International Journal of Information Studies & Libraries
7 (1) 2022, 23-34
http://publishingindia.com/ijisl/

Building Digital Repositories with the Open-Source Software Invenio: Use of SaaS Model Zenodo

Atasi Sinhababu*, Heenam Gakhar, Rupak Chakravarty*****

Abstract

Owing to the prevailing 'unfair' digital divide, libraries worldwide are unable to build and manage institutional repositories for preserving and sharing the research output of academia. This paper evaluates and demonstrates how libraries can practically harness the completely free SaaS service model based on public cloud deployment model infrastructure to fill this divide and achieve the larger goal of open science. The paper highlights the process and steps of using Zenodo, a FREE and OPEN platform, powered by Invenio (Free Open-Source Software) RDM infrastructure, to establish a trusted repository with the provision of self-archiving. To test the KPIs and functionalities, a live online community '*dlispu*' was built on the Zenodo platform and the archiving process was executed. The findings of the study reveal that libraries can easily adopt Green Open Access, thus strengthening the scholarly communication cycle without any upfront and subsequent cost. This immediately outweighs the limitations of the digital divide. The positive outcomes of the study pave the way for the libraries with resource insufficiencies in making research more findable, shareable, and reproducible, with confidence. The findings of the study also reveal that the Zenodo repository is a OAI-PMH-compliant repository supporting metadata harvesting and interoperability.

Keywords: Invenio, Zenodo, FOSS, Open Access, OAI-PMH, Self-Archiving, Curation, CC-Licenses, DOI, Community, FAIR, IR (Institutional Repository)

Study Background and Theoretical Framework

Open Science has been defined as "an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible, and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of scientific knowledge creation, evaluation, and communication to societal actors beyond the traditional scientific community. It includes all scientific disciplines and aspects of scholarly practices, including basic and applied sciences, natural and social sciences, and the humanities, and it builds on the following key pillars: open scientific knowledge, open science infrastructures, science communication, open engagement of societal actors, and open dialogue with other knowledge systems" (https://en.unesco.org/news/draft-recommendation-open-science-its-way-final-adoption, 2021). One integrated concept with open science is open data. Open Data Handbook has defined open data as the "data that can be freely used, re-used and redistributed by anyone – subject only, at most, to the requirement to attribute and share alike" (What is Open Data?, 2021).

One other associated concept could be big data encompassing the ever-growing research output that

* Assistant Librarian, Central Library, Amity University Punjab, Mohali, Punjab, India. Email: asinhababu@pb.amity.edu; ORCID: https://orcid.org/0000-0001-7036-0827
** DLIS, Panjab University, Chandigarh, India. Email: heenamgakhar0111@gmail.com; ORCID: https://orcid.org/0000-0002-0549-9950
*** Professor, DLIS, Panjab University, Chandigarh, India. Email: rupak@pu.ac.in; ORCID: https://orcid.org/0000-0001-5046-1663

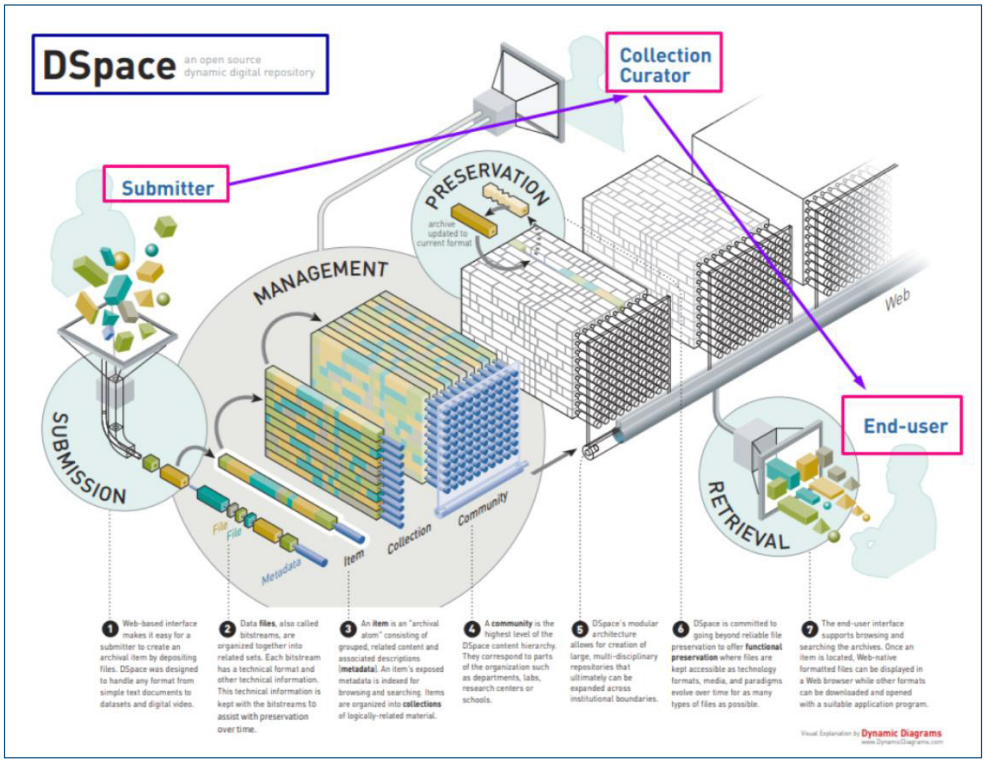
has to be managed sustainably. Big data management (BDM) can be considered a broad term that includes data cleansing, integration, migration, preparation, enrichment, analytics, quality, management, reporting, governance, and planning (Mansouri, 2021). Cloud-based platforms, especially SaaS (software as a service), have emerged as an effective solution for BDM. SaaS provides users with remote access to specific software functions located in the cloud (SaaS – What Is It? Definition – Delante SEO/SEM Glossary, n.d.). Open-source software has the potential to facilitate a large-scale digital repository for research accessibility, visibility, and long-term preservation. A large-scale digital library can service a heterogeneous population and provide diverse digital content (Barifah & Landoni, 2019). A large-scale digital library functions as a portal, aggregating digital resources from disparate collections (Xie & Matusiak, 2016).

