Supporting eosc

Skills 4 eosc

Milestone: Pilot learning path for Data Stewards

Lead Partner:	DTU.DKB, T4.2
Version:	**2.0**
Status:	**Final version incl. feedback (co-creation) from WP4**
Dissemination Level:	**Public**
Document Link:	Skills4EOSC-T4.2-Milestone-evaluation of Pilot DS-20240709.docx - ONLYOFFICE

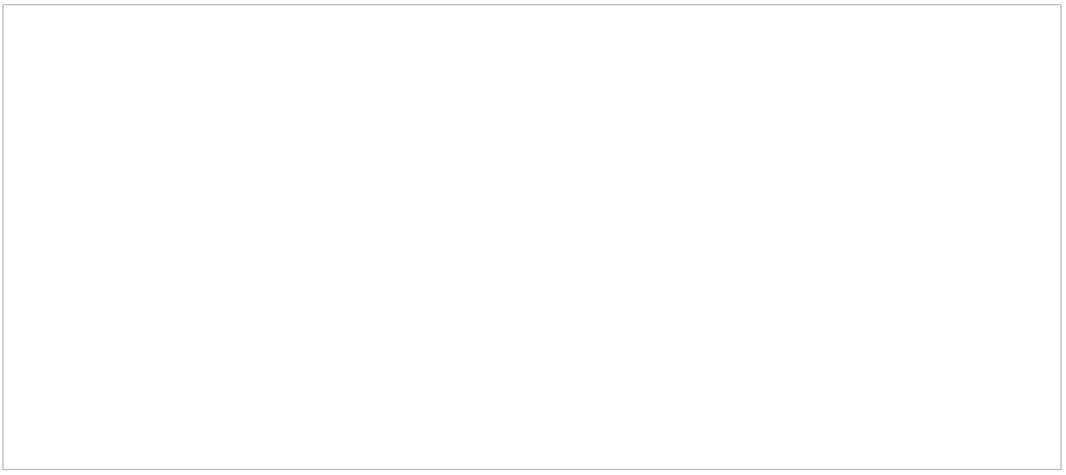
Deliverable Abstract

This is a short report on the pilot ToT Learning Path for Data Stewards designed by the Skills4EOSC project, WP4, T4.2. the EOSC. The report summarises the training and points to areas for improvement regarding the content and delivery of the training and further, alignment with the MVS Data Steward which is a competency profile also developed within Skills4EOSC.

Co-funded by the European Union

co-funded by UK Research and Innovation

Skills4EOSC has received funding from the European Union's Horizon Europe research and innovation Programme under Grant Agreement No. 101058527 and from UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10040140]



COPYRIGHT NOTICE

DELIVERY SLIP

Date	***Name***	***Partner/Activity***	***Date***
From:	Lorna Wildgaard	DKB/T4.2	09.07.2024
Contributors to Learning Path and/or pilot	Benjamin Derksen Karsten Kryger Hansen Christine Hadrossek Joanna Jannik Diane Antoine Mathilde Bernier Saba Sharma Nida van Leersum Agnes Janiska	DKB/T4.2 AAU/T4.2 CNRS/T4.2 CNRS/T4.2 Uni.deLorraine/T4.2 CNRS/T4.2 TU DELFT/T4.2 TU DELFT/T4.2 UDEIN/T4.2	
Reviewed by:			
Approved by:			

DOCUMENT LOG

Issue	*Date*	*Comment*	*Author*
v.1	***09.08.2024***	*Initial draft based on internal outputs of task T4.2 and WP4,*	*Lorna Wildgaard (DKB), Nida, Dominique Green (UDEIN), Agnes Janiska (UDEIN), Joanna Jannik (CNRS)*
…			
…			
v.n			

TERMINOLOGY

https://eosc-portal.eu/glossary

Terminology/Acronym	*Definition*

Contents

1 Executive summary

Overview of the Learning Path Pilot

On the 29th of June, an online pilot for the learning path “Open Licenses for Data, Software, and Code” was conducted using Big Blue Button as the meeting platform. Participants accessed materials published on the Skills4EOSC learning platform. The training was facilitated by Agnes Jasinska (DCC) and Lorna Wildgaard (DKB), with technical support from Andrea Corleto and Gabriella Paolini (GARR). The content was divided into three units covering the importance of licensing research outputs, considerations before applying licenses, and the application process of licenses. Post-training, learners completed a homework assignment advising a fictive project PI on licensing their research outputs, fostering a collaborative learning environment.

