



Adaptation for Climate Change

Data Repository Rules

Deliverable 10.4

Status: Public document from 21 February 2018

Version: 3.0

Date: 21/02/2018



EC contract no 654110, HYDRALAB+



DOCUMENT INFORMATION

Title	Data Repository Rules
Lead Authors	Richard Holmes, James Sutherland, Quillon Harpham
Contributors	LNEC, UPorto, Deltares, CNRS.
Distribution	PU
Document Reference	DOI: 10.5281/zenodo.1182549

DOCUMENT HISTORY

Date	Revision	Prepared by	Organisation	Approved by	Status
20/06/2017	1.0	Richard Holmes	SAMUI		Restricted
02/11/2017	2.0	Richard Holmes	SAMUI		Restricted
21/2/2018	3.0	Richard Holmes	SAMUI	Hamer – HYDRALAB+ Coordinator-	Public

ACKNOWLEDGEMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654110, HYDRALAB+.

DISCLAIMER

This document reflects only the authors' views and not those of the European Community. This work may rely on data from sources external to the HYDRALAB project Consortium. Members of the Consortium do not accept liability for loss or damage suffered by any third party as a result of errors or inaccuracies in such data. The information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and neither the European Community nor any member of the HYDRALAB Consortium is liable for any use that may be made of the information.

LICENSE

This report is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

EXECUTIVE SUMMARY

This report is part of HYDRALAB+ Work Package 10, which seeks to strengthen the coherence of experimental hydraulic and hydrodynamic research undertaken by the HYDRALAB+ partner organisations. It outlines the work carried out by Samui Design & Management Ltd, supported by HR Wallingford, in relation to the development of a data repository structure that can be adopted by participants of the HYDRALAB+ project.

There is a requirement to ensure that HYDRALAB+ scientific publications and research data are open access. As such, the EU funded, free-to-use Zenodo repository (<https://zenodo.org/>) was selected for storing the outputs of HYDRALAB+ experiments. It has an Application Programming Interface (API) to allow integration with the project website. Participants therefore only need to make one step to complete both the publication to the project website and to make it open access beyond the life of the project. Zenodo provides a unique DOI for each dataset which can be cited in publications and papers and will remain valid ensuring publications remain relevant for many years. In addition, it allows for communities to be created to group datasets.

A video presentation and screenshot walkthrough are given by way of user instructions.

CONTENTS

Document Information.....	2
Document History	2
Acknowledgement.....	2
Disclaimer	2
License	2
Executive Summary.....	3
Contents.....	4
1 Introduction	5
2 Background	5
3 Key Issues.....	5
3.1 Protection Requirement.....	5
3.2 Openly Accessible	6
3.3 Free of Charge.....	6
4 Repository	6
5 How To Use the HYDRALAB+ Zenodo Data Repository	7
5.1 Video Presentation for a Transnational Access Project	7
5.2 Screenshot Walkthrough for a Transnational Access Project.....	7
5.2.1 Login to the TA User Area.....	7
5.2.2 Dashboard.....	8
5.2.3 Add Dataset Details	8
5.2.4 DOI Generation	11
5.2.5 Zenodo Platform Preview	12
5.2.6 View Datasets.....	13
5.2.7 Ready to Publish.....	13
5.2.8 Publish Confirmation.....	14
5.2.9 Post Publishing	14
5.3 Application to a Joint Research Activity or Networking.....	15

1 INTRODUCTION

This report is HYDRALAB+ project deliverable D10.4 'Data Repository Rules (including DOIs)'. It is part of Work Package 10, which seeks to strengthen the coherence of experimental hydraulic and hydrodynamic research undertaken by the HYDRALAB+ partner organisations. It has been produced in conjunction with D10.1 'Data Management Plan', D10.2 'Critical Review' and D10.3 'Data Standards Report'. In particular, it outlines the work carried out by Samui Design & Management Ltd, supported by HR Wallingford, in relation to the development of a data repository structure that can be adopted by participants of the HYDRALAB+ project. The document will begin with the background behind the open access requirement, followed by the decision process undertaken and finishing with the rules and procedure for the data repository system adopted.

2 BACKGROUND

As HYDRALAB+ is part of the Horizon 2020 programme there is a requirement to ensure that scientific publications and research data are open access¹. This is defined as "the practice of providing online access to scientific information that is free of charge to the end-user and reusable. 'Scientific' refers to all academic disciplines." Furthermore access must also allow the right to copy, distribute, search, link, crawl and mine the data.

The open access requirement only becomes applicable if the decision is made to publish the data. The decision to publish is up to the grant beneficiaries and possible reasons for not publishing maybe to seek some form of protection, such as patents, to ensure the data remains commercially exploitable².

The key issues for consideration are:

- The degree of protection required for the submission;
- Whether or not it is openly accessible;
- Whether or not it is offered free of charge.

3 KEY ISSUES

3.1 PROTECTION REQUIREMENT

The decision as to whether data may remain protected must be discussed with the Management Team (MT), however, it will ultimately be at the discretion of the author. HYDRALAB+ has committed to the open access requirement and therefore encourages all participants to publish data and publications. The process by which this can be done is demonstrated in section 5 of this document.

