RItrainPlus
info@ritrainplus.eu
www.ritrainplus.eu

RESEARCH DATA MANAGEMENT

LEVEL   EQF6 ☐   EQF7 ☐   EQF8 ▉

ECTS	2
LA Aims	The aim of this course is to provide students with the skills needed to manage the research data of a research project, taking into account the data policy of an institution.
Learning Goals	Upon completion of this LA students are expected to: I. Learn the fundamentals of the Data Lifecycle II. Understand the FAIR principles III. Acquire the skills to write a Data Management Plan. At institutional and European level. IV. Identify sensitive data in your datasets and manage it.
Selected Learning Outcomes	Upon the completion of this LA, students will be able: ● to manage the research data generated in projects, ● to write a Data Management Plan and ● to understand the data policy defined by an institution.

Main characteristics of the LA

Format	The format of the course will be blended. ● Theoretical lessons will be resources published for the community. ● Use cases will be lectures given by invited speakers by online. ● The practical exercises will be a face to face activity.
Pedagogical methods/tools:	Presentations, lectures and group exercises
Assessment methods:	Each theoretical lesson will have a short questionnaire to assess the knowledge acquired by the student. The practical exercises will be evaluated based on the quality of the work presented and in the oral presentation. Also, the participation of the student (making questions) during the other students' presentations will be taken into account.
Evaluation:	The student will be evaluated based on quantitative and qualitative metrics. The practical exercises will be evaluated based on public rubrics.
LA Leader	**UPM**
LA Participants	**UNIBO**

Course Description

Learning Goal	Learning model	Time
Introduction to Research Data Management: Data Life Cycle.		**1 hour**
Students learn the role of data in the research process and in research projects.	Lectures: Esteban González - UPM	10 min
Understand the different research artifacts (datasets, software, publications, etc.) involved in a research.	Lectures: Esteban González - UPM	15 min
Comprehend the evolution of the data in a research project, from raw data to processed data.	Lectures: Esteban González - UPM	25 min
Discussion with students about the data managed in their thesis.	Case discussion	10 min
Introduction to FAIR principles.		**1 hour**
Reason why FAIR data is important and its role in Europe.	Lectures: Silvio Peroni UniBO	10 min
Understand each principle of the FAIR data. We will analyze each principle and its meaning.	Lectures: Silvio Peroni UniBO	30 min
Know the reference paper about FAIR data. Students will have to read the paper before the session.	Case discussion	10 min
Discussion with students about how they can apply FAIR principles in their thesis and the value to do it.	Case discussion	10 min
Ethical and legal considerations. ● *Students will understand the ethical and legal aspects of the data and how to manage them. This will include the different licences used for data.* ● *St*udents will learn how to handle sensitive data.		**2 hours**
Understand what is a license and why is important. Students will learn the different parts of a license.	Lectures: Victor Rodriguez - UPM	10 min

Identify the main licenses used in researching both open and private.	Lectures: Victor Rodriguez - UPM	20 min
Apply a license to your work. Students will analyze their data and will select a license to apply them.	Lecture, case discussion: Victor Rodriguez - UPM	20 min
Identify any kind of sensitive data present in your research. Different representative cases will be exposed. For example: personal data, protected species, etc.	Lectures: Victor Rodriguez - UPM	30 min
Understand the main points in the GDPR and learn how to manage them	Lectures: Victor Rodriguez - UPM	20 min
Apply the GDPR to the data generated in their research	Case discussion	20 min
Practical exercise: Amnesia. ● *In this unit, students will learn how to use the tool Amnesia, from OpenAire, to anonymize sensitive data.*		**2 hours**
Acquire the skills to use tools like AMNESIA. We will use a video created by the OpenAIRE team.	Video (https://www.youtube.com/watch?v=_0Io6c1MPOY)	60 min
Apply pseudonymization techniques to their data using the tool AMNESIA.	Case discussion	60 min
Data publication & Metadata standards. ● *Students will learn how to share and publish their data, following the most popular standards.*		**1 hour**
Understand what is metadata	Lectures: Esteban González - UPM	5 min
Understand what is a schema and the most popular (DCAT, Dublin Core, etc .)	Lectures: Esteban González - UPM	55 min
Practical exercise: Zenodo ● *In this unit, students will learn how to share and publish their data in Zenodo, a general-purpose repository.*		**2 hours**