Participant Evaluation Results

Thirty-three learners participated, with eighteen completing the evaluation survey (54% response rate). Seventeen respondents were satisfied or very satisfied with the pilot, and one was neutral. Sixteen respondents would recommend the learning path to colleagues. Most participants had intermediate to expert knowledge of Open Science principles, FAIR principles, and research data, but were novices to intermediates in research software, copyright, and licensing. Post-training evaluations indicated an improvement in knowledge about research software, copyright, and particularly licensing, which was the core focus of the training.

Positive Aspects and Areas for Improvement of the ToT

Participants praised the professional knowledge and communication skills of the instructor, the clarity and conciseness of the content, and the useful discussions that enhanced learning. Highlights included a congenial atmosphere, practical tools, and effective group activities. However, the two-hour duration was deemed too short, and more time was suggested for discussions and exercises. To optimise future training, it is recommended to reduce content to fit the allocated time and increase focus on discussions.

Additional training materials should be developed to address software and code licensing in more depth.

Technical Platform Evaluation

Technical difficulties with Big Blue Button significantly impacted the learning experience, necessitating better training and practice with the platform. The Big Blue Button platform added complexity to the training, with issues like slide upload problems, formatting changes, and navigation challenges between break-out rooms and the main room. Recommendations include sufficient training and practice with both Big Blue Button and the Learning Platform to avoid disruptions in future sessions.

Alignment with MVS for Data Stewards

The training covered essential skills and competencies for Data Stewards, particularly in open licensing for research data, software, and code. The pilot highlighted the need for including more detailed knowledge of copyright and licensing in the MVS for Data Stewards.

The MVS Data Steward should include a section dedicated to legal aspects of RDM, especially regarding FAIR data dissemination and sharing in compliance with legal recommendations. The pilot demonstrated the need for comprehensive knowledge for each concept listed in the MVS and the importance of operational training materials that go beyond the MVS, relevant also for legal experts and ELSI professionals.

Sources for Learning Path Development

The learning path was hosted on the Skills4EOSC Learning Platform and was adapted from content availible in an editable library of training materials on Github developed by Task 4.2. Together, these sources provide a foundation for creating a comprehensive and effective learning path on open licenses for data, software, and code in local contexts.

2 A brief description of the pilot training (ToT) and descriptive feedback from the evaluation

2.1 Brief description of Learning Path pilot

On the 29th of June, the pilot for the learning path "Open Licenses for Data, Software and Code was held". The pilot was held online using the the Big Blue Button as the online meeting room and participants were asked to access materials published on the Skills4EOSC learning platform during the pilot. The pilot was facilitated learning, with the content divided into three units each containing discussions and learning activities. :

1. Why you should license your research output and how to overcome resistance (slides 7 - 22)
2. What to consider before applying licenses to research outputs (slides 23-73)
3. How to apply licenses to research output (slides 75 - 105)

The slides used in the pilot can be found on the Skills4EOSC learning platform and are published in the Skills4EOSC Zenodo library, please refer to Section 5 of this report for more information. After the training, the learners were invited to complete a homework assignment. The assignment was to use knowledge learned during the pilot to advise a fictive project PI on how to license their research output. The learners were requested to post their advice on the Learning Platform, thus creating a forum where others can read and learn from the responses.

Materials used in the pilot were adapted from the WP4, T4.2 milestone which contains a library of slides for the trainer to choose from dependent on the focus of the training event, instructor notes, glossary, activities, suggestions for evaluation questions and competency badge (Section 5).

Agnes Jasinska (DCC) was the instructor and designed the materials used in the pilot. Lorna Wildgaard (DKB) was the facilitator and Agnes' sparring

partner in the lead-up to the pilot. Andrea Corleto (GARR) and Gabriella Paolini (GARR) provided technical support prior and during the pilot.

