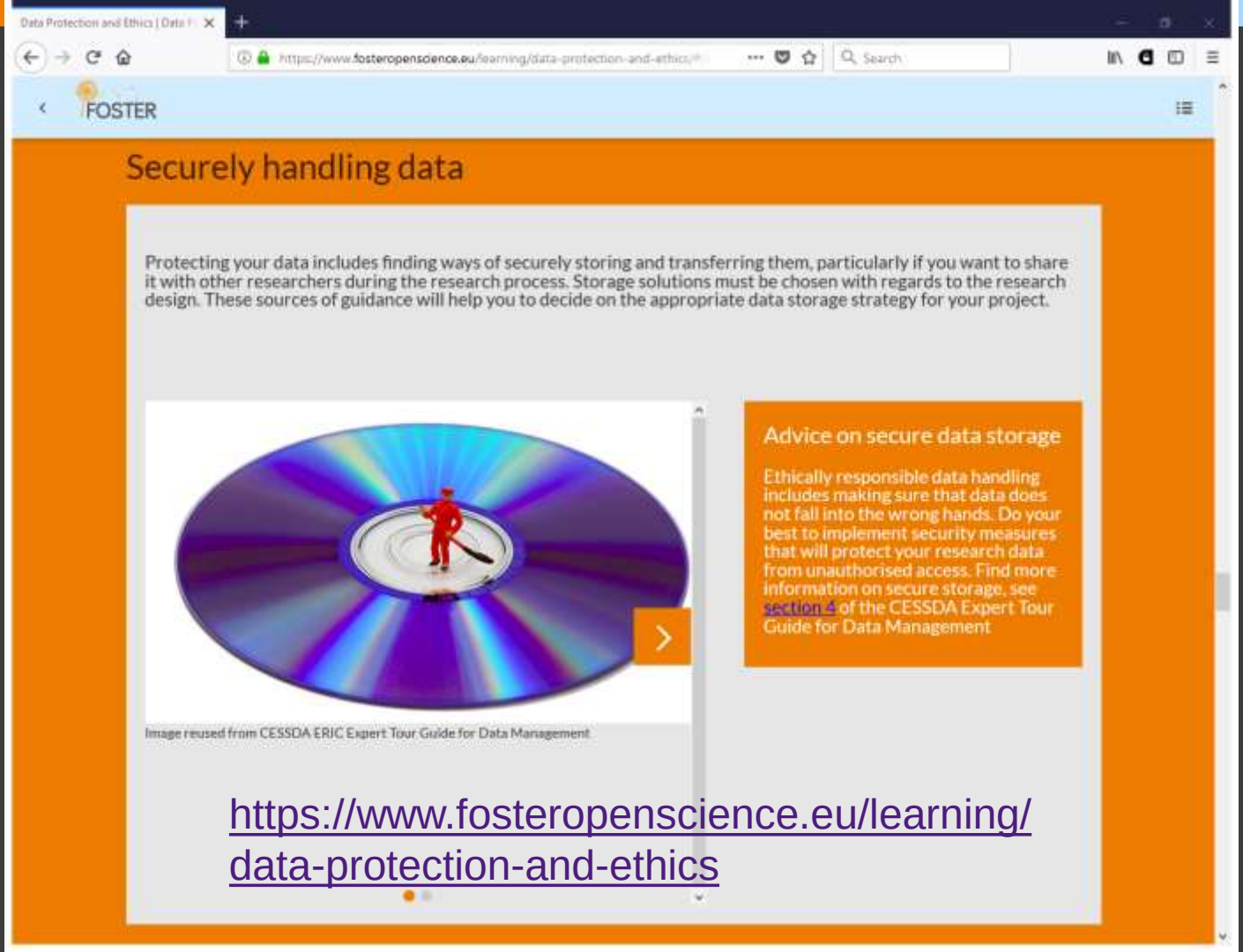


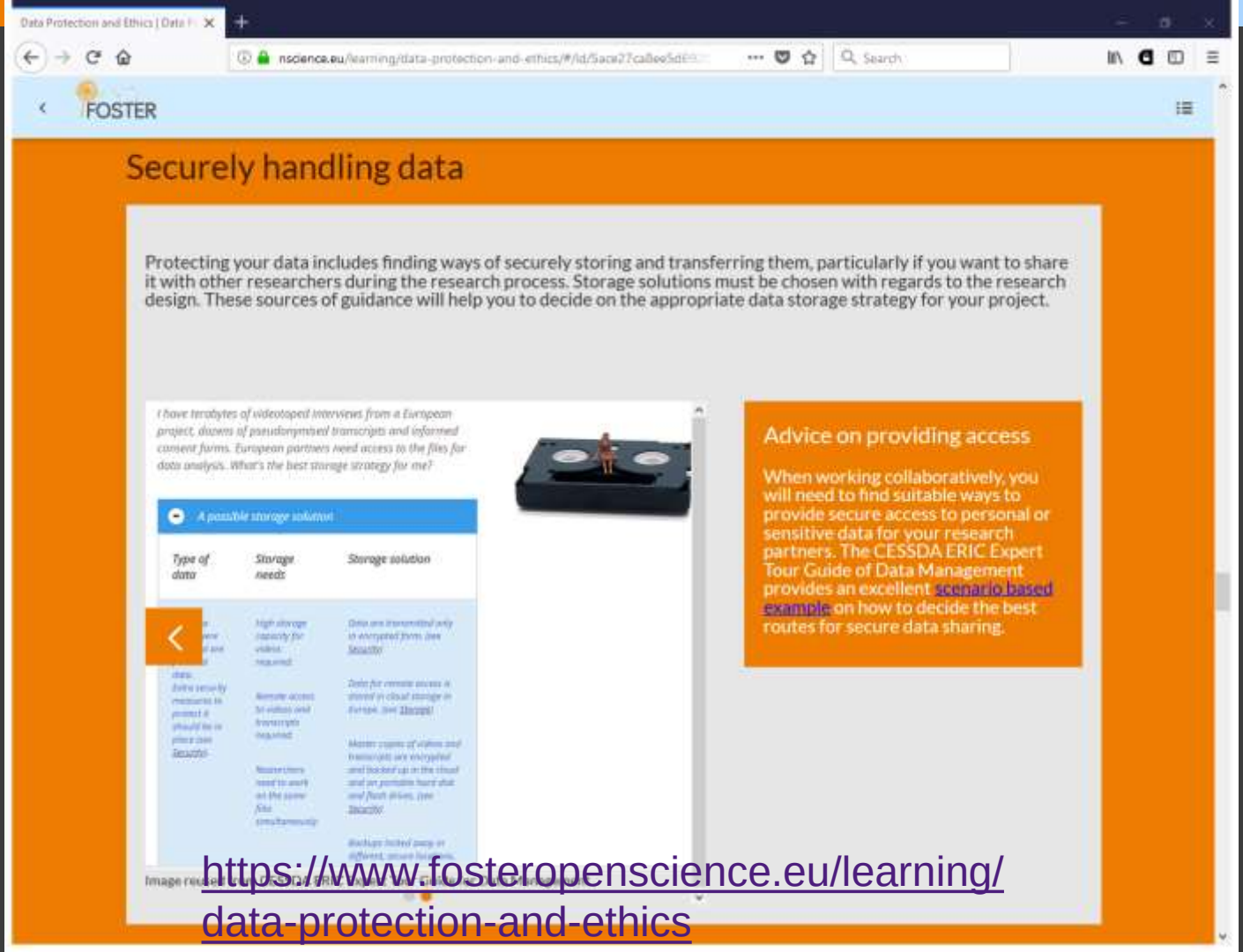
UK DATA ARCHIVE

Consent Form for [name of project]