¹ http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

² Horizon 2020 Programme, 2017. Guidelines on the rules on open access to scientific publications and open access to research data in Horizon 2020. Version 3.2.

3.2 OPENLY ACCESSIBLE

Openly accessible means there must be no barriers to be able to access and thereafter view the data. This means there must be no password protections in place either before access or after download of the data. There must be no financial barriers to be able to access and use the data and no steps should be taken to knowingly discourage the data from being found via searches. A suitable open license should be used³.

As part of the research process and discussions, it was determined that a Digital Object Identifier (DOI) would offer us with a unique referencing system for each dataset and provide us with a basis for any search facility.

3.3 FREE OF CHARGE

To be open access the data must be accessible free of charge (FOC). This only means FOC to the end user wishing to use the data and does not necessarily need apply to HYDRALAB+ project, in this case. The system used to distribute the data may charge for the service and this cost would need to be paid by HYDRALAB+ and not the end user.

As HYDRALAB+ is a project with a life span it was decided that free options should be considered first as it may be difficult to seek agreement for costs that would then be occurred after the project has finished. External and in-house options formed the basis for the research with partners approached, initially, to establish any DOI systems used by their organisations and what, if any, services they could offer our project for the foreseeable future. The University of Hull is currently developing their own DOI system, however, it is unlikely to be ready in time for HYDRALAB+ and as other organisations didn't have robust systems in place for allowing external submissions the research turned to external options.

4 REPOSITORY

Having exhausted any in-house options, researching external routes proved to be fruitful at the first step. Zenodo (<https://zenodo.org/>) is a DOI provider that was developed, in particular, to make science more open. It has been funded by OpenAire (<https://www.openaire.eu/>), which is an EU funded project and has therefore been tailored with Horizon2020 projects in mind. The platform is free to use and provides an Application Programming Interface (API) to allow us to integrate the platform with the project website. This has great significance in that we can combine the uploading and publication of data using the system, that was due for development, on the HYDRALAB+ website. Participants will therefore only need to make one step to complete both the publication to the project website and to make it open access beyond the life of the project.

As Zenodo provides a unique DOI for each dataset⁴. The DOI reference can be cited in publications and papers and it will remain valid ensuring publications remain relevant for many years. In addition, Zenodo allows for communities to be created to group datasets, making it easier to search for

³ For more details see HYDRALAB+ Deliverable D10.3 Data Standards Report.

⁴ A new versioning system for Zenodo creates 2 DOIs <http://blog.zenodo.org/2017/05/30/doi-versioning-launched/>, but this feature has not been used here.

outputs from one location. Having this facility will make the search process easier, for end users, and therefore increase the datasets openness.

For these reasons it was felt Zenodo was a good fit for the open access requirements of HYDRALAB+ so a HYDRALAB+ community has been created in Zenodo (<https://zenodo.org/communities/hydralab>) and the API has been integrated with the HYDRALAB+ project tools. The use of Zenodo is not compulsory, as other suitable repositories exist. Any other repository used should be on the list of OpenAire Compatible data providers found at <https://www.openaire.eu/search/data-providers>.

An interesting new feature in Zenodo (which we have not yet exploited) is DOI versioning⁵. This allows you to update a record's files after they have been published by creating two DOIs for each upload:

- A DOI representing all versions of the record and
- A DOI representing each specific version of the record.

This could be used, for example, to produce different DOIs for the raw data and for each version of processed data. Storing data in a version control system, such as GitHub, can assist in the development of processed data from raw data files. Zenodo is set up to link to GitHub, which provides a means of preserving on Zenodo the version of data used in a journal paper.

5 HOW TO USE THE HYDRALAB+ ZENODO DATA REPOSITORY

5.1 VIDEO PRESENTATION FOR A TRANSNATIONAL ACCESS PROJECT

The following video provides instruction on how to use the data repository system for HYDRALAB+ Transnational Access project:

<http://hydralab.eu/DOI-setup-instructions-480.mov>

5.2 SCREENSHOT WALKTHROUGH FOR A TRANSNATIONAL ACCESS PROJECT

The following walkthrough gives a written set of instructions for the basic steps for using Zenodo as part of HYDRALAB+. The following example is for a Transnational Access (TA) project, where a drop-down list of project names has been created. Other documents, such as journal papers and datasets from Joint Research Activities (JRAs) can also be deposited in Zenodo – see Section 5.3.

5.2.1 Login to the TA User Area

Access the Zenodo repository via the HYDRALAB+ website through this link - <http://hydralab.eu/taking-part/transnational-access-projects/>, using your email address and password. If you have not yet registered you may use the Register option.