Open Access – Green OA

Open access is a broad international movement that seeks to grant free and open online access to academic information, such as publication and data. A publication is defined as 'open access', when there are no financial, legal, or technical barriers to accessing it, that is to say, when anyone can read, download, copy, distribute, print, search for, and search within the information, or use it in education or in any other way within the legal agreements. Open access is a publication model for scholarly communication that makes research information available to readers at no cost, as opposed to the traditional subscription model in which readers have access to scholarly information by paying a subscription, usually via libraries ("What is open access?", 2020). Green open access, also called self-archiving, means that a version of the article is deposited in an open repository, often in the institutional repository of a university or a subject repository (Danielsson, n.d.).

Why Invenio?

There are several open-source software available freely for building institutional repositories (IR). Among them, the two most popular software are DSpace (https://duraspace.org/dspace/) and Eprints (https://www.eprints.org/uk/).



Source: https://duraspace.org/wp-content/uploads/dspace-files/DSpace_Diagram.pdf

Fig. 1

However, none of these FOSS provide 'FREE hosted' repository solutions for the libraries. While libraries can freely download the software, they need the server and other infrastructure to develop and sustain the IR. On account of limited resources, including the financial and complexities of server management, many libraries still do not have their own IR for their research community. This puts libraries, research communities, as well as the academic institution at a loss. They are devoid of all the benefits an IR can offer.

Invenio is a free open-source software licensed under the MIT license, supported by a committed community of multidisciplinary institutions. It has been developed by CERN. The European Organisation for Nuclear Research, known as CERN, is a European research organisation that operates the largest particle physics laboratory in the world. Invenio has been created with security and long-term preservation in mind. Invenio is fast and has been designed to manage 100+ million records and petabytes of files. All research data can be archived independently of the size (InvenioFramework — inveniosoftware.org, 2016). There are many cases wherein massive and large-scale digital repositories are running on Invenio. One such global research repository which has drawn the attention of the scholars worldwide is Zenodo.

Zenodo

The word Zenodo has been derived from "*Zenodotus, the first librarian of the Ancient Library of Alexandria and father of the first recorded use of metadata*", a landmark in library history. Zenodo is an interdisciplinary open dissemination research data repository for the preservation and making available of research, and educational and informational content. Zenodo is hosted by CERN, which has existed since 1954 and is an intergovernmental organisation funded by the European Commission via the openAIRE projects, CERN, Alfred P. Sloan foundation, and Acradia fund. Donations are via CERN and the society foundation, and currently has an experimental programme defined for the next 20+ years. Zenodo is powered by CERN Data Centre and the Invenio digital library framework and is fully run on open-source products all the way through. Physically, Zenodo's entire technical infrastructure is located on CERN's premises, and is subject to CERN's legal status. All files uploaded to Zenodo are stored in CERN's EOS service in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers. All data files are stored in CERN data centres, primarily in Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape on a nightly basis. OpenAIRE Orphan Record Repository got a make-over and was re-branded as Zenodo. If we deposit an article in OpenAIRE Orphan Record Repository, it is also available in Zenodo. Zenodo is open to all research outputs from all fields of science, regardless of the funding source.

