





Open Science and IPR

Online training for researchers

11 November 2021



Miroslav Trajanović, Nikola Korunović, Dragan Mišić University of Niš, Faculty of Mechanical Engineering

Goals

- 1. An introduction to Open Science (OS)
- 2. Deeper insight into elements of OS through real-life examples and exercises

Instructions:

- Mute your microphone
- Switch-off your camera
- Use Q/A option to post a questions.



Why take a training on Open Science?



- The term "Open Science" relates to the movement and the approach to science it promotes, that became extremely important in last decade
- In Europe, this is recognized through a large number of projects and initiatives that support and promote the application of principles of Open Science
- It was mandatory to apply some of these principles within H2020 projects
- The principle of 'Open Science' became the modus operandi of Horizon Europe, requiring open access to publications and data



Topics

- 1. Principles and concepts of Open Science
- 2. Reproducible Research and Data Analysis
- 3. Open Science Policies
- 4. Citizen Science
- 5. Open Peer Review, Metrics and Evaluation
- 6. Open Access
- 7. Open Licensing, Intellectual Property Rights, Open File Formats
- 8. Open Research Data and Materials
- 9. Open Research Software and Open Source
- 10. Collaborative Platforms
- 11. Open educational resources
- 12. Open advocacy



	-
_	
	- 1
	-



Principles and concepts of Open Science

- What is Open Science?
- What are the aims of Open Science?
- Principles of Open Science
- History and motivation

- Open Science practices and pillars
- Open Science in the research process





What is openness?

 "Open means anyone can freely access, use, modify, and share for any purpose

(subject, at most, to requirements that preserve provenance and openness)."

(Open Definition)





What is Open Science?



- "The movement to make scientific research, data and dissemination accessible to all levels of an inquiring society." (<u>FOSTER Taxonomy</u>)
- "The practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods." (FOSTER)
- Open Science aims at transforming science through ICT tools, networks and media, to make research more open, global, collaborative, creative and closer to society. (<u>EC – Open Science Policy</u>)
- "Research simply done properly"



Good Research Practice (according to Medical Research Council)

Good researchers

- Strive for excellence and take responsibility
- Respect the law, research ethics, and professional standards
- Support a culture of <u>transparency</u>, openness, and honesty towards other researchers and the public
- Maximize public benefit and avoid resource waste
- Continue learning and mentor others



Open Science

OSC - LMU Open science center / <u>CC BY 4.0</u> Medical Research Council (2012) Picture from freepik.com by @macrovector





- Academic publishing began in the 17th century first academic journals.
- There exists increasing motivation to share resources between research disciplines and increase transparency, for greater efficiency, exactness, accountability, sustainability for future generations, and reproducibility.
- There are ethical cases whereby increased transparency can reduce fraud, data manipulation, and selective reporting of results.

Drivers of open science

- Pressure from research academies and governments for publicly-funded research to be shared more openly, often for the purpose of accelerated societal or economic growth and innovation.
- Need to **drive cultural change** in research and amongst researchers.
- Embracing of **Web-based tools and technologies** to facilitate scientific collaboration.



What are the aims of open science?



- To make research more open to participation, review/recognizing false research, improvement and (re)use for the world to benefit.
- To change the way research is done, who is involved and how it is valued.
- To transform science through ICT tools, networks and media in order to make research more open, global, collaborative, creative and closer to society.





- Increased transparency, re-use, participation, cooperation, accountability and reproducibility for research.
- Inclusion, fairness, equity, and sharing, to improve the quality and reliability of research.

Pros / Cons of Open Science?



- What is your opinion?
- Frequently mentioned Pros:
 - Easy and quick to get/share information / knowledge with wide audience
 - ...
- Frequently mentioned Cons:
 - Someone else getting detailed results of your work for free
 - IPR

• ...



Open Science and researchers

• Let's try and see Open Science from the point of view of the researchers!



This Photo by Unknown Author is licensed under CC BY-SA



Why should a researcher practice open science?



Wagenmakers et al. (2012)

Why transparency?

The Confirmatory Research Process





Why transparency?

How can you know that it does not look like this?



The science hamster wheel





Arslan (2018)

OSC - LMU Open science center / <u>CC BY 4.0</u>

get all material here: https://osf.io/zjrhu/

Open Science in the research process





Why transparency?





OSC - LMU Open science center / <u>CC BY 4.0</u> Pictures from freepik.com by @brgfx, @makyzz; flaticon.com by Icon Pond, Dimitry Miroliubov

Pillars of Open Science

	Open Data
	Open Material
_	Open Access
	Open Source (Software)
	Open Peer Review
	Open Educational Resources

Important practices

- Transparency in research methods
- Reproducibility of research
- Open Metrics
- Open Licensing and File Formats
- Citizen Science
- Open Science Policies
- Open Advocacy



Science

ben

Exercise – Principles of Open Science

 Q1: What are the benefits of applying Open Science principles to your project?

✓ Principles: transparency, re-use, participation, cooperation, accountability, reproducibility for research, inclusion, fairness, equity, sharing.

Note:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 5 minutes







Reproducible Research and Data Analysis

Reproducibility

- Reproducibility: the ability of a researcher to duplicate the results of a prior study using the same materials (or raw data) as used by original investigator.
- Directly applied to the scientific method, the cornerstone of Science, particularly to the following five steps:
 - 1. Formulating a hypothesis
 - 2. Designing the study
 - 3. Running the study and collecting the data
 - 4. Analyzing the data
 - 5. Reporting the study
- Each step should be clearly and openly documented, making the study transparent and reproducible.





Replicability

- Replicability: the ability of a researcher to duplicate the results of a prior study if the same procedures are followed but new data are collected.
- Reproducibility is methodsoriented, whereas replicability is results-oriented





Common issues with research reproducibility

- Data dredging (p-hacking) repeatedly searching a dataset or trying alternative analyses until a 'significant result is found.
- **Omitting null results** scientists or journals decide not to publish studies unless results are statistically significant.
- Underpowered study study too small to reliably indicate whether or not an effect exists no statistical power.
- Errors technical errors exist within a study, e.g. misidentified reagents or computational errors.
- Underspecified methods methods not shared with other scientists in enough detail, so others cannot precisely replicate the study.
- Weak experimental design study has methodological flaws, so it is unlikely to produce reliable or valid results.



Replication crisis (reproducibility crisis)

 Reproducibility crisis is a methodological crisis primarily affecting parts of the social and life sciences in which scholars have found that the results of many scientific studies are difficult or impossible to replicate or reproduce on subsequent investigation, either by independent researchers or by the original researchers themselves.

• Causes of the crisis

- Generation of new data/publications at an unprecedented rate.
- Compelling evidence that majority of these discoveries will not stand test of time.
- Failure to adhere to good scientific practice & the desperation to <u>publish or perish</u>.
- This is a multifaceted, multi stakeholder problem.
- No single party is solely responsible, and no single solution will suffice.



Implementing reproducibility into a research workflow – a medical science example

- 1. Plan for reproducibility before you start
- 2. Keep track of things
- 3. Share and license your research
- 4. Report your research transparently





1. Planning for reproducibility

- 1. Create a study plan or protocol with study design and methods
 - a) Use a **reporting guideline** (e.g. listed in <u>Equator Network</u> in medical science)
 - b) Track changes to your study plan or protocol using version control
 - **c)** Calculate the power or sample size needed and report the calculation (e.g. <u>clincalc.com</u>)
- 2. Choose **reproducible tools and materials** (open source)
- 3. Set-up a reproducible project
 - a) Centralize and organize project management using an online platform, a central repository, or folder for all research files
 - b) Use electronic lab notebook such as <u>Benchling</u>, <u>Labguru</u>, or <u>SciNote</u>



2. Keeping track of things

1. Registration

- a) Preregister important study design and analysis information to increase transparency and counter publication bias of negative results
- b) Free tools include <u>AsPredicted</u>, <u>Open Science Framework</u>, and <u>Registered Reports</u>

2. Version control

(see Open Research Data and Materials)

3. Documentation

- a) **README** file
- b) Data dictionary (codebook) example: <u>Karl Broman's Data Organization module</u>.

4. Literate programming

a) E.g. use <u>Jupyter Notebooks</u>, <u>KnitR</u>, <u>Sweave</u>, to integrate code and documentation



3. Sharing and licensing the research

1. Data

- a) Avoid supplementary files
- b) Decide on an acceptable permissive license
- c) Share data using a repository (see Open Research Data and Materials)

2. Materials

- a) Share materials so they can be reused
- b) Deposit reagents with repositories like <u>Addgene</u>, <u>The Bloomington Drosophila</u> <u>Stock Center</u>, and <u>ATCC</u> to make them easily accessible

3. Software, notebooks, and containers

- a) License the code to inform how it may be (re)used
- b) Share notebooks via services (e.g. <u>mybinder</u>) that allow public viewing and exec
- c) Share containers or notebooks with services such as Rocker or Code Ocean



4. Transparently reporting the research

- **1. Report and publish methods and interventions** explicitly and transparently and fully to allow for replication.
- 2. Guidelines from the <u>Equator Network</u>, tools like <u>Protocols.io</u>, or processes like <u>Registered Reports</u> can help in reporting reproducibly.
- **3.** Post results to public registration platform (such as <u>ClinicalTrials.gov</u> or the <u>SocialScienceRegistry</u>) within a year of finishing the study, no matter the nature or direction of your results.



Exercise - reproducibility and replicability

• Q1: Do you see any obstacles to reproducibility and replicability of your project contents? If yes, which? If no, why?

 Transparency is important for reproducibility in formulating a hypothesis, designing the study, running the study and collecting the data, analyzing the data, reporting the study

✓ Try to avoid p-hacking, omitting null results, underpowered studies, errors, unspecified methods, weak experimental design

Note:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 5 minutes







Open Science Policies

What are Open Science Policies?



- Strategies and actions aimed at promoting Open Science principles and at acknowledging Open Science practices
- Usually established by research institutions, research funders, governments or publishers
- Initial policies were based on the idea that results achieved from publicly funded research should be available to the public without any restriction
- Now the scope of the policies has grown:
 - National policies foster Open Science practices at any point of the research level
 - There are specific provisions in new and existing laws, regulations or directives



Stakeholders and relevant OS politics topics

• Researchers

- How Open Science policies affect them
- Policy makers
 - Designing and implement a policy to foster Open Science
- Funders or policy makers within an institution
 - How to design, develop, implement and monitor a policy


8 ambitions of the EU's open science policy

- Open Data. FAIR (Findable, Accessible, Interoperable and Re-usable data)
- European Open Science Cloud (EOSC) to store, share, process and reuse research digital objects that are FAIR
- New generation metrics. New indicators for research quality and impact, so as to do justice to open science practices.
 - Mutual learning exercise on open science altmetrics and rewards This exercise focused on defining:
 - alternative metrics to measure the qualities and impact of research outcomes
 - rewards for researchers to engage in open science activities



8 ambitions of the EU's open science policy (2)

- Future of scholarly communication. All scientific publications should be freely accessible, and the early sharing of research outputs should be encouraged.
- **Rewards.** Research career evaluation systems should fully acknowledge open science activities.
- **Research integrity & reproducibility of scientific results.** All publicly funded research in the EU should adhere to commonly agreed standards of research integrity.
- Education and skills. All scientists in Europe should have the necessary skills and support to apply open science research routines and practices.
- **Citizen science.** The general public should be able to make significant contributions and be recognised as valid European science knowledge producers.