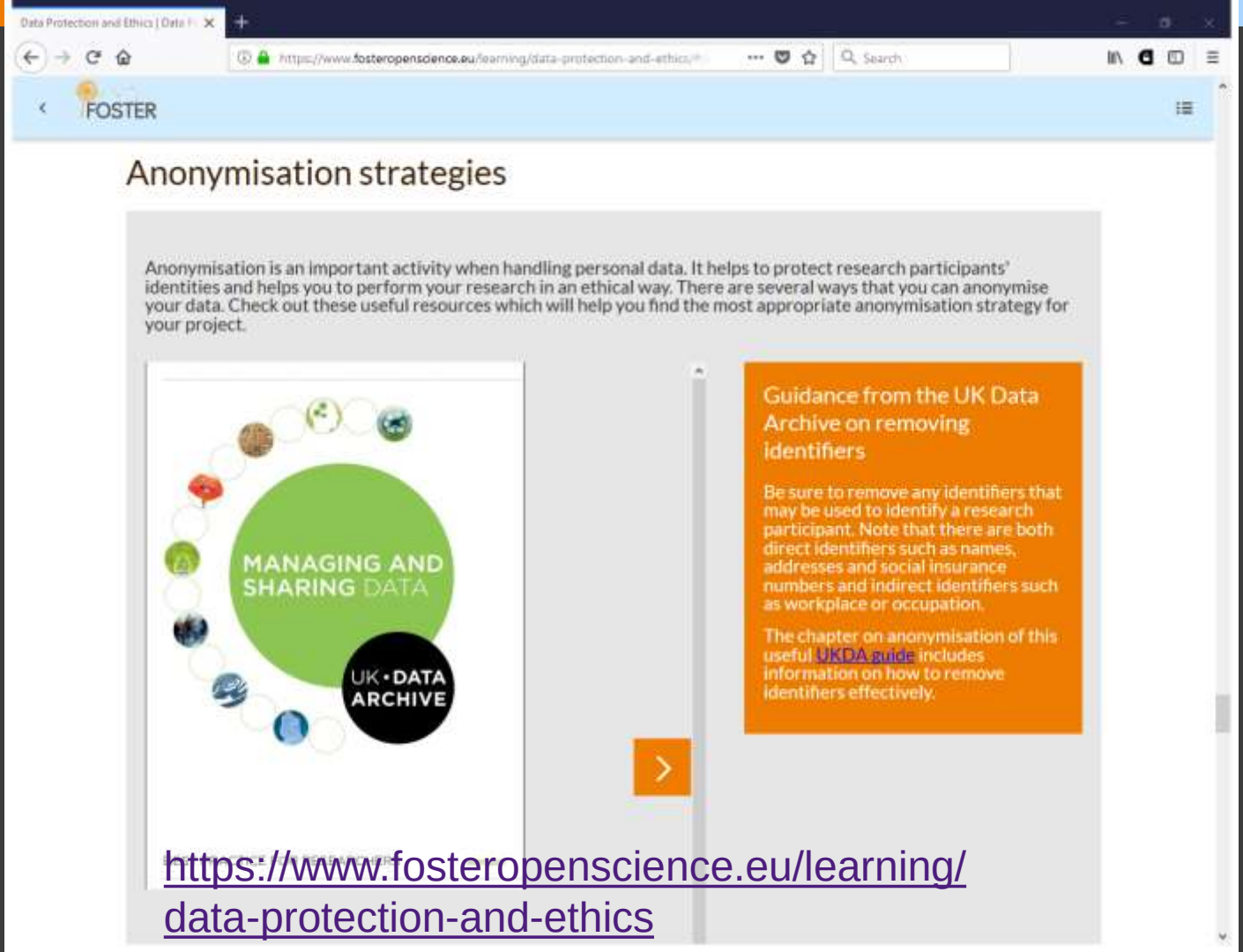
Please tick the appropriate boxes	Yes	No
Taking Part		
I have read and understood the project information sheet dated DD/MM/YYYY.	☐	☐
I have been given the opportunity to ask questions about the project.	☐	☐
I agree to take part in the project. Taking part in the project will include being interviewed and recorded (audio or video).	☐	☐
I understand that my taking part is voluntary; I can withdraw from the study at any time and I do not have to give any reasons for why I no longer want to take part.	☐	☐
Use of the information I provide for this project only		
I understand my personal details such as phone number and address will not be revealed to people outside the project.	☐	☐
I understand that my words may be quoted in publications, reports, web pages, and other research outputs.	☐	☐
Please choose one of the following two options:		
I would like my real name used in the above	☐	
I would not like my real name to be used in the above	☐	
Use of the information I provide beyond this project		
I agree for the data I provide to be archived at the UK Data Archive.	☐	☐
I understand that other authenticated researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I understand that other authenticated researchers may use my words in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
So we can use the information you provide legally		
I agree to assign the copyright I hold in any materials related to this project to [name of researcher].	☐	☐