2.2 Results of the participants evaluation of the pilot

Thirty-three learners participated in the pilot of the learning path. Eighteen learners completed the evaluation, giving a response rate of 54%. Seventeen of the respondents to the evaluation survey were either satisfied or very satisfied and 1 participant expressed neutrality. The professional profiles of the learners is presented in Table 1, the "other" category was indicated by a data steward:

Table 1: Professional profiles of the 18 learners who responded to the evaluation survey

Response	Average	Total
Researcher	17%	3
Research support staff	44%	8
Research infrastructure professional	11%	2
Trainer (on data related topics)	17%	3
Service provider	6%	1
Other	6%	1
Total responses to question	100%	18/18

Sixteen of the respondents would recommend the learning path to a colleague. The majority of the learners described themselves as having intermediate to expert knowledge in Open Science principles, FAIR principles

and research data, and novice to intermediate knowledge of research software, copyright and licensing (Table 2).

Table 2: Participants knowledge of each area prior to participating in learning path

Responses	Novice	Intermediate	Expert	Total
Open Science principles	1 (6%)	5 (29%)	**11** (65%)	**17**
FAIR principles	1 (6%)	3 (18%)	**13** (76%)	**17**
Research data	1 (6%)	**9** (53%)	7 (41%)	**17**
Research software	7 (41%)	**9** (53%)	1 (6%)	**17**
Copyright	6 (35%)	**10** (59%)	1 (6%)	**17**
Licensing	4 (24%)	**11** (65%)	2 (12%)	**17**

After the training, the participants rated if their knowledge of the above topics had improved (Table 3, next page). Unsurprisingly, no vast improvement in their knowledge about Open Science principles, FAIR principles or Research data was acknowledged as the majority of learners were confident in these topics prior to training. The learners confirmed a somewhat increase in knowledge about research software and copyright and a considerable increase in their knowledge about licensing which was the core learning objectives of the training.

Table 3: Participants knowledge of each area after participating in learning path

Responses	Not really	Somewhat	Considerably	Total
Open Science principles	**9** (53%)	8 (47%)	0	**17**
FAIR principles	**10** (59%)	6 (35%)	1 (6%)	**17**
Research data	**9** (53%)	7 (41%)	1 (6%)	**17**
Research software	6 (35%)	**10** (59%)	1 (6%)	**17**
Copyright	1 (6%)	**12** (71%)	4 (24%)	**17**
Licensing	0	8 (47%)	**9** (53%)	**17**

The instructor was complimented for her professional knowledge and skills as a communicator of the complex topic of Open Licenses. The content was presented in a simple, clear and concise manner with useful discussions between participants that framed the flow of the materials and enhanced deeper learning. Even participants with existing knowledge of the subject came away from the training with new insights. All respondents to the evaluation scored the content as highly relevant, present clearly, good pace, useful activities, an effective instructor who created a good learning environment.

We asked the pilot participants in the evaluations: Were there any related topics not covered that you think should be added? In their responses, rather than proposing other, related topics, our participants mostly expressed a desire to examine the topic of licensing in more depth (and to have more time to do so). They asked for more examples of different projects and licenses; to learn more about software licenses in particular; to better understand the role and support of repositories in shaping the choice of the license; and to discuss dual licensing. The question of the impact of open licensing on a researcher's career also came up, evoking the broader discussion about academic research culture and the still often insufficient

incentives and rewards for researchers who openly share their research outputs and engage in Open Science.

Below, is a summary of what the respondents considered went well and where there is room for improvement.

What went well and useful content

- The congenial atmosphere
- Mitigation strategies
- Arguments to convince researchers to licences their work and dispelling misconceptions
- Exchanges in break-out rooms
- The richness of the content that can be adapted locally
- Links to and advice on practical tools
- The examples and cases used throughout the training
- Highlighting learning objectives at the start of the training/unit and wrapping up each unit with key take-aways and activities

Room for improvement

- The two hour learning path is too short for such a big topic. Reduce the content to avoid superficially addressing some content
- More time on group discussions. Some groups worked really well as a small group, others didn't interact as well. More time would allow the instructor to visit the groups and ensure the members have understood the assignment and perhaps kick-start the conversation
- More time dedicated to licensing software and code
- More exercises
- More time on activities – activity 2 could be placed after the unit about different licensing options. This could result in deeper and less abstract discussions
- Using multiple platforms was not a success, the shift from the learning platform, break out groups and main group on the Big Blue Button was not easy.

The major criticism from all participants were technical difficulties using the Big Blue Button which disrupted the teaching, causing delays and confusion. These are addressed in the next section.

