InterViews: An Interdisciplinary Journal in Social Sciences
2024, Vol. 11, No. 1, pp. 24-38

ISSN: 2349-400X ©Saint Claret College, Ziro
https://doi.org/10.36061/IV.11.1.24.24.38

EXPLORING THE LANDSCAPE OF OPEN SCIENCE LEARNING RESOURCES AND DATA REPOSITORIES: A COMPARATIVE ANALYSIS AND RANKING STUDY IN THE DIGITAL ERA

Ranjan Samanta[1] & Nabanita Dutta[2]

This article aims to assess the effectiveness and utility of selected open educational resources (OER) and open data repositories, exploring their emerging role in education and research practices. The study investigates how these resources have become increasingly essential in the digital age, particularly as demand for accessible, open resources has surged due to the pandemic's shift to online learning and research. The analysis offers a detailed ranking of repositories based on an established marking system, where each criterion carries one mark, resulting in a total score that reflects the repository's overall quality. Key findings highlight strengths in transparency, ease of access, and the technical capabilities of each repository, as well as areas needing further enhancement to meet academic standards for openness and reliability. It emphasizes how these tools foster transparency, accountability, and accessibility, aiding researchers, educators, and students in a more interconnected and resource-sharing academic ecosystem. The paper offers practical insights into the current state of open educational and scientific resources, providing recommendations for improvement and highlighting best practices.

***Keywords*:** Digital reference sources, open data repository, open science, open educational resources, OER, online data repository, open learning repository

[1]*Ranjan Samanta* is a Librarian, Naba Ballygunge Mahavidyalaya, Kolkata, India. Correspondence regarding this article may be directed to him at: *ranjansamanta5@gmail.com*

[2] *Nabanita Dutta* is working as a Librarian, De Paul School, Kolkata, India

Introduction

According to UNESCO (2024a), "Open Educational Resources (OER) are learning, teaching and research materials in any format and medium that reside in the public domain or are under the copyright that has been released under an open license, that permits no-cost access, re-use, re-purpose, adaptation and redistribution by others."

In the evolving landscape of education and research, Open Educational Resources (OER) and open data repositories have gained increasing significance. These resources play a critical role in promoting accessible, collaborative, and transparent knowledge-sharing practices, which have become even more crucial with the rise of online learning and research during the COVID-19 pandemic. As institutions and individuals worldwide seek reliable, accessible sources for educational content and research data, OER and open data repositories offer essential support for equitable access to knowledge.

This article aims to assess the effectiveness and utility of selected OER and open science repositories, evaluating their impact on academic and research practices in the digital era. The study conducts a detailed content analysis of various repositories, focusing on aspects such as accessibility, authenticity, licensing, and opportunities for improvement. By comparing widely used repositories—including Harvard Dataverse, Zenodo, EUDAT, Dryad, GitHub, Figshare, Protocols.io, ICPSR, and the Open Science Framework—this research seeks to uncover how these tools contribute to knowledge-sharing and collaboration across the global academic community.

Scope and Coverage

The repositories selected for this study: Harvard Dataverse, Zenodo, EUDAT, Dryad, GitHub, Figshare, Protocols.io, and ICPSR - have been chosen due to open education and global academic collaboration exemplify the evolving landscape of knowledge dissemination in the digital age. These repositories represent a diverse range of functionalities, disciplines, content types, technological approaches, and user bases, making them ideal for comprehensive analysis. By analyzing these platforms, the study aims to provide insights into best practices, challenges, and the overall impact of OER repositories on academic progress.

Methodology

Through content analysis, this study examines various open science tools and data repositories, evaluating them based on criteria such as accessibility, authenticity, licensing, and potential areas for improvement. Additionally,

it includes a comparative analysis of online data repositories—including Harvard Dataverse, Zenodo, EUDAT, Dryad, GitHub, Figshare, Protocols.io, ICPSR, and Open Science Framework—measuring their efficacy in knowledge sharing and collaboration within the academic community.