EU policy and Open Science



- EU has recognized that linking science and society leads to advance in all areas.
- The importance of **open access to data and research results** has been present in Europe for a long time but is particularly prominent in the calls and recommendations related to H2020 projects. Within Horizon Europe, EC requires that all participants have to ensure Open Access to the publications generated during the research, while in the domain of open data it is recommended to enable Open Access whenever possible.
- The principle of 'Open Science' are the modus operandi of Horizon Europe, requiring open access to publications and data.
- In order for current recommendations to become reality in Europe, EC has implemented various projects and initiatives in last ten years.



OpenAIRE project

- **OpenAIRE (Open Access Infrastructure for Research in Europe)** is one of the pioneering projects, which started in 2006, under the name "**Driver**". Goals:
 - creating a network of digital repositories, for storing digital data
 - to support the implementation of open access in Europe
- **OpenAIREPlus** expanded the network of repositories, including specific scientific fields related ones
- OpenAIRE2020:
 - possibilities of finding and using the results of projects realized under the H2020 program
 - creation of services that enable the exchange of scientific data, literature and methods.
- **OpenAIREAdvance** continues to support Open Access and Open Data in Europe



Other Open Science related EU projects

- FOSTER, 2014, accelerating acceptance of an open approach in the context of the ERA. Showed that OA principles can be integrated into the research process
- FosterPlus to help accept the principles of open science in Europe
- The most recent and probably the most comprehensive initiative: European Open Science Cloud
 - Launched in 2015, opened 23.11.2018.
 - From 2020, it enables the creation of a virtual environment for all researchers in which they can store, manage, analyze and reuse their data, related to research, innovation and education



European Open Science Cloud (EOSC)

- European Cloud Initiative, a part of the package of measures for <u>Digitising</u> <u>European industry</u>, provides European science, industry and public authorities with:
 - a world-class data infrastructure to store and manage data (<u>European Open Science Cloud</u>);
 - high-speed connectivity to transport data (European Data Infrastructure); and
 - ever more powerful High Performance Computers to process data (<u>HPC</u>)
- A European <u>Open Science Cloud (EOSC)</u> offers Europe's researchers and science and technology professionals a virtual environment to store, share and re-use the large volumes of information generated by the big data revolution
- This is underpinned by the <u>European Data Infrastructure (EDI)</u>, deploying the high-bandwidth networks and the supercomputing capacity necessary to effectively access and process large datasets stored in the Cloud



cOALition S

- On 4 September 2018, a group of national research funding organizations, with the support of EC and the ERC, announced the launch of <u>cOAlition S</u>
- It is an initiative to make full and immediate Open Access to research publications a reality.
- It is built around <u>Plan S</u>, which consists of one target and <u>10 principles</u>.
- "With effect from 2021, all scholarly publications on the results from research funded by public or private grants provided by national, regional and international research councils and funding bodies, must be published in Open Access Journals, on Open Access Platforms, or made immediately available through Open Access Repositories without embargo"



Horizon Results Platform

- The newly revamped <u>Horizon Results Platform</u> is the central pillar of the EC-funded research results "exploitation ecosystem"
- Is Free you can promote the Key Exploitable Results* (KER's) of your projects
- Is Easy to use hosted under the Funding & Tenders Portal where you manage everything else for your projects
- Is a Matchmaking tool use the wide range of *flags* and *attention-grabbing features* to attract your target audiences
- Is Managed by you you can publish and update KERs whenever they are available,
- **Reaches out to many audiences** politicians, investors, researchers, scientists, scholars, entrepreneurs, financing experts, IP specialists, and other stakeholders visit regularly
- Triggers services and opportunities at no cost to you depending on the nature and needs of your result(s) it allows EC to inform you about relevant upcoming calls for proposals, pitching events with investors, possibilities for assistance with your dissemination plans, business development plans, innovation management, IP management, and many more.



Open Science and Horizon Europe

- General Model Grant Agreement (HE MGA)
- Version 1.0 01 June 2021
- ANNEX 5 : SPECIFIC RULES
- COMMUNICATION, DISSEMINATION, OPEN SCIENCE AND VISIBILITY (- ARTICLE 17)
- Among other things, this article specifies:
 - open access to scientific publications
 - research data management



Source: Exploitation & Open science in Horizon Europe

https://ec.europa.eu/research/participants/data/ref/h2020/other/events/2020-10-09/3_exploitation-ipr-

open_science_en.pdf

Open Science and Horizon Europe Peer-reviewed scientific publications

- Immediate open access through trusted repository (at the latest at the time of publication);
- publications licensed under CC BY (or equivalent); CC BY-NC/ND (or equivalent) allowed for long-text formats;
- Information provided via the repository about any research output, tool, or instrument needed to validate the conclusions of a publication;
- Beneficiaries/authors must retain sufficient IPR to comply with their OA requirements;
- Metadata licensed under CCO or equivalent, in line with FAIR principles (particularly machineactionable); PIDs (publication, authors, if possible their organizations and the grant).
- Only publication fees (if any) in full open access venues for peer-reviewed scientific publications are eligible for reimbursement



Open Science and Horizon Europe Research Data Management (RDM)

- Emphasis shifts from open research data to RDM
- No opting out of RDM. Projects generating research data MUST manage their data responsibly and in line with FAIR principles
- Open access to research data 'as open as possible as closed as necessary', i. e. there can be exceptions to open access to research data.
- Costs for RDM (for example data storage, processing and preservation) are eligible



Open Science and Horizon Europe Research Data Management (RDM)

- Establish and regularly update a Data Management Plan (DMP)
- Deposit data in a trusted repository and provide open access through it
- Deposit and open access ASAP and per DMP
- For some actions, additional obligation to deposit in a repository that is federated under EOSC
- CC BY or CC 0 (or equivalent) license required to open data
- Exceptions to open access (duly justified in the DMP; legitimate interests or constraints);
- Information via the repository about any other research output or any other tools and instruments needed to re-use or validate the data;
- Metadata requirements same as for publications (i.e. CCO and PIDs)

Open Science and Horizon Europe Open science: access for validation

- Obligatory provision of physical or digital access to data or other results needed for validation of conclusions scientific publications
- Legitimate interests/concerns must still be safeguarded



Open Research Europe

- open access publishing platform

- Developed through F1000Research project (https://f1000research.com/)
- Only for Horizon 2020 and Horizon Europe funded projects results across all subject areas
- The fee for publication on Open Research Europe is 780 EUR. All fees are paid centrally by the Commission.
- The transparent open peer review
- Platinum access, published under a CC BY license
- Open Research Europe sends peer-reviewed versions of publications to Zenodo,
- Article level metrics page demonstrating the individual article's reach, interest and 'quality, article citation data, views, downloads, social media and wider engagement.



Publishing process



Exercise – Open Science Policies

Q1: What is the web address of the page that describes the Open Science Policy of your institution or your country?

Note:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 5 minutes







Citizen Science

What is citizen science?



- Involvement of the non-academic public in the process of scientific research (<u>citizenscience.org</u>)
- Open science means sharing knowledge, data and tools as early as possible, not only between researchers and between disciplines, but also with society at large (EC definition)
- **Citizens** do scientific work—often working together with experts or scientific institutions
- Citizen science is **both an aim and enabler of Open Science**
 - Reading open access articles and using open research data leads to greater public understanding of science
 - Citizens support the collection, analysis or description of research data



Citizen science activities

Type of activity	Example
Passive sensing	Collecting data from smartphones and smartwatches of citizens
Volunteer Computing	Use of free capacities of citizens' computers for massive data processing
Volunteer Thinking	Classification all different animals caught in millions of camera trap images (Snapshot Serengeti project)
Environmental and Ecological Observation	Observation and identification of birds
Participatory Sensing	Citizens use their smartphone to monitor noise in different location
Civic/Community science	Citizens, bottom-up, proposal for research into the reasons for the reduction of biodiversity in the local community
Participation in studies	Pharmaceutical efficacy studies of new drugs
Participation in experiments	Crowdsourcing. Measuring speed of typing (OpenClick project)



Adapted from: Muki Haklay, Citizen Science and Policy: A European Perspective. Washington, DC: Woodrow Wilson International Center for Scholars, 2015

Motivations for participation in Citizen Science projects



Adapted from: Muki Haklay, Citizen Science and Policy: A European Perspective. Washington, DC: Woodrow Wilson International Center for Scholars, 2015

Intrinsic Motivation Factors	Extrinsic Motivation Factors
Unique ethos	Career
Learning	Strengthen social relations
Personal enrichment	Project goal
Self-actualization	Community
Self-expression	Identity
Self-image	Reputation
Fun	Monetary return
Recreation	Reciprocity
Instrumentality	System trust
Self-efficacy	Networking
Meeting own need	Socio-political
Freedom to express	
Altruism	

Further info on Citizen science

- Collaborative science collaboration between citizens and scientists.
- **Do-It-Yourself Science** (using research data by public) examples: patient innovation, patient activism/advocacy, NGOs and Civil Rights Groups.
- Classification by distinguishing scientist and non-scientist led activities (see <u>Outside the Academy – DIY Science Communities</u>).
- The public can also be engaged in **policy making**, e.g. agenda-setting for research systems' (see the <u>European Commission's Open Science Monitor</u>).
- <u>"Ten principles of citizen science</u>" by The European Citizen Science Association (ECSA)



Key elements when starting a citizen science project

- How are you going to engage citizens?
- How are you going to ensure data quality?
- How are you going to deal with ethics and legal issues?
- <u>Cases studies</u> extracted from the UK Research Excellence Framework, may serve as examples of how citizen science activities can be included as societal impact in evaluation reports



Open Science in practice: OpenClick project

- Examples in this presentation are based on <u>OpenClick</u> project by Faculty of Mechanical Engineering Niš
- Project OpenClick aims to discover physical status, skills and abilities of people based on human – computer interaction tests
- OpenClick is research project fully based on Open Science principles
- Registered persons can use OpenClick project infrastructure (tools, methodology, data sets) to conduct their own research as long as they respect project's principles, rules, IPR and licensing policy



Example: OpenClick and citizen science

- Go to <u>www.openclick.rs</u> portal
- Explore the portal structure



- Run OpenClick: Project OpenClick > Run OpenClick, or go directly to <u>http://openclick.masfak.ni.ac.rs/#/</u>
- Click on **Register New Account** and enter your data (registration is anonymized)
- Tick the box to **confirm data collection** and processing



Test yourself as a citizen

- Sign in with your username and password
- You can see the total number of participants and the total number of tests
- One participant can be tested unlimited number of times
- Carefully read the instructions for each test before you perform it



Available tests

- There are nine tests available
- Each test relates to an elementary operation in interaction with a computer using mouse and keyboard
- Our motto: Keep it simple
- All you need to test someone's physical status are computer with mouse and keyboard



Your first test

- Click on Mouse clicks
- Carefully read the instructions
- Perform the tests
- If you did not do a regular test, then **quit** the test (you were confused, your phone rang, any other disturbance)
- Save test results
- Select: Show as histogram
- Compare your data with average results



This test is stupid and useless?