Learn how to use a repository as Zenodo to publish your research. We will use a video created by the OpenAIRE team.	Video (https://www.youtube.com/watch?v=yj2r8RayIX8)	60 min
Identify repositories where you can publish your data and how to do it (metadata that should be completed to be FAIR).	Case discussion	60 min
Introduction to a Data Management Plan. ● *In this lesson, the different sections of a DMP will be described.*		**1 hour**
Understand the different sections of a DMP. We will focus on the template used for European projects (Horizon).	Lectures: Esteban González - UPM	60 min
Practical exercise: Argos. ● *Students will see a real implementation of a Data Management Plan. The tool Argos (OpenAIRE) will be used during the exercise.*		**2 hours**
Learn how to use the tool Argos to create DMPs. We will use a video created by the OpenAIRE team.	Video (https://www.youtube.com/watch?v=xOdcmankdj4)	60 min
Students will know representative use cases of DMPs (different domains). These examples will help them to create a DMP for their data.	Lecture, case discussion: Esteban González - UPM	60 min
Practical exercise. ● *Students will be divided in groups and they will have to create a Data Management Plan. After this, students will have to present their work to the other students.*	(2 hours of team work + 2 hours presentations depending on the number of groups)	**4 hours**
Learn how to write a DMP. The students will be grouped in teams and they will work in a real case.	Practical exercise	120 min
Students will present their DMP, where the other students will have to identify potential problems.	Case discussion	120 min

Teaching materials

Presentations + notes for theoretical lessons

Teamwork material for practical exercises

Questionnaires

All the PowerPoint material used for the lecture will be made available to the students at the dedicated website https://drive.google.com/drive/u/0/folders/1kn8jaQbqbw5sE1W3IB6JE9niQAVinPoH (the folder will be populated during the piloting of the courses)