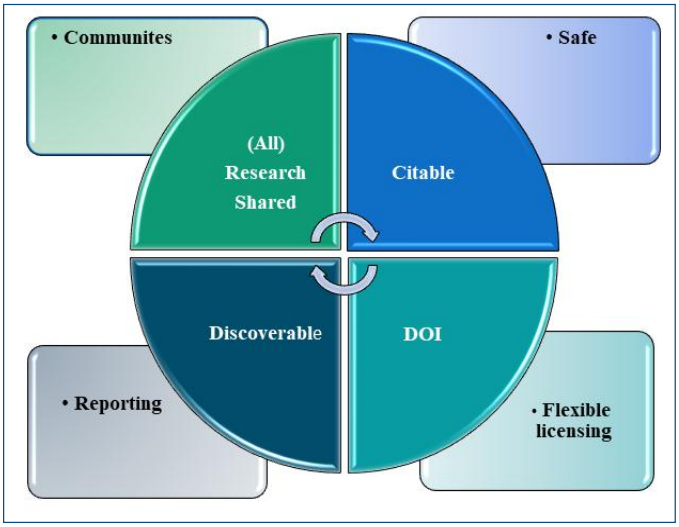


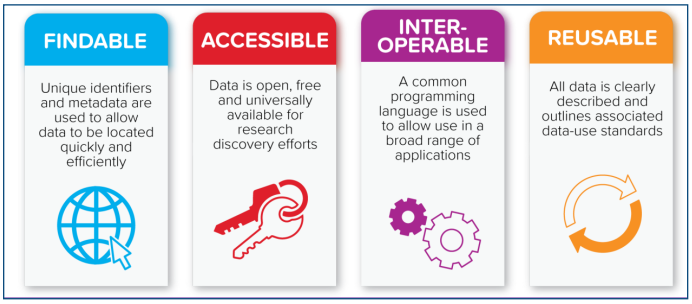
Fig. 2

"Researchers can upload files to Zenodo and there's minimal validation of what goes in there, but these community collections essentially allow everyone to create and curate the content and this solves the issue of us otherwise having to validate everything that's uploaded" (Lars Holm Nielsen, n.d). Zenodo has attracted a large number of communities in its few years of existence. Because Zenodo does not limit registered users from creating communities, their establishment and operation are solely dependent on the desire of individuals and communities who interact with the repository. This makes the repository an intriguing example of a data curation repository in which researcher behaviour is manifested both in the repository's growth and real use, as well as in community selection. Metadata are assigned a globally unique and persistent identifier, i.e., a DOI which is issued to every published record on Zenodo. It is a top-level and mandatory field in the metadata of each record. Data

are described with rich metadata. Zenodo metadata is compliant with data cite Metadata schema minimum and recommended terms, with few additional enrichments.

Zenodo and FAIR Principles

The FAIR Data Principles (Findable, Accessible, Interoperable, and Reusable), published on Scientific Data in 2016, is a set of guiding principles proposed by a consortium of scientists and organisations to support the reusability of digital assets (*What are the FAIR Data Principles? | Augustus C. Long Health Sciences Library*, n.d.).



Source: https://kidsfirstdrc.org/assets/images/f554d2c0-7ae0-11e8-8def- ddba9c8697d1.png

Fig. 3

Findable means data and metadata are online and openly searchable with a persistent link that is uniquely attached to each specific dataset. Accessible means data and metadata are retrievable in a machine-actionable form, with downloading options clearly described (including any needed authentication). Interoperable signifies that data and metadata are consistently structured and described, both syntactically and semantically, so that algorithms can parse and ensure that like data are accurately compared to like. Reusable indicates that data and metadata are sufficiently annotated so machine and human users can determine fit-for-purpose in the context of their analysis.

Machine actionable leads to structuring data and content to make it possible for computational systems to find, access, interoperate, and reuse data without significant human intervention. Data interoperability is the capacity to which data can be analysed and/or merged with similar data. Data interoperability relies on data standards, data documentation, and metadata to indicate to researchers which data sets or variables are comparable (NLM data thesaurus) (Preparing FAIR data for reuse and reproducibility | Research Data Management Service Group, 2020).

Related Studies

The study states that the Open-Source system Invenio helped the library remain active, and its unique collection managed to survive through effective transformation. It helped tackle the issues of lack of methodological and technical support, and problems in research management through digital repositories. Invenio helped a really small specialised library provide high standard services and grey literature open access with a limited budget and minimal staffing. It has low demand on technical, financial, and personal resources, and stable development was ensured. The main aim of the library is to fully satisfy both internal and external users, and to always provide accessible and transparent information about the library's collection, subscribed and freely available online resources, and other information materials that the library can offer all its users. The library is trying to achieve this goal by using the INVENIO library system, which represents a single interface enabling searches in all the library's information sources (Drozda et al., 2015). CERN (European Organization for Nuclear Research), as the world's largest physics laboratory, has always been facing the challenge of distributing and archiving grey material. Invenio is an integrated digital library system originally developed at CERN to run the CERN document server (CDS), and the Invenio software was born in a rich grey literature-producing environment. Its modular design enables it to serve a wide variety of requirements, from a multimedia digital object repository to a Web journal, to a fully functional digital library, with its flexible nature. Grey literature has historically played a key role for researchers in the field of high-energy physics (HEP). The paper highlights how the particular context of grey literature within the HEP community shaped the development of Invenio and focused on the process of grey material within the software, and analysed how it is used in a real production environment, the CERN document server (CDS) (Caffaro & Kaplun, 2010).

Research Problem

Digital divide is a well-accepted phenomenon. This 'unfair' divide may be attributed to several factors, including lack of ICT infrastructure, financial resource crunch, and lack of proficient and skilled professionals. Digital divide leads to so many other divides between the 'haves' and the 'have-nots'. Many libraries in the world are at the wrong end of this digital divide as they find themselves helpless in building an institutional repository (IR) for the benefit of the members they serve. Due to unavailability of funds and skilled human resources, they cannot afford on-prem (on-premise) or hosted solutions or any cloud computing deployment model. In the absence of any systematic and permanent solution, the research productivity of the campus researches possesses high risk of obsolescence behind the paywall. The limited visibility also adversely affects the likelihood of potential citations. Moreover, the unpreserved research is at a risk of future unavailability. This also deviates from the F.A.I.R. data principles. Research unshared is a loss to the humanity and social welfare. The root-cause analysis reveals that libraries are not unwilling to take the initiative, rather they lack the motivation in terms of availability of a FREE and OPEN platform, built on the robust Invenio RDM infrastructure having global exposure. This study highlights the potential of Zenodo as an enabler and empowering solution for libraries to build a reliable castle in the empire of open access. Researchers have been sharing their research via academic social networks (ASNs) or scientific collaborative networks (SCNs) for a long time. The basic problems with these platforms include lack of global visibility, one platform being more popular in certain parts of the world, well-defined long-term research preservation commitment, and explicit mention on their website. More often than not, they are not governed by not-for-profit organisations. Instead, they are partially commercialised.