- The physiotherapist told us that this is a great test that can be used for monitoring patient's recovery after a hand injury
- Neuropsychiatrist claims this is the ideal way to test whether he gave the right dose of a medicine to a patient with diagnosed schizophrenia
- The sports **coach** wants to use this test to track sportsmans' reflexes
- Each test has dozens of applications
- It is possible to create a mix of tests for specific research



More tests

- Mouse positioning speed
- Selecting two keys simultaneously
- Text input using the keyboard
- Alternate selection of the mouse buttons



Exercise – Citizen science

- Q1 How are you going to engage citizens?
- Q2 How are you going to ensure data quality?
- Q3 How are you going to deal with ethics and legal issues?
- Notice:
 - Use Exercise_template.docx to type the answers
 - Give short answers one to two rows long
 - Duration of exercise 10 minutes



Back to Contents



Open Peer Review, Metrics and Evaluation

Evaluation of academic work

- To be a **researcher** is to find oneself **under constant evaluation**
- Evaluation:
 - 1. Of a piece of work
 - 2. Of researchers themselves
- Both research and researcher are evaluated through two primary methods:
 - **1. Peer review** (qualitative)
 - 2. Metrics (quantitative)



Peer review

- Peer review is used primarily to judge pieces of research
- It is the formal quality assurance mechanism whereby scholarly manuscripts (e.g., journal articles, books, grant applications and conference papers) are made subject to the scrutiny of others
- Reviewers' feedback and judgements are used to improve works and make final decisions regarding selection (for publication, grant allocation or speaking time)



Open peer review

- Being a peer reviewer presents researchers with opportunities for:
 - engaging with novel research,
 - building academic networks and expertise, and
 - refining their own writing skills
- In general, researchers do not often receive formal training in how to do peer review
- Open peer review (OPR) aims to bring greater transparency and participation to formal and informal peer review processes
- Even where researchers believe themselves confident with traditional peer review, open peer review present new challenges and opportunities
- An example: <u>https://royalsocietypublishing.org/rspa/open-peer-review</u>



Aspects of open peer review

- There exists a number of **overlapping ways** in which peer review models (double-blind, single-blind) may be adapted to the aims of open science.
- Various features of open review are (first 2 are the most important):
 - 1. "open identities", where both authors and reviewers are aware of each other's identities (i.e., non-blinded)
 - 2. "open reports", where review reports are published alongside the relevant article
 - 3. **"open participation"**, where members of the wider community are able to contribute to the review process,
 - 4. "open interaction", where direct reciprocal discussion between author(s) and reviewers, and/or between reviewers, is allowed and encouraged,
 - 5. **"open pre-review manuscripts"**, where manuscripts are made immediately available in advance of any formal peer review procedures (either internally as part of journal workflows or externally via preprint servers)



Potential advantages for reviewers and authors

- In combination, open identities and open reports are theorized to lead to better reviews, as the thought of having their name publicly connected to a work or seeing their review published encourages reviewers to be more thorough
- Open identities and open reports enable reviewers to gain public credit for their review work, encouraging this activity and allowing review work to be cited in other publications and linked to promotion and tenure
- Open participation could overcome problems associated with editorial selection of reviewers (e.g., biases, closed-networks, elitism). For early career researchers, such open processes may also present a chance to build their research reputation and practice their review skills


Potential pitfalls of OPR

- Open identities removes anonymity conditions for reviewers (single-blind) or authors and reviewers (double-blind) which are traditionally in place to counteract social biases. It's important for reviewers to constantly question their assumptions to ensure their judgements reflect only the quality of the manuscript, and not the status, history, or affiliations of the author(s). Authors should do the same in receiving peer review comments.
- Giving and receiving criticism is often a process filled with unavoidably emotional reactions authors and reviewers may subjectively agree or disagree. The transparency could inflame such difficulties. Reviewers should communicate their points in a clear and civil way, in order to maximize the chances that it will be received as valuable feedback by the author(s).
- Lack of reviewers' anonymity in open identities might subvert the process by discouraging reviewers from making strong criticisms, especially against higher-status colleagues.
- Given these issues, potential reviewers may be more likely to decline to review.



Journal ranking

- Journal ranking is widely used in academic circles in the evaluation of an academic journal's impact and quality.
- Journal rankings are intended to
 - reflect the place of a journal within its field,
 - the relative difficulty of being published in that journal, and
 - the prestige associated with it.
- They have been introduced as **official research evaluation tools in several countries**.



Journal-level metrics

- Impact factor (IF) and <u>CiteScore</u> reflecting the average number of citations to articles published in science and social science journals.
- <u>SCImago Journal Rank</u> (SJR)— a measure of scientific influence of scholarly journals that accounts for both the number of citations received by a journal and the importance or prestige of the journals where such citations come from.
- <u>h-index</u> usually used as a measure of scientific productivity and the scientific impact of an individual scientist, but can also be used to rank journals.



Metrics and open science

- Research publications are often the primary measure of a researcher's work
- General assessment is often based on **metrics** such as:
 - the number of citations that researcher's publications collect (h-index).
 - perceived level of prestige of the journal the paper was published in (quantified by the Journal Impact Factor).
- Metrics used to evaluate research (e.g. JIF, h-index) do not measure and do not reward - open research practices, are not as open and transparent as the community would like.
- The <u>San Francisco Declaration on Research Assessment (DORA)</u> recommends considering all types of output and using various forms of metrics and narrative assessment in parallel.



Alternative metrics

- Alternative Metrics" or <u>altmetrics</u> a proposal for balanced assessment of research efforts
- Citation counting is complemented by other online measures of research impact, including bookmarks, links, blog posts, tweets, likes, shares, press coverage...
- A particular advantage to early-career researchers, whose researchimpact may not yet be reflected in significant numbers of citations
- Helps in identification of influential research and potential connections between researchers



Challenges of altmetrics

- Lack of robustness and susceptibility to 'gaming'.
- Any measure ceases to be a good measure once it becomes a target ('Goodhart's Law').
- Relative lack of social media uptake in some disciplines and geographical regions.
- A reliance on commercial entities for the underlying data.



Exercise – Metrics and Evaluation

Q1: What type of metrics would you prefer to be used for the evaluation of the papers related to your project?

Notice:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 3 minutes







to Published Research results

Image by art designer at PLoS, modified by Wikipedia users Nina, Beao, and JakobVoss - <u>http://www.plos.org/</u> / <u>CC0</u>

Journals and copyright



- Many journals still require for publication that authors transfer full copyright
- Transfer of rights implies that authors must ask for permission to reuse their own work beyond what is allowed by the applicable law unless there are some uses already granted
- Granted uses include teaching purposes, sharing with colleagues, and especially how researchers can self-archive their papers in repositories



Image by art designer at PLoS, modified by Wikipedia users Nina, Beao, and JakobVoss -<u>http://www.plos.org/</u> / <u>CCO</u>

What is Open Access?

- Research publications (like articles and books) can be accessed online, free of charge by any user, with no technical obstacles (like mandatory registration or login to specific platforms)
- Ideally, additional rights should also be provided: to copy, distribute, search, link, crawl and mine
- Various modes of open access: gold, green, etc.





Open Access journals



- According to the Directory of Open Access Journals (DOAJ), in June 2021 there are 16,477 open access journals
- An open access journal **must provide free access** to its contents, but it **also must license them to allow reusability**. No legal notice must be legally understood as "all rights reserved" (see <u>DOAJ basic criteria</u> for inclusion).
- The definition of an OA journal does not include any condition about the business model, but these are commonly known as journal where you have to pay to publish.
 - This misconception is due to the fact that the **most successful journals and the ones** that got the highest impact follow this model.
 - A study shows that the majority of journals registered in DOAJ do not charge any fee for publication (data available <u>here</u>).



Gold Open Access (Open Access publishing)



- The published work is made available in Open Access mode by the publisher immediately upon publication
- The most common business model is based on one-off payments by authors (commonly called **APCs article processing charges**)
- The publisher charges a fee directly from the author or more often indirectly from the institution or the financier of the project



Green Open Access (self-archiving)

- Due to the increase of policies requiring access to research results
- The **published work** or the **final peer-reviewed manuscript**, **accepted for publication** is made freely and openly accessible by the author, or a representative, **in an online repository**
- Usually, it is not allowed to publish the final manuscript, a.k.a. the version of record, prepared for printing by the publisher, but only pre-prints (pre-refereeing) or post-prints (final draft post-refereeing)
- Some publishers request that Open Access be granted only after an embargo period (several months to several years)



Preprints

- Preprints are documents that have not been peer reviewed but are considered as a complete scientific publication in a first stage.
- Publication through institutional repositories and preprint servers.
- Some of the preprints servers include **open peer review services** and the availability to post new versions of the paper once reviewed by peers.
- Open peer review processes of preprints enabled on new publishing platforms (see <u>Wellcome Trust</u> or the <u>Bill and Melinda Gates Foundation</u>).
- Some publishers distinguish between **preprints** and **author-submitted articles** and do not consider the preprints as a form of prior publication.



Embargo period

- In relation to the moment to make the paper publicly available, many journals establish a period from its original publication: the embargo period, which can range from zero to 60 months
- Some journals include or exclude embargoes depending on the versions
 - For instance, the accepted version could be made publicly available after publication, but the published version must wait 12 months



Repositories and self-archiving



- In October 2020 more than 4700 repositories available, according to the <u>Registry of Open Access Repositories</u>.
 - **1.** Institutional repositories managed by research institutions to provide a place to archive and share openly papers and other research outputs
 - 2. Subject based repositories usually managed by research communities and most of the contents are related to a certain discipline.
 - **3.** Harvesters aggregate content from different repositories. Sites to perform general searches and build other value-added services.
- A good choice is EU-cofounded repository <u>ZENODO</u>



Example: OpenClick related papers

- OpenClick portal > Project OpenClick > Published Papers
- Link: <u>http://www.openclick.rs/index.php/en/project-openclick/published-papers</u>
- Two self –archived conference papers



Other Open Access modes

- Hybrid Open Access. Journals charge a subscription or the cost of downloading articles, but allow authors open access to their article by additional payment of processing fees
 - Institution may pay twice: for the subscription to the entire magazine and for publishing one or more articles
- Platinum Open Access. Does not charge either a subscription or a fee from the author. Expenses are covered by other means, such as volunteer work, donations, subsidies, grants, etc.





"Bronze" open access



- Free to read on the publisher page, but without a clearly identifiable license
- Like Gold and Hybrid OA, Bronze OA articles are publisher-hosted
- Unlike Gold OA, Bronze articles are not published in journals considered open access in the <u>Directory of Open Access Journals</u> (DOAJ)
- Unlike Hybrid, Bronze articles carry no license information. Although this may not be intentional, without an identifiable license the articles are free to read but do not allow extended reuse rights beyond reading
- It is also not clear if Bronze articles are temporarily or permanently available to read for free



"Predatory" publishers

- This is a proposition by American librarian <u>Jeffrey Beall</u> to label publishers publishing in an open-access regime, charging mainly high publishing fees but not providing a proper review process and quality control.
- Beall's list was removed under the pressure of publishers who thought they were unjustly accused. A similar list is currently available at <u>this address</u>.
- Predatory editorial policy may be illustrated by detailed analysis of two journals from Bosnia and Herzegovina that were once removed from the Web of Science: <u>Legitimacy of citations in predatory publishing: The case of proliferation of papers</u> <u>by Serbian authors in two Bosnian WoS-indexed journals</u> by Center for Evaluation in Education and Science.



How to determine self-archiving policy of a journal?



- 1. At any of the portals that provide general information as <u>SHERPA/RoMEO</u>
 - Sherpa Romeo is an online service that aggregates and presents publisher and journal open access policies from around the world.
 - Used by researchers, repository staff and research support teams across the world, to help users understand complex publisher and journal open access policies.
- 2. Based on the information available at the journal or publisher website
- There are still some researchers reluctant to deposit other versions than the final published version. It is important to inform them about the copyright implications when they sign a transfer document.