2.3 Evaluation of the Big Blue Button and the Skills4EOSC Learning Platform

The Big Blue Button (BBB) is the online meeting room recommended in the Skills4EOSC project. However, in this pilot we experienced the BBB as adding an extra layer of complexity to conducting the training and many participants commented negatively impacted their experience. Prior to the training we coordinated with GARR on how to use the BBB. However, even with this preparation and technical support during the ToT, the training suffered technical difficulties. Thankfully the Tech support team were present during the pilot and saved the workshop.

If the training is to happen on the BBB and Learning Platform, then everyone in Skills4EOSC needs:

- sufficient training on the BBB
- sufficient practice on the BBB - for example: the WP meetings could be on the BBB, to let all of us get familiar with it and practice using the various features
- Sufficient practice using the Learning Platform
- Awareness that creating learning activities and adding content to the Learning Platform takes time that needs to be allocated in the development of the learning path

Difficulties experienced during the pilot are listed below and can be used to inform training in the BBB:

- Not being able to upload slides to the BBB
- Changes to slide content during import, formatting issues on the uploaded slides, missing animations and missing text,

- Not being able to to set up polls or break-out groups prior to the ToT
- Not being able to reuse break-out rooms
- Navigation for learners between break-out rooms and the main room (Note: this is an experience issue as most people are used to Teams or Zoom. On the BBB the break-out room opens in a new tab, so learners switch between the break-out room and the main room by clicking on the tab
- Sound issues when returning to the main group. We found out that reloading the page solved this issue
- Break out groups, with a timer on, end abruptly. No message that the rooms are closing is transmitted. There is a timer at the top of each break-out room – draw attention to it!
- If no timer is set on a break-out room, the break-out rooms have to be closed manually – the instructor/facilitator must broadcast to each group, one at a time, that it is time to return to the main room
- The learners in the break out rooms can click on the message “leave meeting” to return to the main room. This caused confusion, as it was interpreted as “leave the entire meeting”

2.4 Participants evaluation and alignment with the MVS for data stewards

In terms of the skills and competencies covered in the pilot, as the title suggests, the main focus was on open licenses for research data, software, and code. Four learning objectives were shared with the students:

- Explain the different types of licenses that can be applied to FAIR research outputs in accordance with regulations, policies and other requirements.
- Apply an appropriate license to a research output.
- Argue for the choice of appropriate licenses to research output and for the importance of creating machine-readable licenses.
- Engage in the research process and influence it to accommodate for more available and licensed data, software, and code.

The training was organized into 3 units. Unit 1 covered the rationale and benefits of applying open licenses to research products. It also discussed common concerns of researchers on the one hand, and ways to alleviate these concerns and advocate for open licensing on the other. Unit 2 presented a license selection checklist and discussed each of the ten main factors to consider, from identifying the rights holder and the funder, publisher, and repository requirements, to data privacy laws, commercialization potential, and dual use issues. Unit 3 presented an overview of available open licenses, including the Creative Commons licenses, together with recommended licenses for specific research products and tools that help in the license selection process. The importance of metadata and machine-readable licenses was also discussed.

In their evaluations, participants commented on the relevance of the topic in general, on the value of the license selection checklist and the practical guidelines to select the appropriate open license, as well as on the usefulness of the examples of arguments to advocate for open licenses and how such licenses are beneficial to both the creators of the research outputs and the scientific community and the society more broadly.

Below, two examples of what the participants found most valuable about the course:

"The clear description of what licenses are available for what outputs, as well as the practical tools that can be used to help with the license selection process."

"Explanation of licenses and counterpoints for convincing those who are reluctant to deploy them."

Some participants also commented that, even if the content was already familiar, the training would help them teach or explain this content to others:

"For me, being an RDM consultant/trainer and test user, is was not so much the content that was important or new but how it was presented. I am taking some good ideas with me, thank you!"

Finally, participants also remarked that the training was introductory, so most suitable to people new to the topic of licensing, and that a longer training would allow a more in-depth examination of the issues involved, including more real-life examples and more group activities and discussions.

3 Alignment with the MVS data steward

In this section we explore further the alignment between the activities described in the MVS for Data Steward, as a coordinator and embedded, and the training materials used on the pilot. Particular attention is paid to which activities are covered and which, if any, content is missing in the MVS.

3.1 Which skills and competencies from the training materials are missing in the MVS based on feedback from participants?