⁵ <http://blog.zenodo.org/2017/05/30/doi-versioning-launched/>

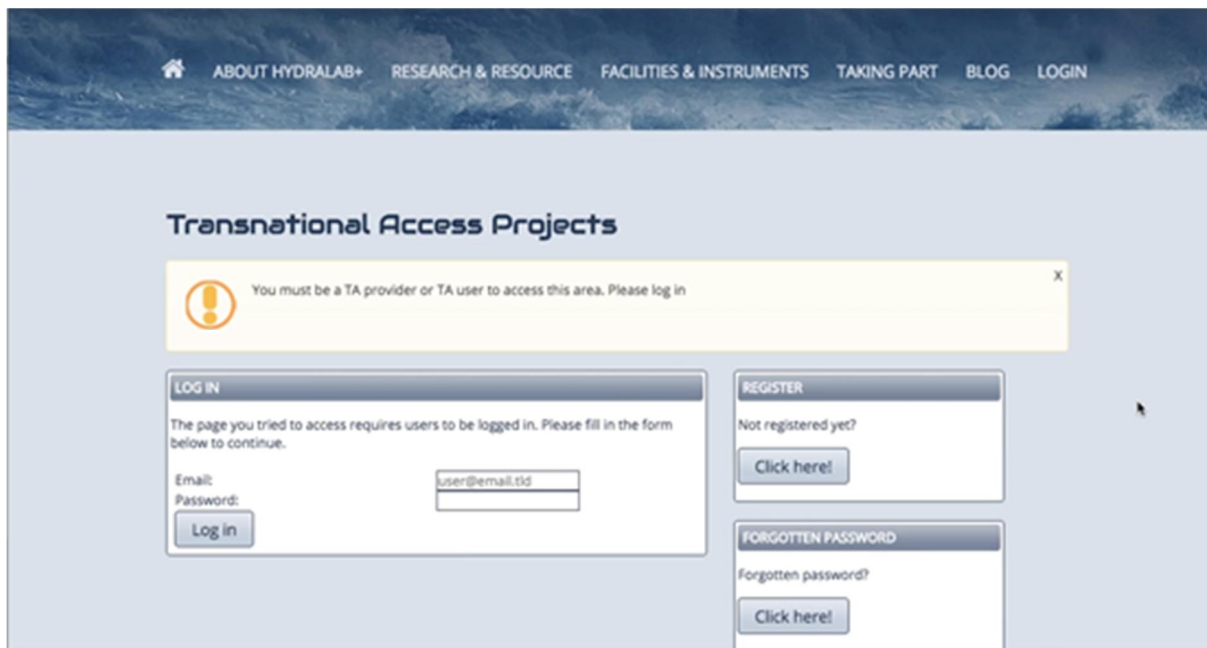


Figure 1. Logging on to the Transnational Access area of the HYDRALAB+ website

5.2.2 Dashboard

The dashboard may vary depending on your access rights but the Add Dataset icon will be present for all, as shown in Figure 2.



Figure 2. Transnational Access Dashboard

5.2.3 Add Dataset Details

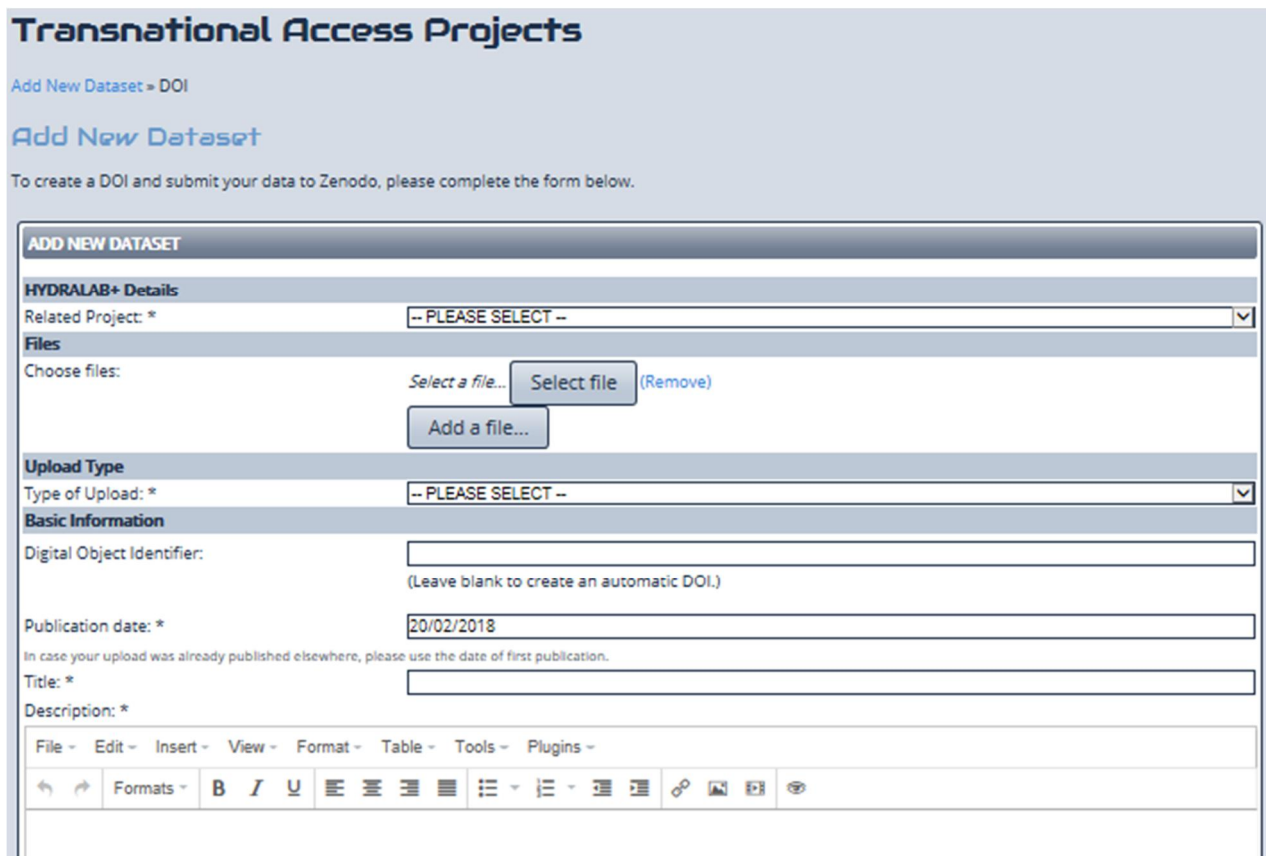
The dataset form requires a number of fields to be completed (Figure 3 and Figure 4). It is not necessary to upload a file on first submission, you can simple enter the basic details in order to generate a DOI reference. You can then add the DOI into the documentation / metadata in the