The evaluation parameters for Dataverse, Zenodo, EUDAT, Dryad, GitHub, Figshare, Protocols.io, and ICPSR have been meticulously designed to provide a comprehensive and balanced assessment of their functionality, accessibility, and impact as repositories for open educational resources (OER). These parameters are based on a thorough review of existing frameworks for evaluating digital repositories and have been tailored to align with the specific objectives of this study.

In line with UNESCO's (2024b) Guidelines for OER (https://en.unesco.org), evaluation frameworks for open educational resources often share commonalities with those for digital repositories.

Literature Review

The literature on Open Science Learning Resources and Data Repositories has evolved over the years to capture its theoretical underpinnings, challenges, and emerging practices, as well as the broader implications for education systems globally. Yuan, MacNeill, and Kraan (2008) provided foundational insights into OER development and the future trajectory of this movement. Their work aimed to stimulate debate within the JISC community and outlined key areas for research, recognizing OER's potential to reshape educational resources. D'Antoni (2009) offered a comprehensive introduction to the OER movement, noting it as relatively young yet impactful across diverse settings. She highlighted the key issues, including the importance of addressing accessibility, quality, and sustainability to ensure that OER can significantly benefit educators, learners, and educational institutions. Ehlers (2011) explored a shift in focus from mere access to resources to the adoption of open educational practices (OEP). Based on the OPAL report, his work emphasized the importance of practices alongside resources, introducing a framework to support OEP and advocating for open access as essential but insufficient for the full potential of open education. This perspective suggested a move toward transforming educational architectures to enhance OER adoption. Tuomi (2012) contributed to the conceptual framework of OER, identifying four OER types and their alignment with learning theories. His analysis suggested that the

widespread adoption of OER has the potential to transform learning and education within a knowledge society, both constraining and accelerating this transformation. McKerlich, Ives, and McGreal (2013) focused on the faculty and staff readiness to adopt OER at Athabasca University through a survey, providing quantitative insights into OER use and creation. With a ratio of OER use to creation, their findings introduced a new metric to assess adoption levels across institutions, contributing to understanding the factors necessary to encourage OER engagement. Atenas and Havemann (2014) reviewed the literature on OER and repositories of open educational resources (ROER), identifying four themes crucial to repository design and outlining ten quality indicators for effective ROER. These indicators—such as peer review, standardized metadata, and social media integration—provide a framework to enhance the quality and accessibility of ROER and ensure it meets the goals of the OER movement. Clements, Pawlowski, and Manouselis (2015) developed a theoretical framework for LOR quality (LORQAF) to classify and guide quality assurance approaches for learning object repositories (LOR). Their work underscored the need for collaborative quality assurance tools to support the sustainable development of OER repositories. Mishra (2017) revisited the fundamental goals of OER, cautioning against conflating OER with broader open education movements. He argued for flexibility in defining OER to support mainstream adoption and reduce internal barriers, advocating for a more open and inclusive approach within the movement. Santos-Hermosa, Ferran-Ferrer, and Abadal (2017) assessed the effectiveness of ROER in meeting educational needs at the international level, analyzing educational indicators to gauge ROER suitability in academic contexts. Their study found that while many ROERs focused on educational resources, the integration of educational aspects remains limited. The few successful cases used educational standards and metadata, indicating room for improvement in how ROER meets educational goals. Mirowski (2018) presented a critical perspective, suggesting that the open science movement, which intersects with OER, may be co-opted by platform capitalism. He argued that the push for openness in science is sometimes misrepresented, with motives aligned more with commercial interests than with the public good. Finally, Hylén (2021) explored the OECD's OER project, questioning the motivations and implications of OER in the context of intellectual property within higher education. His preliminary findings highlighted the rising trend of institutions openly sharing educational resources online, challenging

traditional views on intellectual property and offering insights into the potential long-term impacts on the educational landscape.