Name of participant [printed] _____ Signature _____ Date _____


Data Protection and Ethics | Data
https://www.fosteropenscience.eu/learning/data-protection-and-ethics/
Search
FOSTER
Securely handling data
Protecting your data includes finding ways of securely storing and transferring them, particularly if you want to share it with other researchers during the research process. Storage solutions must be chosen with regards to the research design. These sources of guidance will help you to decide on the appropriate data storage strategy for your project.
Image reused from CESSDA ERIC Expert Tour Guide for Data Management
Advice on secure data storage
Ethically responsible data handling includes making sure that data does not fall into the wrong hands. Do your best to implement security measures that will protect your research data from unauthorised access. Find more information on secure storage, see section 4 of the CESSDA Expert Tour Guide for Data Management
https://www.fosteropenscience.eu/learning/data-protection-and-ethics



https://www.fosteropenscience.eu/learning/data-protection-and-ethics


Data Protection and Ethics | Data
https://www.fosteropenscience.eu/learning/data-protection-and-ethics/
Search
FOSTER
Anonymisation strategies
Anonymisation is an important activity when handling personal data. It helps to protect research participants' identities and helps you to perform your research in an ethical way. There are several ways that you can anonymise your data. Check out these useful resources which will help you find the most appropriate anonymisation strategy for your project.
MANAGING AND SHARING DATA
UK • DATA ARCHIVE
Guidance from the UK Data Archive on removing identifiers
Be sure to remove any identifiers that may be used to identify a research participant. Note that there are both direct identifiers such as names, addresses and social insurance numbers and indirect identifiers such as workplace or occupation.
The chapter on anonymisation of this useful UKDA guide includes information on how to remove identifiers effectively.
https://www.fosteropenscience.eu/learning/data-protection-and-ethics