Study Objectives

The main objective of the study is to evaluate and demonstrate the usability of Zenodo repository (powered by Invenio software) for research curation and sustainable long-term preservation through archiving (both self-archiving and mediated-archiving), as per the Green Open Access. The other objectives revolving around the main objectives are enumerated below.

- To critically evaluate the key performance indicators (KPIs) of the Zenodo platform as the trusted global research archiving platform.
- To showcase the utility/efficacy/suitability/appropriateness of a digital repository for libraries lacking ICT infrastructures like on-premise or SaaS-based cloud platforms.
- To examine and execute the steps of a building community as a repository administrator.
- To investigate the process and steps of archiving with Dublin Core metadata schema, open access status, and supported Creative Commons (CC) licenses.
- To observe the download pattern of archived papers.
- To recommend best practices based on the findings of the study.

Research Methodology

The authors used their ORCID to sign in to Zenodo. ORCID provides a persistent and unique digital identifier (an ORCID ID) that researchers own and control, and that distinguishes them from every other researcher (ORCID, 2021). To achieve the study objectives, a community '*dlispu*' was built by the authors. The community was built with the focus of maintaining it as a real and live repository in the discipline of library and information science. The collection URL and repository community details were communicated to several depositors for self-archiving. The depositors were instructed to carefully fill the Dublin Core metadata elements as per the requirement of the open repository, after uploading the content. Published research papers were given priority with the immediate availability of full-text PDFs.

Upon successful submission, the repository administrator curated the same by accepting the submissions. Post acceptance the papers were immediately visible on the Zenodo platform and were further examined for their accuracy and completeness. Any discrepancy noted was intimated to the submitters for corrective measures. Submissions with inaccurate and inadequate information were rejected at the curation stage. Duplicate submissions were also rejected by the community administrator.

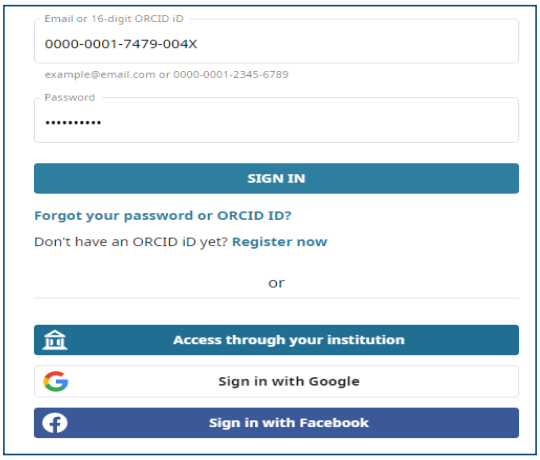
After the successful appearance of the research papers in the global open repository platform Zenodo under the community dlispu (https://zenodo.org/communities/dlispu/), the download data was closely monitored for several weeks for any positive growth rate.

Steps for Building IR using Zenodo SaaS

Sign-Up and Login

Researchers/librarians must register as a user of Zenodo. This will allow them to access as well as deposit content in all possible formats for which they possess the appropriate rights.

- Go to Zenodo.org.
- Create an account using an e-mail or sign up using a GitHub account or by using an ORCID account.
- After logging in, users can have authorised access to different Zenodo functions.



Building Community

- After logging in to Zenodo, click on Communities, shown in the red box.

- After that, click on 'New' to create a community. To create a community, the user account must be verified at least an hour prior.

- Identifier: (Required)

Only letters, numbers, and dash are allowed. The identifier is used in the URL for the community collection, and cannot be modified later.

- Title: (Required)

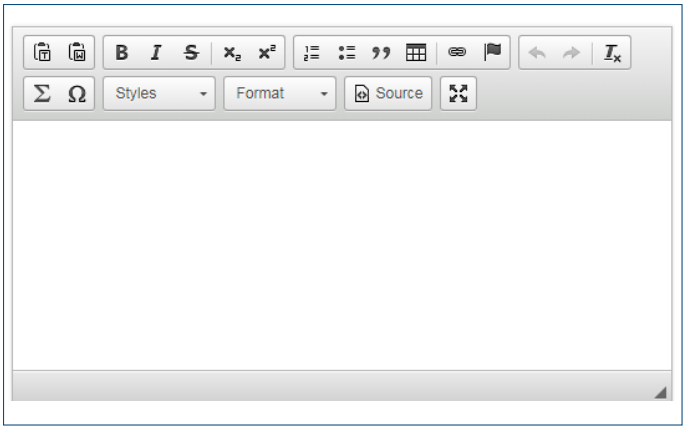
Enter the title of the community in the provided text box.

- Description: (Optional)

A short description of the community collection can be entered in the text box provided, which will be displayed on the index page of the community.

- Curation Policy: (Optional)

The curator describes briefly and precisely the policy by which he/she accepted/rejected new uploads in this community.



- Page: (Optional)

A long description of the community collection will be displayed on a separate page linked from the index page.

- Logo: (Optional)

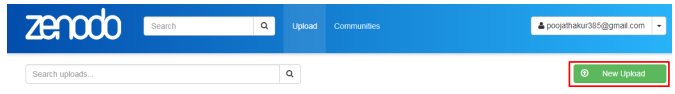
This is the image file used to aid and promote instant public recognition. Supported formats are: PNG, JPG, and SVG. Max. file size: 1.5MB.