dataset (including in the data storage report or a data paper) before uploading. Should you not wish for your dataset to be open access, please contact a HYDRALAB+ administrator for agreement.

For transnational access projects there is a drop-down list of related projects. Various types of data can be uploaded:

- Publication;
- Poster;
- Presentation;
- Dataset;
- Image;
- Video/audio;
- Software; and
- Lesson.

Leaving the 'Digital Object Identifier' field blank will cause Zenodo to allocate a new DOI for you. You have free choice of title, description (Figure 3) and keywords (Figure 4).



Transnational Access Projects

Add New Dataset = DOI

Add New Dataset

To create a DOI and submit your data to Zenodo, please complete the form below.

ADD NEW DATASET

HYDRALAB+ Details

Related Project: *

Files

Choose files:

Upload Type

Type of Upload: *

Basic Information

Digital Object Identifier:
(Leave blank to create an automatic DOI.)

Publication date: *
In case your upload was already published elsewhere, please use the date of first publication.

Title: *

Description: *

File - Edit - Insert - View - Format - Table - Tools - Plugins -

← → Formats - **B** *I* U [List icons] [Link icon] [Image icon] [Media icon]

Figure 3. Transnational Access project data fields(top)

Authors should enter their name, current affiliation and an ORCID number⁶, which is a persistent digital identifier for each individual that they retain wherever their career takes them (Figure 4). The use of ORCID numbers is becoming widespread in academia and is increasingly being asked for in grant applications and publications⁷. We encourage the use of ORCID numbers, as many researchers involved in HYDRALAB+ are likely to move to different institutions as their careers progress. In these cases an individual's affiliation and email address will change, but their ORCID number will not. They act as the equivalent of a DOI for a person.

An important field for many HYDRALAB+ researchers will be "Access right" where the following options are given:

- Open Access – where data becomes freely available as soon as it is published;
- Embargoed Access – where data is deposited on Zenodo, but is only made open access at a set date. The embargo date can be up to two years after the experiments took place (but should be as soon as possible). Metadata will be published immediately, so people can find out about your data and when it will become available.
- Restricted Access - where you specify the conditions under which you grant users access to your files. User requesting access will be asked to justify how they fulfil the conditions and you decide if they can have access. This option should not be selected as it contravenes the principles of Open Access in Horizon 2020.
- Closed Access. This option can be selected if you seek some form of protection, such as a patent, to ensure the data remains commercially exploitable⁸.

An open access license should be chosen. Data entered here will automatically be uploaded into the HYDRALAB+ community on Zenodo.

⁶ <https://orcid.org/>

⁷ For example, <https://www.elsevier.com/authors-update/story/innovation-in-publishing/new-orcid-id-aims-to-resolve-authorship-confusion> and <http://www.rcuk.ac.uk/research/orcid/>

⁸ Horizon 2020 Programme, 2017. Guidelines on the rules on open access to scientific publications and open access to research data in Horizon 2020. Version 3.2.

POWERED BY TINYMCE

Keywords:
 Separate each keyword with a ; (semi-colon)

Additional Notes:

Authors *

Name (Family Name, Given Name)	Affiliation	ORCID
<input style="width: 95%;" type="text"/>	<input style="width: 95%;" type="text"/>	<input style="width: 95%;" type="text"/>

License

Access Right: *

- Open Access
- Embargoed Access
- Restricted Access
- Closed Access

Required. Open access uploads have considerably higher visibility on Zenodo.

For more information on each of the license options please [click here](#)

License: *

Communities

Select Communities: *

- EC Funded
 This submission will automatically be allocated to the 'Hydralab+' community.