Open Access publication example (1)

Task: check Open Access options for the paper "Mišić, D., Zdravković, M., Mitković, M., Vitković, N., & Mitković, M. (2018). Real-Time Monitoring of Bone Fracture Recovery by Using Aware, Sensing, Smart, and Active Orthopedic Devices. *IEEE Internet of Things Journal*, *5*(6), 4466-4473."

Solution:

- Google the paper (using Google Scholar). Only the abstract is available for free on publisher's site. Login is required for download.
- Use <u>SHERPA/RoMEO</u> (v2 available from 20th June 2020) to find about *"IEEE Internet of Things Journal"* policies:
 - Type or paste "IEEE Internet of Things Journal" in "Journal Title or ISSN" field and press Search button. You will get the information page related to the journal.
 - First section, **Publication Information**, contains basic data about the journal: **Title**, **ISSN**, **URL** and **Publishers**.
 - Next, and the most important section, **Publisher Policy**, contains a set of possible scenarios through which the article can be made open access. Those scenarios are called pathways.



Open Access publication example (2)

- Pathways are sorted by paper versions: Published Version, Accepted Version (postprint) and Submitted Version (preprint sent to publisher).
- For some publication types, more than one pathway may appear, which are mutually different regarding funder's request, location of publication or some other condition.
- Pathways are initially shown in one row each, where the rows contain the icons describing various aspects of publisher policy (see the image on the right or <u>Sherpa Romeo user guide</u>).
- By clicking the "+" sign on the right, pathway description gets expanded to show detailed info.

Icon	Name	Description
Ē	Open access publishing	The pathway includes open access publishing
£	Additional open access fee	The pathway requires the payment of a fee (in addition to any normal publication fees that may be required) to make the article open access
×	Not permitted	No open access pathway exists for the article version
Ô	Copyright owner	The copyright owner that the pathway requires
: : :	Conditions	Conditions that apply to the pathway
റ ീ	Licence	The licence that the pathway requires
!	Prerequisites	Requirements that must be met to allow the pathway to be used. These may include prerequisite funders, subjects, or permissions from the publisher
	Location	The websites on which the pathway allows the article version to be available. This includes self-archiving and publisher-deposit locations, including the website of the journal
Ø	Notes	Additional notes on the policy
Ŧ	Publisher deposit	The publisher will deposit on your behalf in the location specified
X	Embargo	The embargo that the pathway requires. Unless stated otherwise, the embargo starts on the date of publication



Open Access publication example (3)

- In this case, there are five pathways, one for accepted version, two for postprints and two for preprints.
- Published version pathway is shown as follows:

Published Version	£ 🖹 🗏 None 🔂 CC BY 🗹 ☞ Journal Website
£ OA Fee	This pathway has an Open Access fee associated with it
🖹 OA Publishing	This pathway includes Open Access publishing
🛛 Embargo	No Embargo
🔁 Licence	CC BY
🔁 Location	Journal Website
ピ Notes	Authors must contact publisher stating Creative Commons Attribution requirement

 In short, published version may be made Open Access if a fee is payed (the amount must be sought on publisher's website), which makes this journal "hybrid", as normally there exists a journal subscription fee. CC BY license is included, and publisher defines the attribution.



Open Access publication example (4)

• Accepted version pathway (a) is shown as follows:

Accepted Version [pathway a]

Ξ	Embargo
©	Copyright Owner
B	Location

None Image: Second state
 Any Repository, arXiv, Institutional Website, +3

No Embargo

Publishers

Any Repository Author's Homepage Institutional Repository Institutional Website

Named Repository (arXiv, TechRxiv)

- Accepted versions may be put in any repository, institutional repository of website or author's homepage, without embargo.
- Publisher retains copyright. After acceptation publisher's statement must be included as well as a link to publisher's versions via DOI.
- Accepted version pathway (b) is relevant if funder requests a specific publication location. Embargo of 24 months for posting in funders repository may be shortened, which depends on the funder and location.

When accepted for publication, set statement to accompany deposit (see policy) Must link to publisher version with DOI Publisher copyright and source must be acknowledged



Open Access publication example (5)

• Submitted version pathway (a) is shown as follows:

Submitted Version [pathway a]	None 🔚 None Tesignated Location, TechRxiv, +3	 Submitted version must be replaced by accepted version,
🛛 Embargo	No Embargo	when it becomes available.
► Location	Author's Homepage Funder Designated Location Institutional Website Named Repository (arXiv, TechRxiv) Preprint Repository	 Statements must be added after submission and acceptation, details must be sought on publisher's website.
¥∃ Conditions	Must be replaced with accepted version When submitted, set statement to accompany deposit When accepted for publication, set statement to accom IEEE must be informed as to the electronic address of t Publisher copyright and source must be acknowledged	(see policy) npany deposit (see policy) he pre-print 1

• Submitted versions may be put in listed locations, without

Open Access publication example (6)

• Submitted version pathway (b) is shown as follows:

Submitted Version [pathway b]	 None E Academic Social Network 	 Academic be signate Principles
🛛 Embargo	No Embargo	a signator
🔁 Location	Academic Social Network	
₹ ≡ Conditions	Must be removed upon acceptance for publication	
🕜 Notes	Academic Social Network must be a signatory of STM Sharing	Principles

- Submitted versions may be put on Academic Social Networks, without embargo, but must be removed after acceptance for publication.
 - Academic Social Network must be signatory of STM Sharing Principles (ResearchGate is not a signatory)



Open Access publication example (7)

- If journal policy is explored further (links below pathways) (<u>Author Posting</u> of IEEE Copyrighted papers online, IEEE Copyright Policy, Policy: Posting Your Article) it may be found that the accepted version may be posted only on author's personal website, author's employer's website, arXiv.org or funder's repository (with embargo period of 24 months that may be shortened if the funder requests so).
- Authors should ask the publisher for article version that includes the Digital Object Identifier, IEEE's copyright notice, and a notice showing the article has been accepted for publication.
- Preprint could have been published on Scholarly Collaboration Networks (SCNs) before acceptance. It could not have been published on ResearchGate, as it is not the signatory to the International Association of Scientific, Technical and Medical Publishers' "Sharing Principles".



Open Access publication example (8)

• Submitted or accepted versions must be accompanied with set phrase (see <u>Policy: Posting Your Article</u>):

"© 20XX IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works."

- Preprints have to be removed after the accepted paper is published.
- Open Access APCs are \$2,045 as of 1 January 2019.



Open Access Exercise

- For a given paper from the list on next page, do the following:
 - 1. Google the paper and find its original journal page and DOI. Use Google Scholar if needed to filter the search results.
 - 2. Check if the journal is listed in DOAJ is it an Open Access journal?
 - 3. Find Open Access policy for the journal using <u>SHERPA/RoMEO</u>. Answer the following questions (using main course presentation as an example):
 - a. Which paper versions may be published as open access (submitted version, accepted version, publisher's version/PDF) and under which conditions?
 - b. If there is OA publishing cost (APC), how much is it? If the journal is not purely Open Access, is there a paid Open Access option? If yes, use given link to find the fee on publisher's site. Which kind of Open Access journal is this?
 - c. Using the links in "Copyright" section find further information related to self-archiving, if applicable:
 - What is the embargo period for the paper (journal)?
 - Is the article free to access after the embargo period (from publisher's site)?
 - How to attach a user license to a post-print?
 - 4. Use the site <u>How can I share It?</u> and paper DOI to find out where and how the paper from the journal can be shared.
 - 5. If the journal is not fully Open Access, find the journal (or publisher) page defining paper publishing options. Where can the post prints be shared immediately?



Papers for Open Access exercise

- Korunovic, N., Marinkovic, D., Trajanovic, M., Zehn, M., Mitkovic, M., & Affatato, S. (2019). In Silico Optimization of Femoral Fixator Position and Configuration by Parametric CAD Model. *Materials*, 12(14), 2326.
- Korunović, N., Fragassa, C., Marinković, D., Vitković, N., & Trajanović, M. (2019). Performance evaluation of cord material models applied to structural analysis of tires. *Composite Structures*, 224, 111006.



Answers (1)

Korunovic, N., Marinkovic, D., Trajanovic, M., Zehn, M., Mitkovic, M., & Affatato, S. (2019). In Silico Optimization of Femoral Fixator Position and Configuration by Parametric CAD Model. Materials, 12(14), 2326.

- 1. URL: <u>https://www.mdpi.com/1996-1944/12/14/2326</u> DOI: <u>10.3390/ma12142326</u>
- 2. The journal is listed in DOAJ it is an Open Access journal
- 3. Open Access policy for the journal
 - a. Submitted version, accepted version or publisher's version/PDF may be published.
 - b. Open Access Conditions
 - Submitted and accepted versions may be self-published in any repository, without embargo. License is CC BY 4.0 and authors retain copyright. Both versions must contain a link to published article. Authors are encouraged to submit their published articles to institutional repositories.
 - Published version may me self-published under same conditions as above. In addition, those are automatically published in journal website and PubMed Central. Published source must be acknowledged with citation.



Answers (2)

- c. From DOAJ, the publishing fee is 2000 CHF (updated 9 Mart 2021). On publisher's site (<u>https://www.mdpi.com/journal/materials/apc</u>), the same price may be found (June 2020). This is a gold open access journal. Further information related to self-archiving (section Copyright > Policy, redirecting to <u>https://www.mdpi.com/about/openaccess</u>)
 - This is a full open access journal.
 - The article is free to access from publisher's site.
 - All paper versions may be self-published immediately, no embargo period exists.
 - Everyone is free to re-use the published material if proper accreditation/citation of the original publication is given.
 - The paper is published under CC BY license (details in course presentation)



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

4. <u>How can I share It?</u> still does not have info on MDPI journals.



Answers (3)

Korunović, N., Fragassa, C., Marinković, D., Vitković, N., & Trajanović, M. (2019). Performance evaluation of cord material models applied to structural analysis of tires. *Composite Structures*, 224, 111006.

- 1. URL: <u>https://www.sciencedirect.com/science/article/pii/S0263822319308840</u> DOI: <u>10.1016/j.compstruct.2019.111006</u>
- 2. The journal is **not** listed in DOAJ it is not an Open Access journal
- 3. Open Access policy for the journal (according to Sherpa Romeo v1, please compare to v2 output)
 - a. Only pre-prints and post-prints may be published, not publisher's version/PDF, under the following conditions:
 - Authors pre-print on any website, including arXiv and RePEC
 - Author's post-print on author's personal website immediately
 - Author's post-print on open access repository after an embargo period of between 12 months and 48 months
 - Permitted deposit due to Funding Body, Institutional and Governmental policy or mandate, may be required to comply with embargo periods of 12 months to 48 months
 - Author's post-print may be used to update arXiv and RepEC
 - Publisher's version/PDF cannot be used
 - Must link to publisher version with DOI
 - Author's post-print must be released with a Creative Commons Attribution Non-Commercial No Derivatives License



Answers (4)

- b. There is a paid Open Access option, The fee is 2850 USD. This is hybrid open access journal.
- c. Further information related to self-archiving
 - Embargo period is 24 months (Journal Embargo Period List (pdf))
 - The article is not free to access from publisher's site after the embargo period, it is not on the list find in <u>"Green open access</u> > Select a journal that features an <u>open archive</u>"
 - To attach a user license to a post-print:

On the accepted manuscript add the following to the title page, copyright information page, or header /footer: © YEAR. Licensed under the Creative Commons [insert license details and URL]. For example: © 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/ (Sharing Policy > How to attach a user license)