Data Protection and Ethics | Data F

https://www.fosteropenscience.eu/learning/data-protection-and-ethics/

Search

FOSTER

Anonymisation strategies

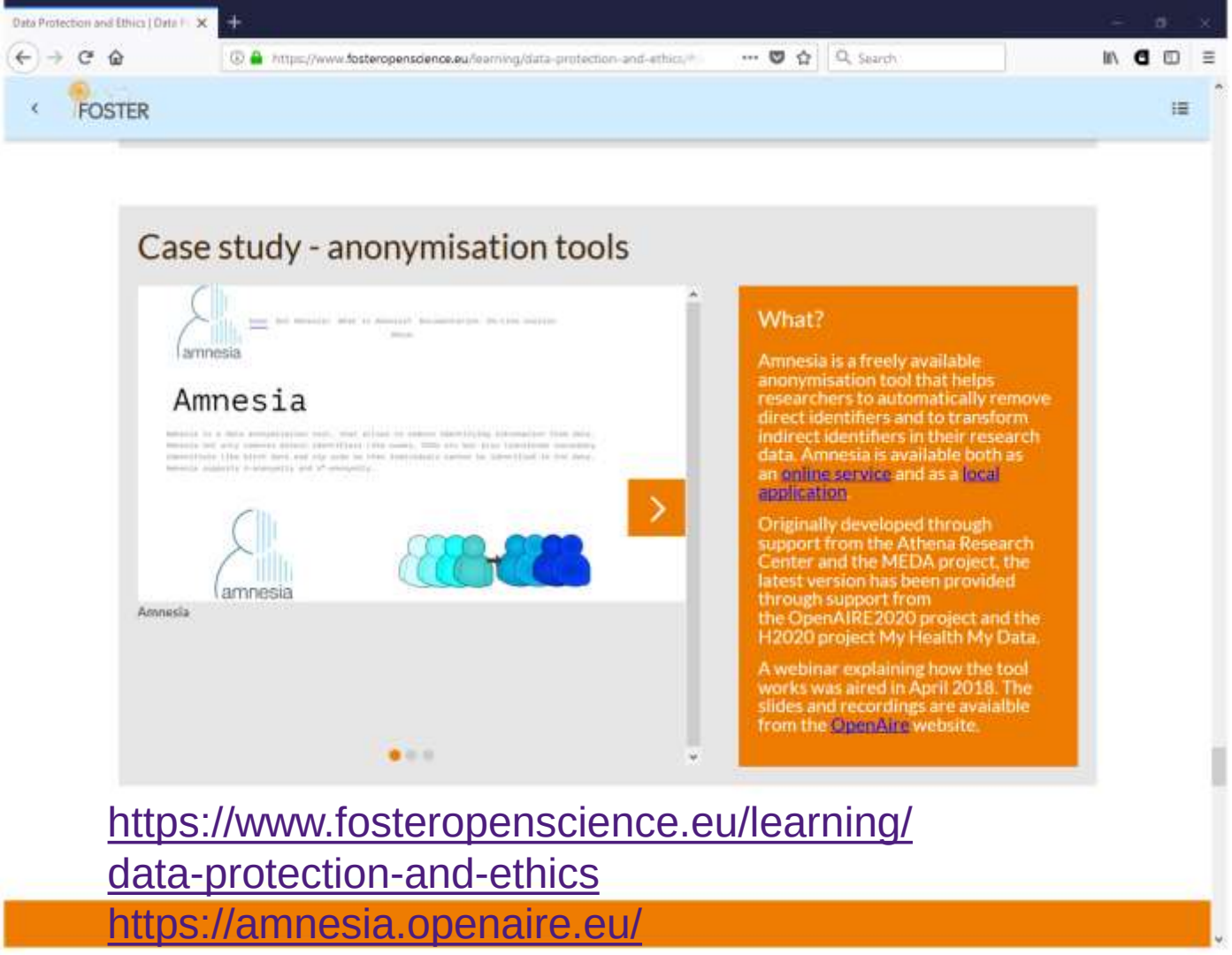
Anonymisation is an important activity when handling personal data. It helps to protect research participants' identities and helps you to perform your research in an ethical way. There are several ways that you can anonymise your data. Check out these useful resources which will help you find the most appropriate anonymisation strategy for your project.

Guidance from CESSDA ERIC on anonymising qualitative and quantitative data

The CESSDA ERIC Expert Tour Guide for Data Management provides very helpful advice about how to anonymise quantitative and qualitative data including aggregating variables, reducing precision, and the use of pseudonyms.

The CESSDA tutorial also provides excellent case studies, practical tips, and suggested approaches for anonymising data.

https://www.fosteropenscience.eu/learning/data-protection-and-ethics



https://www.fosteropenscience.eu/learning/data-protection-and-ethics

https://amnesia.openaire.eu/

FOSTER

Case study: working with messy data in the Humanities

Huge amounts of data are being collected and used by researchers in all disciplines - including the Humanities. Before researchers can analyse these large datasets for trends, they need to be able to make sense of the data.

− OpenRefine

What?