This study collection illustrates the OER movement's multifaceted nature, challenges, and potential to transform educational access and practice.

Data Collection and Analysis

From Table 1, we gather some basic information about the websites, including their starting year, publishers, URLs, subjects they deal with, etc. It is to be noted that none of these repositories mention the frequency of website updates. We have used an 'X' symbol to represent the absent parameters and a 'ü' symbol to indicate that the features are present in the repositories.

Table 1: General Information of Repositories

Parameters	**Harvard Dataverse**	**Zenodo**	**EUDAT**	**Dryad**	**GitHub**	**Figshare**	**Protocols.io**	**ICPSR**
Abbreviation / other name	Dataverse	OpenAire orphan records repository(2013)	EUDAT collaborative data infrastructure (CDI)	DRYAD	GitHub	Figshare	Protocols.io	Inter-university Consortium for Political and Social Research
Year of starting	-	2013	2016	2008	2007	2011	2012	1963
URL of website	https://dataverse.harvard.edu/	https://zenodo.org//	https://www.eudat.eu/	https://datadryad.org/stash	https://github.com/	https://figshare.com/	https://www.protocols.io/	https://www.icpsr.umich.edu/web/pages/
Publisher	Harvard University	CERN	EUDAT CDI, CyI, GWGD, KIT	American academy for advancement of science, American gene association, etc.	Microsoft (since 2018)	Digital science		University of Michigan
Frequency of update	X	X	X	X	X	X	X	X

Broad subject matter	All scientific domains, mostly Agricultural sciences, Arts and Humanities, Astronomy and Astrophysics, Business and Management, Chemistry, Computer and Information Science, Earth and Environmental Sciences, Engineering, Law, Mathematical Sciences, Medicine, Health and Life Sciences, Physics, Social Sciences	Almost all disciplines	Covers several research disciplines	Multidisciplinary in natureespecially data underlying scientific and medical publications (mainly of evolutionary, genetic, and ecology biology)	It allows users to browse public repositories on the site. It facilitates the version control and issue tracking aspects of software development.	Biology, chemistry, biotechnology, computational workflow, etc.	Major academic disciplines	Social and political sciences. ICPSR advances and expands social and behavioural research.

Table 2a: Features

Parameters	Harvard Dataverse	Zenodo	EUDAT	Dryad	GitHub	Figshare	Protocols.io	ICPSR
Features	Metadata, datasets	Trusted, safe, citeable, no waiting time, collab with GitHub, user statistics	Data safety, EUDAT service catalogue	Open, easy-to-use, not-for-profit, community-governed data infrastructure. USP-finding data by name place, journal, institution, subject and selecting location directly from map	The most advanced and animated website that allows user interaction on the interface	Figshare allows researchers to publish negative data. Figshare also tracks the download statistics for hosted materials, acting in turn as a source for altmetrics	Protocols.io is open access and there are no fees charged to users for sharing or accessing the public content. Protocols.io charges for private non-academic premium group accounts. The company also charges reagent vendors and publishers for access to aggregated anonymized analytics on the use of protocols and reagents.	It acts as a global leader in data stewardship and provides rich data resources and responsive educational opportunities for present and future generationsICPSR provides a number of tools as classroom aids for college-level instructors:Bibliography of Data-related Literature, Social Science Variables Database, Survey Documentation and Analysis online analysis package, Data Driven Learning Guides

In Table 2b, we delve into the features and content available in the repositories. For this, we consider parameters such as keyword search, presence of site map, declaration about the objectives of the projects, copyright issues, availability of resources for members, institutional login facilities, etc. It must be noted that none of these repositories have multiple-language features or provisions for translation, this is a hindrance in the path of universal access to these resources.