- 4. According to <u>How can I share It?</u>:
 - a. You can share the metadata and abstract as well as a link to the article on the respective publishers' platforms on: Academia.edu, ArXIV, BioRxiv, Center for Open Science, EndNote, Figshare, Institutional Repository, LabArchives, LOOP, Mendeley, Paperhive, Papers; PubMed Central, ReadCube, RefWorks, REPEC, SSRN and Trellis
 - b. You can share your **author's original (preprint)** on: **Academia.edu**, ArXIV, BioRxiv, Center for Open Science, EndNote, Figshare, Institutional Repository, LabArchives, LOOP, **Mendeley**, Paperhive, Papers, PubMed Central, ReadCube, RefWorks, REPEC, **ResearchGate**, SSRN adn Trellis
 - c. You can share the **accepted manuscript (post-print)** on: ArXIV, BioRxiv after the embargo expired, Center for Open Science after the embargo expired, **Institutional repository author showcasing (public)**, Institutional Repository group collaboration (private), **Mendeley author showcasing (public)**, REPEC and SSRN after the embargo expired
 - d. You can share the version of record on: Mendeley group collaboration (private)
 - e. Additional full-text sharing options may exist please check this directly with publisher Elsevier



Answers (5)

- Journal (or publisher) page defining paper publishing options: <u>https://www.elsevier.com/about/policies/sharing</u>. Read this in addition to <u>How can I share It?</u> info. If any conflict exists with previously found data, this is the most relevant, as it comes directly from publisher's site.
 - The post-prints can be shared immediately:
 - via their non-commercial personal homepage or blog
 - by updating a <u>preprint</u> in arXiv or RePEc with the <u>accepted manuscript</u>
 - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
 - directly by providing copies to their students or to research collaborators for their personal use
 - for private scholarly sharing as part of an invitation-only work group on <u>commercial sites with which Elsevier has an agreement</u>


Exercise – Open Access

- Q1: Select an OA journal in which you would like to publish the results of your project. Explain the choice shortly.
 - Use <u>DOAJ</u> and some of the following portals: <u>WebOfScience</u> (requires free account), <u>Scopus</u> (requires free account) or <u>Scimago Journal & Country Rank</u>. You may also use other portals, including country-specific ones.

Notice:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 5 minutes







Open Licensing, Intellectual Property Rights, Open File Formats

What is Intellectual property?

- Intellectual property (IP) is a category of property that includes intangible creations of the human intellect
- Intellectual property encompasses two types of intellectual property rights (IPR) :
 - industrial property rights (trademarks, patents, designations of origin, industrial designs and models)
 - copyright



Source: Wikipedia

Copyright ③

- Copyright is a form of intellectual property that grants the creator of an original creative work an exclusive legal right to determine whether and under what conditions this original work may be copied and used by others, usually for a limited term of years
- As **animal-made art**, this <u>monkey selfie</u> is ineligible for copyright in the United States





What is open licensing?

- A license is a legal document that grants specific rights to user to reuse and redistribute a material under some conditions
- Licenses can be applied to any material where some exploitation or usage rights exist (e.g., sound, text, image, multimedia, software)
- Free content licenses are licenses that grant permission to access, re-use, and redistribute material with few or no restrictions



Open license and copyright

- Any creative work is automatically copyrighted, i.e. under copyright terms "all rights reserved" - copyright holder *reserves*, or holds for its own use, all the rights provided by copyright law
- By specifying an open license copyright terms are changed to "some rights reserved", i.e. copyright holder can choose how the work is shared
- Applying an open license to a scientific work (an article, dataset or other type of research output) is a way for the copyright holder to express the conditions under which the work can be accessed, re-used and modified
- Since copyright laws are not internationally harmonized one must refer to the applicable laws in current context



Creative Commons (CC) licenses



- The most used licenses for scientific content are <u>Creative Commons</u> <u>licenses</u>
- Depending on CC <u>license type</u>:
 - one is always obliged to attribute the authors (give credit, indicate changes)
 - material may be free or not free to adapt (remix, transform, and build upon the material)
 - **Commercial use** may be allowed or not allowed
- There are also <u>public domain related CC licenses</u> "all rights granted"



CC licenses – License Conditions (1)

• All CC licenses require that others who use your work in any way must give you credit the way you request, but not in a way that suggests you endorse them or their use. If they want to use your work without giving you credit or for endorsement purposes, they must get your permission first.



Attribution (by)

• You let others copy, distribute, display, perform, and modify your work, as long as they distribute any modified work on the same terms. If they want to distribute modified works under other terms, they must get your permission first.



Source: Creative Commons licenses, licensed under CC BY 4.0.

CC licenses – License Conditions (2)

- NonCommercial (nc)
- You let others copy, distribute, display, perform, and (unless you have chosen NoDerivatives) modify and use your work for any purpose other than commercially unless they get your permission first.



• You let others copy, distribute, display and perform only original copies of your work. If they want to modify your work, they must get your permission first.

This movie explains the effect of various license conditions on sharing and reuse of CC licensed work.



Types of CC licenses

Attribution CC BY

Attribution — ShareAlike CC BY-SA

Attribution — NoDerivs CC BY-ND

.

Attribution — NonCommercial CC BY-NC

Attribution — NonCommercial — ShareAlike CC BY-NC-SA

CC BY-NC-ND

Attribution — NonCommercial — NoDerivs

Attribution (CC BY)

- Attribution ShareAlike (CC BY-SA)
- Attribution-NoDerivs (CC BY-ND)
- Attribution-NonCommercial (CC BY-NC)
- Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)
- Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)



Public domain CC licenses

"No Rights Reserved"



No Known Copyright"

- **CCO**. Use this universal tool if you are a holder of copyright or database rights, and you wish to waive all your interests that may exist in your work worldwide. Because copyright laws differ around the world, you may use this tool even though you may not have copyright in your jurisdiction but want to be sure to eliminate any copyrights you may have in other jurisdictions.
- **Public Domain Mark**. Use this tool if you have identified a work that is free of known copyright restrictions. Creative Commons does not recommend this tool for works that are restricted by copyright laws in one or more jurisdictions.



Compatible CC licenses



Two, differently CC licensed, works may be combined if corresponding rows and columns cross at the field containing the green checkmark.

Use at least the most restrictive licensing of the two.



CC licenses used for scientific work

- <u>CC BY license</u> (attribution only) a good option for scientific works like articles, books, working papers, and reports
 - Lets others distribute, remix, tweak, and build upon the work, even commercially, as long as they credit the author for the original creation
 - Recommended for maximum dissemination and use of licensed materials
- <u>CCO</u> recommended **for datasets and databases**
 - Built for jurisdictions where a full public domain dedication is not possible (e.g. in many continental Europe countries).
 - Creative Commons CC0 Public Domain Dedication waives copyright interest in a work you've created and dedicates it to the world-wide public domain
 - Used to opt out of copyright entirely and ensure the work has the widest reach



An example of CC license in Open Science (1)

Task: based on your scientific work, you (John Smith) created a web site "Internet of things and electric cars". You want to enable intensive sharing of your ideas without copyright issues and to be credited by others for your work.

Solution

- Go to Creative Commons site to choose a license (Share your work). (you may also want to try beta version, <u>Chooser beta</u>, which may be more comprehensive)
- Below "Allow adaptations of your work to be shared?" choose Yes.
- Below "Allow commercial uses of your work?" choose Yes.
- You should have selected <u>Attribution 4.0 International</u> license. Click on license link to see the simple description, named "the **Commons Deed** (also known as the "human readable" version of the license)". There are two more layers that describe the same license: Legal Code, written in the formal language of law, and Machine Readable, a summary that software systems can understand (see details here).



An example of CC license in Open Science (2)

- Solution (continued)
 - Examine license elements (that explain what others should be aware of or what should they do if they share or use your work):
 - You are free to:
 - **Share** copy and redistribute the material in any medium or format.
 - Adapt remix, transform, and build upon the material for any purpose, even commercially.
 - Under the following terms:
 - Attribution You must give <u>appropriate credit</u>, provide a link to the license, and <u>indicate if</u> <u>changes were made</u>. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
 - No additional restrictions You may not apply legal terms or <u>technological measures</u> that legally restrict others from doing anything the license permits.
 - On the bottom of page click on <u>use the license</u>.



An example of CC license in Open Science (3)

- Solution (continued) (also visible on <u>choose a license</u> page)
 - Choose which style of button you'd like on your webpage (choose left icon)
 - If you are satisfied with preview of the license, copy the code from grey window to your web site, to be shown in an appropriate place (e.g. on the bottom of each page).
 - The code to copy should look like:

This work is licensed under a Creative Commons Attribution 4.0 International License.

• The license should appear like:





Attributing CC licensed work

- Most CC licenses require that the authors of the shared work are attributed
- Some authors note how the attribution should be made, but most of the times the one that is attributing should know to create the attribution
- The wiki-how article <u>How to Attribute a Creative Commons Licensed</u> <u>Work</u> gives step-by-step instructions on how to:
 - **1.** Find information for attribution
 - 2. Create the attribution
 - 3. Provide data to help others attribute someone's own work



An example – attributing CC licensed work (1)

Task: You are a scientists who likes the work of John Smith from the <u>previous example</u>, and want to use it on you web site. On one page, you will feature the exact copy of John's work. On the other, you will remix his material, by extracting some chapters and using them as bulleted points. You want to create and use a proper attribution for both pages.

Solution:

- Find licensing information on John Smith's site. You should find the text "This work is licensed under a <u>Creative Commons Attribution 4.0 International License</u>."
- Click on license link and **check license elements**. You should see that you are free to share and adapt, if you attribute the author and indicate if changes were made.
- Go to wikiHOW page How to Attribute a Creative Commons Licensed Work.
- Click on "Sample Attributions" and on <u>General Creative Commons Attribution</u>, where you can find attribution example for the case where no changes were made.



An example – attributing CC licensed work (2)

Solution (continued):

- Read the text. On the bottom of the page, find the **attribution template**: "[Title, including link to original source]" by [author, including profile page link if possible] is licensed under [license, linked to license deed if possible].
- Use the template to create attribution. If you need further explanations, read the content of <u>How to Attribute a Creative Commons Licensed Work</u>.
- The attribution should look like: "Internet of things and electric cars by John Smith is licensed under <u>CC BY 4.0</u>". In previous sentence, the title of the work should be linked to John's web page about electric cars and his name should be linked to his profile web page.
- Return to wikiHOW page <u>How to Attribute a Creative Commons Licensed</u> <u>Work</u>.



An example – attributing CC licensed work (3)

Solution (continued):

- Click on "Sample Attributions" and on <u>Creative Commons Attribution for</u> <u>Modified Material</u>, where you can find attribution example (if changes were made).
- Read the text. On the bottom of the text, find the **attribution sample**, which is similar to previous one but contains a "/" in the end followed by indication of what was changed.
- Use the template to create attribution. If you need further explanations, read the content of <u>How to Attribute a Creative Commons Licensed Work</u>.
- The attribution should look like: "Internet of things and electric cars by John Smith is licensed under CC BY 4.0 / Original chapter separated into bulleted points"



Helping others attribute your work

• You may use <u>Help others</u> <u>attribute you!</u> link on <u>choose a license</u> page of Creative Commons site to help others attribute your work.