The topic of open licenses for data, software, and code (and copyright and intellectual property rights more generally) is considered by the participants as important and useful to data stewards and others working with research data and research software. This is worth noting because currently the Minimum Viable Skillset (MVS) for Data Stewards licensing is only explicitly acknowledged in the Embedded Data Steward profile under main activities in regards to data management planning. Understanding how to select, apply, and advocate for open licenses for various types of research outputs is an essential skill of data stewards. This is especially true in the context of Open Science, in which open licenses both enable and encourage reuse of research outputs. This critical role of licensing is acknowledged in the FAIR principle R1.1 for both research data *("R1.1. (Meta)data are released with a clear and accessible data usage license") and research software ("R1.1. Software is given a clear and accessible license").*

One of the pilot's participants confirmed the relevance of open licensing to Data Stewards in their evaluation:

"As a Data Steward, I must remain well-informed about the areas you've mentioned. However, my primary focus on addressing the needs of researchers often limits my capacity to update my knowledge base. Consequently, your course is critical in helping me refresh and expand my

understanding. The tips and insights you provide are invaluable, enabling me to enhance my interactions and discussions with the researchers I support."

Subsequently, a suggestion would be to include additional text regarding licensing as a part of the essential skills for data stewards. It is recognised that the MVS is the minimum skills needed, so the additional text should reflect that. An example of this could be: Working knowledge of copyright and licensing as it relates Open Science and other research products.

3.2 Comparison of the MVS Data Steward and Learning Path for DS

The pilot learning path provides more intermediary material that goes beyond the MVS. Table 4, below, lists how the learning path is supporting elements of the MVS.

Table 4: How the learning path supports the MVS Data Steward

This learning path contributes to the following open science outcome for a Data Steward: Research data and related digital objects are effectively managed to ensure their suitability for curating, sharing, and reuse, and potential impacts towards advancement of research methods appropriate to the discipline(s). Digital research objects are made as FAIR and open as possible, and as closed as necessary
The learning path provides material that would help the Coordinator and Embedded Data steward in the following main activities: *Coordinator Data Steward:* Develops institutional guidance on Data management Planning, e.g. templates offering cross-domain knowledge to contextualise data handling and **advice on planning** how to use local services or infrastructure. **Understands research stakeholder needs** and contributes to developing, implementing and monitoring Data Policy and Governance, along with

service level management to support this.

Embedded Data Steward:

Develops Data Management Plans templates tailored for research teams, offering support in writing a DMP according to the relevant template. Includes **provision for archiving and FAIR sharing** (standards, metadata exposure, PIDs, **licensing,** data repository management/selection)

Implements **good practice on data and/or software/code** during proposal development for funders, and as a regular aspect of doing research, and liaises with other experts inside and outside the institute to adopt effective solutions to challenges.

Identifies gaps and takes action to ensure **ethical conduct and awareness of the potential impacts of data reuse**, management and sharing on wider society.

Supports researchers on **legal and regulatory compliance** aligning local practices with these through connections with the institutional privacy officers, legal advisers, and research ethics bodies.

Essential skills and competencies covered are:

Service provision to support specific Open Science practices including: applying **FAIR** and CARE principles, **Open Access (publishing) , data curation and preservation**

Knowledge brokering about Research Data Management, (personal) data governance and ethics, including to understand information security challenges, and provide access risk assessment and mitigation

Mentoring on open and fair methods, to develop professional practice including knowledge/awareness of programming, FAIR code and FAIR software and use of standards and ontologies.

Advocacy, analysis and assessment on FAIR data criteria, FAIR code and software preservation.

Stakeholder engagement and collaboration building strategic relationships, bridging needs, and speaking and presenting to data creators, users, and research stakeholders about the value of good data management.

Creativity, critical and analytical thinking, curiosity, openness, and cultural competence with a willingness to learn.

As indicated above, the proposed learning path offers the possibility of acquiring intermediate skills, beyond those proposed in the MVS. Below, Table 5, is a list of skills from MVS's Essential skills and competences that are aligned with the skills acquired as a result of the learning path.