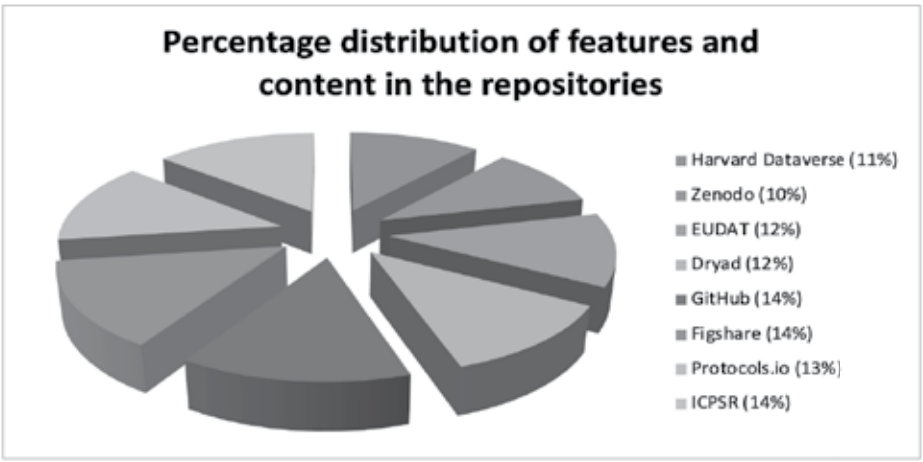
Table 2b: Content Available in Repositories

Parameters	Harvard Dataverse	Zenodo	EUDAT	Dryad	GitHub	Figshare	Protocols.io	ICPSR
About	✓	✓	✓	✓	✓	✓	✓	✓
Mission and vision	X	X	✓	✓	✓	✓	✓	✓

Logo	HARVARD Dataverse	zenodo	ENDVL	DRYAD		figshare	protocols.io	ICPSR
Keyword search	✓	✓	✓	✓	✓	✓	✓	✓
Collections	✓	✓	✓	✓	✓	✓	✓	✓
Member-ship	✓	✓	✓	✓	✓	✓	✓	✓
Member institutions	✓	✓	✓	✓	✓	✓	✓	✓
Provision for login of members	✓	✓	✓	✓	✓	✓	✓	✓
News and events	✓	✓	✓	✓	✓	✓	✓	✓
External links	✓	✓	✓	✓	✓	✓	✓	✓
Copyright	✓	X	✓	✓	✓	✓	✓	✓
FAQ	✓	✓	✓	✓	✓	✓	✓	✓
Contact informa-tion	✓	✓	✓	✓	✓	✓	✓	✓
Visitor count	✓	✓	✓	✓	✓	✓	✓	✓
Site map	✓	✓	✓	✓	✓	✓	✓	✓
Website in multiple languages	X	X	X	X	X	X	X	X
Demo/ trials	✓	X	X	X	✓	✓	✓	✓
Last update	✓	✓	X	X	✓	✓	✓	✓
Experts' review/ opinion	X	X	✓	X	✓	✓	✓	✓
Accessi-bility	✓	✓	✓	✓	✓	✓	✓	✓
Career opportuni-ties	X	X	X	X	✓	✓	✓	✓
Customiza-tion	✓	✓	✓	✓	✓	✓	✓	✓
Publica-tions	X	X	✓	✓	✓	✓	X	✓
Total points (out of 22)	**17**	**15**	**18**	18	**21**	**21**	**20**	**21**

The following 3-D pie chart shows that GitHub, Figshare, and ICPSR have almost all of the features that have been enlisted, making them the highest-scoring repositories in this section. Protocols.io is close behind, while Harvard Dataverse scores the lowest due to the lack of certain general features present in the other repositories. For calculating the percentage, the marks obtained out of 22 have been divided by the total points, i.e., 22, and multiplied by 100. The color index has been provided alongside for better understanding.

Figure 1: Percentage-distribution of features and content in the repositories



In Table 3, the technical features such as authenticity, default license, download, printing facility, etc., are dealt with. In this section, all the repositories score a common point of 8 marks as they do not have printing facilities except EUDAT. Thus, all the concerned repositories have similar technical features to serve their needs.