During the 'New community' creation, description, curation policy, and page are all optional. They provide the complete text formatting features, such as bold, italics, strikethrough, subscript, insert/remove, numbered list, insert/remove bulleted list, block quota, table, link, anchor, undo, redo, remove format, math, insert special character, formatting styles, paragraph format, source, maximise, and so on.

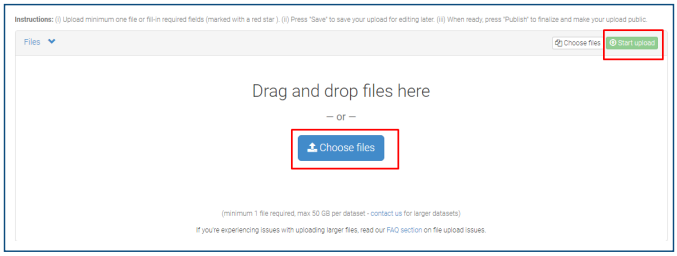
Content Deposit

- After creating a community, log in to your account and select the 'upload' tab at the top of the page. You will be brought to an upload landing page.

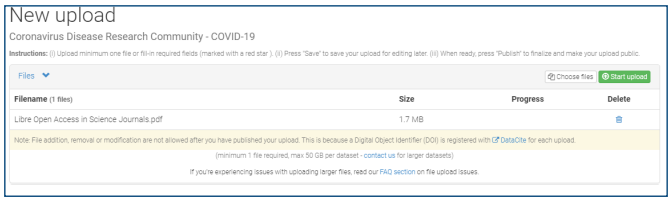
- Now click on the 'New Upload' green button on the top right.



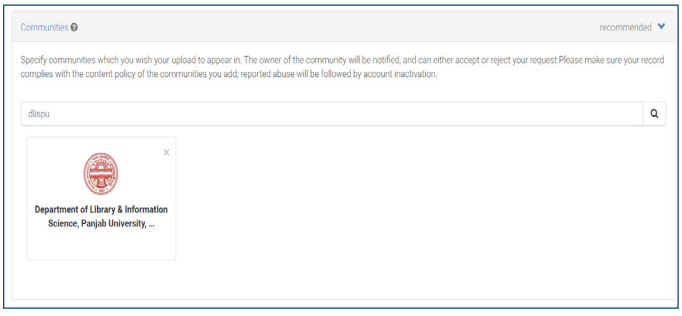
- To upload a new file, click on the 'choose files' button. Once the file appears, click on the 'Start upload' button.



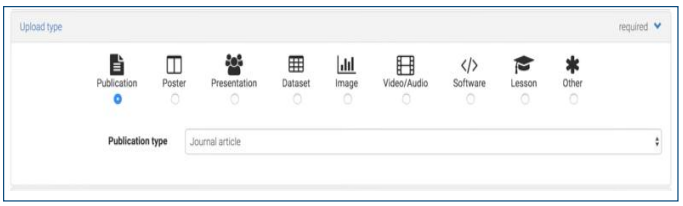
- The content scope includes all types of research content in all fields of research adhering to copyright and privacy policy originating from any stage of the research lifecycle. Users are encouraged to use preservation-friendly formats, with a total file size limit per record of 50GB. Higher quotas can be requested and granted on a case-by-case basis.



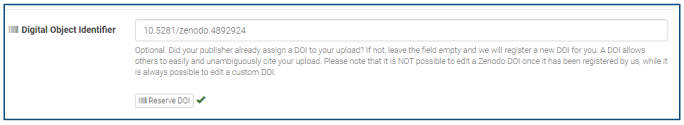
- There are several metadata fields that can be selected in Zenodo. Some of them are recommended and some of them are required. Among the recommended ones, there are communities. Select the community name, i.e. 'dlispu'.



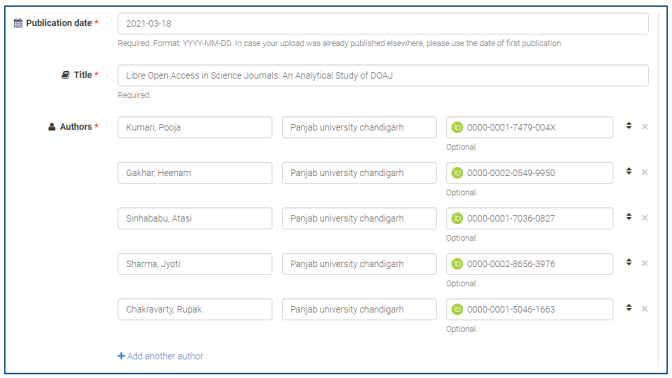
- Select the type of file that the user wants to upload. The user is free to upload any type of file, like publication, poster, presentation, dataset, image, video/audio, software, lesson, and others. Select the publication type, like a journal article.



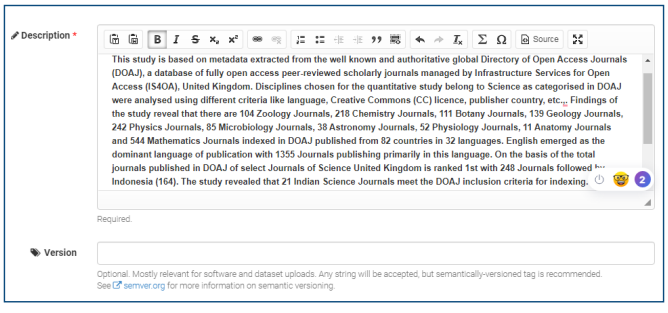
- Click the button 'Reserve DOI' to generate a DOI for your work instantly, i.e. before you submit the work. This allows you to know the DOI prior to submission, though it will not be registered until the submission is complete.