Funding

Select Grants: *

This submission will automatically be allocated to the 'Hydralab+ Adapting to climate change' grant.

Figure 4. Transnational access project data fields (bottom)

5.2.4 DOI Generation

When you save the form details you will receive a success message (Figure 5) including your allocated DOI reference. Please make a note of it.

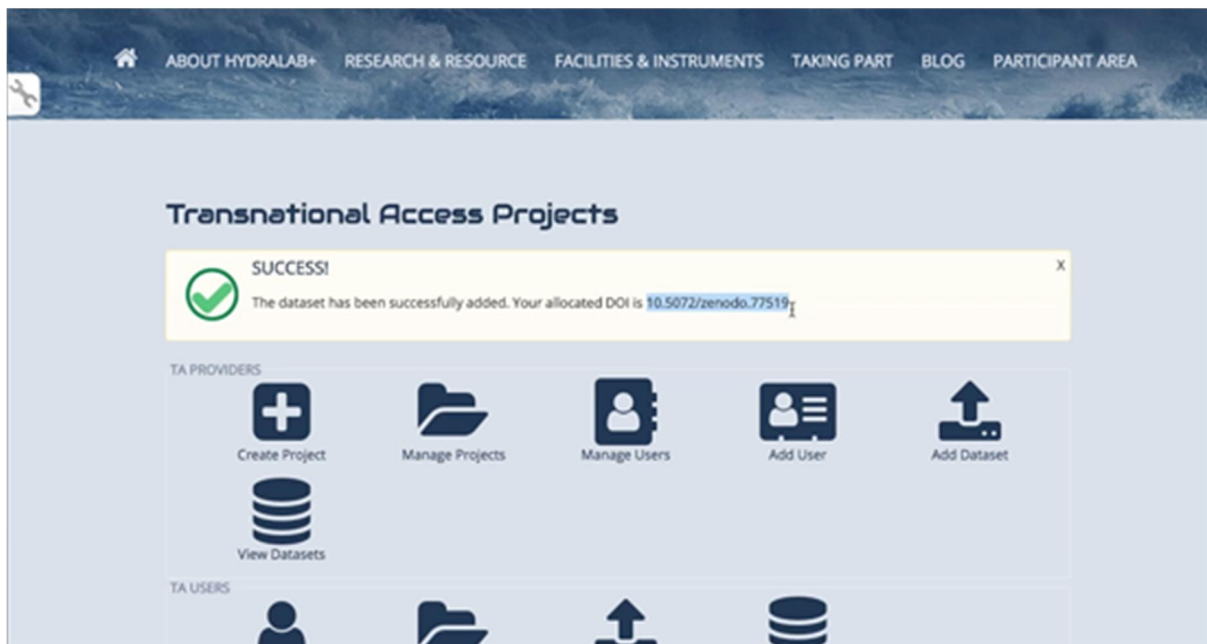


Figure 5. DOI Generation

5.2.5 Zenodo Platform Preview

There will be no need for you to access the Zenodo platform itself, however, when you press save a screen like Figure 6 will appear on Zenodo for users who do access it directly.

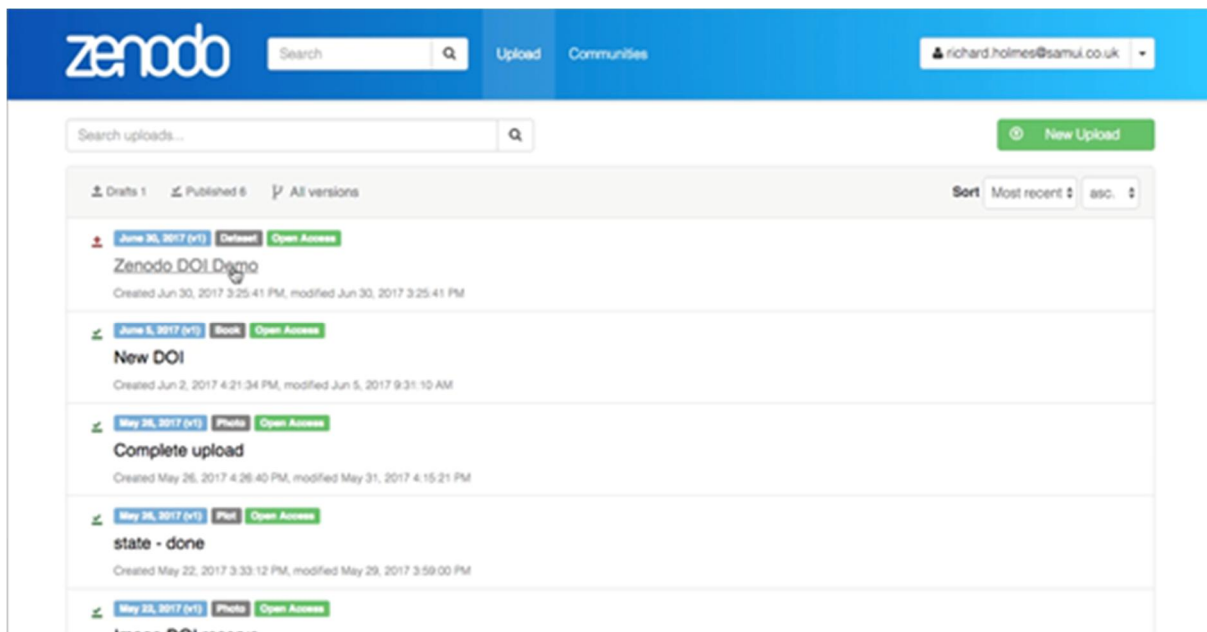


Figure 6. New DOI on Zenodo

5.2.6 View Datasets

From the dashboard you can view datasets, in order to manage those you have submitted (Figure 7). Here you may Update or Delete datasets you have entered.