Bibliography/ Sitography/ Video

- Bahlai, C., Bartlett, L., Burgio, K., Fournier, A., Keiser, C., Poisot, T., & Whitney, K. (2019). Open Science Isn't Always Open to All Scientists. American Scientist, 107(2), 78. https://doi.org/10.1511/2019.107.2.78
- Hendricks, G., Tkaczyk, D., Lin, J., & Feeney, P. (2020). Crossref: The sustainable source of community-owned scholarly metadata. Quantitative Science Studies, 1(1), 414–427. https://doi.org/10.1162/qss_a_00022
- Avanço, K., Balula, A., Błaszczyńska, M., Buchner, A., Caliman, L., Clivaz, C., Costa, C., Franczak, M., Gatti, R., Giglia, E., Gingold, A., Jarmelo, S., Padez, M. J., Leão, D., Maryl, M., Melinščak Zlodi, I., Mojsak, K., Morka, A., Mosterd, T., … Wieneke, L. (2021). Future of Scholarly Communication—Forging an inclusive and innovative research infrastructure for scholarly communication in Social Sciences and Humanities (p. 46). Digital Humanities Centre at the Institute of Literary Research of the Polish Academy of Sciences. https://doi.org/10.5281/zenodo.5017705
- Belhajjame, K., B'Far, R., Cheney, J., Coppens, S., Cresswell, S., Gil, Y., Groth, P., Klyne, G., Lebo, T., McCusker, J., Miles, S., Myers, J., Sahoo, S., & Tilmes, C. (2013). PROV-DM: The PROV Data Model (L. Moreau & P. Missier, Eds.). World Wide Web Consortium. https://www.w3.org/TR/prov-dm/
- Chue Hong, N. P., Katz, D. S., Barker, M., Lamprecht, A.-L., Martinez, C., Psomopoulos, F. E., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., & Honeyman, T. (2022). FAIR Principles for Research Software (FAIR4RS Principles) [Recommendations with RDA Endorsement in Process]. Research Data Alliance. https://doi.org/10.15497/RDA00068
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), no. L119, Official Journal of the European Union (2016). http://data.europa.eu/eli/reg/2016/679/oj
- GO FAIR. (2018). FAIR Principles. https://www.go-fair.org/fair-principles/
- Gualandi, B. (2022). What do we mean by 'data' in the arts and humanities? [Master thesis, University of Bologna]. https://doi.org/10.5281/zenodo.6385132
- Haak, L. L., Fenner, M., Paglione, L., Pentz, E., & Ratner, H. (2012). ORCID: A system to uniquely identify researchers. Learned Publishing, 25(4), 259–264. https://doi.org/10.1087/20120404
- Kramer, B., & Bosman, J. (2015, June 18). The good, the efficient and the open—Changing research workflows and the need to move from Open Access to Open Science. CERN Workshop on Innovations in Scholarly Communication (OAI9), University of Geneva, Geneva, Switzerland. https://www.slideshare.net/bmkramer/the-good-the-efficient-and-the-open-oai9
- Landi, A., Thompson, M., Giannuzzi, V., Bonifazi, F., Labastida, I., da Silva Santos, L. O. B., & Roos, M. (2020). The "A" of FAIR – As Open as Possible, as Closed as Necessary. Data Intelligence, 2(1–2), 47–55. https://doi.org/10.1162/dint_a_00027
- Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhause, M., & Westbrook, J. (2020). The TRUST Principles for digital repositories. Scientific Data, 7(1), 144. https://doi.org/10.1038/s41597-020-0486-7
- Michener, W. K. (2015). Ten Simple Rules for Creating a Good Data Management Plan. PLOS Computational Biology, 11(10), e1004525. https://doi.org/10.1371/journal.pcbi.1004525
- Open Knowledge Foundation. (2015). Open Definition 2.1. https://opendefinition.org/od/2.1/en/
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2017). Data From: The State Of Oa: A Large-Scale Analysis Of The Prevalence And Impact Of Open Access Articles [Data set]. Zenodo. https://doi.org/10.5281/zenodo.837901
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18
- Wolfe, M. (2017, August 9). CC0 and Data Citation. https://www.library.ucdavis.edu/news/cc0-and-data-citation/

RESEARCH DATA POLICY

LEVEL   EQF6 ☐   EQF7 ☐   EQF8 ☒

ECTS	2
LA Aims	The aim of this course is to provide students with the skills needed to define and to implement the data policies of an institution.
Learning Goals	Upon the completion of this LA, students will be able to understand and to define the data policy of an institution.
Selected Learning Outcomes	Upon the completion of this LA, students will be able: ● to define and implement a data policy for an institution, including: ▪ FAIR principles ▪ Licences ▪ Infrastructure needed (repositories & human resources) ▪ Data Management Plan ▪ Personal & Sensitive data

Main characteristics of the LA

Format	The format of the course will be blended. ● Theoretical lessons will be resources published for the community. ● Use cases will be lectures given by invited speakers by online. ● The practical exercises will be a face-to-face activity.
Pedagogical methods/tools:	Presentations, lectures and group exercises
Assessment methods:	Each theoretical lesson will have a short questionnaire to assess the knowledge acquired by the student. The practical exercises (DMP and data policy) will be evaluated based on the quality of the work and the oral presentation. Also, the student participation (making questions) during the other students' presentations will be taken into account.
Evaluation:	The student will be evaluated based on quantitative and qualitative metrics. The practical exercises will be evaluated based on public rubrics.
LA Leader	**UPM**
LA Participants	**UNIBO**