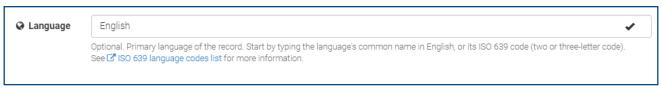
- Enter the original publication date in the text box provided. Enter the title of the publication and then enter each of the authors of the publication, their affiliation, and ORCID ID where possible. To add multiple authors, click '+ add another author'.



- Add description of data type in the provided description box. Version is optional here and is mostly relevant for software and dataset uploads.

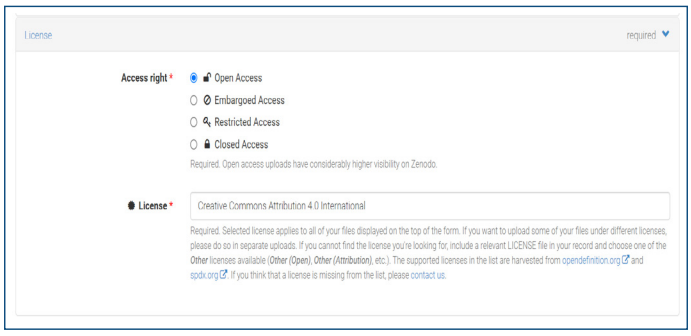


- Select the language of the article. Language codes are provided according to ISI 639 code in Zenodo. For textual items, English is preferred; however, all languages are accepted.



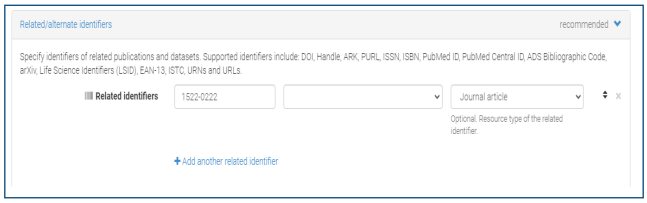
- Add keywords from the article in the text box. To add more than one keyword, click on '+ add another keyword'. One can also add additional notes; it is optional.

- Users must specify a license for all publicly available files. Files may be deposited under closed, open, or embargoed access. Access to metadata and data files is provided over standard protocols such as HTTP and OAI-PMH. Users may deposit content under an embargo status and provide an end date for the embargo. The repository will restrict access to the data until the end of the embargo period, at which time the content will become publically available automatically. Users may deposit restricted files with the ability to share access with others if certain requirements are met. These files will not be made publicly available and sharing will be made possible only with the approval of the depositor of the original file. If we choose open access, we must specify the license under which the publication is distributed. Usually, the default Creative Commons Attribution 4.0 license is appropriate. However, check with any existing publisher. Files deposited under closed access are protected against unauthorised access at all levels.

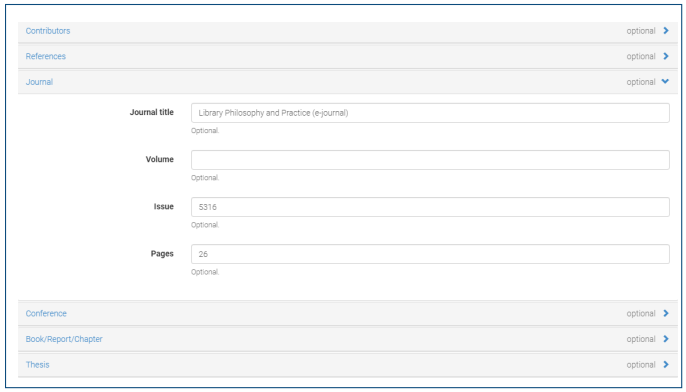


- Enter the grant number for the grant/project in the neighbouring text box.

- Enter related/alternate identifiers, such as DOI, handle, ARK, PURL, ISSN, ISBN, PubMed ID, PubMed Central ID, ADS Bibliographic Code, Life Science Identifiers (LSID), and resource type of the related identifier.



- Enter contributors, references, journal, conference, book/report/chapter, thesis, and subject in the text boxes. These are optional.

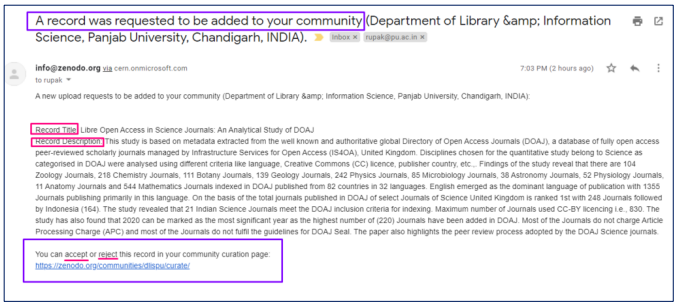


- Work in progress can be saved by clicking 'Save' at the top or bottom of the page. Then click 'Publish' to publish your work on Zenodo.

Items will be retained for the lifetime of the repository. Zenodo makes no promises of usability and understandability of deposited objects over time.