Reusing the contents of CC BY licensed scientific papers

- Figures, tables or photographs: may be reused, remixed, built upon etc., for any purpose, if you give credit to the original authors
- Text, may be reused and built upon, but there is still academic code of good conduct. Unmarked literal quotes and missing references in a conference paper, for example, are clear cases of plagiarism and lacking originality (Wild&Wilson 2013)



Fig. 17.9 Design variables (trochanteric bar length (a) and clamp spacing (b)) and specific points on CAD model of femur (Korunovic et al. 2019)¹

¹ In Silico Optimization of Femoral Fixator Position and Configuration by Parametric CAD Model by N. Korunovic et al., is licensed under <u>CC BY</u> / "Dynamic hip screws" changed to "Lag screws"



Software licensing

- Creative Commons licenses should not be used for licensing software because they were not designed for that purpose, as the organization states
- Software developers should use appropriate licenses like those collected by the Open Source Initiative or Free Software Foundation (<u>https://choosealicense.com</u>)



Example: OpenClick License

- License text : <u>http://www.apache.org/licenses/LICENSE-2.0</u>
- License information: <u>http://www.openclick.rs/index.php/en/resources/licensing</u>
- Provides rights to use and modify the software
 - all changed files must be marked
- Attributation only and patent rights granted
 - patent rights are withdrawn if litigation is initiated



An example of software attribution – attributing OpenClick

Simple attribution of OpenClick portal, if used within other content, should be written as:

"Openclick.rs" and included data are provided by Faculty of Mechanical Engineering, University of Nis, Serbia and licensed under <u>APACHE 2.0</u>.



Useful links related to CC licensing

- Considerations for Licensors if you are licensing your own work
- Considerations for Licensees if you are using someone else's work
- Open Content A Practical Guide to Using Creative Commons
 Licenses/Guide
 - How do open content licenses work? How do I choose the most suitable license for my individual needs? Where can I find open content online? And more...



File formats and open science



- A file format is a standard way that information is encoded for storage in a computer file
- Not all formats have freely available specification documents, partly because some developers view their specification documents as trade secrets
- Within the context of Open Science, files should not be compressed or encrypted and should avoid proprietary or patent-encumbered formats in favor of **open formats**. This ensures the access and reusability of the content



Examples for open file formats

- Text: TXT, ODT, PDF/A, XML
- Tabular data: CSV, TSV
- Image: TIFF, PNG, JPG 2000, SVG, WebP
- Audio: WAV, FLAC, OPUS
- Video: MPEG2, Theora, VP8, VP9, AV1, Motion JPG 2000 (MJ2)
- Binary hierarchical data: HDF5





Intellectual property rights (IPR) and Open Science

- The copyright holder is the one who **can decide to lift restrictions** if they are not lifted by default through the licenses
- Regarding research outputs, the copyright holder can be a researcher, a publisher, a scientific society, a research institution, a funder, etc.
- It is important to define who holds copyright or any other related rights of the research output
- There exist different intellectual property policies in universities and public research institutions



IPR and Open Science conflict

- Throughout their careers, researchers receive different messages from their institutions:
 - research education and utilization, which refers to the basic responsibility of making knowledge available to the public;
 - making knowledge accessible to very specific stakeholders, such as in the case of contracted research, which can lead to patent applications;
 - transmitting knowledge through innovation projects such as start-ups or other commercial or non-commercial projects;
 - making knowledge **available through Open Science** infrastructures, such as databases, massive open online courses and open software.



Source: IPR, Technology Transfer & Open Science, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC106998/kj1a28661enn.pdf

IPR and Open Science conflict (2)

- The current approach of the European Commission: make the data *as open as possible and as closed as necessary*
- This approach requires that a line is drawn to determine what can reasonably be considered as an acceptable argument for opting out of the by-default disclosure of research outputs
- Exceptions defined through consortium agreements
- Case-by-case solution
- Universities to provide IPR stewardship



Source: IPR, Technology Transfer & Open Science, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC106998/kj1a28661enn.pdf

Exercise – Open Licensing and IPR

 Q1: How will the paper in the journal you selected be licensed? What will you be allowed and what will you be not allowed to do with the paper?

Notice:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 5 minutes



Image by Jonathan Gray - Photo on flickr: https://www.flickr.com/photos/jwyg/4528443760/ CC0





Open Research Data and Materials

What is open research data?

- Data that can be freely accessed, reused, remixed and redistributed, for purposes of academic research and teaching and beyond
- Ideally, open data have no restrictions on reuse or redistribution, and are appropriately licensed as such
- At most, the requirement to **attribute** and **share alike** are present





Metadata

• Metadata is "data [information] that provides information about other data" (Wikipedia).

Errors per ten digits, Time for one digit, User ID, Test date, Gender, Right hand, Birth year

0.0000,0.4199,1,2019-03-20 17:07:49,male,1,1964 0.3333,0.4922,1,2019-02-04 00:00:00,male,1,1964 0.0000,0.2975,2,2019-02-06 00:00:00,female,1,1965 0.0000,0.7907,4,2019-03-09 00:00:00,male,1,1954 0.0000,1.1121,4,2019-02-26 00:00:00,male,1,1954 0.0000,0.6516,4,2019-04-15 23:02:24,male,1,1954 0.0000,0.6259,4,2019-03-11 11:59:39,male,1,1954 0.0000,0.6259,4,2019-02-06 00:00:00,male,1,1974 0.0000,0.6056,6,2019-02-06 00:00:00,male,1,1974 0.0000,0.6013,7,2019-02-06 00:00:00,female,1,1997 0.0000,0.3179,7,2019-02-06 00:00:00,female,1,1997 0.0000,0.3435,8,2019-02-06 00:00:00,male,0,1976





Metadata for typing numbers test

Attribute	Description
Time required for typing numbers	
Errors per ten digits	number of errors per 10 entered digits
Time for one digit	the time needed for entering of one digit
User ID	user identification
Test date	Date and time of the test
Gender	User gender. This field can have two values: male and female
Right hand	this field can have values 0 and 1. Value 1 indicates that the user's dominant hand is right. 0 indicates that the dominant arm is left
Birth year	year of birth of the user


FAIR data principles

- Findable: easy to find the data and the metadata for both humans and computers. Enabled by machine-readable persistent identifiers (PIDs) and metadata
- Accessible: data can be retrieved using open protocols, possibly including authentication and authorization
- Interoperable: can be combined and used with other data or tools
- **Re-usable:** well-described so that they can be replicated and/or combined in different settings



Difference between FAIR data and Open data

- Open data should be available to everyone to access, use, and share, without licences, copyright, or patents. At most, it should be subject to attribution/share-alike licenses
- FAIR data, uses the term "Accessible" to mean accessible by appropriate people, at an appropriate time, in an appropriate way. Data can be FAIR when it is private, when it is accessible by a defined group of people, or when it is accessible by everyone (open data). For example:
 - New experimental data accessible by the generator and their group to start, with consortia partners as the findings become refined, with the public upon publication
 - Personally sensitive data may never be publicly accessible and usable
 - **Commercially sensitive data** may be held privately for stretches of time after collection and interpretation. Users are also free to use more restrictive licenses to govern how the data may be reused



Sharing sensitive and proprietary data

- With appropriate **data management planning**, much sensitive and proprietary data can be shared, reused, and FAIR
- The metadata can almost always be shared
- Guidance and best practices for sharing sensitive data are regionspecific because of differing regulations
- How to share personal data, copyright and database (issues across some European countries): <u>The CESSDA Expert Tour Guide on Data</u> <u>Management</u>



Sensitive personal data

- Within the General Data Protection Regulation (<u>GDPR</u>, European Union, 2016a) personal data is defined as any information relating to an identified or identifiable natural person known as 'a data subject'.
- 'Special categories of personal data' according to GDPR:
 - Racial or ethnic origin;
 - Political opinions;
 - Religious or philosophical beliefs;
 - Trade union membership;
 - Genetic data;
 - Biometric data;
 - Data concerning health;
 - Data concerning a natural person's sex life or sexual orientation.



GDPR example: OpenClick Data Protection Policy

- According to EU General Data Protection Regulation
- Contents:
 - 1. Data protection principles
 - 2. General provisions
 - 3. Lawful, fair and transparent processing
 - 4. Lawful purposes
 - 5. Data minimization
 - 6. Accuracy
 - 7. Archiving / removal
 - 8. Security
 - 9. Breach

Data Protection Policy

Faculty of Mechanical Engineering, University of Nis, Serbia

Last updated 11.0	4.2019.
-------------------	---------

Definitions

Institution	means Faculty of Mechanical Engineering, University of Nis, Serbia, a registered national Institution of Serbia.		
GDPR	means the General Data Protection Regulation.		
Responsible Person	means dr Miroslav Trajanovic, full time professor		
Register of Systems	means a register of all systems or contexts in which personal data is processed by the Institution.		



Data brokers

- **Data brokers** are knowledgeable, independent parties who act as data stewards for sensitive data
- Researchers can transfer their sensitive data and jurisdiction over access to that data to the broker
- Especially common with patient-level data from clinical studies
- Provide a level of independence in the evaluation of whose data requests are scientifically valid, will not violate the privacy of research participants
- Examples of data brokers include <u>The YODA Project</u>, <u>ClinicalStudyDataRequest.com</u>, <u>National Sleep Research Resource</u> and <u>Supporting Open Access for Researchers (SOAR)</u>



Analysis portals

- Platforms that allow approved analysis of data without allowing full access (viewing or downloading) or access control
- Control what additional datasets can be pooled with the sensitive data as well as what analyses can be run to ensure that personal information is not revealed during reanalysis
- Examples include <u>Project Data Sphere</u>, <u>Vivli</u>, <u>RAIRD</u>, <u>Corpuscle</u>, and <u>INESS</u>



Data management plan

- A Data Management Plan (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 and preserved
- DMPs are often submitted as part of grant applications, but are useful whenever researchers are creating data
- Further questions: <u>CESSDA Data Management Expert Guide, part 6</u>
- DMP templates across EU: <u>CESSDA European diversity</u>, <u>Horizon Europe</u> <u>DMP template</u>



DMP template (Horizon Europe)





Data Summary questions

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



Fair data (findable) questions

- Will data be identified by a **persistent identifier**?
- Will rich **metadata** be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
 - <u>RDA Metadata Standards Directory</u> lists hundreds of standards, extensions, tools, and use cases. The directory can be browsed by discipline and subject area.
- Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
- Will metadata be offered in such a way that it can be harvested and indexed?



Fair data (accessible) questions

Repository

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



Fair data (accessible) questions

• Data

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



Fair data (accessible) questions

Metadata

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



Fair data (interoperable) questions

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



Fair data (re-use) questions

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



Other research outputs

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



Allocation of resources

- What will the **costs** be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, reuse, security, etc.) ?
- How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)
- Who will be responsible for data management in your project?
- How will **long term preservation** be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?



Data security

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



Ethics

- Are there, or could there be, any **ethics or legal issues** that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).
- Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?



Other issues

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



Data management plan example: OpenClick

- http://www.openclick.rs/index.php/en/project-openclick/data
- <u>http://www.openclick.rs/images/files/openclickdtm.pdf</u>

EUROPEAN RESEARCH COUNCIL (ERC): ERC DMP

SUMMARY

Project Acronym: OpenClick Project Number

DATASET SUMMARY

OpenClick project aims to discover physical status, skills and abilities of people based on human – computer interaction tests. There are currently 9 different sets of data within the OpenClick project. Each set of data corresponds to a single test that exists within the project.

Testing tools measure specific user performance in human – computer interaction. Each test is designed to measure one, exactly defined, type of interaction. During the test, several samples are taken, the average value is calculated, and then this average value is recorded in the database as a result of the test for specific person. Recorded data sets are stored in text files, in the CSV (comma separated value) format.