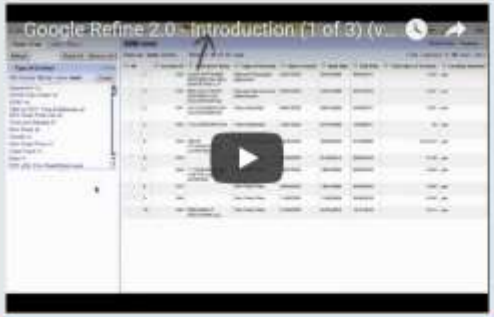
OpenRefine (formerly Google Refine) has been fully supported by volunteers since 2012. OpenRefine is a free and open source tool for working with messy data. Data can be cleaned, transformed into different formats, extended with other tools or linked to databases.

Why?

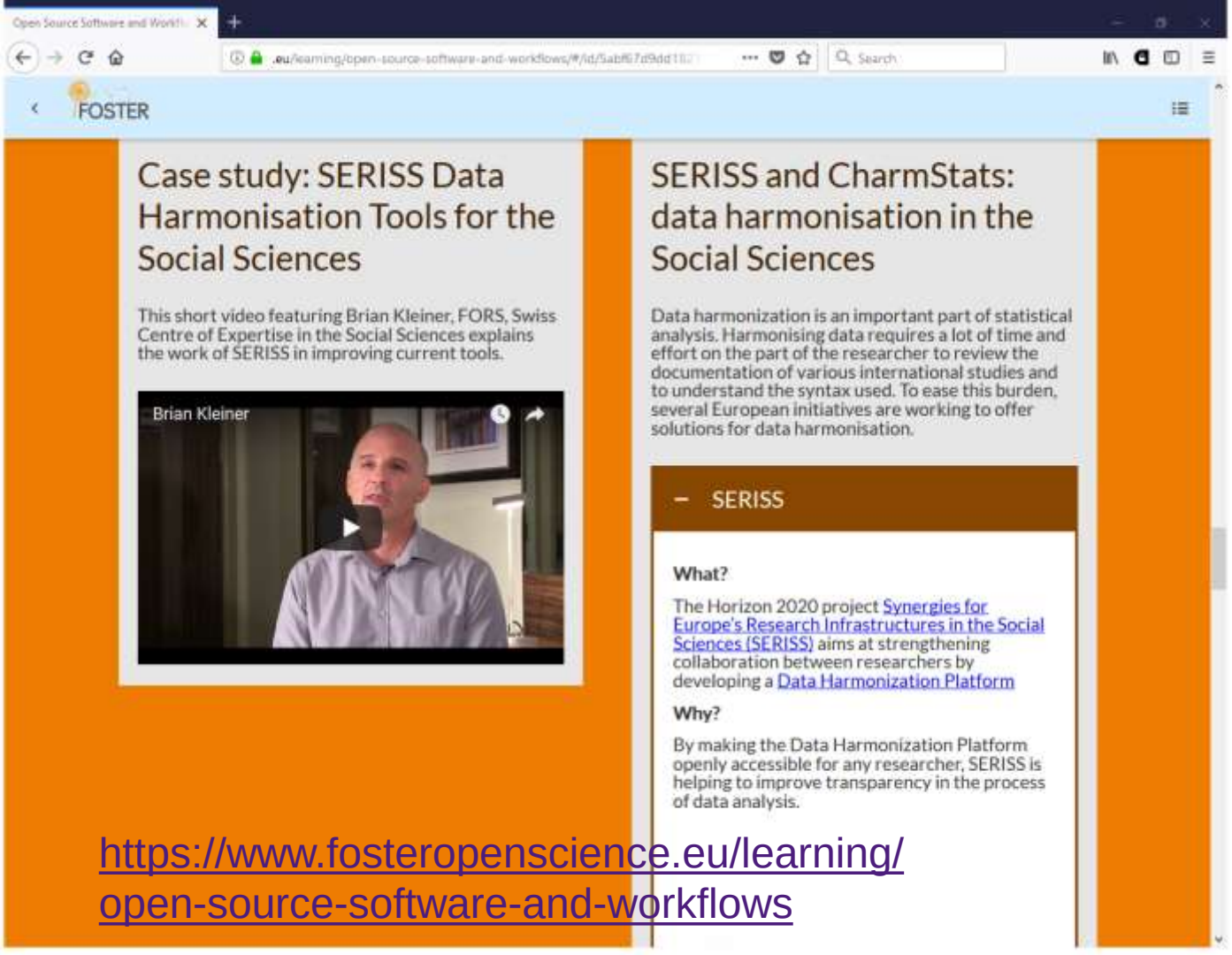
OpenRefine helps researchers clean and deal with large datasets which is essential when managing data and making them reusable.