Table 3: Technical Aspects of Repositories

Parameters	Harvard Dataverse	Zenodo	EUDAT	Dryad	GitHub	Figshare	Protocols.io	ICPSR
Default license	✓ License/ data user agreement and terms of use provided	✓	✓	✓	✓	✓	✓	✓

Data sharing	✓	✓	✓	✓	✓	✓	✓	✓
Information retrieval	✓	✓	✓	✓	✓	✓	✓	✓
Authenticity	✓	✓	✓	✓	✓	✓	✓	✓
Downloads	✓ (Original and archival format available)	✓	✓	✓	✓	✓	✓	✓
Printing facility	X	X	✓	X	X	X	X	X
Collaboration	✓	✓	✓	✓	✓	✓	✓	✓
Data transparency	✓	✓	✓	✓	✓	✓	✓	✓
Quality assurance	✓	✓	✓	✓	✓	✓	✓	✓
Total points (out of 9)	8	8	9	8	8	8	8	8

Table 4 comprises the domain and display features of the repositories, considering parameters like graphics, background color, etc. It is noteworthy that GitHub has the most attractive and appealing website, with an in-built interface for customer interaction and multiple accessibility features that allow for better understanding. Thus, the maximum score in this section was obtained. On the other hand, Harvard Dataverse and ICPSR have very basic websites that do not score much in the case of universal appeal.

Table 4: Domain and Display Recognition

Parameters	**Harvard Dataverse**	**Zenodo**	**EUDAT**	**Dryad**	**GitHub**	**Figshare**	**Protocols.io**	**ICPSR**
Domain name	.edu	.org	.eu	.org	.com	.com	.io	.edu
Graphics (out of 5)	2	3	3	4	5	3	3	2
Quality (out of 5)	3	3	3	3	5	4	4	3
Background colour (out of 5)	2	3	2	3	4	4	4	2

Overall rating (out of 5)	2	2	3	3	5	4	4	2
Total points (out of 20)	9	11	11	13	19	15	15	9

In Figure 2, the total points in the pie chart amount to 160 (as there are 8 columns of 20 points each). For our convenience, these 160 points correspond to 360 points or degrees, as is the norm for a pie chart. Thus, 160/360 for each section is multiplied by the marks obtained to arrive at the final result. Considering GitHub as an example, (160/360)*21= 42.75, which corresponds to the sky blue area in this 2-D pie chart.

Figure 2: Distribution of marks in the domain and display recognition

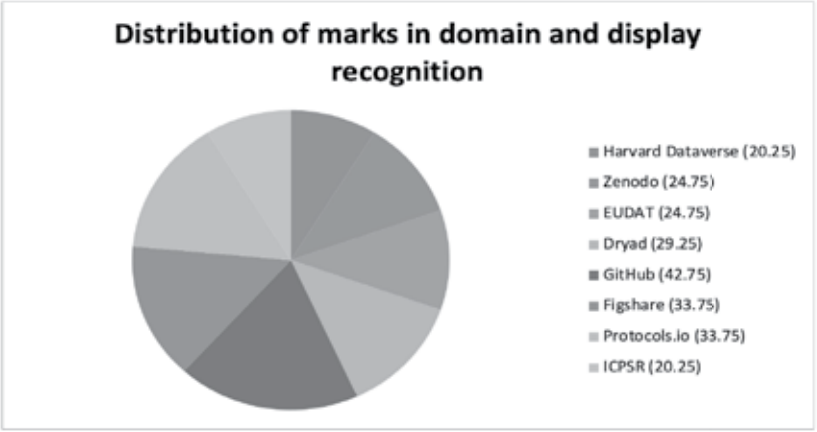


Table 5 shows the total points and subsequent ranking of the repositories. Here, we have added the total points obtained at the end of each table and reached a total of 51 points. The following bar graph shows the total of the points obtained by each repository. The X-axis shows each repository in color-coded form as a bar or column occupying unit space in this graph, and the Y-axis represents the marks obtained out of 51.