Pros and Cons of open data - an example (medical)

• Cons

- Consent and ethics:
 - real anonymisation is difficult,
 - losing control over other conclusions taken from your data and for which patients gave consent.
- Intentions of the people asking for the data
 - e.g. a pharmaceutic company that finds no link of their product to cancer when re-analysing data
- Competence of the people re-analysing your data

• Pros

 the data freed by pharma scrutinised by community, hidden side effects found



How to make research data accessible

- 1. As supplemental material **with a research article**, **hosted by the publisher** of the article
- 2. Hosting data on a **publicly-available website**, with files available for download
- Depositing data in a repository developed to support data publication (e.g. <u>Zenodo</u>)
- 4. Publishing a data paper about the dataset (as a preprint, in a journal, or in a data journal dedicated to supporting data papers). Hosted by the journal or hosted separately in a data repository



Selecting a data repository



Order of preference recommended by **OpenAIRE**:

- 1. Use an **external data archive** or repository established for your research domain to preserve the data. Some <u>recommendations</u> are given by Nature
- 2. If available, use an **institutional repository**, or your research group's established data management facilities
- 3. Use a cost-free data repository such as <u>Dataverse</u>, <u>Dryad</u>, <u>figshare</u> or <u>Zenodo</u>.
- 4. Search for **other data repositories** in <u>re3data</u>. Filter options that will help you find FAIR-compatible repositories: access categories, data usage licenses, trustworthy data repositories and data a persistent identifier (PID). Consider whether the repository supports versioning



How to make data open?

Key steps in approximate order (can be done simultaneously):

- 1. Choose your dataset(s) Choose the dataset(s) you plan to make open. You may need to return to this step if you encounter problems at a later stage.
- 2. Apply an open license (more info, details in Open Licensing and File Formats)
 - Determine what intellectual property rights exist in the data.
 - Apply a suitable 'open' license that licenses all of these rights and supports openness.
 - If you can't do this go back to step 1 and try a different dataset.
- 3. Make the data available in bulk and in a useful format or via an API.
- 4. Make it discoverable post on the web, optionally organize a central catalog to list open datasets.



Exercise – upload data to Zenodo

- Zenodo open access repository
- OpenAIRE program (CERN)
- Deposit
 - Research papers
 - Data sets
 - Research software
 - Reports
- Persistent DOI (Digital Object Identifier)
- Free and open, no file format restrictions, up to 50 GB



Zenodo advantages

- Your data will be assigned a DOI.
- You can control how accessible you want your dataset to be.
- Your data will be preserved!



Zenodo sandbox

- This is a test version of the real Zenodo platform
- http://sandbox.zenodo.org



Signing up

Research. Shared! Sign up today.

Citeable. Discoverable.

Uploads get a Digital Object Identifier (DOI) to make them easily and uniquely citeable.

Communities

Accept or reject uploads to your own community (e.g workshops, EU projects, institutions or entire disciplines).

Trusted Research Data Management

Built on top of CERN's expertise in managing 100s of petabytes of research data from the Large Hadron Collider.



O Sign up with GitHub

D Sign up with ORCID



Uploading a file

New upload

Chicago COVID-19 Response

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.



Communities

Specify communities which you wish your upload to appear in. The owner of the community will be notified, and can either accept or reject your request. Please make sure your record complies with the content policy of the communities you add; reported abuse will be followed by account inactivation.

Start typing a co	typing a community name			
LORY Locase Open Reportery Universität Locase	LORY - Lucerne Open Repository	•		
COOL 19 (The 19th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun			
In it has non a paperies, toucher Luczennes University of Applied Sciences and Artis	Proceedings from the 19th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, hosted by Uppsala University in Uppsala, Sweden from 06 – 10 June 2016.			
HOCHSCHULE LUZERN	Hochschule Luzern			



Format of upload





Basic information

Basic information

RESEARCHERS IN MOTION

Digital Object Identifier	10.5281/zenodo.5654906	nodo.5654906				
	Optional. Did your publisher already assign a DOI to your upload? If not, leave the field empty and we will register a new DOI for you. A DOI allows others to easily and unambiguously cite your upload. Please note that it is NOT possible to edit a Zenodo DOI once it has been registered by us, while it is always possible to edit a custom DOI.					
Publication date *	2021-11-08					
	Required. Format: YYYY-MM-DD. In case your upload was already published elsewhere, please use the date of first publication.					
┛ Title *	* Time required for typing numbers					
	Required.					
Authors *	Misic Dragan	University of Nis	D ORCID (e.g.: 0000-0002-1825-0097)			
			Optional.			
	Trajanovic Miroslav	University of Nis	● ORCID (e.g.: 0000-0002-1825-0097)			

Description *

Description, keywords ...

The goal of this test is to see how much time it is needed for one keystroke of a number on the keyboard. By random principle the program displays a number made of three whole digits and one or two decimals separated by comma. Upon starting the test program displays the first number. When the user finishes pressing all number keys (and a comma that separates whole from decimal digits), user has to press Enter so the next number can be displayed. The process is repeated for 5 different numbers.

In addition, the number of errors per ten typed numbers entered for each user is counted. The Damerau–Levenshtein algorithm is used to calculate the number of errors. The Damerau–Levenshtein distance between two words is the minimum number of operations (consisting of insertions, deletions or substitutions of a single character, or transposition two adjacent characters) required to change one word into the other. Since this number is usually too small, we counted the number of errors per 10 typed

Required.

1.0.0

Version

Optional. Mostly relevant for software and dataset uploads. Any string will be accepted, but semantically-versioned tag is recommended. See C semver.org for more information on semantic versioning.

🚱 Language

English

Optional. Primary language of the record. Start by typing the language's common name in English, or its ISO 639 code (two or three-letter code). See 🕑 ISO 639 language codes list for more information.

Seywords

\$ ×

+ Add another keyword

Additional notes



Optional.

License




Funding

Funding		recommended	d 🗙					
Zenodo is integrated into reporting lines for research funded by the European Commission via 🕝 OpenAIRE. Specify grants which have funded your research, and we will let your funding agency know!								
• Grants	European Commission (EU)	Start typing a grant number, name or abbreviation	×					
	Optional. OpenAIRE-supported projects only. For other funding acknowledgements, please use the Additional Notes field. Note: a human Zenodo curator will need to validate your upload - you may experience a delay before it is available in OpenAIRE. Add another grant							



Related/alternate identifiers

Related/alternate identifiers	Related/	alternate	identifiers
-------------------------------	----------	-----------	-------------

Specify identifiers of related publications and datasets. Supported identifiers include: DOI, Handle, ARK, PURL, ISSN, ISBN, PubMed IC arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs and URLs.

Related identifiers	e.g. 10.1234/foobar.56789		~
	♣ Add another related identifier	cites this upload is cited by this upload is supplemented by this upload is a supplement to this upload is referenced by this upload references this upload	
Contributors		published this upload is previous version of this upload	
References		is new version of this upload continues this upload	
Journal		describes this upload is described by this upload	
Conference		has this upload as part is part of this upload	
Book/Report/Chapter		is reviewed by this upload documents this upload	
Thesis		is documented by this upload	•



Contributors

Contributors					optic
	Contributors	Family name, given name	Affiliation	(D) ORCID (e.g.: 0000-0002-1	
		+ Add another contributor		Optional.	Contact person Data collector Data curator Data manager
References					Editor C
Journal					Other Producer



Save and publish

💼 Delete





Warning

Warning

Once the record is published you will no longer be able to change the files in this upload.

This is because a Digital Object Identifier (DOI) will be registered immediately after publishing. You will still be able to update the record's metadata later.

If you only want to create a test upload, please do so on Zenodo Sandbox.



х



Open data example: OpenClick

- Click Resources > Download data sets
- Click on Mouse clicks test
- Click on

"OpenClick_MouseClickTest_08-05-2019 09_51_11.csv

X	OpenClic File Edit	k_Mo Insert	USE For	ClickTe mat Help	st_	08-05- All ch	- 2019 anges s	09_5 aved at	5 1_11 . t ~\Docu	. xlsx ıments\Open
	r ~ @	ર્ ૦	Cal	ibri	•	12 -	В	ΙU	<u>ት</u> A	• 🇞 •
f^{χ}										
	А	E	3		С		D		E	F
1	Time for one clie	ck User	ID	Test date			Gende	r Righ	t hand	Birth year
2	0.185	52	1	2019-02-	04 0	0:00:00	male		1	1964
3	0.306	51	4	2019-02-	06 0	0:00:00	male		1	1954
4	0.186	53	6	2019-02-	06 0	0:00:00	male		1	1974
5	0.188	37	7	2019-02-	06 0	0:00:00	female		1	1997
6	0.168	35	5	2019-02-	06 0	0:00:00	male		1	1972
7	0.182	9	8	2019-02-	06 0	0:00:00	male		0	1976
8	0.208	33	2	2019-02-	06 0	0:00:00	female		1	1965
9	0.189	9	11	2019-02-	07 0	0:00:00	female	9	1	1974
10	0.186	3	11	2019-02-	07.0	00.00.0	female	2	1	1974



Data processing examples

- Mouse clicks test
 - Time for one mouse click (only left button): 0.2469
- Alternate mouse button selection

	Left button time	Right button time	Middle button time	Double click time
All	1.029506	0.910285	0.997979	1.389282
Female	1.006681	0.806984	1.023247	1.343797
Male	1.046136	0.985546	0.979568	1.422421
All <20 years	1.014585	0.752221	0.881470	1.404601



More results – Show history





More results – Show radar chart







Data citation

- Data citation services help research communities discover, identify, and cite research data (and often other research objects) with confidence
- Typically involves the creation and allocation of Digital Object Identifiers (DOIs) and accompanying metadata through services such as DataCite (<u>https://www.datacite.org</u>) and can be integrated with research workflows and standards
- An emerging field, that involves:
 - conveying to journal publishers the importance of appropriate data citation in articles,
 - enabling research articles themselves to be linked to any underlying data



Open Materials

- In addition to data sharing, the openness of research relies on sharing of materials.
- Examples of materials that can be shared:
 - **Reagents** (a substance, compound or mixture that can be added to a system in order to create a chemical or other reaction)
 - **Protocols** (a formal or official record of scientific experimental observations in a structured format). Example repository: <u>Protocols.io</u>.
 - Notebooks, containers, software, and hardware.
 - Containers and notebooks can be shared with <u>Rocker</u> or <u>Code Ocean</u>.
 - Software and hardware used in research should be shared following best practices for documentation as outlined in <u>Open Research Software and Open Source</u> chapter.



Exercise – Open data

- Q1 What data will generate your research? Answer:
- Q2 In what repository will data be openly available? Answer:
- Q3 How will you ensure re-use of the data? Answer:

Notice:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 6 minutes







Open Research Software and Open Source

<u>Photo</u> by The Open Source Initiative is licensed under <u>CC BY-SA</u>

What is Open Methodology?

- The term "**open methodology**" includes procedural, technical and technological solutions intended for use in scientific research, which are openly available to the scientific community
- These solutions can be available in their final form, but are **usually open to further development**, which implies that each member of the scientific community can contribute to their improvement
- The key feature of open methodologies is their transparency, that is, their direct applicability by the end user
- At this point, probably the best illustration of open methodologies are different software solutions, such as the statistical software environment <u>R</u> or software solutions such as statistical packages <u>JASP</u> or <u>jamovi</u>
- Example: OpenClick methodology: <u>http://openclick.rs/index.php/en/project-openclick/methodology</u>



What is Open Research Software?