Table 5. The alignment between the MVS skills and competencies and the training materials

MVS's Essential skills and competencies	Covered in training material skills
Service provision to support cross-domain/domain specific Open Science practices including: use of FAIR and CARE principles, Open Access, data optimization, data preservation, archiving and responsible re-use	Data & software policy; Knowledge to contextualise FAIR principles to legal domain (licensing)
Knowledge brokering about Research Data Management (personal) data governance and ethics, Open Science data publication and exchange(sharing) services, intellectual property rights, information security and risk management.	Knowledge to contextualise FAIR principles to legal domain; Information security
Awareness raising among data creators and users, researchers, organisational colleagues, and decision-makers of the value of good data management.	Knowledge of preservation for FAIR research outputs (selection of appropriate licenses)
Knowledge/awareness of	Software preservation; Metadata

programming, FAIR code and FAIR software and use of standards and ontologies	exposure (knowledge of machine-readable licenses)
Advocacy, analysis and assessment on FAIR data criteria, FAIR code and software preservation.	Awareness about advocacy for licensing

4 Lessons learned and recommendations

In this section, we summarise our lessons learned from the pilot with the aim to optimise the efficiency of the hosting a ToT event. Further, we recommend adaptions to the training materials and ultimately additions to the MVS Steward.

4.1 Operational efficiency

The BBB adds a layer of complexity to teaching online. The instructor and facilitator need to invest time in using the BBB and guiding participants in how to link to and from the learning management system, where learning materials are published. When teaching online, it is important to create a comfortable learning environment, where the learners can focus on the materials. The technical platform must not distract from the learning experience or the quality of the training for learners and the pedagogical strategies of the instructor.

Even though we technical challenges, the pilot has contributed to our better understanding of the use and limitations of the BBB. We are confident in, that through our collaboration with GARR on this ToT processes around the use of BBB will be tested further and technical glitches reduced in the long run. We therefore recommend that everyone in Skills4EOSC needs:

- sufficient training on the BBB
- sufficient practice on the BBB - for example: the WP meetings could be on the BBB, to let all of us get familiar with it and practice using the various features
- Sufficient practice using the Learning Platform
- Awareness that creating learning activities and adding content to the Learning Platform takes time that needs to be allocated in the development of the learning path

4.2 Adaption of Learning Path

Further adaptions of the training materials are recommended to reduce the amount of content and use more time on discussions. Most of the content addressed licensing data, develop more indepth material about software and code.

As learners come to the training with different levels of knowledge of the topics, use effort facilitating the discussion groups to ensure assignments are understood and conversations are flowing.

4.3 Revision of the MVS Data Steward

The ToT pilot addressed only a small sub-set of skills described in the MVS Data Steward. The MVS DS deals with skills and competencies in a fairly general way. The training material on the other hand is more operational, showing use cases where the arguments for convincing people to use the licenses are clearly identified (slides 16 and 20). We recommend including additional text in the MVS regarding licensing as a part of the essential skills for data stewards, for example "Working knowledge of copyright and licensing as it relates Open Science and other research products". Hence, the MVS would gain in precision by proposing a section dedicated to the legal aspects of the RDM, in particular concerning the dissemination and sharing of FAIR data in compliance with legal recommendations (obligations).

If anything, the training exemplifies how much knowledge is needed for each single concept listed in the MVS as this learning path provides more intermediary material that goes beyond the MVS and is also relevant for MVS for Legal Experts.

We consider the training materials also to be in alignment with the MVS for Legal Experts as a supplement to their technical skills and competencies in copyright licensing rules and existing frameworks. Likewise the materials,

especially unit 1 concerning why researchers should license their research outputs and common myths why not to do so, are highly relevant for ELSI professionals to further their understanding of the psychosocial-disciplinary context and concerns of the researcher.

5 Sources for Learning Path Development

The learning path described in this report builds on the following sources of information:

The Moodle created for the pilot, published on the Skills4EOSC Learning Platform: https://learning.skills4eosc.eu/course/view.php?id=29

The Moodle contains slides and learning activities used during the pilot. The training materials used in the pilot were adapted from the library of suggested materials created in the T4.2, below.

Library of suggested materials created by T4.2

An editable version of the library of training materials including instructor notes published on Github and availible via the following link: https://task-4-2.github.io/Open-Licenses-data-code-and-software/latest/

And also in Zenodo: https://zenodo.org/records/12703494

These materials are intended to be used as a “pick and mix” slide deck that covers Open Licences for Data, Software and Code in a greater depth that is possible to communicate within a 2 hour learning path. The intention is, that traininers select content and adapt to their local environment.