OpenRefine

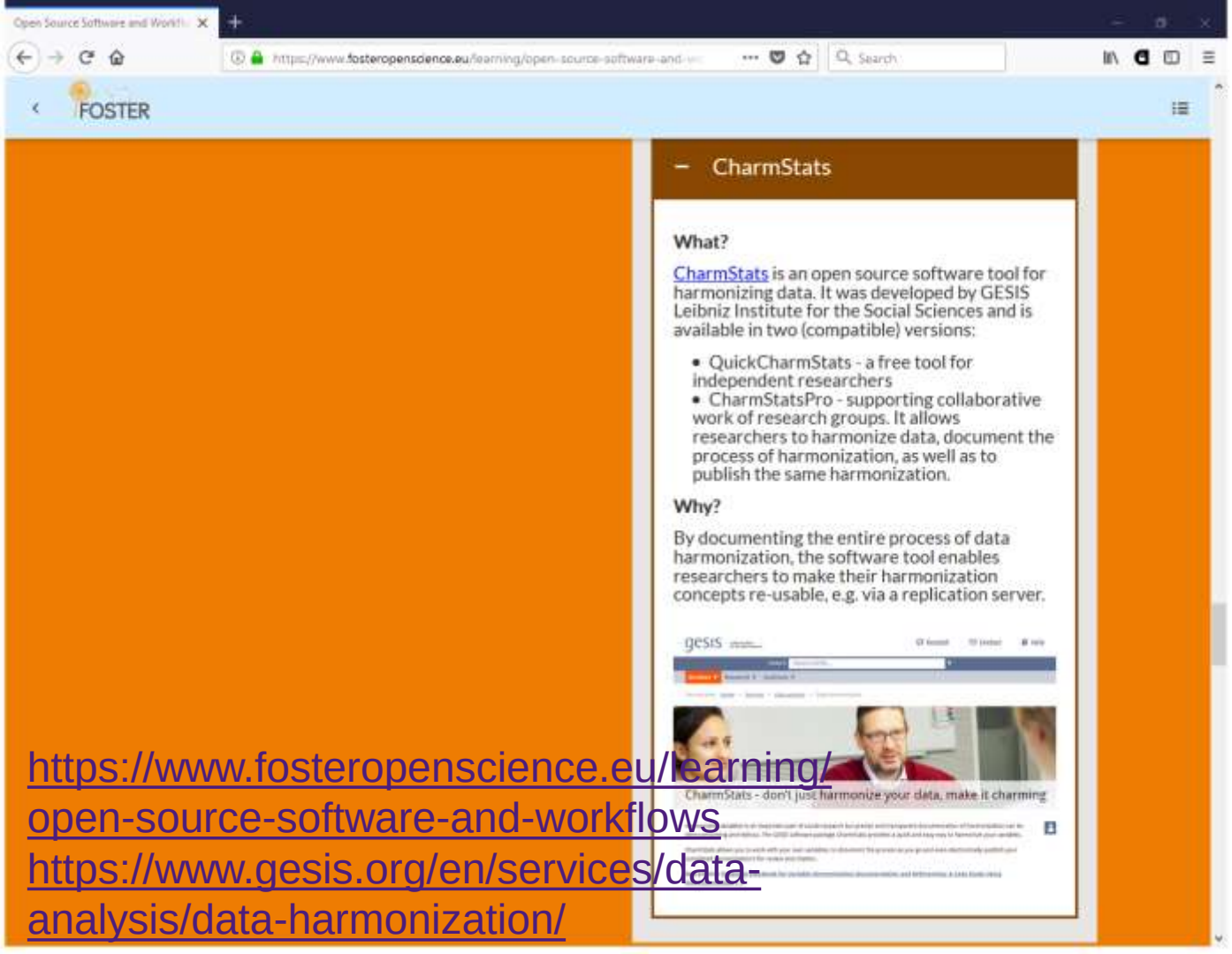
This short video introduces OpenRefine (formerly Google Refine).