Figure 7. View datasets

5.2.7 Ready to Publish

In order to publish your dataset you must have at least one file uploaded. You can add files using the Update option, seen in 5.2.6 and then saving the amended information. Returning to the View Datasets page will then show a new 'Publish' button (Figure 8) allowing you to publish to Zenodo.

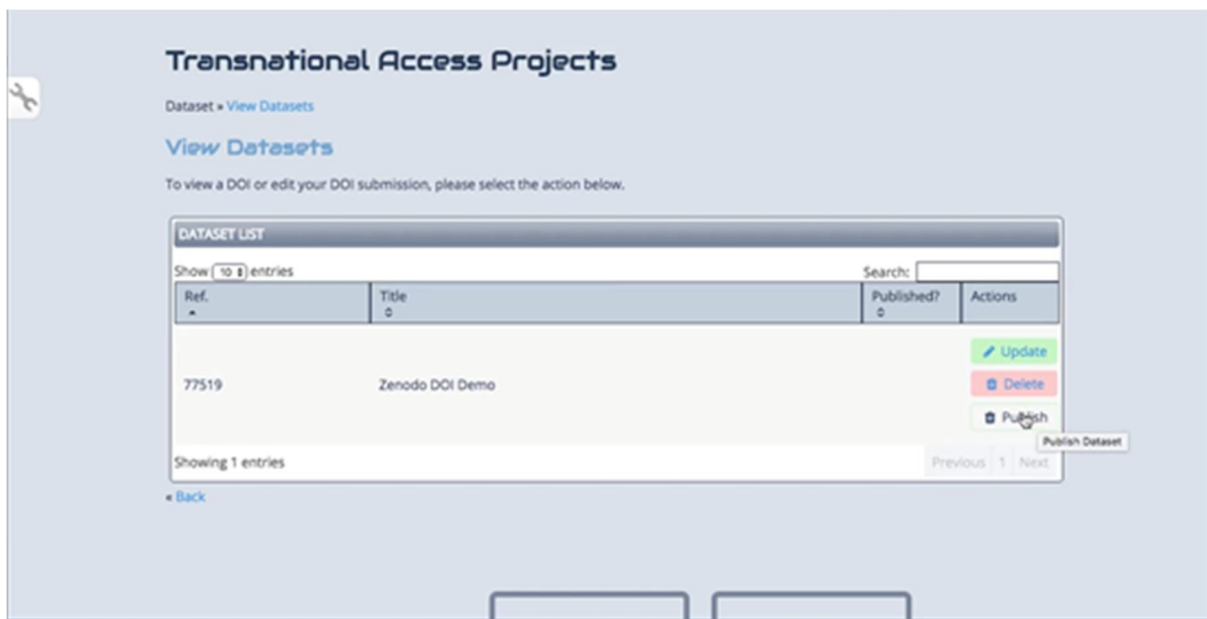


Figure 8. Updating and publishing a dataset

5.2.8 Publish Confirmation

Please be aware that publishing a dataset is not reversible and should only be carried out once you are happy all information is correct and all files are present.

A warning screen, like Figure 9, appears when you click 'publish', giving you the option to reconsider.