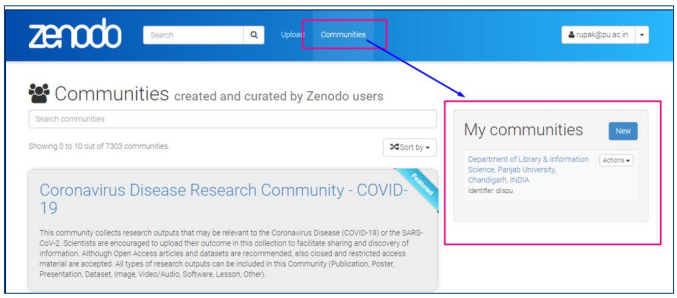
Content Curation (Post Submission)

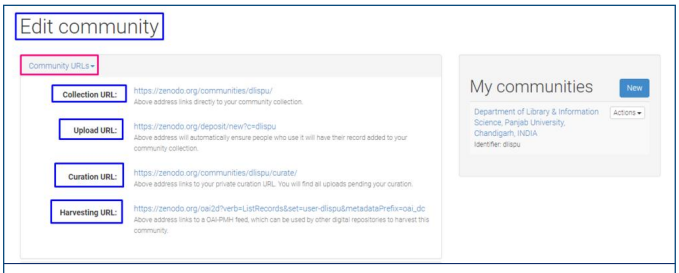
- *Curation e-mail Alert/Intimation*

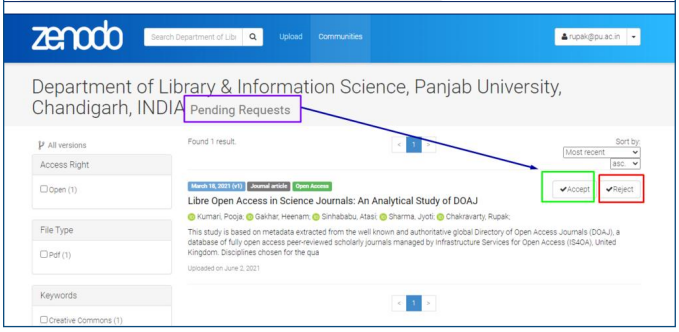


- *Repository Community Curation Service*

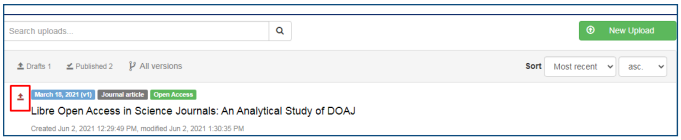
The pending submissions for curation can be viewed by the repository administrator (RA) after login, wherein the submission may either be accepted or rejected.



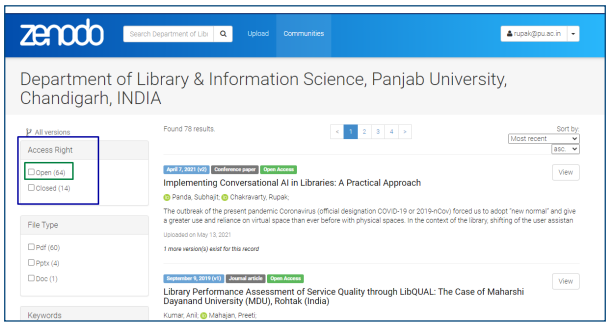




Until a submission is accepted or rejected by the curator, the documents are saved in drafts with an indication of a red arrow for the submitter.



- *View Collection*



As soon as a submission is accepted or approved by the RA, the same becomes part of the repository and is readily available for viewing.

Zenodo Compliance with OAI-PMH

Zenodo allows IR administrators to harvest our entire repository via the Open Archives Initiative Protocol for Metadata Harvesting (OAI- PMH). OAI-PMH is a widely used protocol for harvesting metadata and the most popular repository software provides support for this protocol. OAI-PMH is developed by the Open Archives Initiative. It is a low-barrier mechanism for repository interoperability. Data providers are repositories that expose structured metadata via OAI-PMH. Service providers then make OAI-PMH service requests to harvest that metadata. The base_URLof Zenodo is https://zenodo.org/oai2d. The validation test was executed using the link https://www.openarchives.org/Register/ValidateSite. The interface of this validator and the results generated are reproduced below.

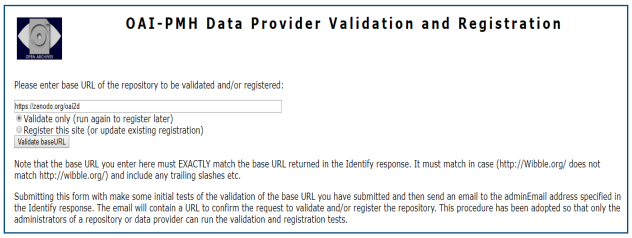


Fig. 4: OAI-PMH Validation Test with Zenodo Base URL

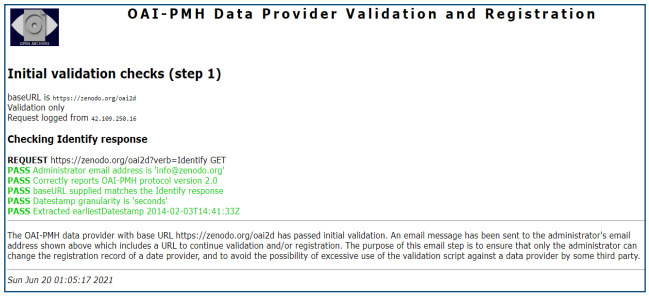


Fig. 5: OAI-PMH Validation Test Result

The result of the validation test indicates that the Zenodo archive successfully passed the validation test with full conformity.