Course Description

Learning Goal	Learning model	Time (minutes)
Introduction to Research Data Management: Data Life Cycle. ● *Brief introduction to research data life cycle, which stages must be defined in the data policy*		**1 hour**
The role of data in the research process and in research projects.	Lectures: Esteban González - UPM	10 min
Understand the different research artifacts (datasets, software, publications, etc.) involved in a research.	l Lectures: Esteban González - UPM	15 min
Comprehend the evolution of the data in a research project, from raw data to processed data.	Lectures: Esteban González - UPM	25 min
DIscussion with students about the data managed in their thesis.	Case discussion	10 min
Open & Private data. ● *The students will learn the difference between the two models, and how to implement them in a research environment. We will focus on FAIR principles and the main licences used for both sharing policies.*		**1 hour**
Understand each principle of the FAIR data.	Lectures: Victor Rodriguez - UPM	20 min
Identify the main licenses used in researching both open and private.	Lectures: Victor Rodriguez - UPM	20 min
Understand the main points in the GDPR and learn how to manage them	Lectures: Victor Rodriguez - UPM	20 min
Main data infrastructures elements in an organisation. ● *The student will learn what is a data repository, a data management plan tool, a data storage service, etc.*		**1 hour**

Understand what is a data repository and the different policies applied to publish on it. We analyze concepts like metadata schemas and persistent identifiers.	Lectures: Esteban González - UPM	30 min
Understand what is a data management plan and the tools that can be used to create it	Lectures: Esteban González - UPM	30 min
Use case: Institutional and European data infrastructures. ● *In this unit, we present a representative collection of data infrastructures based on partners' institutions infrastructures and European as EOSC services.*		**2 hours**
Analyze the data infrastructures of students' institutions	Case discussion	50 min
Identify EOSC data tools such as Zenodo, Argos, B2Share, Amnesia, B2Handle, AMNESIA, etc .	Lectures: Esteban González - UPM	50 min
Learn how EOSC tools can be adopted in the students' institutions	Case discussion	20 min
Introduction to Data Governance. ● *Students will learn the main aspects of an institutional data policy, which includes how the data is gathered, stored, processed, etc.*		**1 hour**
Learn the different sections of a data policy. It includes, where data can be published, how to manage the GDPR, where and how private data can be stored, etc.	Lectures: Esteban González - UPM	60 min
Roles and Responsibilities of a data stewardship organisation. ● *This unit will show how to create a data stewardship and how to integrate it in the institution's structure. Also, we analyse the different roles of a data steward.*		**1 hour**

Learn how data managers are structured in different organizations.	Lectures: Esteban González - UPM	20 min
Learn the functions of data stewardship.	Lectures: Esteban González - UPM	20 min
Learn how to acquire the skills to be a data stewardship.	Lectures: Esteban González - UPM	20 min
Data governance tools. ● *In this unit, students will learn the use of some tools to communicate the data policies to the staff. This includes websites, communication plans, training courses, networking and roadmaps.*		**2 hours**
Understand the different tools the use of some tools to communicate the data policies to the staff.	Lectures: Esteban González - UPM	15 min
Learn how to structure a website to show the data policies to researchers.	Lectures: Esteban González - UPM	15 min
Define a communication plan to assure that researchers know the data policy.	Lectures: Esteban González - UPM	15 min
Learn how to prepare training courses for researchers	Lectures: Esteban González - UPM	15 min
Define a roadmap to apply data policies in a research infrastructure	Lectures: Esteban González - UPM	60 min
Data policies use cases. ● *In this unit, we will analyse the different data policies defined in several institutions.*		**2 hours**
Know the data policies landcape in europe. We will focus on universities and research facilities.	Lectures: Esteban González - UPM	60 min
Comprehend the data policies implemented in students' institutions	Case discussion	60 min
Practical exercise. ● *Students will be divided in groups and they will have to elaborate a data policy for a research institution. They will present the work to the other students*	(2 hours of team work + 2 hours presentations depending on the number of the group)	**4 hours**

Learn how to write a data policy document. The students will be grouped in groups and they will work in a real case	Practical exercise	120 min
Students will present their data policies, where the other students will have to identify potential problems.	Case discussion	120 min

Teaching Materials

All the PowerPoint material used for the lecture will be made available to the students at the dedicated website
https://drive.google.com/drive/u/0/folders/1kn8jaQbqbw5sE1W3IB6JE9niQAVinPoH (the folder will be populated during the piloting of the courses)