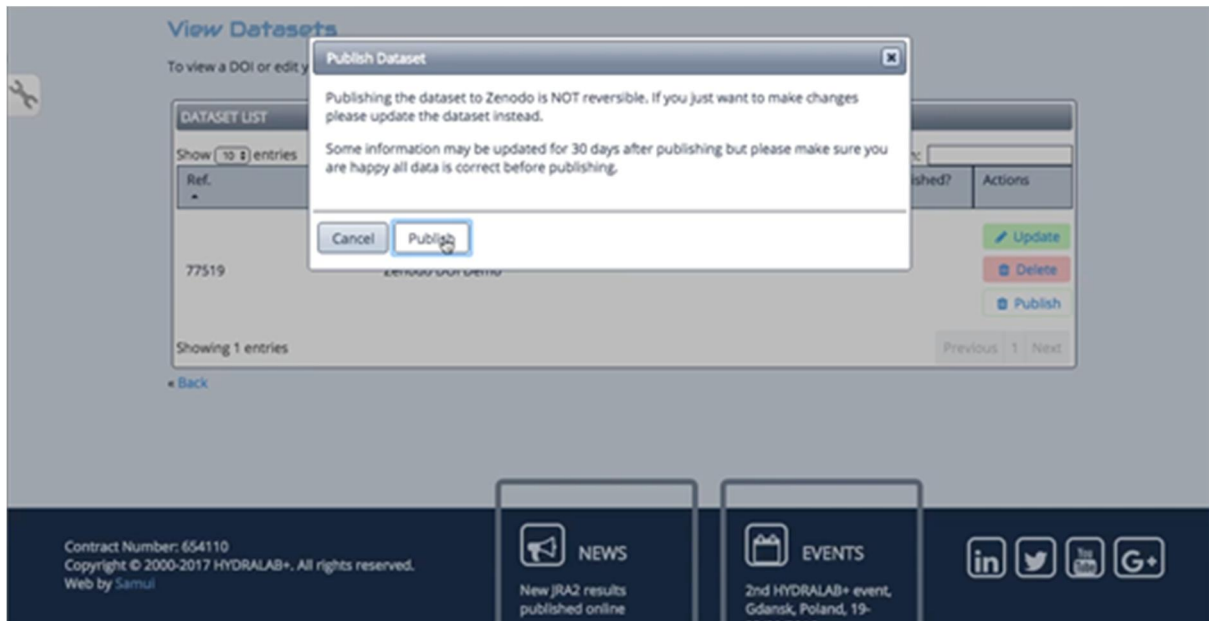


Figure 9. Warning message when publishing data on Zenodo

5.2.9 Post Publishing

Having published your dataset you will no longer see the Delete or Publish option under the View Datasets page. A new View button is present (Figure 10) which you will direct you straight to the DOI page on Zenodo, where you can see it in its published state (Figure 11).