Findings

- Zenodo is ORCID-enabled, providing convenient access with a futuristic vision.
- Records are indexed immediately in OpenAIRE (able to see the record within a few minutes in OpenAIRE).
- Building a repository (individual/community) in Zenodo is easy, fast, and quick, since it is a SaaS deployment model of cloud computing.
- Disciplinary/subject/topical repository can also be implemented quite easily.
- The platform (powered by FOSS Invenio) provides accessibility, visibility, and long-term preservation of research with vast storage and adequate security.
- Usage statistics are readily available for downloading and viewing.
- It is for all researchers, scientific communities, and research institutions, and is open to all research outputs regardless of the funding source.
- Zenodo provides free DOI (prefix: 10.5281) in case the publisher has not assigned any to the archived content.
- Zenodo allows small modifications to the record's file by providing the edit option.
- It adheres to the FAIR Data Principles, and thus encourages the sharing of research openly for maximum utilisation and re-use of research results (research re-use and reproducibility).

- Zenodo repository was found to adhere to OAI-PMH in terms of interoperability for metadata harvesting.

Recommendations and Suggestions

- As of now, Zenodo does not allow "sub-community" under a given community. It is highly recommended that provision of sub-communities must be offered.
- Social sharing links also need to be integrated for instant sharing of research articles.
- Libraries that still do not have their IR should create the same using the open Zenodo platform for research curation, preservation, and visibility.
- Metadata creation workshop and training programmes may be conducted for making independent self-archiving in Zenodo platform using Dublin-Core metadata schema.
- Depositors should be encouraged to deposit content under Attribution 4.0 International (CC BY 4.0) License wherever possible.
- The benefits of the IR in terms of download statistics and citations must be shared/communicated to the stakeholders for marketing and promotion of the service.
- Authors/researchers must be encouraged and motivated to deposit the full-text version (published or author-accepted VOR – Version-of-Record) of the content (wherever possible) as per the Sherpa Romeo (available at https://v2.sherpa.ac.uk/romeo/) analysis of publisher open access policies from around the world on a journal-by-journal basis.
- Libraries can focus on self-archiving by the authors as the first priority, with an additional provision of mediated archiving by a pre-identified team.
- Libraries might explore the possibility of establishing on-prem cloud IR using Invenio software on the SaaS model, for constituting Repository as a Service (RaaS)/Digital Repository Service (DRS).
- Curator should know the credentials of the depositor before accepting or rejecting the document.
- Provision of duplication checking of document should be available, before accepting or rejecting the document.

Conclusion

The successful implementation of a digital repository elucidates and validates the availability of Zenodo as a pragmatic and sustainable platform for research preservation and visibility. Zenodo facilitates metadata access, and reuse by it is licensed under CC0. All metadata is exported via OAI-PMH and can be harvested. It offers longevity to the archived content, with versioning of the data files. The uploaded data is archived as a Submission Information Package. To facilitate file preservation, the files and metadata are backed up nightly and replicated into multiple copies in the online system. Zenodo ensures fixity and authenticity of the content; at regular time intervals files are checked for their integrity. It has a well-planned succession plan. In case of closure of the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject-based repositories.

References

Barifah, M., & Landoni, M. (2019). Exploring usage patterns of a large-scale digital library. *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. doi:https://doi.org/10.1109/jcdl.2019.00020

Caffaro, J., & Kaplun, S. (2010). Invenio: A modern digital library for grey literature. Retrieved from https://cds.cern.ch/record/1312678/files/CERN-OPEN-2010-027.pdf

Caffaro, J., & Kaplun, S. (2011). Invenio: A modern digital library system for grey literature. *Grey Journal (TGJ)*, *7*(2), 105-108.

Cornell University. (n.d.). Research Data Management Service Group Comprehensive Data Management Planning & Services. Preparing FAIR Data for Reuse and Reproducibility | Research Data Management Service Group. Retrieved from https://data.research.cornell.edu/content/preparing-fair-data-reuse-and-reproducibility

Drozda, J., Synková, V., & Pejšová, P. (2014). Publishing geodesy, topography and cartography research via Invenio. *Conference Papers: International Conference on Grey Literature*, *16*, 59-63.

Drozda, J., Synková, V., & Pejšová, P. (2015, March 1). Publishing Geodesy, Topography and Cartography Research via Invenio. www.opengrey.eu. http://hdl.handle.net/10068/1024639

(2021, May 18). Draft recommendation on Open Science on its way to final adoption. UNESCO. Retrieved from https://en.unesco.org/news/draft-recommendation-open-science-its-way-final-adoption

InvenioFramework — inveniosoftware.org. (n.d.). Inveniosoftware.Org. Retrieved from https://inveniosoftware.org/products/framework/ ORCID (2019). *ORCID*. Orcid.org. https://orcid.org/

Mansouri, Y. (2021, May 7). Big data management: Data repository strategies and data warehouses. *Medium; Towards Data Science*. Retrieved from https://towardsdatascience.com/big-data-management-data-repository-strategies-and-data-warehouses-bfa9b45c95de

SaaS – What is it? Definition – Delante SEO/SEM Glossary. (n.d.). Delante. Retrieved from https://delante.co/definitions/saas/

Welcome to Sherpa Romeo – v2.sherpa. (2017). Retrieved from https://v2.sherpa.ac.uk/romeo/

What is Open Data? (2021). Opendatahandbook.org. Retrieved from https://opendatahandbook.org/guide/en/what-is-open-data/

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M.,…Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci Data, 3*. doi:https://doi.org/10.1038/sdata.2016.18

Xie, I., & Matusiak, K. (2016). *Discover digital libraries: Theory and practice*. Elsevier Science Publishing.