- Open (or open-source) research software refers to the use and development of software (for analysis, simulation, visualization, etc.) where the full source code is available
- must be distributed in source and/or compiled form (with the source code available in the latter case)
- must be shared under a license that allows modification, derivation, and redistribution







Open research software

- Sharing software used for research (whether computational in nature, or that relies on any software-based analysis/interpretation) is a necessary, though not sufficient, condition for reproducibility (unavoidable ambiguity arises when trying to fully describe software using natural language)
- Sharing software openly allows developers to receive career credit for their efforts, either through direct citation (Smith et al., 2016) or via software meta-articles
 - Software meta article journal examples: <u>Journal of Open Research Software</u> or the <u>Journal of Open Source Software</u> (Smith et al., 2018)
 - A <u>list of many domain-specific journals</u> that publish software articles by Neil Chue Hong



Storage and unique identification of open software

- **Git** is a popular tool that allows version control: management and overall tracking of changes in a particular piece of software
- Services such as <u>GitHub</u>, <u>GitLab</u>, <u>Bitbucket</u>, and others provide an interface to Git as well as remote storage services that can be used to maintain, share, and collaborate on research software
- Apart from software publishing and version control, it is equally important to have a published and persistent identifier associated with it, such as a DOI
- There are several ways of associating a DOI with a GitHub repository; the easiest one is to employ <u>Zenodo</u> (a free, open catch-all repository created by OpenAIRE and CERN)



Open source licensing

- Publicly shared software is not actually open source unless accompanied by a suitable license, as any other creative work it falls under exclusive copyright to the creators no one else can use, copy, distribute, or modify your work (<u>choosealicense.com</u>)
- An appropriate license should be used for software, based on what is preferred to let others do (or prevent them from doing) with the code
- If the code is to be shared with no restrictions dedicate it to the public domain (<u>The</u> <u>Unilicense</u>)
- The <u>choosealicense.org</u> site is a helpful resource to **differentiate between licenses**, although it does not feature <u>every available or popular open-source license</u>
- Once a license is selected, the text—edited to include the author name(s) and year—is put in the software repository as a **plaintext LICENSE file**



If you choose no licence...

- When you make a creative work (which includes code), the work is under exclusive copyright by default
- Unless you include a license that specifies otherwise, nobody else can copy, distribute, or modify your work without being at risk of takedowns, shake-downs, or litigation
- Once the work has other contributors (each a copyright holder), "nobody" starts including you!



Documenting open research software

- Documented software will have more impact and be more easily used by others (and the creator)
- Documentation includes:
 - comments in the code **why** something is done
 - a README file e.g., how to install, cite, run, important dependencies,
 - tutorials/examples, and/or
 - API documentation
- Missing or inaccessible dependencies or insufficient documentation very common barriers to reuse and reproducibility
- Containers package the code with the dependencies and computational environment so others can more easily run your analysis. Examples: <u>Docker</u>, <u>Rocker</u>, <u>Binder</u>, and <u>Code Ocean</u>



Open source publishing procedure example

- 1. Create a repository on **GitHub** and enable the integration with **Zenodo**. Mint the first release of the software
- 2. Choose a software license using (e.g.) <u>choosealicense</u> or the <u>Open</u> <u>Source Initiative</u>
- 3. Create documentation for a software package, including README, comments, and examples
- 4. Appropriately cite software used for a paper



Open Source Hardware

- The open source principles extend to hardware
- Scientific hardware includes everything from sequencing tools and microscopes to specialized testing equipment and particle colliders
- The ability to use, study, replicate, and improve scientific instrumentation is a central part of experimental science
- Above activities are currently restricted by proprietary instrumentation, which is difficult and expensive to obtain and maintain
- Open Science Hardware (OScH) community, is leading a push for the open source movement to include scientific tools, hardware, and research infrastructures through their <u>Global Open Science Hardware Roadmap</u>







Collaborative Platforms

What are collaborative platforms?

- Collaborative platforms are usually online services that provide a virtual environment to which multiple people can concurrently connect and work on the same task
- Online collaborative platforms connect geographically-dispersed researchers to enable them
 - to cooperate seamlessly on their research,
 - to share research objects as well and ideas and experiences
- Platform types range from extensive virtual research environments (VREs) to single specific tools



Virtual research environments (VREs)

- Facilitate sharing and collaboration tools (e.g. forums and wikis), collaborative document hosting and discipline-specific tools (e.g. data analysis or visualization)
- Many of these tools are **disciplinary-specific**
- The European Commission has funded a range of community-specific VREs, like:
 - <u>VI-SEEM</u> VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean,
 - <u>MuG</u> Multi-Scale Complex Genomics
 - <u>OpenDreamKit</u> Open Digital Research Environment Toolkit for the Advancement of Mathematics



Open Science Framework (OSF)

- OSF is an especially important collaborative platform in the context of Open Science
- Based on open source technologies and created by the not-for-profit <u>Center for Open Science</u>, the OSF brands itself as "a scholarly commons to connect the entire research cycle"
- Enables researchers to work on projects privately with a limited number of collaborators and make any part or the whole of their project public
- Connects directly with many other collaborative systems like dropbox, GitHub and Google Docs
- Can be used to store and archive research data, protocols, and materials



Single specific tools

- Enable researchers to work together in real time on specific aspects of research (such as writing or analysis)
- The examples of single specific tools are:
 - Collaborative writing platforms
 - Reference management & discovery
 - Annotation and review
 - Academic social networks



Collaborative writing platforms

- Enable researchers to work together on documents in real-time and avoid emailing and managing document versions.
- Include <u>Overleaf</u>, <u>Authorea</u>, <u>Fidus Writer</u>, <u>ShareLaTeX</u> and <u>Google</u> <u>Docs</u>.
- Many of these tools are based on proprietary technologies and some require payment for advanced features.



Reference management & discovery

- Tools which enable groups to store and manage references
- Examples include <u>Zotero</u>, <u>Citavi</u> and <u>CiteUlike</u>
- <u>Mendeley</u> incorporates a sharable reference manager, as well as a social network and article visualization tools



Annotation and review

- Web based modes of post-publication collaborative review.
- Services like <u>PubPeer</u> and <u>Academic Karma.</u>
- Annotation tools like <u>Hypothes.is</u> and <u>PaperHive</u>.



Academic social networks

- Mainstream social networks like Twitter, Facebook and Linkedin.
- **Dedicated** academic social networks like <u>ResearchGate</u>, <u>Academia.edu</u> and <u>Loop</u>.



What is ORCID?

- The ORCID iD (Open Researcher and Contributor ID) is a nonproprietary <u>alphanumeric code</u> to uniquely identify <u>scientific</u> and other <u>academic authors</u> and contributors
- ORCID iD addresses the problem that a particular author's contributions to the scientific literature can be hard to recognize as most personal names are not unique, they can change (such as with marriage), have cultural differences in name order, contain inconsistent use of first-name abbreviations and employ different writing systems.
- It provides a persistent identity for humans, similar to that created for contentrelated entities on digital networks by <u>digital object identifiers</u> (DOIs)
- The ORCID organization, ORCID Inc., offers an open and independent registry intended to be the *de facto* standard for contributor identification in research and <u>academic publishing</u>.



Exercise – Collaborative platforms

• Q1: Name the collaborative platforms that you intend to use for your project.

Notice:

- Use Exercise_template.docx to type the answers
- Give short answers one to two rows long
- Duration of exercise 3 minutes



Back to Contents



Open educational resources

What are Open educational resources?

 "Teaching, learning and research materials in any medium – digital or otherwise – that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions" (William and Flora Hewlett Foundation definition)

• Open educational resources include:

- full courses, course materials, modules,
- textbooks, streaming videos, images,
- tests, software,
- any other tools, materials, or techniques used to support access to knowledge.





Open educational resources and Open Science

- Open educational resources are often built upon research findings
- An Open Science practitioner will probably want his educational resources to maintain the level of openness as his research
- Other teachers can use the open material to elaborate new resources or adapt existing ones
- The creation of educational resources can be seen as a cycle similar to the research cycle: find, compose, adapt, use, and share


Licensing and open educational resources

- Open Educational Resources (OER) are only OER, if they have an open license
- There is **no clear guideline for the choice of license** for open educational resources
- Open Creative Commons licenses are <u>CCO (Public Domain Dedication)</u>, <u>CC BY (Attribution)</u> and <u>CC BY-SA (Attribution-ShareAlike)</u>, which can be used for most educational resources
- Licenses should be explained in detail to properly attribute authors and to create true OER. This also includes the combination of different license types and its consequences



OER platforms and their intended use

- <u>OpenCourseWare (OCW)</u> is one of the first open educational resource platforms and key initiators of the OER movement
- <u>Open Education Consortium</u>, initiated at MIT in 2002 provides materials from all over the world as **courses under free licenses**
- Other examples of OER platforms:
 - <u>Creative Commons Search</u> for image, audio, and video files
 - <u>OERCommons</u> for educational resources



OER and quality of materials

- Just as with printed text materials, quality can not be guaranteed
- A first indication of quality may be:
 - Open user comments,
 - peer review,
 - publication of materials on platforms of established institutions, e.g. universities
- Finally, it is personal judgement whether the selected material is suitable for the intended purpose and whether its content is correct



<u>Photo</u> by Manco Capac is licensed under the <u>Creative Commons</u> Attribution-Share Alike <u>4.0 International</u>, <u>3.0 Unported</u>, <u>2.5 Generic</u>, <u>2.0 Generic</u> and <u>1.0 Generic</u> license.





Open Advocacy

What is Open Advocacy?



- Advocacy in all its forms seeks to ensure that people, particularly those who are most vulnerable in society, are able to:
 - Have their voice heard on issues important to them
 - Defend and safeguard their rights
 - Have their views and wishes genuinely considered when decisions are being made about their lives
- Includes defending, influencing, changing, decision-making, persuading, lobbying, attracting attention
- Open Advocacy focuses on the movement to promote Open Science at various levels of stakeholders, highlighting and stressing the societal, professional and personal advantages that it entails



Advocacy and EURAXESS trainings

- Trainings (workshops, seminars, presentations) can be used as advocacy tools
- Training here is considered as a tool for effecting specific changes, and for building an Open Science advocate community





How to organize OpenScience training

- Use this and the extended presentation available on EURAXESS Extranet and OpenClick portal and create your own training materials
 - The materials should be adapted to the audience, be interactive and include specific examples and exercises
- 2. Book the trainer over EURAXESS Extranet
 - If the contents of the training is similar to this one
 - The training should be at least 1 day long, if held to researchers and/or institutional staff, to contain all the necessary information (and, if possible, interactive)



Literature

- The Open Science Training Handbook: <u>https://open-science-training-handbook.gitbook.io/book/</u>
- Open science introduction, OSC LMU Open science center: <u>https://mfr.osf.io/render?url=https://osf.io/z7954/?action=download%26mode=render</u>
- Vicente-Sáez, R., & Martínez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of business research*, 88, 428-436.
- Fecher, B., & Friesike, S. (2014). Open science: one term, five schools of thought. In *Opening science* (pp. 17-47). Springer, Cham.
- Open Definition: <u>http://opendefinition.org/</u>
- Smith, Arfon M., Daniel S. Katz, Kyle E. Niemeyer, and FORCE11 Software Citation Working Group. 'Software Citation Principles'. PeerJ Computer Science 2 (19 September 2016): e86. <u>https://doi.org/10/bw3g</u>.
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., ... & Haustein, S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, *6*, e4375.
- Candela, L., Castelli, D., & Pagano, P. (2013). Virtual research environments: an overview and a research agenda. *Data Science Journal*, GRDI-013.

Other used literature is referenced throughout the text.



Thank you for your attention!



This Photo by Unknown Author is licensed under CC BY-SA