https://www.fosteropenscience.eu/learning/open-source-software-and-workflows



https://www.fosteropenscience.eu/learning/open-source-software-and-workflows



https://www.fosteropenscience.eu/learning/open-source-software-and-workflows
https://www.gesis.org/en/services/data-analysis/data-harmonization/

Dissemination Material

For training events and conferences

Moo cards

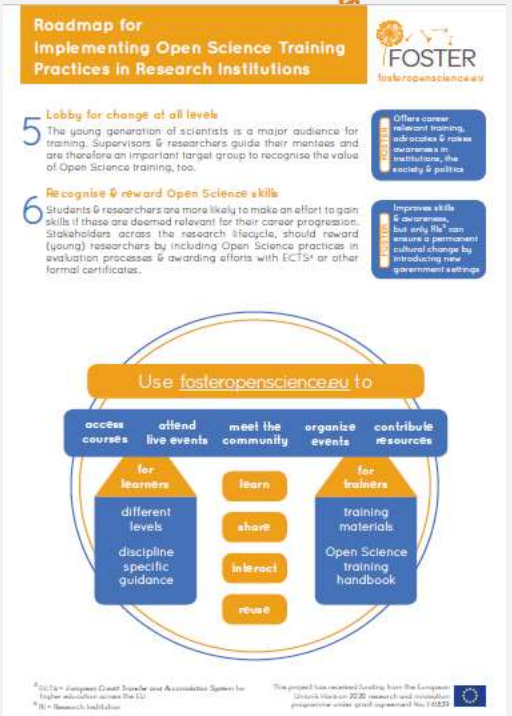
Fact sheets

Advocacy stickers

FOSTER

2-page policy briefing

- ‘Roadmap for implementing Open Science training practices in research institutions’
- 6 practical steps how to modernise OS qualification framework
- e.g. career stage appropriate & discipline specific training, incentives, regular and standardized training in researcher training



https://doi.org/10.5281/zenodo.1209175

FOSTER

Open Science Training Handbook

- Book sprint: ensure a finished book in only a few days
- In five days a book of 200 pages was written
- On last day pre-release online available for community to comment

FOSTER

Open Science Training Handbook

- Open Science Basics
  - Open Concepts & Principles
  - Open Research Data & Materials
  - Open Research Software & Open Source
  - Reproducible Research & Data Analysis
  - Open Access to Published Research Results
  - Open Licensing & File Formats
  - Collaborative Platforms
  - Open Peer Review, Metrics & Evaluation
  - Open Science Policies
  - Citizen Science
  - Open Advocacy
- Introduction
- On Learning & Training
- Organizational Aspects
- Examples & Practical Guidance
- Glossary
- References
- About the Authors & Facilitators

FOSTER

Open Science Training Handbook

- Now available as GitBook & for download (PDF, epub, mobi)
- CC 0 license to enable simple re-use

book.fosteropenscience.eu

https://doi.org/10.5281/zenodo.1212496

FOSTER

FOSTER

Thank you! Questions?

Facebook: @fosteropenscience

Twitter: @fosterscience

Youtube: FOSTER Open Science