Table 5: Total Points and Ranking of the Repositories

Parameters	**Harvard Dataverse**	**Zenodo**	**EUDAT**	**Dryad**	**GitHub**	**Figshare**	**Protocols.io**	**ICPSR**
General information	17	15	18	18	21	21	20	21
Technical aspects	8	8	9	8	8	8	8	8

Domain and display	9	11	11	13	19	15	15	9
Total points (out of 51)	34	34	38	39	48	44	43	38
Rank	(7+8)/2= 7.5	7.5	(5+6)/2= 5.5	4	1	2	3	5.5

Table 6: Ranked Sequence of Repositories

1	**GitHub**
2	Figshare
3	Protocols.io
4	Dryad
5.5	EUDAT and ICPSR
7.5	Harvard Dataverse and Zenodo

From Figure 3, we conclude that GitHub has scored the highest points, followed by Figshare and Protocols.io at a close third position. The ranking was based on the total marks obtained. In the case of a tie, the average of the two positions of the tie has been calculated, and both of them have been attributed the same ranking. For example, if two positions, 4 and 5, have the same marks, the position is calculated as (4+5)/2=3.5. Thus, the rankings are provided in Table 6.

Figure 3: Bar graph showing total points obtained by each repository

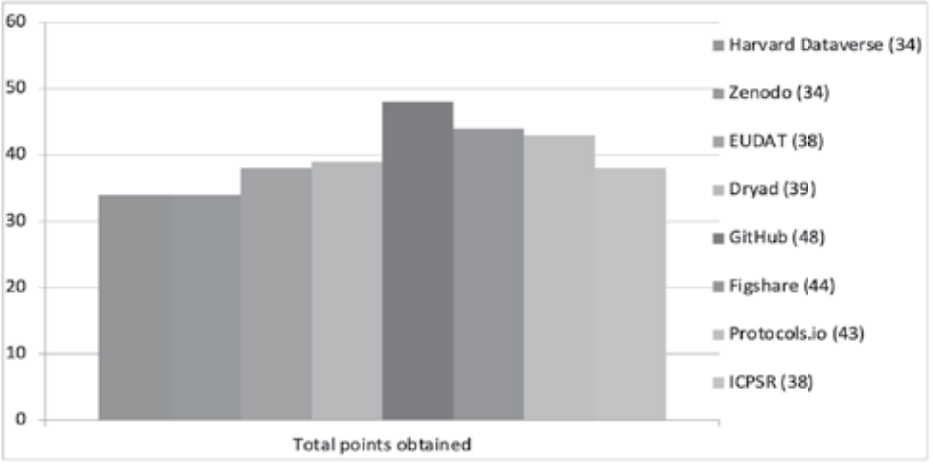


Figure 4 shows the ranking in ascending order, wherein rank 1 denotes the highest marks obtained, and rank 8 denotes the lowest. Thus, among the

repositories, GitHub has scored the highest, and Harvard Dataverse and Zenodo have tied for the lowest positions.

Figure 4: Bar graph showing the ranking in ascending order

Critical Analysis

The analysis provides a structured evaluation of repositories based on specific features. The use of defined parameters, such as keyword search, site map availability, and copyright issues, offers a systematic approach to comparing repositories. These criteria ensure a structured framework for analysis. The calculation method (marks obtained out of 22) is straightforward and ensures transparency in how the repositories are ranked. The inclusion of percentages makes the results accessible and easy to interpret. The evaluation provides a helpful overview of repository features, ranking GitHub, Figshare, and ICPSR as leaders, with Protocols.io following closely behind and Harvard Dataverse lagging. The analysis highlights some crucial gaps, particularly the lack of multilingual support, which impedes universal access. However, the reliance on a purely quantitative approach limits the depth of insight. The analysis highlights a structured evaluation of repositories based on technical, domain, and display features. GitHub emerges as the clear leader due to its user-friendly interface, attractive design, and multiple accessibility features, while Harvard Dataverse and Zenodo score the lowest, primarily due to their basic interfaces and limited universal appeal. However, the reliance on uniform scoring criteria, such as assigning equal weight to all parameters, may oversimplify the assessment. Additionally, the pie chart and bar graph methodologies, while mathematically precise, are described in a way that