Bibliography/Sitography/Videos

- Kramer, B., & Bosman, J. (2015, June 18). The good, the efficient and the open—Changing research workflows and the need to move from Open Access to Open Science. CERN Workshop on Innovations in Scholarly Communication (OAI9), University of Geneva, Geneva, Switzerland. https://www.slideshare.net/bmkramer/the-good-the-efficient-and-the-open-oai9
- Meng, X.-L. (2020). Reproducibility, Replicability, and Reliability. Harvard Data Science Review, 2(4). https://doi.org/10.1162/99608f92.dbfce7f9
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), no. L119, Official Journal of the European Union (2016). http://data.europa.eu/eli/reg/2016/679/oj
- GO FAIR. (2018). FAIR Principles. https://www.go-fair.org/fair-principles/
- Perneger, T. V. (2004). Writing a research article: Advice to beginners. International Journal for Quality in Health Care, 16(3), 191–192. https://doi.org/10.1093/intqhc/mzh053
- Lamprecht, A.-L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., Dominguez Del Angel, V., van de Sandt, S., Ison, J., Martinez, P. A., McQuilton, P., Valencia, A., Harrow, J., Psomopoulos, F., Gelpi, J. Ll., Chue Hong, N., Goble, C., & Capella-Gutierrez, S. (2020). Towards FAIR principles for research software. Data Science, 3(1), 37–59. https://doi.org/10.3233/DS-190026
- Alexandrov, A. V., & Hennerici, M. G. (2013). How to Prepare and Deliver a Scientific Presentation. Cerebrovascular Diseases, 35(3), 202–208. https://doi.org/10.1159/000346077
- Allen, L., O'Connell, A., & Kiermer, V. (2019). How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. Learned Publishing, 32(1), 71–74. https://doi.org/10.1002/leap.1210
- Aspesi, C., & Brand, A. (2020). In pursuit of open science, open access is not enough. Science, 368(6491), 574–577. https://doi.org/10.1126/science.aba3763
- Bilder, G., Lin, J., & Neylon, C. (2020). The Principles of Open Scholarly Infrastructure. The Principles of Open Scholarly Infrastructure. https://doi.org/10.24343/C34W2H
- Burgelman, J.-C. (2021). Politics and Open Science: How the European Open Science Cloud Became Reality (the Untold Story). Data Intelligence, 3(1), 5–19. https://doi.org/10.1162/dint_a_00069
- Doumont, J. (2011). Creating Effective Presentation Slides. Optics & Photonics News, 22(3), 12–14. https://www.osa-opn.org/home/articles/volume_22/issue_3/departments/career_focus/creating_effective_presentation_slides/
- Ficarra, V., Fosci, M., Chiarelli, A., Kramer, B., & Proudman, V. (2020). Scoping the Open Science Infrastructure Landscape in Europe. SPARC Europe. https://doi.org/10.5281/zenodo.4159838

- Kramer, B., & Bosman, J. (2015, June 18). The good, the efficient and the open—Changing research workflows and the need to move from Open Access to Open Science. CERN Workshop on Innovations in Scholarly Communication (OAI9), University of Geneva, Geneva, Switzerland. https://www.slideshare.net/bmkramer/the-good-the-efficient-and-the-open-oai9
- Skinner, K., & Lippincott, S. (2020). Assessment Checklist (Commonplace). Knowledge Futures Group. https://doi.org/10.21428/6ffd8432.5175bab1/00710d8a
- Teal, T., Lowenberg, D., Smith, T., Gonzales, J. B., Nielsen, L. H., & Ioannidis, A. (2020, August 26). Sustainable, Open Source Alternatives Exist [Blog]. Dryad News and Views. https://blog.datadryad.org/2020/08/26/sustainable-infrastructure-exists/
- Teperek, M., & Dunning, A. (2020, August 18). Why figshare? Choosing a new technical infrastructure for 4TU.ResearchData [Blog]. Open Working. https://openworking.wordpress.com/2020/08/18/why-figshare-choosing-a-new-technical-infrastructure-for-4tu-researchdata/
- Aalto university (Research Data Management Videos). https://www.youtube.com/watch?v=jZn1d2DZpro&list=PLa0sKMV88Js4La7rqVBxzW4NwRWygVdYj