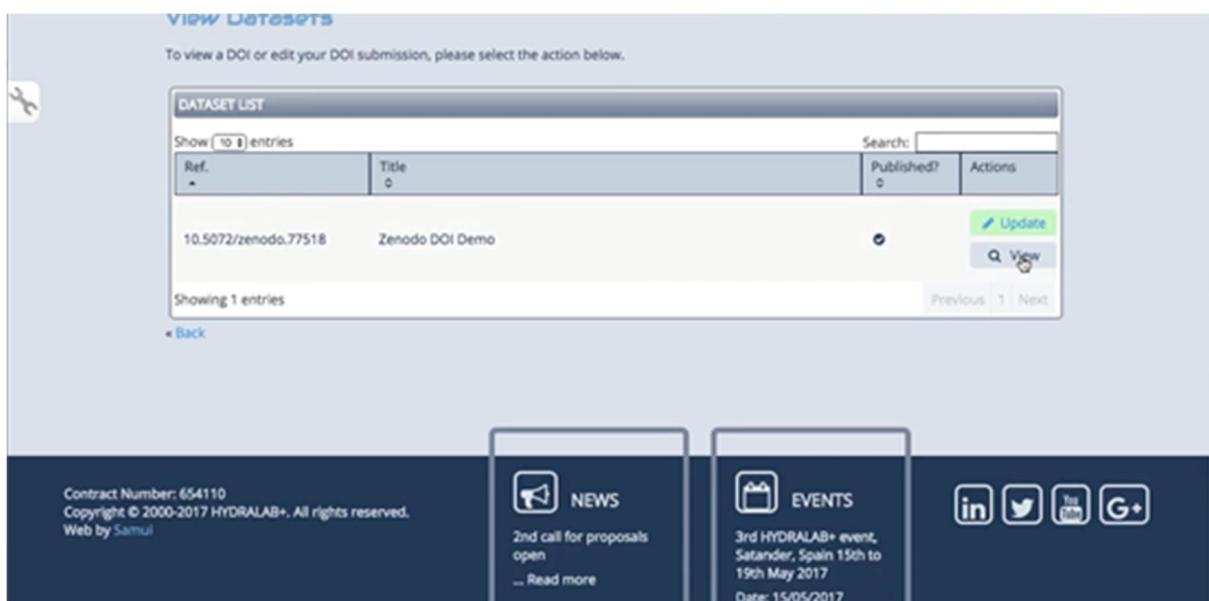


Figure 10. Post-publication screen view showing new 'view' button

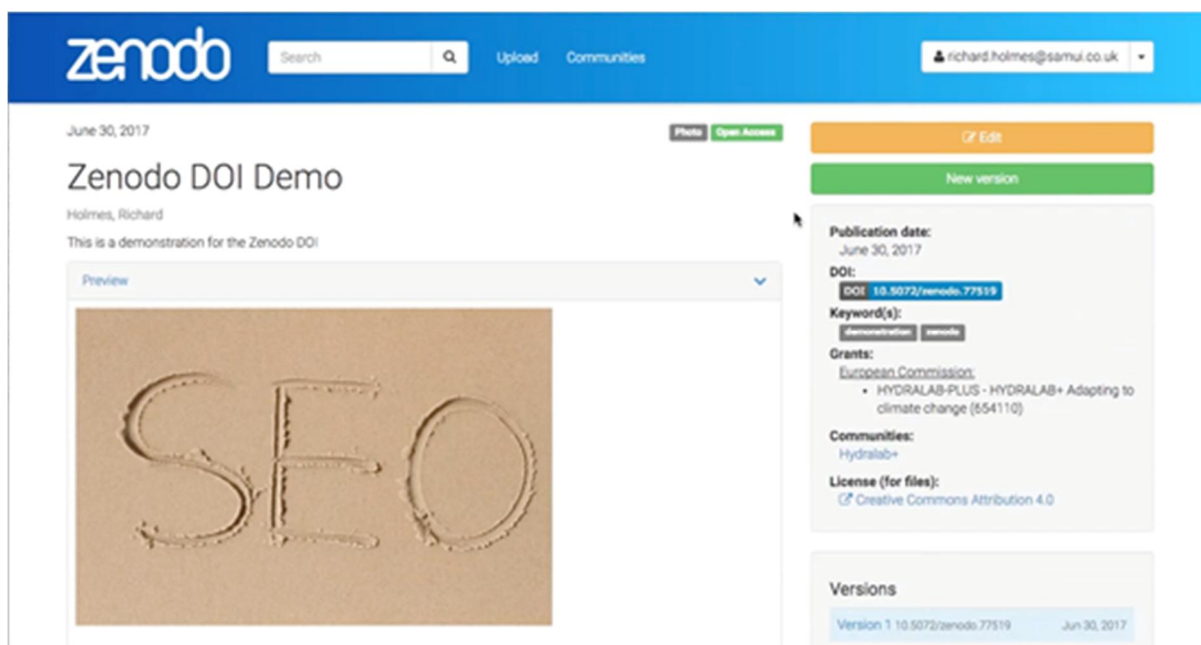


Figure 11. View of demonstration dataset on Zenodo

5.3 APPLICATION TO A JOINT RESEARCH ACTIVITY OR NETWORKING

Other experiments are being carried out as part of the HYDRALAB+ Joint Research Activities (JRAs) which will be producing datasets, public deliverables and conference / journal papers. Further public reports and deliverables, as well as other products like papers and posters can be produced in our Networking Activities. The option for preserving these datasets on Zenodo has been provided through the 'DOI datasets' link at <http://hydralab.eu/participant-area/>.

Registered participants on the HYDRALAB+ website can log in to the website on that page and will find the dashboard shown in Figure 12. Selecting 'DOI datasets' takes you to a new window, shown in Figure 13, that allows you to edit your DOI submissions or add a new one.

Clicking on add new dataset takes you to a form that is very similar to the one shown in Figure 3 and Figure 4, only with the drop-down list of TA project names replaced with a free-text 'Hydralab reference' box. If you have a project deliverable, report or paper to upload, select 'Publication' from the 'Type of Upload' drop-down list, then select the appropriate form of deliverable ('project deliverable', 'journal article', etc) from the 'Publication type' drop-down list that appears when 'Publication' is selected.

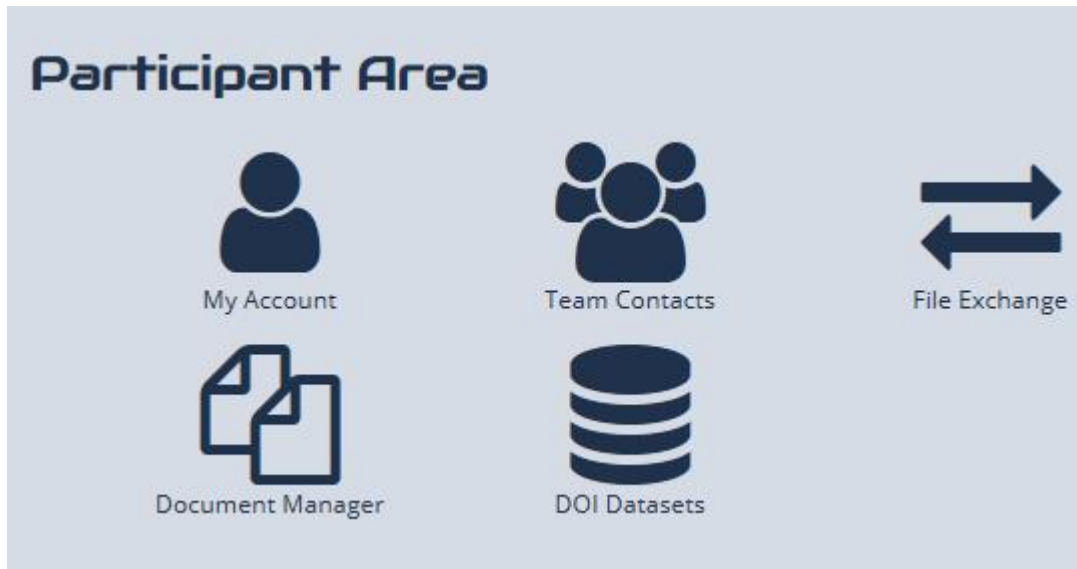


Figure 12. Dashboard for uploading data to Zenodo from JRAs and Networking



Figure 13. Edit or add a new dataset