could clear readers familiar with such calculations. Future evaluations should provide a nuanced analysis of feature significance, consider a weighted scoring system, and offer insights into improving low-ranking repositories to ensure equitable usability across platforms.

Conclusion

The open educational repositories have been able to fulfill their objectives of bringing cohesiveness to the universal academic community by offering authentic, shareable, and high-quality resources to academicians worldwide. Ever since the open science movement, these resources have been amplified on the Internet and continue to provide trusted and diverse resources to those who seek knowledge. From the detailed study of the repositories, we gathered data about the general and technical aspects that have been represented through the tables. A number of facts have been highlighted through this, and we have analyzed them through charts and graphs. The selected repositories have more or less similar technical parameters such as accessibility, default license, download facilities, etc. In these cases, the attractiveness and ease of use of the repositories can make a huge difference. GitHub provides an interactive user interface and a website with lots of animation and click-ons that draw users' attention on the first attempt. This is an important factor these days, considering the ever-increasing number of open resource repositories and the cut-throat competition in contemporary times.

References

Atenas, J., & Havemann, L. (2014). Questions of quality in repositories of open educational resources: A literature review. *Research in Learning Technology*, *22*, 20889.

Clements, K., Pawlowski, J., & Manouselis, N. (2015). Open educational resources repositories literature review–Towards a comprehensive quality approaches framework. *Computers in Human Behavior*, *51*, 1098-1106.

Colvard, N. B., Watson, C. E., & Park, H. (2018). The impact of open educational resources on various student success metrics. *International Journal of Teaching and Learning in Higher Education*, 30(2), 262-276.

D'Antoni, S. (2009). Open educational resources: Reviewing initiatives and issues. *Open Learning: The Journal of Open, Distance and E-Learning*, 24(1), 3-10.

Ehlers, U. D. (2011). Extending the territory: From open educational resources to open educational practices. *Journal of Open, Flexible and Distance Learning*, 15(2), 1-10.

Hylén, J. (2021). Open educational resources: Opportunities and challenges.

In *OECD's Centre for Educational Research and Innovation*. https://docs.prosentient.com.au/ prosentientjspui/bitstream/10137/17756/1/interpublish41675.pdf

McKerlich, R., Ives, C., & McGreal, R. (2013). Measuring use and creation of open educational resources in higher education. *International Review of Research in Open and Distributed Learning*, 14(4), 90-103.

Mirowski, P. (2018). The future (s) of open science. *Social Studies of Science*, 48(2), 171-203.

Mishra, S. (2017). Open educational resources: Removing barriers from within. *Distance Education*, *38*(3), 369-380.

Santos-Hermosa, G., Ferran-Ferrer, N., & Abadal, E. (2017). Repositories of open educational resources: An assessment of reuse and educational aspects. *International Review of Research in Open and Distributed Learning: IRRODL*, 18(5), 84-120.

Tuomi, I. (2013). Open educational resources and the transformation of education. *European Journal of Education*, 48(1), 58-78.

UNESCO. (2024a). *Open educational resources*. https://www.unesco.org/en/communication-information/open solutions/open educational-resources

UNESCO. (2024b). Guidelines for OER. https://en.unesco.org

Yuan, L., MacNeill, S., & Kraan, W.G. (2008). Open Educational Resources-Opportunities and challenges for higher education.